**Universidade de Aveiro** Departamento de Biologia
**Ano** 2015

**JOANA MANUEL GONÇALVES TEIXEIRA COUTO**

# ANÁLISE TRANSCRIPTÓMICA DE GLÂNDULAS SALIVARES DE *ANOPHELES STEPHENSI* DURANTE A INFEÇÃO POR *PLASMODIUM BERGHEI*

# TRANSCRIPTOMIC ANALYSIS OF *ANOPHELES STEPHENSI* SALIVARY GLANDS DURING THE INFECTION WITH *PLASMODIUM BERGHEI*

# DECLARAÇÃO

Declaro que este relatório é integralmente da minha autoria, estando devidamente referenciadas as fontes e obras consultadas, bem como identificadas de modo claro as citações dessas obras. Não contém, por isso, qualquer tipo de plágio quer de textos publicados, qualquer que seja o meio dessa publicação, incluindo meios eletrónicos, quer de trabalhos académicos.

**Universidade de Aveiro** Departamento de Biologia
**Ano 2015**

**JOANA MANUEL GONÇALVES TEIXEIRA COUTO**

# ANÁLISE TRANSCRIPTÓMICA DE GLÂNDULAS SALIVARES DE *ANOPHELES STEPHENSI* DURANTE A INFEÇÃO POR *PLASMODIUM BERGHEI*

# TRANSCRIPTOMIC ANALYSIS OF *ANOPHELES STEPHENSI* SALIVARY GLANDS DURING THE INFECTION WITH *PLASMODIUM BERGHEI*

Dissertação apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Biologia Molecular e Celular, realizada sob a orientação científica da Doutora Ana Domingos, Investigadora do Instituto de Higiene e Medicina Tropical da Universidade Nova de Lisboa (IHMT-UNL), e co-orientação da Doutora Maria Lourdes Pereira, Professora Associada com Agregação do Departamento de Biologia da Universidade de Aveiro.

Dedico este trabalho à minha avó, a minha guerreira

**o júri**

| | |
|---|---|
| presidente | Prof. Doutora Helena Silva<br>professora auxiliar do Departamento de Biologia da Universidade de Aveiro |
| vogal- arguente | Doutor Bruno José Fernandes Oliveira Manadas<br>investigador auxiliar no Centro de Neurociências e Biologia Celular da Universidade de Coimbra |
| vogal-orientadora | Doutora Ana Isabel Amaro Gonçalves Domingos<br>investigadora auxiliar do Instituto de Higiene e Medicina Tropical da Universidade Nova de Lisboa |

**agradecimentos**

Os meus agradecimentos são dirigidos à orientadora Doutora Ana Domingos, pela motivação, conhecimento e auxílio em tempos críticos quer a nível profissional quer pessoal. Também agradeço ao Renato Silva, pela disponibilidade e conhecimentos transmitidos. Às colegas de laboratório, Sandra Antunes, Joana Ferrolho e Sara Pardal pelo apoio, acolhimento, saber e disponibilidade quer laboratorial quer pessoal. Todos foram, à sua maneira, cruciais para me tornar mais crítica e independente.

Agradeço o apoio e disponibilidade da equipa do Biotério do Instituto de Higiene e Medicina Tropical; e à Joana Gomes pelo apoio e transmissão de conhecimentos relacionados com os insectos deste estudo e de um modo geral.

Agradeço de um modo geral a todos que, directa ou indirectamente, me apoiaram na execução deste trabalho.

Por fim, agradeço à minha família por todo o apoio incondicional, motivação, e por serem o meu alicerce em momentos bons e menos bons. Um sincero obrigado.

Ao meu João Pedro, por ser o meu porto de abrigo em todas as horas, por me fazer sorrir e amar por amar.

**palavras-chave**

malária, *Anopheles stephensi*, glândulas salivares, RNAseq, transcriptómica, qPCR, RNAi

**resumo**

A malária continua a ser a principal causa de morbilidade e mortalidade nas regiões tropicais e subtropicais, contribuindo para o surgimento de 198 milhões de casos clínicos no ano de 2013. O mosquito *Anopheles stephensi* é um dos vectores de malária mais prevalentes na região asiática, tendo sido recentemente implicado no ressurgimento de malária em Djibouti.
Através de técnicas como sequenciação de RNA, genes diferenciadamente expressos nas glândulas salivares deste mosquito em resposta à infecção por *Plasmodium berghei* foram identificados. Alguns destes genes podem ser selecionados para avaliar a sua potencialidade como alvos para bloqueio da transmissão da malária.
Entre os genes com expressão diferencial resultante da análise dos resultados de RNA-seq e confirmação por qPCR, um gene relacionado com o transporte de $Cl^-$ e $HCO_3^{2-}$, *prestin*, estava sobrexpresso após infeção com *P. berghei*. Este gene tem um papel crucial na invasão do parasita no intestino médio e na optimização do meio em que o parasita se desenvolve. Por esse motivo, o silenciamento deste transcrito foi efectuado para averiguar o papel funcional nas glândulas salivares.
O silenciamento de genes utilizando a técnica de RNA de interferência permite inferir sobre o seu papel ou função num dado processo metabólico ou fisiológico. Após o silenciamento do gene *prestin*, o número de mosquitos viáveis apresentou um decréscimo significativo em comparação com o controlo (*β2M*). Também houve uma queda significativa entre o número de mosquitos antes da injeção e no último dia após injecção. O número de esporozoítos em geral não foi afectado pelo silenciamento da prestin quando comparado com o controlo.
Para esclarecer resultados obtidos durante o estudo, tais como a influência do silenciamento da *prestin* na sobrevivência dos mosquitos e a presença e número de esporozoítos na glândulas salivares, seria fundamental realizar ensaios de qPCR para determinar a expressão diferencial deste gene após silenciamento. Além disso, será também importante analisar a expressão diferencial do *off-target* (ASTE006714) após silenciamento da *prestin*, uma vez que a semelhança da sequência entre este e a *prestin* é elevada.

**keywords**

malaria, *Anopheles stephensi*, salivary glands, RNAseq, transcriptomics, qPCR, RNAi

**abstract**

Malaria remains the leading cause of morbidity and mortality in tropical and subtropical regions, contributing to the emergence of 198 million clinical cases in 2013. The mosquito *Anopheles stephensi* is one of the most prevalent malaria vectors in the Asian region having recently been implicated in malaria resurgence in Djibouti.

Using techniques as RNA sequencing, differentially expressed genes in the salivary glands of the mosquito in response to infection by *Plasmodium berghei* were identified. Some of these genes can be selected to evaluate their potential as targets for malaria transmission blocking.

Among the genes with differentially expression resulting from the analysis of RNA-seq results and confirmation by qPCR, a gene related transport of $Cl^-$ and $HCO_3^{2-}$, *prestin*, was upregulated after infection with *P. berghei*. This gene plays a crucial role in parasite invasion in the midgut and the optimization of the environment in which the parasite develops. For this reason, the silencing of this transcript was made to evaluate the function of prestin in salivary glands. The gene silencing, using RNA interference technique, allow inferring about the role or function of *prestin* gene in a particular metabolic or physiological process. After *prestin* gene silencing, the number of viable mosquitoes had a significant decrease in comparison with the control (*ß2M*). There was also a significant decrease in the number of mosquitoes before injection and at the last day after injection. The number of sporozoites were not generally affected by silencing of *prestin* when compared with the control.

To clarify other results obtained during the study, as the influence of the silencing of *prestin* in the survival of mosquitoes and the presence and number of sporozoites in the salivary glands, will be essential to perform qPCR to determine differential expression of this gene after silencing. Furthermore, it is also important to examine differential expression of off-target (ASTE006714) after silencing *prestin*, since the sequence of this gene have a high percentage of identity with *prestin*.

universidade de aveiro
departamento de biologia
debio

INSTITUTO DE HIGIENE E MEDICINA TROPICAL
DESDE 1902

# Análise transcriptómica de glândulas salivares de *Anopheles stephensi* durante a infeção por *Plasmodium berghei*

# Transcriptomic analysis of *Anopheles stephensi* salivary glands during the infection with *Plasmodium berghei*

**Joana Manuel Gonçalves Teixeira Couto**

Junho, 2015

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| Abs | Absorbance |
| ACTs | Artemisinin-based combination therapies |
| AIDS | Acquired immune deficiency syndrome |
| AL | Artemether-lumefantrine |
| AM | Artemether, |
| AMA | Apical membrane protein |
| AQ | Amodiaquine |
| ART | Artemisinin |
| AS0 | Adjuvant system |
| AS | Artesunate |
| BCV | Biological Coefficient of Variation |
| bHLH | Basic helix-loop-helix |
| BLAST | Basic Local Alignment Search Tool |
| BLASTX | Basic Local Alignment Search Tool for proteins using a translated nucleotide query |
| bp | Base pairs |
| C | carbamates |
| CD1 | CD-1 mouse |
| CDC | Centers for Disease Control and Prevention |
| cDNA | Complementary deoxyribonucleic acid |
| cGMP | Cyclic guanosine monophosphate |
| $ChSO_4$ | Chondroitin sulfate |
| $Cl^-$ | Chlorine ion |
| CL | Clindamycline |
| CLIP | Clip domain serine proteases |
| CQ | Chloroquine |
| CSA | Chondroitin sulfate A |
| $C_t$ | Cycle threshold |
| D | Doxycycline |
| DAP | Drug-arrested parasites |
| DDT | Dichlorodiphenyltrichloroethane |
| DGAV | Direção-Geral de Alimentação e Veterinária |
| DHA | Dihydroartemisinin |
| DHFR | Dihydrofolate reductase |
| DNA | Deoxyribonucleic acid |
| dNTP | Deoxynucleotide |
| *ds*RNA | Double-stranded RNA |
| EBA | Erythrocyte-binding antigen |
| EC | Enzyme Commission |
| ELF | Elongation factor |
| Enzyme DB | Enzyme Database |
| EXFL | Exflagellation |

| | |
|---|---|
| FPKM | Fragments *per* Kilobase of transcript *per* Million mapped reads |
| G6PD | Glucose-6-phosphate dehydrogenase |
| GAP | Genetically attenuated parasites |
| GFP | Green fluorescent protein |
| GLURP | Glutamate-rich protein |
| GO term | Gene Ontology term |
| GO | Gene Ontology database |
| SG C | Salivary glands Control |
| SG I | Salivary glands Infected |
| HBsAg | Hepatitis B surface antigen |
| $HCO^{3-}$ | Bicarbonate |
| HDP | *Plasmodium* heme detoxification protein |
| HIV | Human immunodeficiency virus |
| HMGB1 | High mobility group box 1 |
| HRP2 | Histidine-rich protein-2 |
| HSP | Heat-shock proteins |
| Imd | Drosophila immune deficiency |
| IRS | Indoor residual spraying |
| ITN | Insecticide-treated nets |
| KD | Knockdown |
| kdr | Knockdown resistance |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| LAMP | Loop-mediated isothermal amplification |
| LAP | Live attenuated parasites |
| lincRNA | Long intergenic non-coding RNA |
| LRR | Leucine-rich repeat |
| LS | Liver stage |
| microRNA | Small non-coding RNA |
| miRNA | microRNA |
| MQ | Mefloquine |
| mRNA | Messenger RNA |
| MSP | Merozoite surface protein |
| NCBI | National Center for Biotechnology Information |
| ND | NanoDrop |
| NDAE1 | $Na^{+}$ driven anion exchanger protein |
| NGS | Next-generation sequencing |
| NQ | Naphroquine |
| OC | Organochlorines |
| oligo(dT) | Short sequence of deoxy-thymine nucleotides |
| OP | Organophosphates |
| OPRT | Orotate phosphoribosyltransferases |
| ORF | Open reading frame |
| PBM | Post-blood meal |
| PBS | Phosphate buffered saline |

| | |
|---|---|
| PCR | Polymerase chain reaction |
| PDH | Pyruvate dehydrogenase |
| Pf HRP2 | *P. falciparum* histidine-rich protein-2 |
| PfEMP1 | *Plasmodium falciparum* erythrocyte membrane protein 1 |
| PfRH5 | *Plasmodium falciparum* reticulocyte-binding protein homologue 5 |
| *Pf*s | *Plasmodium falciparum* surface |
| PfSPZ | *Plasmodium* falciparium sporozoites |
| PG | Proguanil |
| PKG | cGMP-dependent protein kinase |
| pLDH | Plasmodial lactate dehydrogenase |
| polyA | Polyadenylated |
| PP | Percentage of parasitemia |
| PPQ | Piperaquine |
| PQ | Primaquine |
| *Pv*s | *Plasmodium vivax* surface |
| PYR | Pyronaridine |
| QN | Quinine |
| qPCR | Quantitative real-time polymerase chain reaction |
| qRT-PCR | Real time reverse transcription polymerase chain reaction |
| RAS | Radiation-attenuated sporozoites |
| RDT | Rapid diagnostic test |
| RefSeq | Reference Sequence database |
| RNA | Ribonucleic acid |
| RNAi | RNA interference |
| RNase | Ribonuclease |
| RNA-seq | RNA Sequencing |
| RPS7 | Ribosomal protein S7 |
| RT-PCR | Reverse transcription polymerase chain reaction |
| RTS,S | Recombinant subunit vaccine |
| RTT | Replication, transcription and translation |
| SAGE | Serial analysis of gene expression |
| SBS | Sequencing by synthesis |
| SG | Salivary glands |
| shRNA | Small hairpin RNA |
| siRNA | Small interfering RNA |
| SIT | Sterile insect technique |
| SLC | Solute carrier family |
| SM1 | Salivary gland- and midgut- binding peptide 1 |
| SNP | Single nucleotide polymorphism |
| $SO_4^2$ | Sulfate ion |
| SOLiD | Sequencing by Oligonucleotide Ligation and Detection |
| SP | Sulphadoxine-pyrimethamine |
| SPZ | Sporozoites |
| SRPN | Serpin protein |

| | |
|---|---|
| *ss*RNA | Single-stranded RNA |
| STRING | Search Tool for the Retrieval of Interacting Genes/Proteins |
| T | Tetracycline |
| TBE | Tris/Borate/EDTA |
| TEP1 | Telomerase protein component 1 |
| TOR | Target of rapamycin |
| tRNA | Transfer RNA |
| UniGene | Expressed pseudogene |
| UniParc | UniProt Archive |
| UniProt | Universal Protein Resource |
| UniRef | UniProt Reference Clusters |
| URL | Uniform Resource Locator |
| VAR2CSA | Variant Surface antigen 2-CSA |
| WHO | World Health Organization |
| $\beta_2$M | Beta-2 microglobulin |

**Nucleotide Bases**

| | |
|---|---|
| A | Adenine |
| C | Cytosine |
| G | Guanine |
| T | Thymine |
| M | A or C |
| N | Any nucleotide (A, C, G or T) |
| R | Purine (A or G) |
| Y | Pyrimidine (C or T) |
| K | G or T |
| W | A or T |
| D | A, G or T |

# Chapter 1: Introduction

Malaria is a public health problem with high biological complexity and social impact[1]. According to the WHO (2014), it was estimated that 198 million cases of malaria and 584 000 deaths occurred in 2013[2]. It is a disease caused by the bite of *Anopheles* mosquitoes infected with a parasite of the genus *Plasmodium*[3]. Vector control is one of the most common techniques used to reduce malaria[1].

*Anopheles gambie* is the major malaria vector in Africa which has been the subject of numerous research projects and its genome and transcriptome have been already published[4]. *Anopheles stephensi* which is more prevalent in the Asian region has recently been implicated with the malaria resurgence in Djibouti (Horn of Africa)[5], reinforcing the importance to increase the study about this vector, that have had less attention comparing with *A. gambiae*. Based on this resurgence, new transmission blocking targets are required.

The genomics and transcriptomics age provide an interesting opportunity to uncover the molecular mechanisms for the identification of new targets that can reduce the spread of the disease. Identification of differential expressed genes using techniques such as subtractive hybridization, microarrays and RNA sequencing allow selecting potential targets for blocking the transmission of malaria[6].

In view of the impact and life cycle of the vector *Anopheles stephensi* and parasite *Plasmodium berghei,* transcriptome analysis of salivary glands (SG) may be a good approach for the discovery of new targets, since the *Plasmodium* stage that transmits malaria is present in this tissue[7]. To infer about a gene and its importance or function, gene knockdown can be accomplished, reducing the expression of one or more genes from an organism. This can be obtained through genetic modifications or by injection of short DNA or RNA oligonucleotide that align to either a target gene or a mRNA transcript [8].

This work aims to identify differential expressed genes in salivary glands of *Anopheles stephensi* mosquitoes in response to infection by *Plasmodium berghei* by analyzing data obtained from direct RNA sequencing, and confirm gene expression by real-time PCR. Another purpose of this work was to evaluate the importance/function of selected gene(s) in the infection by *P. berghei*, using RNA interference.

These studies will allow a better understanding of the vector- pathogenic agent interface contributing to the development of new methods to control the pathogens transmission by arthropods, including new antigens or potential vaccine candidates.

## 1.1. Malaria

Parasitology is one of the most fascinating branches of biological sciences but, in spite of the great findings that have been achieved in this field, some essential questions about the biology of the interactions between parasites, vectors and hosts remain unclear [9]. Malaria is caused by a parasite of the genus *Plasmodium*, that is transmitted through the bite of mosquitoes from the *Anopheles* genus [3,10]. Early this disease exerted influence on the life of Humanity[11] and, about 10 000 years ago, malaria had a great impact on human survival coinciding with the beginning of agriculture in the Neolithic revolution [12,13].

Malaria has always been associated with sickness and fever[14,15]. Scientific advances on this disease only became possible after the discovery, by Charles Laveran, of the interaction between microorganisms such as protozoan parasites and the human erythrocytes, in 1880[14].

A few years later mosquitoes were identified as disease transmitters or vectors: initially, in the case of avian malaria, caused by *Plasmodium relictum* described by Ronald Ross in 1897; and later for human malaria described by Giovanni Battista and co-workers[14].

Due to the climate of tropical and subtropical areas, *Anopheles* mosquitoes can easily thrive, providing suitable conditions for malaria parasites to complete their growth cycle [16].

Since the year 2000, malaria incidence and mortality rates have fallen down about 30% and 47% respectively. However, it remains significant with 198 million cases and 584 000 deaths caused by malaria in 2013[17].

The vast majority of cases and deaths are from sub-Saharan Africa and South Asia. However, Asia, Latin America, the Middle East and parts of Europe are also affected. According to WHO (2014) and CDC (2012) [2,18], 97 countries and territories still had ongoing malaria transmission (Figure 1).

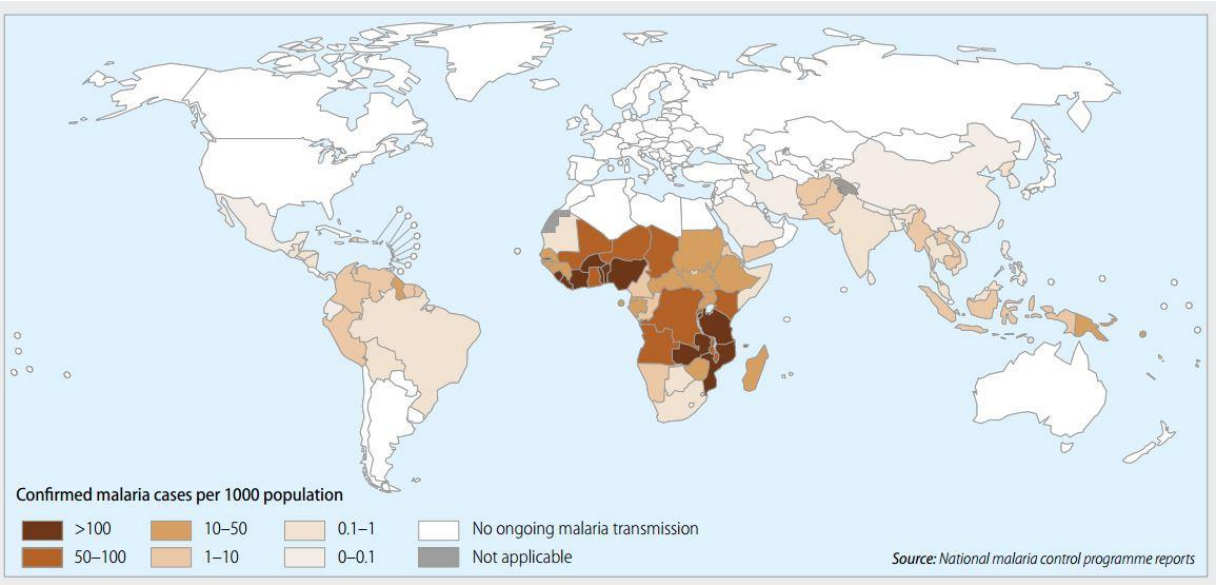



**Figure 1 - Countries with ongoing transmission of malaria, in 2013.** The majority of cases are from tropical and sub-tropical regions[17]**.**

Besides malaria impact in several mammals, reptiles and birds[19]; in humans, some population groups are at higher risk of contracting malaria, and developing severe disease, than others. These include children younger than five years of age, pregnant women, patients with HIV/AIDS, as well as non-immune migrants, mobile populations and travelers[17]. National malaria control programs need to take special measures to protect these population groups from malaria infection, taking into consideration their specificity[17]. In high-transmission areas of the world, 85% of malaria deaths correspond to children less than five years of age, making this a significant vulnerable group[17]. Malaria in pregnant women increases the risk of maternal and fetal anaemia, stillbirth, spontaneous abortion, low birth weight and neonatal death[17]. Infants born from mothers living in endemic areas are vulnerable to malaria from approximately three months of age, when immunity acquired from the mother starts to decline[17].

Furthermore, co-infection and interaction between HIV and malaria have major public health implications. HIV infection increases the risk of malaria infection, severe malaria and death, aggravating clinical status of AIDS[17]. Traveling population groups or migrants lack partial immunity to malaria, and have restriction in prevention, diagnostic and treatment services[17].

## 1.2. *Plasmodium* spp.

### 1.2.1. Taxonomy

Protozoa are unicellular organisms (or acellular) (5-50µm) in which a single cell is able to perform all vital functions. Some diseases caused by protozoa, especially malaria and sleeping sickness, are well known as an important cause of morbidity and mortality in the tropics, and the discovery of etiologic and transmission mechanisms are considered as classical contributions to the history of human medicine[11].

Malaria is caused by protists of phylum Apicomplexa, in which are common to have an organ called apicoplast that allows the entry into the host cell[20]. *Plasmodium* belongs to the class Aconoidasida (although the parasite (*Plasmodium*) lacks a conoid structure, a set of microtubules) that includes the order Haemosporidia, containing the parasites that invade red blood cells[20]. Ciliates, like *Plasmodium*, share several features, including a vertebrate host asexual and sexual reproduction in a definitive host (a mosquito, that infected by the species of *Plasmodium*, infect several mammals, including humans) and that is why they are in the family Plasmodiidae[20, 21, 22].

### 1.2.2. Species

In addition to the human population, *Plasmodium spp.* can also infect other mammals, reptiles and birds. Four species of *Plasmodium* were identified as causing malaria, in humans: *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae*. A fifth, *P. knowlesi*, infects non-human primates, but recently has been found to infect and cause disease in humans [17,21–23].

*P. falciparum* has been the target of most vaccines as this parasite is responsible for the most severe malaria disease and causes many deaths [2,19]. *P. berghei*, does not affect humans but has been used in many studies in rodent animal models to improve the understanding on disease and also to test new anti-malaria drugs and vaccines[24].

## 1.3. *Anopheles* spp.

A vector is any living being capable of transmitting an infecting (or not) agent, actively or passively. The active transmission occurs when the vector is infected and then infects other species. Passive way occurs when the vector is not infected by infectious agents, but causes the infection of other species, which is the case of malaria.

With this in mind, it is important to deeply understand the taxonomy, species, anatomy and life cycle of the malaria vector.

### 1.3.1. Species

Malaria is caused by the bite of a female mosquito from *Anopheles* genus. This genus is part of the phylum Arthropoda which has a hard exoskeleton and several pairs of jointed appendices[25]. The genus *Anopheles* belongs to the class Insecta, the family Culicidae, and the order Diptera characterized by having a pair of wings and a pair of dumbbells [25].

Of 460 described species of *Anopheles*, 60 are able to ensure the transmission of malaria and this number is constantly increasing (Figure 2).

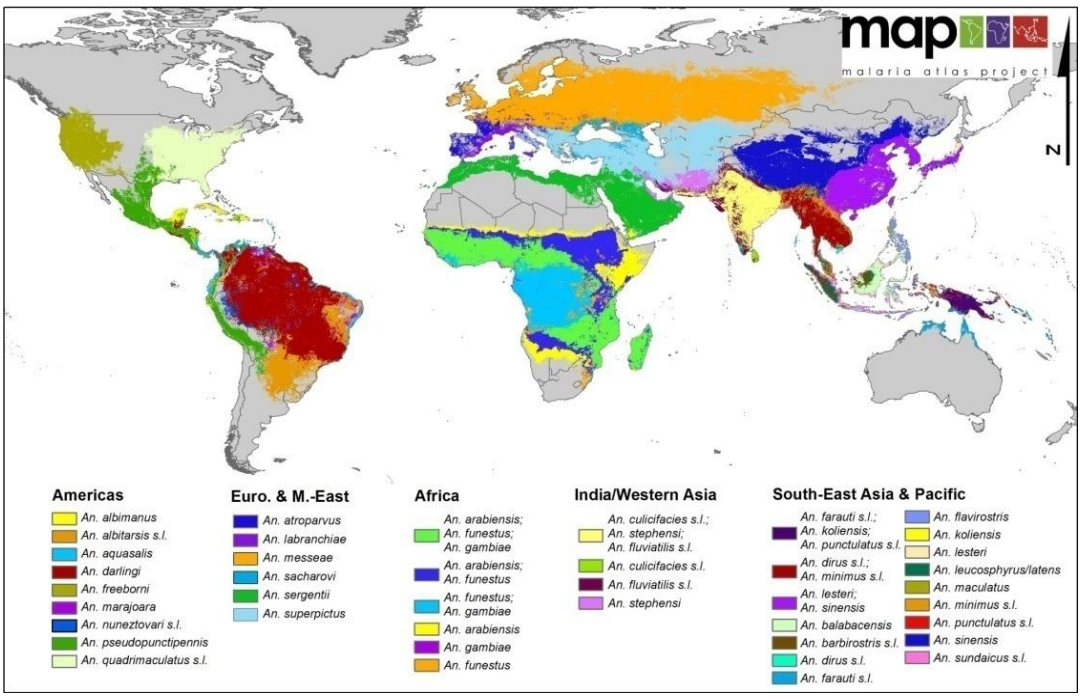



**Figure 2 - Map of global distribution of the dominant vector species of malaria[26].**

Mosquitoes of the genus *Anopheles* are very efficient malaria vectors, and a great diversity of ecosystems in Africa are favorable for their presence and development[27,28]. According to Figure 2, *A. stephensi* is found throughout the Indian subcontinent[26]. Recently, it distribuition extends from the Arabian Peninsula, through Iran and Iraq, to the Bangladesh, southern China, Myanmar and Thailand[29]. *A. stephensi* is an important vector for both *P. falciparum* and *P. vivax*, the most virulent parasites associated with this disease [28].

### 1.3.2. Life cycle and anatomy

Like some other mosquitoes, *Anopheles* goes through four stages in their life cycle: egg, larva, pupa and imago (Figure 3). Imago is the only stage where the insect reaches sexual maturity and, if it is winged; only at this stage wings will be functional. Typically, this life cycle phase is usually referred to as adult stage. The first three stages develop from water and last 5-14 days depending on the species and ambient temperature. The female *Anopheles* mosquito in the adult stage that acts as malaria vector, can live for up to one month or more in captivity, but not longer than two weeks, in nature [30].

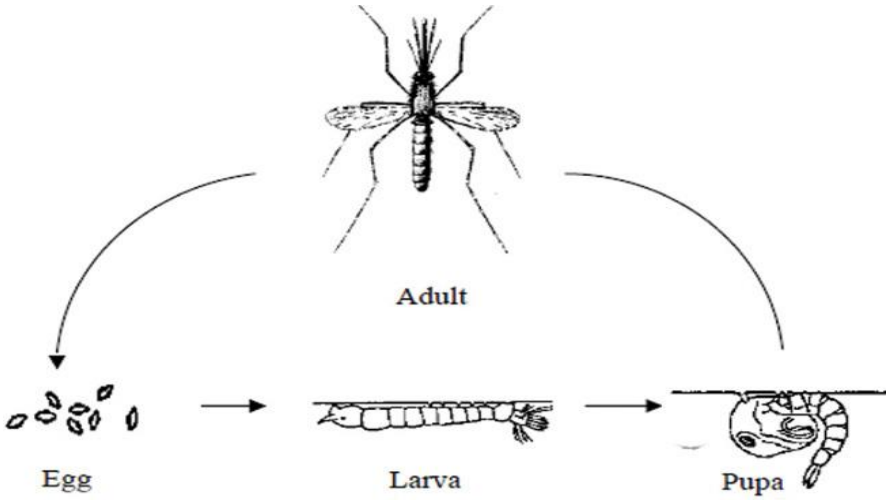



**Figure 3 - Life cycle of *Anopheles* mosquitoes.**
**(**http://labspace.open.ac.uk/file.php/6718/!via/oucontent/course/1151/comms_dis_session5_fig6.jpg)

Adult females lay between 50-200 eggs "*per* oviposition" on the water since their eggs have the ability to float and, after 2-3 days they hatch to larvae. Mosquito larvae have a

well-developed head, a mouth with brushes, a large thorax and a segmented abdomen. This state is characteristic for the absence of legs. In contrast to other mosquitoes, *Anopheles* larvae do not have a breathing siphon (air tube) leading to a parallel positioned body in to the water surface (Figure 4). The larvae feeds on algae, bacteria and other microorganisms, that can be found in different water sources as the surface of fresh or salt water, water agricultural-related areas, rivers, and small pools of temporary rain. Many species prefer habitats with vegetation, and others habitats that have none.

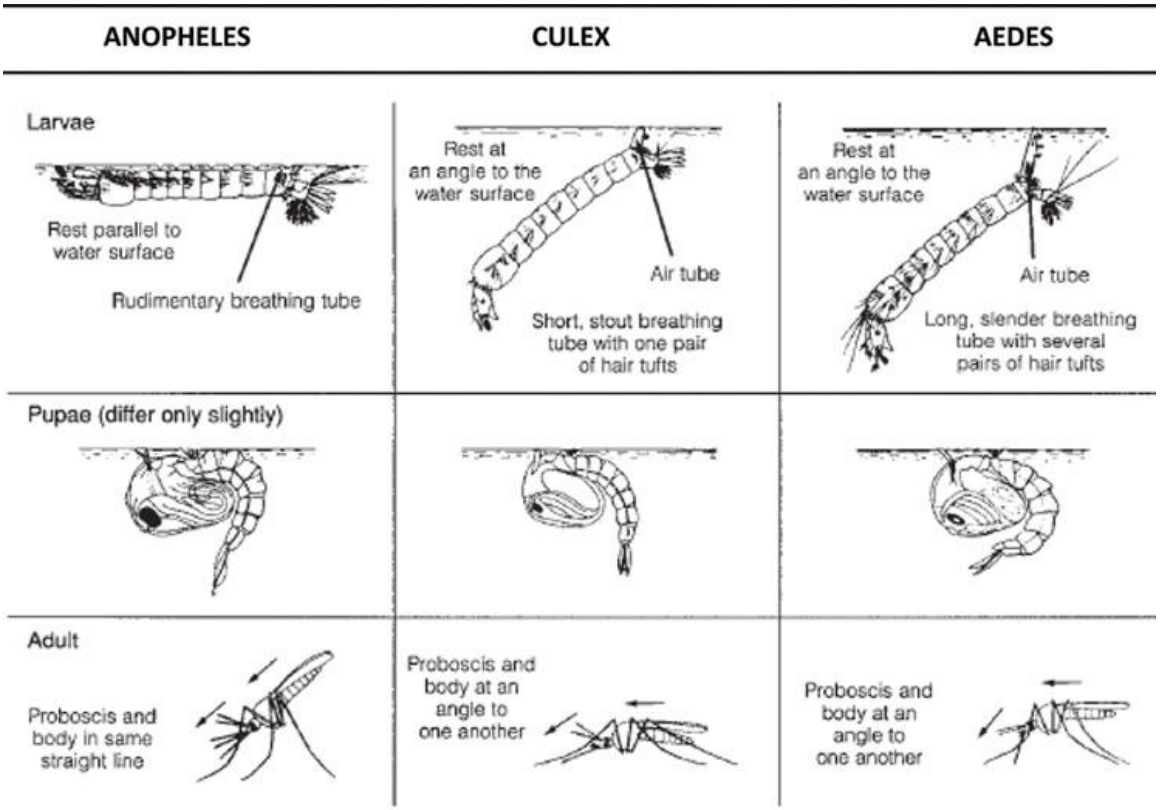



**Figure 4 - Representation of differences in the development of different mosquitoes[31].**

Larvae develop through four stages and after that they metamorphose into pupae. The head and thorax of a pupa are merged into a cephalothorax. In this stage is necessary to come to the surface to breath, with a pair of respiratory trumpets located on the cephalothorax. After a few days as pupae, the dorsal surface of the cephalothorax is divided and adult mosquitoes emerge.

Adult *Anopheles* has slim bodies with three sections: head, thorax and abdomen (Figure 5). The head is specialized for acquiring sensory information and for feeding. It contains the eyes and a pair of antennae, important for the detection of odors from the "host" and odors where breeding females lay eggs[28,30]. The head also has an elongated proboscis used for feeding and two sensory palps[28,30]. The thorax, specialized for locomotion, has three pairs of legs and a pair of wings. Food digestion and egg development occur in the abdomen[28,30]. This part of the segmented body expands noticeably when the female takes a blood meal. The blood is digested to facilitate a source of proteins to the production of eggs, which will gradually fill the abdomen[28,30]. Adult mosquitoes (Figure 5) mating normally occurs few days after emerging from the pupa stage[28,30]. The adult mosquitoes from genus *Anopheles* can be distinguished from other mosquitoes by the palps, the thorax, and the presence of distinct blocks of black and white scales on the wings and by their normal resting position as the abdomens are perpendicular to resting surfaces (Figure 4). Proboscis, abdomen and cercus can differentiate male and female *A. stephensi* mosquitoes [32] (Figure 6).

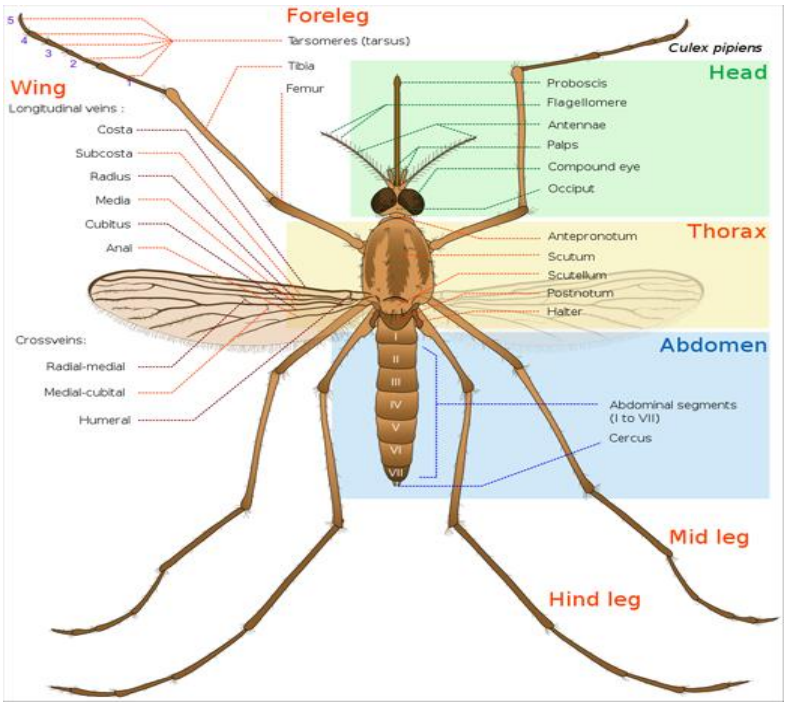



**Figure 5 - Anatomy of an *Anopheles* mosquito.**
http://www.pestcontrolmanagement.org/gfx/anatomy-of-a%20mosquito.png

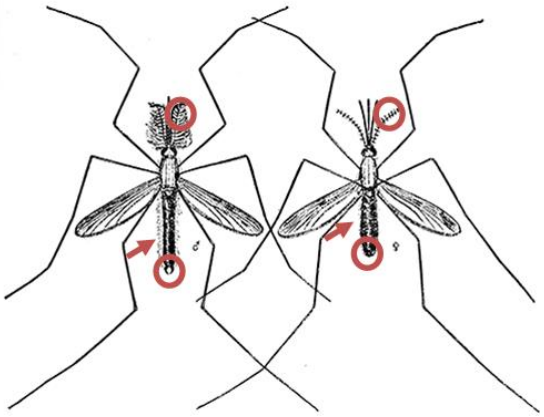


**Figure 6 - Representation of anatomic differences between male and female mosquitoes.** Males have a plumy proboscis and a bifurcated cercus. Females have a large abdomen and a proboscis less plumy. (Adapted by http://www.digilibraries.com/html_ebooks/106843/34279/www.digilibraries.com@34279@34279-h@images@f130.png)

Contrary to males that live about a week, feeding on nectar and other sources of sugar; females can feed on sugar sources for energy, and on a blood meal for the development of the eggs[28,30]. After a blood meal, the female rests for a few days while the eggs develop for 2-3 days (in tropical conditions). Once the eggs are fully developed and layed in a surface, female mosquito resumes search host[28,30]. This cycle of reproduction is repeated until the female dies. Temperature, humidity and the ability to obtain a blood meal while avoiding host defenses are some of the factors that influence survival of mosquitoes [28,30].

### 1.3.3. Life cycle of *Plasmodium* spp.

The life cycle of malaria is particularly complex. Malaria parasites have a heteroxenous life cycle (hetero = different; xenous = host)[33], during which the parasite develops either in a invertebrate definitive host, or in an vertebrate intermediate host. Also these parasites are eurixenous, which means that are found on a large number of host species [34].

*Plasmodium* life cycle begins when hundreds of parasites as an infectious structure, called sporozoites (SPZs), are transmitted by the female *Anopheles* while taking a blood meal in the host [19,23,30]. Females need to take blood meals rich in hemoglobin to carry out the production and enrichment of the eggs [30,35].

The life cycle of malaria disease follows three phases: pré (or exo) - erythrocytic phase, erythrocytic phase and sporogonic phase (Figure 7).

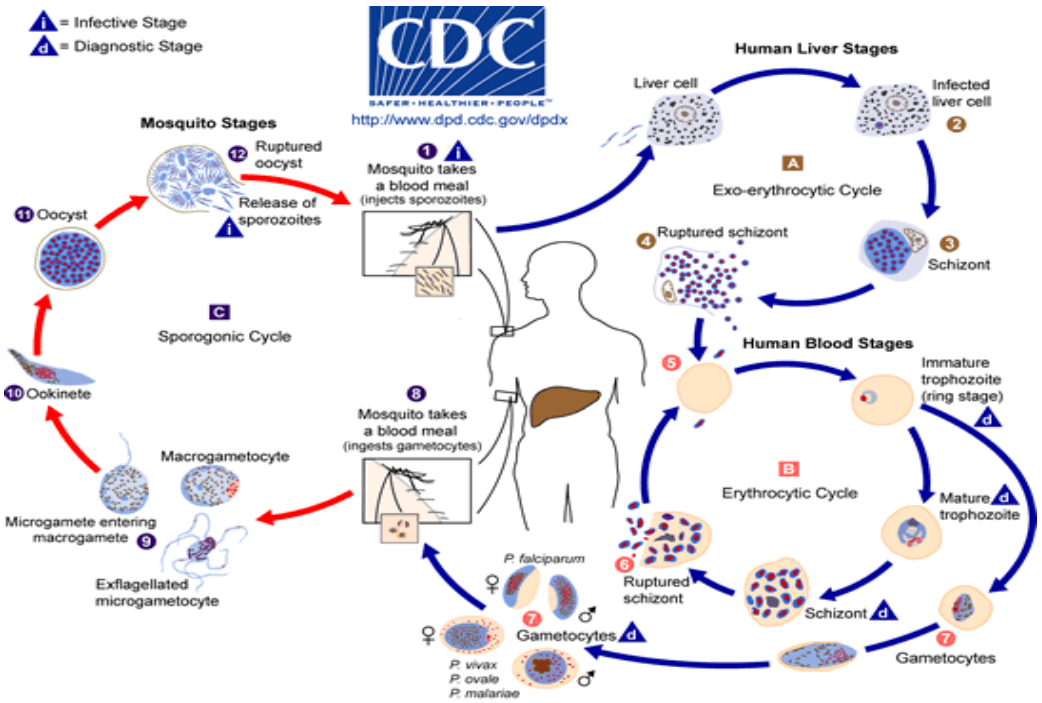



**Figure 7 - Life cycle of malaria parasite in human and in *Anopheles*.** In this picture is represented the sporogonic phase in mosquito-vector; and pré (or exo) - erythrocytic phase and erythrocytic phase in humans[18].

During the pre-erythrocytic cycle, SPZs (Figure 8) invade blood vessels and are transported to the liver, forming the liver stage (LS)[18]. Each SPZ multiplies and grows in a single hepatocyte for 7 days, yielding rounded uninuclear cells known as criptozoites[19]. With the subsequent nuclear division, it forms a multinuclear cell known as schizont. Schizogony is an asexual reproduction process that results in the formation of the schizont, which in turn gives rise to merozoites. The schizogony, which occurs also in hepatocytes, is known as hepatic, pre-erythrocyte or exoerythrocyte schizogony[19]. After 8 to 15 days, the infested liver cells are disrupted, releasing numerous merozoites into the bloodstream[19]. The pre-erythrocytic stage cause no clinical signs or symptoms.

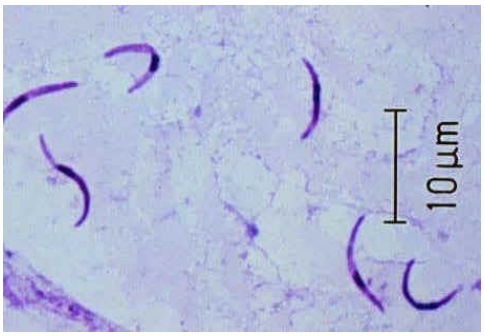



**Figure 8 - Sporozoites of a malaria parasite.** The sporozoites are 10 -15 µm in length and about 1 µm in diameter[36].

The erythrocytic cycle begins when hepatocytes rupture and release numerous merozoites in to the bloodstream [19]. There, this parasite stage develops through several stages, as represented in Figure 9.

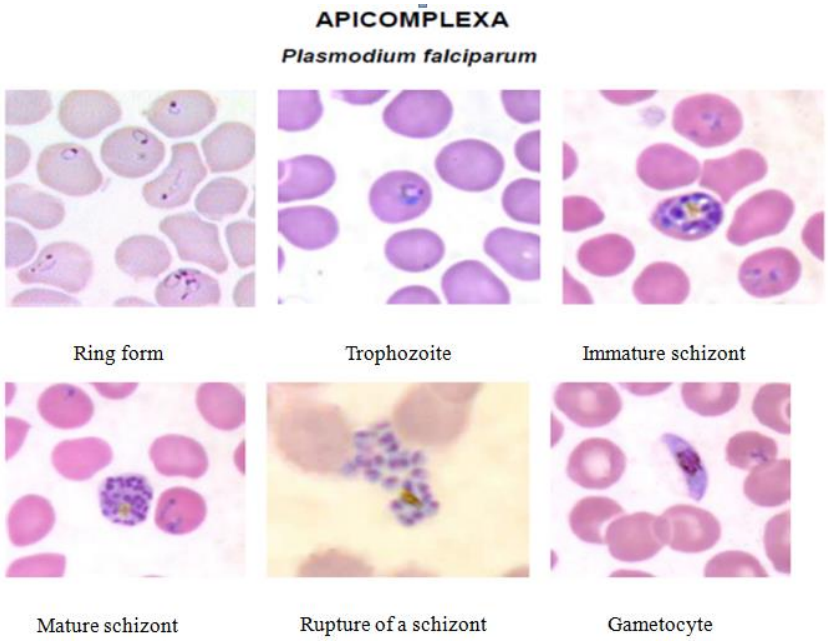



**Figure 9 - The several stages of the malaria parasites.** These stages are sequential since ring form to gametocyte, and they are present in the erythrocytic cycle (Adapted by image from Aurélia Saraiva).

The first intraerythrocytic stage is a ring form followed by a new form called trophozoite. Then, a new schizogony occurs inside the red blood cells (RBCs) where mature erythrocyte schizonts have between 6 and 32 cores[23]. At the end of the erythrocyte schizogony,

erythrocytes lyses and the merozoites are released into the bloodstream, coinciding temporally with the characteristic periodic fever peaks of malaria[23]. Intraerythrocytic stage is diverse between different species of *Plasmodium* (Appendix I). The interval between febrile peaks corresponds to the duration of blood schizogony of each species[23].

The merozoites that invade new red blood cells can turn into trophozoites and later in schizonts or, alternatively, differentiate into sexual forms, male and female gametocytes that can be ingested by the mosquito during the next blood meal[23].

During the sporogonic cycle in the mosquito, parasites go through various developmental stages in the midgut (gametogenesis, fertilization, zygote, ookinete and oocyst) and also in other epithelia as the salivary glands [7] (Figure 10 and 11) taking 2-3 weeks to complete.

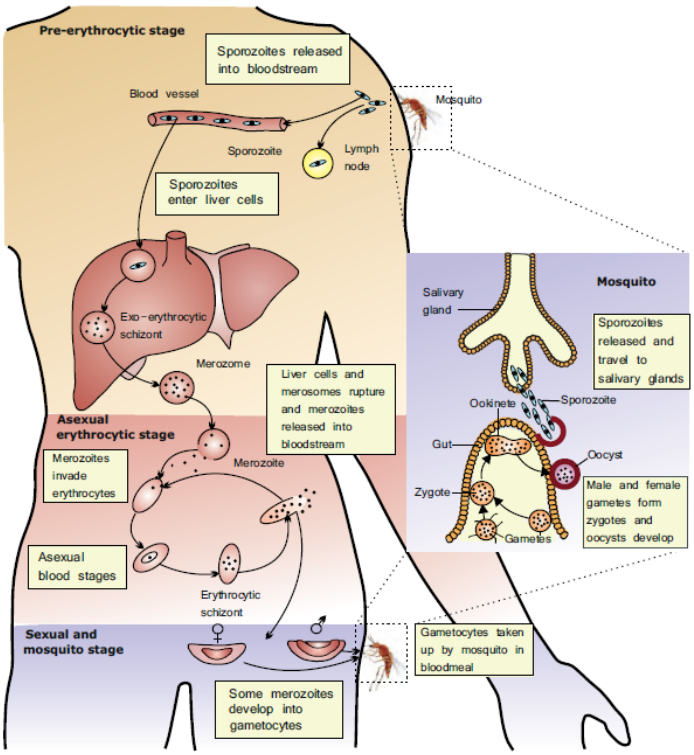



**Figure 10 - *Plasmodium* life cycle showing the three development stages of the parasite.** Since pre-erythrocytic stage, where sporozoites are released into the bloodstream; to erythrocytic stage, where gametocytes can be found; the vertebrate-host can manifest several symptoms[37].

In a few minutes, the male gametocyte undergoes exflagellation releasing 6-8 microgametes by extrusion of nuclear material into peripheral processes resembling flagella[23]. While the female gametocytes turn into macrogametes[23]. As showed in Figure 11, male and female gametes ingested during the blood meal of a female mosquito, fuse with each other, and form a zygote that will initiate infection within the mosquito[23]. The fusion of microgametes and macrogametes form a zygote which is further transformed in a mobile structure named ookinete. To penetrate the mosquito gut wall (midgut), the ookinete becomes an oocyst. This spherical structure accommodates between the epithelium and the basal membrane, and form thousands of sporozoites inside. With the disruption of the oocysts, the sporozoites are released up and then they migrate to the salivary glands. Salivary glands have a high impact, since each pair of coupling gametes will form 1,000 SPZ's [23] and in each blood meal numerous sporozoites will be injected into the bloodstream of the vertebrate host.

In some occasions, during a blood meal, only about 10% of the thousands of gametocytes that are ingested, develop successfully in ookinetes and about five of them succeed in invading the midgut epithelium to form sessile oocysts [38]. However, a dramatic amplification of the number of parasites are followed when each oocyst released thousands of sporozoites[39].

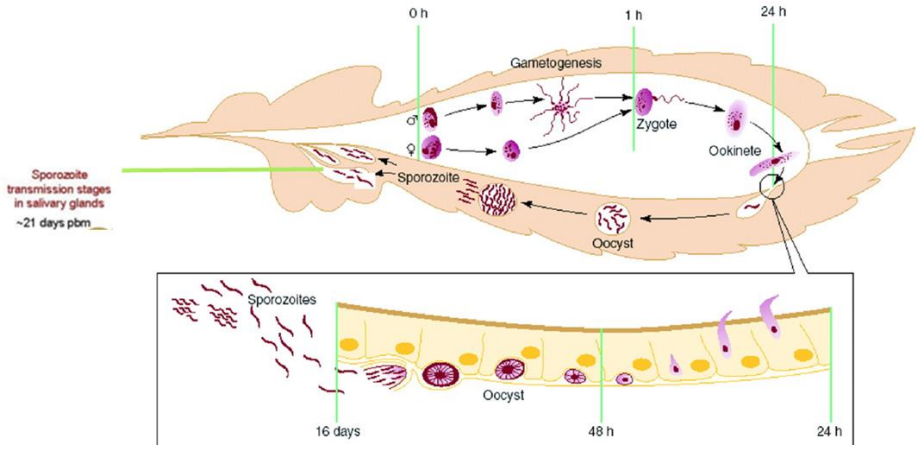



**Figure 11 - Scheme of time duration of each stage of the sporogonic cycle.** After a blood meal the gametes fuse and form a zygote. After one day the ookinete will appear and after ten days an oocyst will be formed. In the 18st day the sporozoites will be released up and they will migrate to the salivary glands[40].

### 1.3.4. Transmission efficiency

As mentioned before, the successful development of the malaria parasite in the mosquito depends on several factors, such as environmental temperature, humidity and feeding.

Host susceptibility is greatly influenced by its immunity, manly in adults living in areas where the transmission conditions are moderate or intense[41]. Years of exposure contribute for the development of partial immunity through the ability of the immune system to generate a memory cells [41]. The host is never completely protected but this response is able to reduce the risk of malaria infection. For this reason, most of Africa's malaria deaths occur in young children that have low immunity since have lower exposure to malaria infection[2].

### 1.3.5. Symptomatology and pathology

Seven days or more (usually 10-15 days) after the bite of infective mosquito, a malaria infection presents an acute febrile illness in a non-immune individual[2,3]. The symptoms such headache, muscle pains, fatigue, fever, chills, nausea, vomiting, can be trivial and hard to recognize as being associated with malaria [2,3]. Besides, involvement of multiple organs in malaria infection is frequent, contributing to severe anaemia, respiratory failure due to metabolic acidosis, or cerebral malaria[2,3]. In addition, a malaria infection caused by *P. falciparum* can progress to a serious illness, often leading to death, if not treated within 24 hours[2,3].

The development of a severe disease depends on the parasite species that causes the infection and the immune status of the infected host. Partial immunity can be developed by people that resides in malaria-endemic areas, allowing asymptomatic infections to occur[2,23]. Sometimes relapses may occur weeks or months after the first infection, due to latent liver forms known as hypnozoites[2,23]. These cases of re-infection are related to parasites species such as *P. vivax* and *P. ovale*; being absent in *P. falciparum* and *P. malariae*. In this situation treatment towards the liver phase is required for a complete cure [2].

### 1.3.6. Malaria diagnosis, treatment and vector control

Factors such as diversity of vectors species and parasites strains, intensity of parasite transmission, and the age of the patient can contribute to several clinical conditions, ranging from asymptomatic malaria, mild uncomplicated disease to life-threatening severe disease[42]. In this context, approaches to regard diagnosis, therapy, vaccination and vector control must be multiple [42].

The diagnosis of malaria is based on the clinical signs and/or on the parasitological diagnosis, that detects evaluate the presence of parasites in the blood[43]. Diagnosis and treatment must be accomplished quickly to reduce and prevent mortality and the transmission of the disease. Diagnostic tests (or microscopy or rapid diagnostic test-RDT (Figure 12) must be performed before administering a treatment in all suspected cases of malaria [43]. These tests can detect several species of malaria parasites. The most common RDTs are specific for *P. falciparum*, having antigens for *P. falciparum* histidine-rich protein-2 (Pf HRP2)[44]. These test can have others antigens, such as enzymes of the parasite glycolytic pathways, namely plasmodial lactate dehydrogenase (LDH) and aldolase[44]. LDH can be specific for *P. falciparum* or *P. vivax*, and aldolase for to all species: *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale*, and *P. knowlesi*[44].

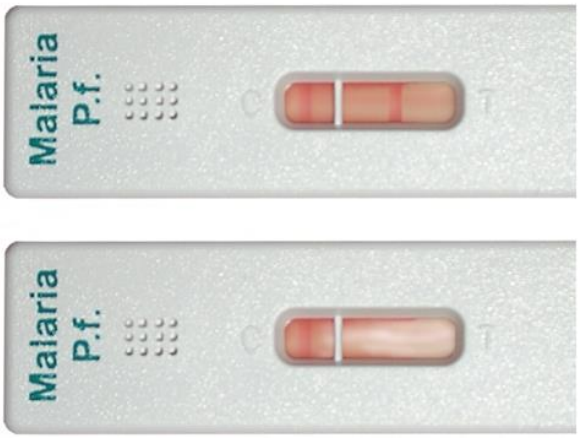



**Figure 12 - Example of two RDTs with HRP-2 (as the antigen) with positive (first) and negative (second) result[45].**

Other diagnostic targets have been assayed including highly conserved proteins [46]. Heat-shock proteins (HSPs) [47–50], *Plasmodium* heme detoxification protein (HDP) [51,52], Dihydrofolate reductase (DHFR)[53–55], Glutamate-rich protein (GLURP)[56] and High mobility group box 1 (HMGB1) protein[57] are some examples of new malaria diagnostic targets.

A different diagnostic approach is based on PCR method, that has been described in the early 1990s and proven to be the most sensitive test able to identify low levels of infection, parasite species, or mixed infections [58,59]. Other similar method is Loop-mediated isothermal amplification (LAMP), which is a specific nucleic acid amplification method that simplify the detection of the malaria human species [60–62].

When a parasitological diagnosis is not possible, WHO (2014) suggest that the treatment based on symptoms should be wisely considered[19]. Based on epidemiological data, the main target populations for malaria vaccines are children, young people, women of childbearing age, malaria-naive travelers and military troops that are in areas of moderate and high transmission[19].

For treatment, current antimalarial chemotherapy is carry out, focusing in the asexual blood stages of the parasite, that are the main targets responsible for the malaria symptoms. However, with the new goal of malaria eradication, the strategy has been changed to produce a drug that should prevent both disease transmission and the relapse of dormant liver stages[63,64]. An example is primaquine that affects both the mature stage V gametocytes and the hypnozoites. However, this antimalarial drug that has been currently used, is contraindicated in pregnancy and in young children and may cause hemolysis in patients with glucose-6-phosphate dehydrogenase (G6PD)-deficiency[63,64].

A growing problem related to malaria treatment is the parasite resistance to almost every class of antimalarial compounds. To overcome this problematic situation, two or more drugs with different mechanism of action are combined and used to delay the development of resistance. Recommended by WHO (2014), this is known as combined therapy[2]. This strategy includes artemisinin-based combination therapies (ACTs) that is a combination of a fast acting artemisinin derivative and a partner drug with a different and longer half-life mechanism of action[2]. For example, in almost all endemic countries of *P. falciparum* malaria, ACT is adopted as the first-line treatment policy (Figure 13).

| | | *P. falciparum* | | | | *P.vivax* |
|---|---|---|---|---|---|---|
| WHO region | Country/area | Uncomplicated unconfirmed | Uncomplicated confirmed | Severe | Prevention during pregnancy | Treatment |
| African | Algeria | - | - | - | - | CQ |
| | Angola | AL | AL | QN | SP(IPT) | - |
| | Benin | AL | AL | QN | SP(IPT) | - |
| | Botswana | AL | AL | QN | CQ+PG | - |
| | Burkina Faso | AL;AS+AQ | AL;AS+AQ | QN | SP(IPT) | - |
| | Burundi | AS+AQ | AS+AQ | QN | - | - |
| | Cameroon | AS+AQ | AS+AQ | AM;QN | SP(IPT) | - |
| | Cape Verde | AL | AL | QN | CQ | - |
| | Central African Republic | AL | AL | AM;QN | SP(IPT) | - |
| | Chad | AL;AS+AQ | AL;AS+AQ | AM;QN | SP(IPT) | - |
| | Comoros | AL | AL | QN | SP(IPT) | - |
| | Congo | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Côte d'Ivoire | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Democratic Republic of the Congo | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Equatorial Guinea | AS+AQ | AS+AQ | QN | - | - |
| | Eritrea | CQ+SP | AS+AQ | QN | - | CQ+PQ |
| | Ethiopia | AL | AL | QN | - | CQ |
| | Gabon | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Gambia | AL | AL | QN | SP(IPT) | - |
| | Ghana | AS+AQ | AL;AS+AQ | QN | SP(IPT) | - |
| | Guinea | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Guinea-Bissau | AL | AL | QN | SP(IPT) | - |
| | Kenya | AL | AL | QN | SP(IPT) | - |
| | Liberia | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Madagascar | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Malawi | AL | AL | QN | SP(IPT) | - |
| | Mali | AS+AQ | AL;AS+AQ | QN | SP(IPT) | - |
| | Mauritania | AS+AQ | AL;AS+AQ | QN | - | - |
| | Mozambique | AL | AL | QN | SP(IPT) | - |
| | Namibia | AL | AL | QN | SP(IPT) | AL |
| | Niger | AL | AL | QN | SP(IPT) | - |
| | Nigeria | AL;AS+AQ | AL;AS+AQ | AM;AS;QN | SP(IPT) | - |
| | Rwanda | AL | AL | AM;QN | SP(IPT) | - |
| | Sao Tome and Principe | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Senegal | AS+AQ | AL;AS+AQ | QN | SP(IPT) | - |
| | Sierra Leone | AS+AQ | AL;AS+AQ | AM;QN | SP(IPT) | - |
| | South Africa | - | AL;QN+CL;QN+D | QN | CQ+PG | AL+PQ;CQ+PQ |
| | Swaziland | - | AL | QN | CQ+PG | - |
| | Togo | AL;AS+AQ | AL;AS+AQ | QN | SP(IPT) | - |
| | Uganda | AL | AL | QN | SP(IPT) | - |
| | United Republic of Tanzania | | | | | |
| | Mainland | AL | AL | QN | SP(IPT) | - |
| | Zanzibar | AS+AQ | AS+AQ | QN | SP(IPT) | - |
| | Zambia | AL | AL | QN | SP(IPT) | - |
| | Zimbabwe | AL | AL | QN | SP(IPT) | - |

**Figure 13 - Representation of each artimisin combinated drugs for each WHO region.** Symbols and their means: AL as Artemether-lumefantrine, AM as Artemether, AQ as Amodiaquine, ART as Artemisinin, AS as Artesunate, CL as Clindamycline, CQ as Chloroquine, D as Doxycycline, DHA as Dihydroartemisinin, MQ as Mefloquine, NQ as Naphroquine, PG as Proguanil, PPQ as Piperaquine, PQ as Primaquine, PYR as Pyronaridine, QN as Quinine, SP as Sulphadoxine-pyrimethamine, T as Tetracycline[65].

As a treatment, WHO (2014) recommends ACTs for uncomplicated malaria caused by *P. falciparum* parasite [2,19] . However, in some cases, the safety of artemisinin derivatives is not yet established, and for this reason other antimalarial drugs have to be administrated[2,19]. For example, quinine and clindamycin are administrated during the first trimester of pregnancy;

chloroquine in combination with primaquine to treat *P. vivax* malaria in endemic areas; and sulfadoxine-pyrimethamine with amodiaquine, to treat pregnant women and children[2,19].

Resistance of parasite to these drugs is also described. In the 1970s and 1980s, *P. falciparum* became resistant to chloroquine and sulfadoxine-pyrimethamine. Sometimes in areas where P. vivax parasite is chloroquine-resistant, ACTs should be used with caution. Considering that resistance to artemisinins and other antimalarial drugs is widely spread in it is estimate that no alternative antimalarial drugs will appear in the market for at least five years [2,19].

Also resulting from malaria infection, severe malaria caused by infection with *P. falciparum* , or *P. vivax* and *P. knowlesi*, can be life-threatening[66,67]. For treatment, quinine has been used to mainstay. However, injectable artesunate (intramuscular or intravenous) has recently became the recommended treatment for severe malaria worldwide[68].

In general, the groups of antimalarial drugs that exist are: 4-Aminoquinolines, Aryl Amino Alcohol, 8-Aminoquinolines, Antifolates, Artemisinin and derivatives. Unfortunately, fake antimalarials drugs are widespread in African and Asian countries (Figure 14) compromising effectiveness and increasing the risk for emergence of resistance [69].

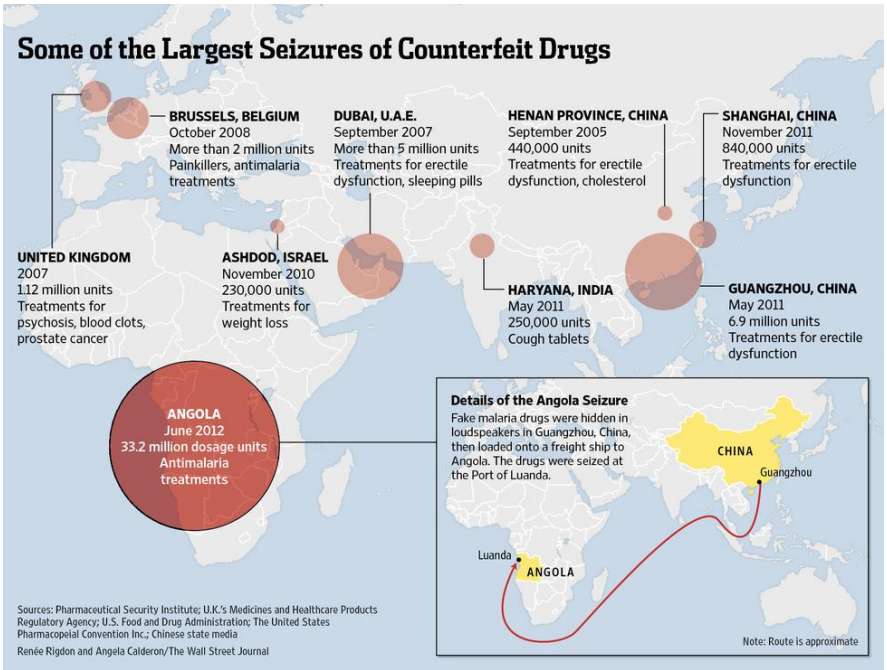



**Figure 14 - Map of the largest seizures of counterfeit drugs [70].**

Counterfeit medicines have a disastrous effect on global health and on individual patient safety and can lead to antimicrobial resistance for several diseases which includes malaria [71].

The goals of WHO (2014) regarding malaria eradication agenda have been limited by the emergence resistance to antimalarial drugs and insecticides. Thus, safe and protective vaccines are needed to control malaria, reaching all species of *Plasmodium*[2].

A suggestion to develop a malaria vaccine is to focus in the people that live in endemic areas and develop clinical protective immunity; and so produce an approach that limits the severity of the disease and prevent mortality [72]. However, the development of vaccines can be harder to achieve since the malaria parasite has strong ability to evade host's immune system by displaying a genetic diversity to express different antigens in its surface. For example, *P. falciparum* erythrocyte membrane protein-1 (PfE MP1) is a vaccine that shows temporal switching of variant expression, and that could affect the efficiency of this approach[72].

To develop a malaria vaccine must be considered if it is a safe treatment to prevent efficiently clinical disease and transmission of parasites to young infants, pregnant women and others. Vaccination against malaria can target different life cycle stages. These vaccines can be indeed divided into three types, as follows

Pre-erythrocytic vaccines, whose target are sporozoites and/or hepatic stages of the parasite[19,23]. They can induce an immune response to prevent invasion (mediated by an antibody) or to attack the infected liver cells (mediated by T cells). More than 40 years ago, sterile immunization in mice[73] and in humans[74] was performed by inoculating irradiated sporozoites. This approach aimed to reproduce the natural process that provides a high parasitaemia but also a protection against severe disease[74]. However, it does not result in sterilizing immunity [75–77], because an asymptomatic individual, can be a reservoir of gametocytes leading to transmission of malaria parasites[78,79].

Contrarily, immunizations with live SPZ preparations served as paradigms as antimalarial vaccines[80], once it was accomplished a sterile protection against infectious parasites[39] in experimental animal models and in malaria-naive humans. Besides, in 2011, Epstein and co-workers [81] demonstrated that an intravenous inoculation of attenuated and cryopreserved *P. falciparum* sporozoites protect the patients, opposing to intradermal inoculation results[82]. On the other hand, malaria vaccine candidates proved less effective in malaria-naïve adults of

field studies, possibly due to the genetic diversity of *P. falciparum* parasites in endemic areas [83], immunologic factors specific of children exposed in endemic areas [84] or the duration of vaccine-induced immunity[84].

In Figure 15, other approaches are represented that aim to arrest parasites stages to growth, and are grouped as live attenuated parasites (LAP) [85]. Three different approaches belong to this group. One of them is whole-sporozoite immunization called Radiation-attenuated sporozoites (RAS). These sporozoites have an arrestment in theirs nuclear division when DNA replication starts (depending on the irradiation dose), contributing to a sterile protection in previously naïve adults with high efficacy. Approaches such as RAS could potentially advance to phase 2b vaccination trials [85].

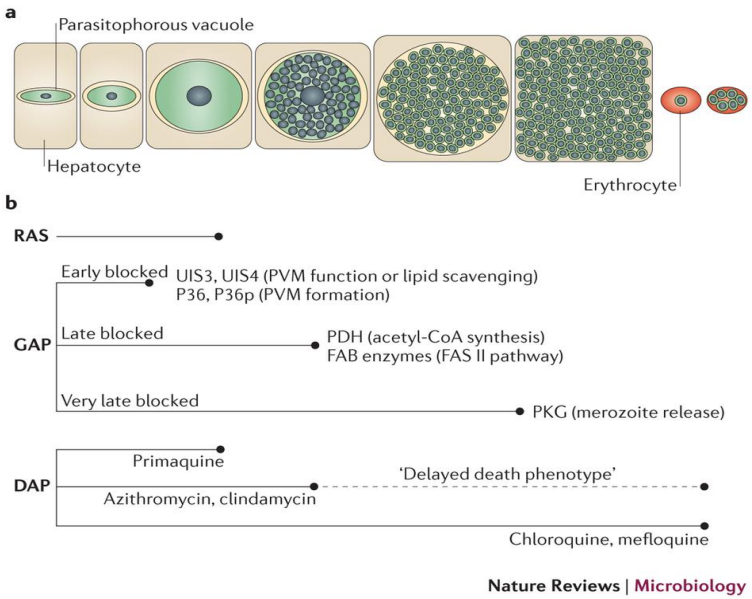



**Figure 15 - Transformation of sporozoites into merozoites in a hepatocyte.** a) Sporozoite transformation into merozoites inside a hepatocyte. b) Stages of growth arrest of various live attenuated parasites (LAP). Radiation-attenuated sporozoites (*RAS*), genetically attenuated parasites (*GAP*), and drug-arrested parasites (DAP) are some of the methods of malaria treatment[86].

In genetically attenuated parasites (GAP), some strategies can cause early liver-stage arrest (depletion of P36 and/or P36p), rapid liver-stage elimination (depletion of UIS3 or UIS4), mid-to-late liver-stage arrest (depletion of PDH or FAB enzymes) or very late liver-stage block (conditional depletion of cyclic GMP-dependent protein kinase (PKG)) [86]. In

drug-arrested parasites (DAP) are included drugs that cause mid liver-stage arrest (such as the antifolate pyrimethamine and the 8-aminoquinoline primaquine), inhibit apicoplast biogenesis in the liver stage contributing to the formation of non-infectious merozoites (such as azithromycin and clindamycin) and inhibit haemozoin (that is an insoluble crystalline form of free-toxic heme, resulting from blood digestion) formation avoiding the growth of intra-erythrocytic asexual parasites (such as Chloroquine, a 4-aminoquinoline, and mefloquine, a 4-methanolquinoline)[86].

RTS,S/AS01 is the most advanced vaccine candidate against *P. falciparum* and results from a GAP method[87]. Basically a genetically attenuated parasite was performed by dual gene deletions, leading to unable to complete liver stage development [87]. The circumsporozoite antigen was identified as the major component of the sporozoite surface, that inhibit the invasion of sporozoites to the hepatocytes, and that induce T cell responses against sporozoite-infected liver cells [88,89]. These findings led the development of RTS,S that formulated with potent adjuvant system, AS02 and AS01, has limited immunogenicity[87]. A Phase III trial began in May 2009 and completed in 2011, with 15 460 children with 5–17 months of age across seven African countries [90]. The results showed a vaccine efficacy for clinical malaria of 50% in older children [91] but only 30% in infants (the target population) without significant protection from severe malaria at 18 months post-vaccination [92]. These vaccination trials did not reach what was expected by the Malaria Vaccine Technology Roadmap for 2015 which were to have a vaccine that could provide 50 % protection against severe disease and death for at least 1 year[92].

Considering pre-erythrocytic vaccine strategies, they are unlikely to achieve complete protection in all individuals, due to the difficulty to neutralize the capacity a single sporozoite or liver-stage form to initiate a blood-stage infection, leading to malaria illness [93]. However, vaccine strategies that target the erythrocyte-invasive form of the parasite (merozoites) in combination with the pre-erythrocytic forms [94,95], could complement each other [93,96].

Blood-stage vaccines aim to prevent or blocking the invasion and adherence of parasites in the red blood cells (RBCs) to several tissues. Exist many asexual blood stage vaccine candidates, but most target merozoite antigens, such as the merozoite surface protein 1 (MSP1), 2 (MSP2), 3 (MSP3), the apical membrane protein (AMA1), the glutamate-rich

protein (GLURP), and the erythrocyte-binding antigen 175 (EBA175) [97–99]. Besides their unclearness in clinical protection, these vaccines attenuate the clinical symptoms of malaria [97–99].

Antigens expressed on the surface of infected RBCs are considered as not good candidates vaccines since they are highly polymorphic. An exception is a variant of the erythrocyte membrane protein 1 (Pf EMP1) vaccine known as VAR2CSA that prevents pregnancy-associated malaria[100]. Complications during pregnancy are associated with maternal anemia, placental malaria infection, birth weight reduction and risk of neonatal mortality[100]. This vaccine are supported by the fact that women acquire immunity after one pregnancy through the production of antibodies against VAR2CSA that inhibits the accumulation and adhesion of infected erythrocytes to placental chondroitin sulfate A protein (CSA)[100]. Vaccine candidates against VAR2CSA are currently under development[101].

The most promising blood-stage vaccine candidate is the merozoite protein of *P. falciparum* reticulocyte-binding protein homologue 5 (PfRH5), due to its non-redundant role during erythrocyte invasion, limited genetic diversity that specifically elicit antibodies to neutralize parasite growth, and protection from febrile malaria in endemic populations [102–104].

In addition, other studies [105,106] have contributed to an improve understanding on blood-stage biology and the potential for a multi-component approach that targets distinct steps during the erythrocytic cycle [93].

Transmission-blocking vaccines contribute to a reduction of incidence of infection, once they can block transmission to an individual person and subsequent spread of parasites in endemic populations. By targeting the sexual stages of the parasite; process like fertilization of gametes, oocyst formation and sporogonic cycle in the mosquito midgut can be blocked by specific host antibodies, complement proteins, and cytokines, preventing parasite development [72,107–109]. These vaccines however, do not protect the vaccinated individual protecting instead other individuals close by. Surface proteins expressed on gametocytes, gametes, zygotes, and ookinetes (such as *Pf*s25, *Pf*s28, *Pf*s48/45, and *Pf*s230 of *P. falciparum*[110]; and *Pv*s25, *Pv*s28 and *Pv*s230 of *P. vivax*[111,112]*)* are the main targets of this category of vaccines that have been shown to block transmission by induction of mosquito antibodies during the blood meal. Among these leading candidates, *Pfs*25 and *Pvs*25 have completed Phase I clinical trials with

limited results[113]. Another perspective, is transmission-blocking targets of mosquito ligands, that can allow the development of vaccines able to interrupt the transmission of more than one plasmodial species [114].

Regarding the necessity of preventive measures against malaria, WHO (2014) include vector control procedures such as indoor residual spraying, the use of long-lasting insecticide-treated bed-nets, and the destruction of larval breeding sites; to accomplish a reduction of transmission and spread of malaria [2]. Insecticide-treated materials can influence the reduction of density, survival, contact with humans, and feeding frequency of vector's population in its habitat[2,27]. Insecticide-treated nets (ITNs) and indoor residual spraying (IRS) have contributed to the decrease in malaria morbidity and mortality in the last decades[43]. ITNs protect the users, not only against mosquitoes but also against other infective insects that carry diseases, and should be used in all endemic areas[19]. Dependent on the insect's behavior, IRS is effective for indoor night biter mosquitoes because the insecticides included are a powerful way to rapidly reduce malaria transmission for 3–6 months. IRS is effective, depending on the insecticide used and the type of surface on which it is sprayed. Considering in some cases a longer-lasting form of IRS insecticides, dichlorodiphenyltrichloroethane (DDT) can be effective for 9–12 months[19]. Due to its high toxicity, other strategies have to be procedure.
Pyrethroids are the only insecticides currently recommended by the WHO (2013) for use on bed nets [43], and as a result of its use in a wide-scale in the last decades, resistant mosquitoes to this insecticide have emerged in some areas of Africa [115].

Besides multiple mechanisms of resistance to insecticides that have been observed in *Anopheles* populations[116], alterations in the target site that are often referred to as knockdown resistance (kdr) provide to insects the ability of with these alleles to survive at prolonged exposure to insecticides. kdr genotype and the resistance phenotype are associated with resistance to pyrethroids and DDT, but it is not evident that the presence of this resistance allele alone is sufficient to result in the unsuccessful vector control [117].

Other insecticides such organochlorines (OCs), organophosphates (OPs) and carbamates (Cs) have been also used in IRS throughout Africa and several reports have shown that resistance to all three had been as well developed in mosquitoes [118]. These alarming observations show the capacity of *Anopheles* to adapt to different and noxious conditions. For

this reason and as referred before, is necessary to discover new strategies/targets to reduce malaria transmission.

Another perspective to control transmission could be the use of genetically modified mosquitoes unable to transmit parasites. Competitor peptides able to bind salivary gland receptors have been tested to analyze their ability to block parasite transmission in mosquitoes. *A. stephensi,* for example, was genetically modified for the expression of the SM1 peptide, to inhibit the ability of mosquito to transmit parasites in the midgut[119]. Other peptides that affects parasites in the vector such as cecropins, defensins, and scorpine have been also described [120–122]. An interesting way to interfere with malaria transmission is the use of paratransgenesis that attempts to eliminate a pathogen from vector populations by using genetically modified symbiotic microorganism of the insect[123]. Lastly, the use of transgenic procedure can improve the sterile insect technique (SIT) for *Anopheles* mosquitoes, by releasing radiation-sterilized males in the field to compete with wild ones. However, this method can result in mosquitos that have poor mating competitiveness [124,125]. To introduce a transgenetic mosquito in the field in order to substitute the wild vector population, imperative aspects have to be analysed; such as the potentially harmful ecological effects and the acceptance of genetically modified organism procedure. Killing the mosquito larvae by spraying of kerosene, pyrethrum oil, and the introduction of larvivorous fishes like *Gambusia* s*p.*, insectivorous plants like *Utricularia* s*p.* into stagnant water places where mosquitoes breed, are some of the biological control methods to avoid the growth of mosquito larvae[124,125].

## 1.4. Next-Generation Sequencing (NGS)

New scientific resources must be available to facilitate studies about the biology, transcriptomic and evolution of mosquito species, making the development and implementation of novel targets to control malaria an easier process [126].

The recent developments of next-generation sequencing (NGS) allow researchers to study biological systems throw rapidly sequencing whole genomes, understand sequence target regions, analyze genome-wide methylation or DNA-protein interactions, sequence large numbers of samples simultaneously using multiplex sequencing with DNA barcode tags,

profile microbial diversity in humans or in the environment, discover novel RNA variants and splice sites, or precisely quantify mRNAs for gene expression analysis by RNA sequencing[127]. The complex genomic research questions request techniques that provide information beyond the capacity of traditional DNA sequencing technologies. To fill this absence, next-generation sequencing became an useful tool to address these questions [128].

As a NGS technique, RNA-seq contributes for many improvements in the transcriptomic characterization and quantification, allowing transcription start site mapping, strand-specific measurements, gene fusion detection, small RNA characterization [129], detection of alternative splicing events [130,131], single nucleotide polymorphisms [132], provide the opportunity to interrogate allele-specific expression and RNA editing. In addition to mRNA transcripts, RNA-seq can evaluate total RNA, small RNA, such as miRNA, tRNA, and ribosomal profiling [133]. Ongoing RNA-seq research includes observing cellular pathway alterations during infection [134], and gene expression level changes in cancer studies [135]. Besides that, it can measure precisely transcripts levels and their isoforms when comparing to other methods [136]. Prior to NGS, transcriptomic and gene expression studies were done with microarrays, which contain thousands of DNA sequences (probes) that possibly match with the complementary sequences in the sample, making available a profile of all transcripts being expressed. This was later done with serial analysis of gene expression (SAGE). Microarrays have limited coverage since only target known common alleles, that only represents approximately 500 000 to 2 000 000 SNPs of the more than 10 000 000 in the genome [137]. As such, libraries are unavailable to detect and to evaluate rare allele variant transcripts [138], preventing a better fully and detailed research results[139]. However, microarrays approach still has meaning to regulatory body-approved diagnostics that target common allele variants already known[139]. Comparison in work-flow between microarrays and RNA-seq are represented in Figure 16.

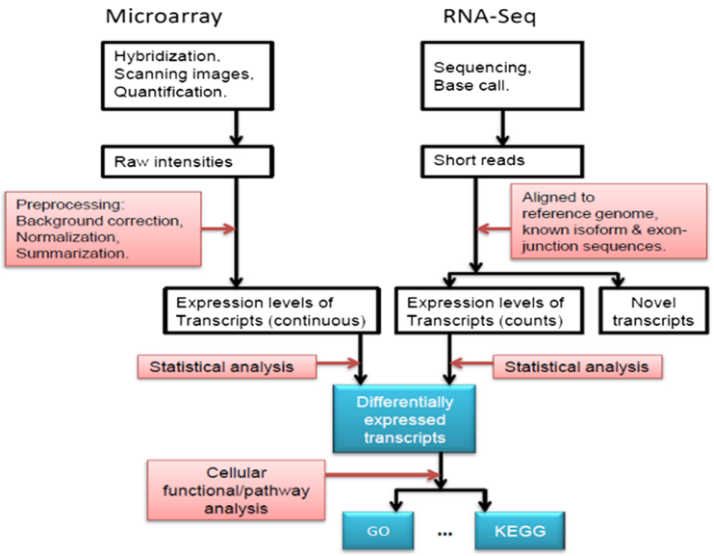



**Figure 16 - Overview of analysis workflow for microarray and RNA-seq transcriptional profiling.**[140]

RNA-seq is becoming increasingly attractive for quantitative studies of differential gene expression in non-model species, for which there is generally a weak knowledge about transcriptome [141–143]. RNA-seq data from non-model species can be successfully *de novo* assembled into transcriptomes with high quality. From non-models species reads *per* gene and annotating assembly contigs can be achieved by using the genome or predicted gene set of a related species as proxy [144,145]. Proxy genomic reference species that have an divergence up to approximately 100 million years from the target species can be utilized to generate results in order to functionally annotate *de novo* transcriptomes[129].

Annotating assembly contigs can then be used with high performance, as scaffolds for mapping RNA-seq read data for quantitative whole gene expression analyses[129].

### 1.4.1. RNA sequencing (RNA-seq)

Once mRNA is extracted and purified from cells, quality of RNA is critical to be analyzed, to proceed to alternative protocols if degraded RNA exist[146]. Further, fragmentation of the RNA in to smaller pieces followed by reverse transcription of fragments using random hexamers or oligo (dT) primers, contributes to ensure that coding RNA (mRNA) is separated from noncoding RNA (magnetic beads may be necessary)(Figure 17) [147].

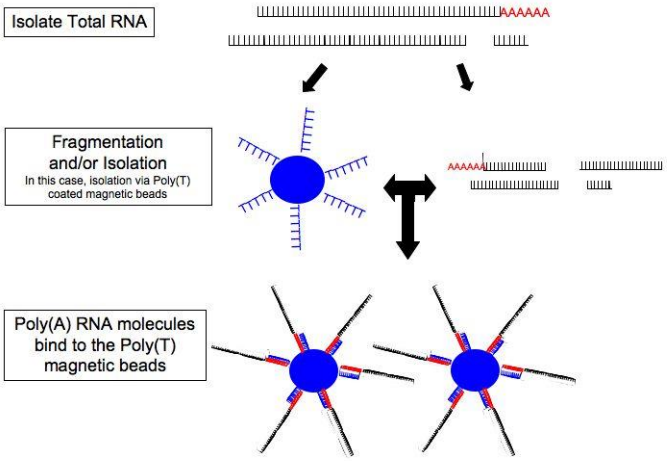



**Figure 17 - Representation of librarys of cDNA.**
(Image from Vincent Montoya)

As represented in the Figure 18, the 5' and/or 3' ends of cDNA are repaired and adapters (containing sequences to allow hybridization to a flow cell) are ligated. PCR are performed to enriched correctly ligated cDNA fragments (Figure 18).

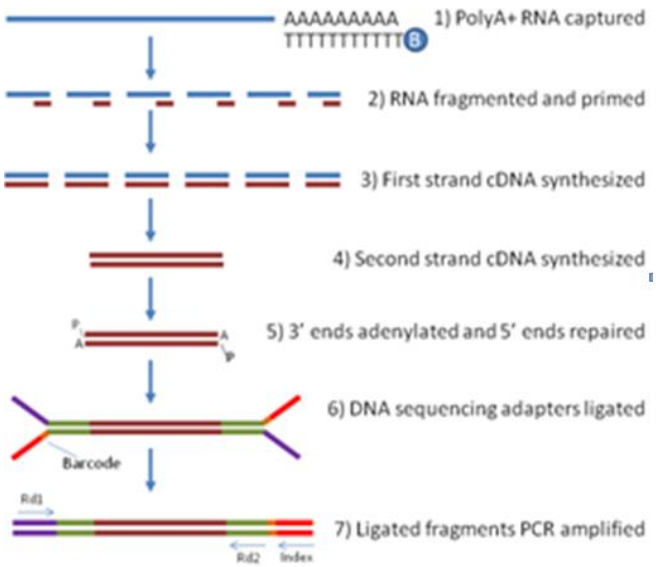



**Figure 18 - Illumina library preparation.** PolyA+ RNA is enriched using oligo (dT) beads followed by fragmentation and reverse transcription. The 5' and 3' ends of cDNA fragments are next prepared to allow efficient ligation of "Y" adapters containing a unique barcode and primer binding sites. Finally, ligated cDNAs are PCR-amplified and ready for cluster generation and sequencing[148].

After assessment of library concentration using qPCR and/or Bioanalyser, it is ready for sequencing on a high-throughput sequencing facility [149]. Each of the following platforms utilizes a different type of technology to sequence millions of different short reads: 454 Sequencing, Illumina, SOLiD, Ion Torrent, and Illumina's TruSeq or HiSeq.

In Illumina TruSeq, for example, dNTPs are incorporated and detected simultaneously at millions of fixed positions on a flow cell [150]. Represented in Figure 19, in this technique, libraries that were prepared are now hybridized to a flow cell which contains a lawn of covalently bound oligonucleotides complementary to the sequencing adapters that were introduced during library preparation. Once hybridized, the capture oligonucleotide primes DNA polymerase extension activity resulting in a covalently bound full-length complementary copy of the cDNA fragment that is subjected to several rounds of PCR amplification to produce discrete clones ~ 1 μm in diameter that can be optically resolved during sequencing[150].

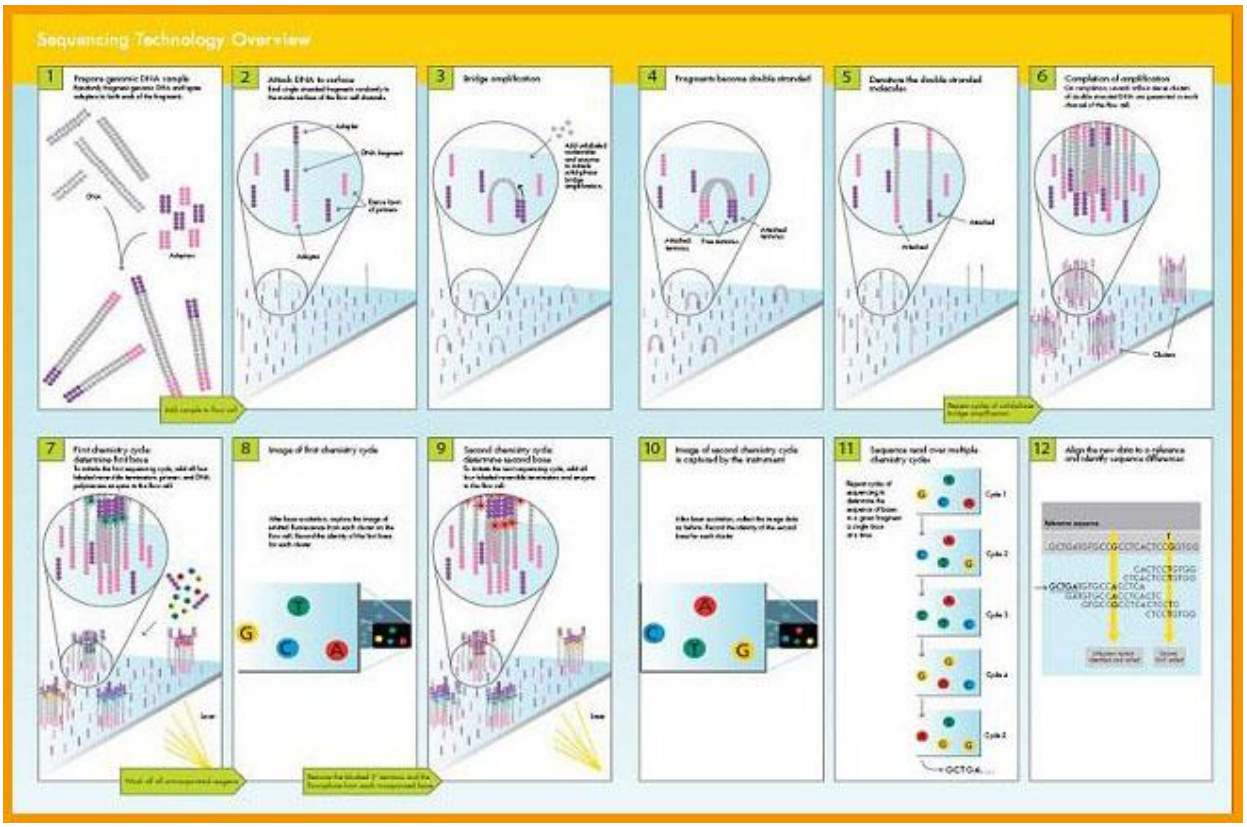



**Figure 19 - Illumina sequencing resume. (**http://www.geneconn.com/images/upload/Image/5-fu-1.jpg)

Manipulating RNA-seq data is computationally intensive and requires access to several databases that provide information about clusters. Two different transcriptomic assembly methods are used for producing a transcriptome from raw sequence reads: *de novo* or genome-guided.

As mentioned before, in order to reconstruct the nucleotide sequence, *de novo* assembly does not rely on the presence of a reference genome described for the organism of interest. Due to the large size of the reads, *de novo* assembly may be easier though some software that exist (Velvet (algorithm), Oases [151], Trinity [152], and others). In *de novo* assembly, the absence of a known genome can be overcome by assembly the sequences through a proxy genomic reference species[129], taking into account possible alternative isoforms and the dynamic range of expression values. To build transcript isoforms, Oases can analyze these results by using paired end read and long read information (Figure 20)[153].

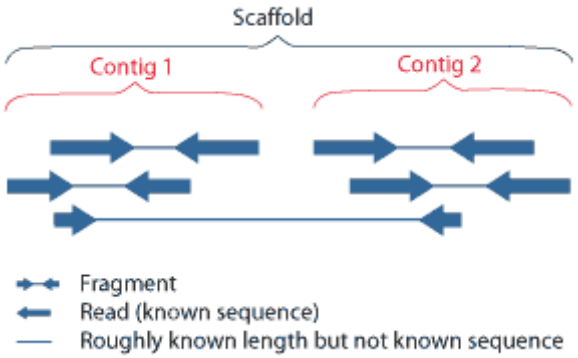



**Figure 20 - Overlapping reads.** Reads from paired-end sequencing form contigs, contigs and gaps of known length form scaffolds. **(**http://genome.jgi.doe.gov/help/scaffolds.html)

Another computationally approach is genome-guided assembly, which aligns the millions of reads to a "reference genome". This are performed for annotated organisms by using several software packages and algorithms for short read alignment, e.g. Bowtie and Samtools for RNA-seq short read alignment [154155], TopHat for aligning reads to a reference genome to discover splice sites [156], Cufflinks to assemble the transcripts and compare/merge them with others [157].

Then is necessary a annotation of reads and for this, the UniProt Reference Clusters (UniRef) is a very useful tool for annotation providing clustered sets of sequences from the UniProt Knowledgebase (including isoforms) and selected UniParc records in order to obtain complete coverage of the sequence space at several resolutions. Using UniRef90 it is possible to analyse each cluster having at least 90% sequence identity to and 80% overlap with the longest sequence (a.k.a. seed sequence) of the cluster [158]. That software gives GO (Gene Ontology) annotations, EC numbers, Interpro motifs and others. After a collection of expressed sequences aligned or located to same position on genome, UniGene, an alignment by BLASTX to protein databases must be performed, retrieving proteins with the highest sequence similarity, to obtain their protein functional annotation.

After mapping and annotation, it is necessary an expression quantification of each UniGene. Quantification of the abundance of target sequences obtained from sampled subsequences, can be done with the eXpress streaming tool which greatly reduces computing infrastructure requirements being more efficient when coupled with a streaming aligner (such as Bowtie).

The characterization of gene expression in cells by measurement of which genes and at what levels of mRNA levels has long been of interest to researchers, was achieved by Fragments *per* Kilobase of transcript *per* Million mapped reads (FPKM) that correspond to the the number of fragmented reads belonging to the particular exon in fragmentation prior to cDNA synthesis, that are normalized by its length in vivo yields gene expression levels[159] and by the total number of mapped reads. Following library normalization, statistical testing for differential expression can be performed. R/Bioconductor statistical computing environment [160], edgeR and DE-Seq are used to perform count normalization using a negative binomial distribution[161,162] and further uses the counts table with biological replicates to calculate variation and test for statistically significant differential expression.

Functional annotation of the assembled transcripts allows for insight into the particular molecular functions, cellular components, and biological processes in which the putative proteins are involved. Blast2GO enables Gene Ontology (GO) based data mining to annotate sequence data for which no GO annotation is available yet. Often employed in functional

genomics research on non-model species [163], Blast2GO works by blasting assembled contigs against a non-redundant protein database (at NCBI), then annotating them based on sequence similarity. Following annotation, KEGG (Kyoto Encyclopedia of Genes and Genomes) [164], STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) [165], WoLF PSORT (Protein Subcellular Localization Prediction)[166], MultiLoc2[167] and TMHMM[168] enables visualization of predicted metabolic pathways and molecular interaction networks captured in the transcriptome.

Tests for enrichment of Gene Ontology (GO) terms associated with a list of significantly over or under expressed genes can be performed with for example, g:Profiler[169,170] or GOstat[171]; depending on which model or non-model organism is used. Sequencing analysis using all the above mentioned tools will contribute to a better understanding of which molecular function or biological process is more or less representative in the transcriptome of interest.

Quality control of the process and the final results has to be controlled by trimming of sequencing adapters and removal of reads with poor quality scores followed by mapping reads, analysis of differential expression, identification of novel transcripts and pathway analysis[172].

Just as for any other technique, both technical and biological replicates must be carefully considered to guarantee a proper replication, randomization and blocking of a well-designed experiment. The required number of replicates will vary depending on amount of biological variability associated with the samples of interest and should be empirically determined to performed RNA-seq experiments[173]. However, in the infancy of RNA-seq, technical replicates (libraries prepared from the same RNA sample) were commonly used and shown that biological variation far outweighs technical variation, at least when coverage of at least 5 reads *per* nucleotide is obtained [174,175].

## 1.5. Real Time Polymerase Chain Reaction (qPCR)

Transcriptional profiling using NGS platforms allow a better understanding about the cell development and diseases progression in a certain tissue under varied circumstances. RNA-seq has shown to be highly accurate for quantifying expression levels, but qPCR can be as well performed in parallel to confirm RNA-seq data results[176].

First RNA is reverse-transcribed and the produced cDNA used as a template in real-time PCR to detect and quantitate expressed gene products and then infer about gene expression or messenger RNA (mRNA) presence. The data output is expressed as a fold-change or a fold-difference of expression levels depending if it is used quantitative or relative quantification[177].

To obtain absolute quantities of gene targets it is necessary to calculate the copy number of the gene usually by relating the PCR signal to a standard curve. Methods for relative quantitation of gene expression present the data of the gene of interest relative to some internal control gene known as reference gene, which will result in changes in gene expression in a given sample relatively when compared with a control sample (e.g. untreated control sample).

The calculation methods used for relative quantification are standard curve method and comparative CT method. This last method is widely used and it compares the Ct value of one target gene to another (using the formula: 2ΔΔCt)—for example, an internal control or reference gene (e.g., housekeeping gene)—in a single sample[178].

## 1.6. RNA interference

In nature, RNA interference (RNAi) is an important pathway that regulates gene expression in many different organisms. In higher eukaryotes, RNAi-mediated *knockdown* is the most common and the best known strategy for depleting cells of a gene product of interest in the cytoplasm [179]. By contrast, genome editing that alters the genetic code; contribute to complete elimination of gene function, typically causing a *knockout*.

RNAi-mediated *knockdown* is preferable when knockout is undesirable, and the aim is only to reduce gene function temporarily[180]. As represented in Figure 21, RNAi process begins when a hairpin-forming precursors (shRNAs) or introduced exogenously (siRNAs) generate short (approximately 20-25 nucleotides) double stranded RNA molecules (*ds*RNA). After processing by Dicer, a *ds*RNA is converted in a single stranded RNA base-pairs (*ss*RNA) with a target mRNA [181]. Further, Argonaute mediate RNA gene knockdown, resulting in mRNA degradation or translation inhibition. This leads to post transcriptional down-regulation of gene expression, without altering the genetic code [182], termed a “knockdown”.

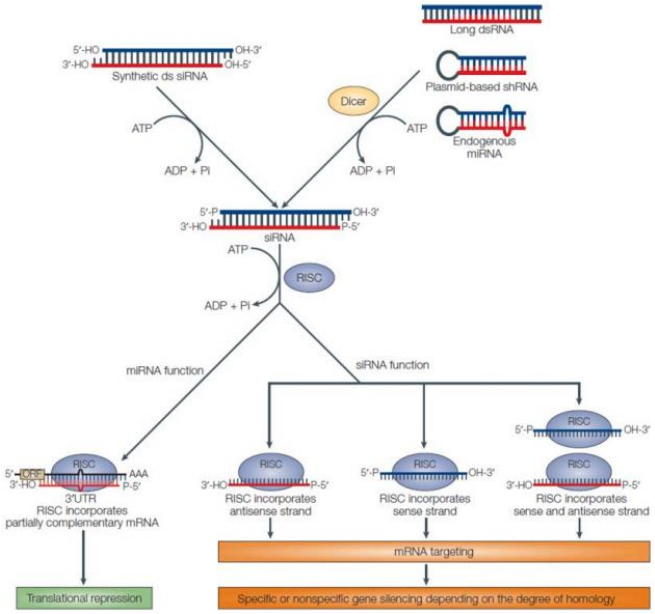



**Figure 21 - General scheme of RNAi pathways.**[182]

# Chapter 2: Material and Methods

## 2.1. Identification of genes differentially expressed in salivary glands (SG) of *A. stephensi* in response to infection by *P. berghei*

### 2.1.1. RNA-seq

The transcriptomic analysis by sequencing RNA transcripts (RNA-seq) was achieved by Era7 Company, where reads were pre-processed to have the high quality required for the *de novo* assembly. In the report received from Era7 were described the methods used to obtain this RNA-seq data, and finally the results were presented for four Biological Coefficient of Variation (BCV). These values inferred from how much the variance of the counts exceeds the variance that would arise from Poisson counts, and BCV values that stabilized between genes are selected for forward analysis. According to Robinson, McCarthy and co-workers, a BCV value for technical replicates most be 0.01[183]. In this work, we did not use such low value, to focus in higher significantly expressed transcripts; and either not choose a high value as 0.4, which reduces analysis of transcripts that can be potentially interesting but with have a bit less significant expression. For these reasons, were chosen a BCV of 0.2 and a $p$ value $<0.05$ to filter significant data.

The RNA-seq result compiles information of UniProt, RefSeq, Gene Ontology and NCBI Taxonomy Enzyme databases. Each transcript has information about Uniprot ID, log of concentration, log of fold-change, $p$ value, protein names, gene names, length of the sequence, Enzyme Commission number (EC number), gene ontology IDs (GO terms) and URL UniRef cluster. Data were further analysed and the transcripts organised into classes according to their function. The 14 functional classes chosen were: cell function, chemosensory, detoxification, diverse, immunity, metabolism, proteolysis, recognition, regulatory, RTT (replication, transcription and translation), signalling, structural, transport and unknown. The amount of transcripts that showed a significant differential expression were analysed and grouped

depending on the over or under expression in each functional class and the distribution of each class in total. Over and under-represented functional classes were identified using GO ID's for each transcript by a hypergeometric test named GOstat[171]. Transcripts without GO ID's or without Gene ID were excluded.

### 2.1.2. Confirmation of differential expression by qPCR

All animals procedures were conducted in accordance with the Council of Ethics of Institute of Hygiene and Tropical Medicine, performed on anaesthesia, minimizing animal suffering. The maintenance and care of experimental animals was as well carried out in accordance to the Europe Directive 86/609/EEC and Portuguese law (Decreto-Lei 129/92) recommendations and protocol approved by the Direção-Geral de Alimentação e Veterinária (DGAV), Portugal, under Portaria 8 n°1005/92 from 23rd October.

*A. stephensi* mosquitoes were obtained in the insectary of IHMT, reared at 20°C in 70% humidity, under a 12h light/dark photoperiod and fed *ad libitum* on a 10% glucose solution.
Female CD1 mice were intraperitoneally inoculated with $10^7$ *P. berghei* ANKA parasitized red blood cells. The levels of parasitaemia were measured from blood samples of the mouse tail using Hemacolor® rapid staining kit (EMD Millipore, Germany). When the parasitaemia reached 5-10% and the presence of microgametocytes capable of exflagellation was observed [184], mice were used to infect mosquitoes. Four to five days-old female mosquitoes were allowed to feed directly on naïve (control) and *P. berghei* infected mice up to one hour, with regular monitoring to certify mice were anesthetised. Unfed mosquitoes were removed after each blood meal. Fully engorged mosquitoes were kept at 19-21°C and 80% relative humidity for *P. berghei* development.

After anesthetization of mosquitoes by subjection/exposition to a temperature of 4°C, they were submerged in 75% (v/v) ethanol, and then in phosphate buffer (PBS) *A. stephensi* tissues were dissected on a microscope slide using fine needles under a stereomicroscope at 4x magnification (Motic SMZ-171B, China).
Midguts from the infected female mosquitoes were dissected 8-9 days post-blood meal (PBM) to confirm the presence of oocysts for the confirmation of infection. Once *P. berghei* was

genetically modified and constitutively expressing a fluorescent protein, GFP (Green Fluorescent Protein), the oocysts were observed, using fluorescent microscopy. Salivary glands (SG) were dissected 18-19 days PBM in PBS, subsequently transferred to RNA later (Life Technologies, USA) and maintained at 4ºC[185] for further RNA extraction.

RNA was quickly extracted using Ribo*Zol*™ (AMRESCO, USA). For qPCR, a standard curve was performed with total RNA extracted from a pool of eight females kept exclusively on a sugar diet; and analysis of differential expression was performed with RNA from both control and infected SGs. After extraction, the RNA samples were quantified and analysed for purity on a ND-1000 *NanoDrop*® spectrophotometer (NanoDrop ND1000, Thermo Fisher Scientific, MA). Integrity of RNA was evaluated by an *Agilent 2100 Bioanalyser*® system (Agilent Technologies, USA). Later, samples were stored at -80 ºC.

Total RNA extracted from SG was used to synthesize cDNA using the *iScript™ Reverse Transcription Supermix for RT-qPCR* kit (Bio-Rad, USA) using the following protocol: a phase of priming at 25$^{o}$C for 5 min, followed by a reverse transcription at 42$^{o}$C for 30 min and reverse transcriptase inactivation at 85$^{o}$C, for 5 minutes.

For validation of RNA-seq data of *A. stephensi* infected and non-infected, relative qPCR was performed. A total of 20 transcripts identified by RNA-seq, which were expressed differentially, were evaluated to confirmed RNA-seq results. In qPCR were used three biological replicates of each condition. The Primer 3 platform (http://bioinfo.ut.ee/primer3-0.4.0/) was utilized to design all primers. The qPCR reactions (total volume of 20 µl) were performed in triplicate using *IQTM kit SYBR ® Green super mix* kit (Biorad, Singapore) in a thermocycler *CFX96 Touch Real-time PCR* (Bio-Rad, USA). qPCR conditions were: an initial cycle of denaturation at 95° C for 10 min, followed by 40 cycles of 95ºC for 15 seconds and temperature of each primer set for 45 seconds. A melting curve (60-95°C) was performed to determinate the quality of the amplicon.

The sequences of the primers used in these experiments are listed in Appendix II.

Data normalisation in real-time qPCR can be improved by including an invariant endogenous control (reference gene) in the assay[186]. Normalization against a single reference gene is not acceptable, unless experiments confirm its invariant expression under the experimental conditions[187]. For this reason, two reference genes candidates were chosen, the

gene for ribosomal protein S7 (Vectorbase: AGAP010592, because primers were similar between species of *A. gambie* and *A. stephensi*) and elongation factor (Vectorbase: ASTE005097), also indicated in the Appendix II.

Standard curve was performed from serial dilutions of cDNA to calculate the PCR reaction efficiency and determine the best reaction conditions. The same curve as the reference for the relative quantification of expression levels was used. Starting from a concentration of 100 ng/µl of total RNA cDNA, serial dilutions of 1:5, 1:25, 1:125 and 1:625 were achieved. No primer dimer was detected when inspecting the melting curves and primer pairs were chosen that displayed greater than 90% amplification efficiency.

Relative expression results were normalized with ribosomal protein S7 (RPS7) as reference gene and analysed by the 2 delta Ct (ΔΔCt) method [188]. The same was performed for an elongation factor (ELF).

The correlation between the expression values detected by RNA-seq and qPCR for the 13 genes confirmed was determined by calculating Pearson's correlation, comparing the expression values for each transcript in the two methods [189]. A *P* value between 0.25 and 0.50 was considered to be weakly correlated, a *P* value between 0.51 and 0.75 was considered to be moderately correlated, and P value over 0.76 was considered to be strongly correlated.

## 2.2. *In vivo* gene knockdown in *Anopheles stephensi* by RNA interference

### 2.2.1. *ds*RNA synthesis

After analysing protein-protein-interations, ontological characteristics and expression level, RNAi was used to knockdown the gene ASTE009391 (*prestin*) in *A. stephensi* SG and the effect of *prestin* gene knockdown in the malaria vector further analysed.

In order to knockdown the gene, double stranded RNA was produced. First specific primers and cDNA from SG as template, was used to amplify fragments of interest, using iProof™ High-Fidelity DNA Polymerase (BioRad, Hercules, CA). Specific primers containing T7 promoter sequences at the 5´-end that were synthesized (Appendix III), only will generate

PCR products with T7 promoter. Further, this PCR products with T7 promoter will be necessary to produce a lot of *ds*RNA of interest, when added T7 RNA polymerase[190].

PCR reactions of 50 µL included: 1X iProof HF Buffer, 10 mM of dNTP mix, 0.5 µM of each primer (Appendix III), 1mM of MgCl2, 0.02 U/µL Proof DNA Polymerase and 50-500ng of template DNA. The conditions were: initial denaturation performed at 98ºC for 3 minutes; following for denaturation at 98ºC for 10 seconds, annealing at 64ºC for 30 seconds and extension at 72ºC for 15 seconds cycling for 35 cycles; and then a final extension at 72ºC for 10 minutes.

Amplification results were analysed on a 0.5X TBE, 1.2% (w/v) agarose gel and purified using Zigmoclean™ Gel DNA Recovery Kit (Zymo Research, USA). *ds*RNA was produced using the MEGAscript T7 kit (Ambion, Austin, TX, USA), according to manufacturer's instructions. The resulting *ds*RNA was purified, quantified and checked on a 0.5X TBE, 1.5% (w/v) agarose gel.

### 2.2.2. *ds*RNA injection in *Anopheles stephensi*

For gene knockdown, three experiments of 300 *P. berghei*-infected female mosquitoes for each gene (*β2M* and *prestin*) were conducted. Three to five-days-old (14 days PBM) cold anesthetized female mosquitoes were injected intrathoraxically with 69 nl (4µg/ml) of *ds*RNA using a nano-injector (Nanoject, Drummond Scientific, Broomall, PA, USA). The control group was injected with *dsβ2M*. Before and after injection viable mosquitoes were counted.

### 2.2.3. Statistical analysis

Data were analysed across entire experiments (all time points) using Student's t-test (Microsoft ® Excel® 2013). A $p$ value of $<0.05$ was considered to be statistically significant. All $p$ values are two-sided.

### 2.2.4. Gene knockdown assessment

Mosquitoes were dissected 4 days post injection. Pools of 15 SG per treatment were used to perform sporozoite quantification and the remainder SG were used to extract RNA using RiboZol™ (AMRESCO, USA), as previously describe, to perform downstream qPCR in order to verify the knockdown effect.

Each SG pool was homogenized in a total volume of 100µl of phosphate-buffered saline using a mini glass tissue homogenizer (Kontes Glass Co., Vineland, NJ, USA). Sporozoites were counted by light microscopy using a haemocytometer as described by Pinto and colleagues [191].

# Chapter 3: Results

## 3.1. Identify genes differentially expressed in salivary glands (SG) of *A. stephensi* in response to infection by *P. berghei*

### 3.1.1. RNA-seq

RNA sequencing is a powerful tool to analyse a transcriptome across different tissues or conditions[192–194], having clear advantages over existing approaches[195]. Based on this technique comparison of transcriptome from an organism in different situations such as infected and non-infected (tissues) allows the identification of transcripts that are more related with the infection process. In this work we performed infected *vs* control transcriptomic analysis of *Anopheles* SG in order to identify genes that can be related to transmission of malaria. In this work plan the start point was data from RNA sequencing of infected *vs* control *Anopheles* SG. RNA-seq results compile information of UniProt, RefSeq, Gene Ontology and NCBI Taxonomy Enzyme databases. Information of each transcript was proceeded, considering *p* value <0.05 as significant data and BCV value of 0.2 considering a report from Era7 company and the research data obtained by Kokoza and Ahmed [183].

Transcripts were organised into classes according to their molecular function, cellular component, biological process and domains obtained by comparison with the mentioned databases. The most represented functional classes were those discriminated below:

- Cell function: having transcripts that play a role in cell adhesion, cell migration, apoptosis, etc;
- Chemosensory: including transcripts having sensorial functions, such as odorant proteins;
- Detoxification: containing genes/proteins implicated in the removal of toxic substances;
- Diverse: containing transcripts that have molecular functions that can be associated to more than one functional class;

- Immunity: related to transcripts that play a role in biological defences to fight infection, biological invasion, inflammation;
- Metabolism: correlated to transcripts that are associated to chemical transformations or enzyme-catalyzed reactions that form metabolites;
- Proteolysis: linked to transcripts that allow the breakdown of proteins into smaller polypeptides or amino acids, i.e., hydrolysis of the peptide bond by proteases, and others;
- Recognition: including transcripts that recognize molecules of signalling pathways, and others;
- Regulatory: containing transcripts related to processes that occur within a cell to maintain homeostasis, either at signalling or at genetic level;
- RTT: transcripts in this class appear related to processes such as replication, transcription and translation;
- Signalling: contains transcripts related to emission of signals that allow cellular activities and coordinates cell actions;
- Structural: has transcripts associated to cell structure and other components;
- Transport: includes transcripts that play a role in transport of molecules by channels, motor proteins, or others;
- Unknown: in this functional class are represented transcripts that do not have information to infer about it.

The amount of significant transcripts were analysed depending on general distribution of each functional class (Figure 22A) and on its over and under expression in each functional class (Figure 22B).

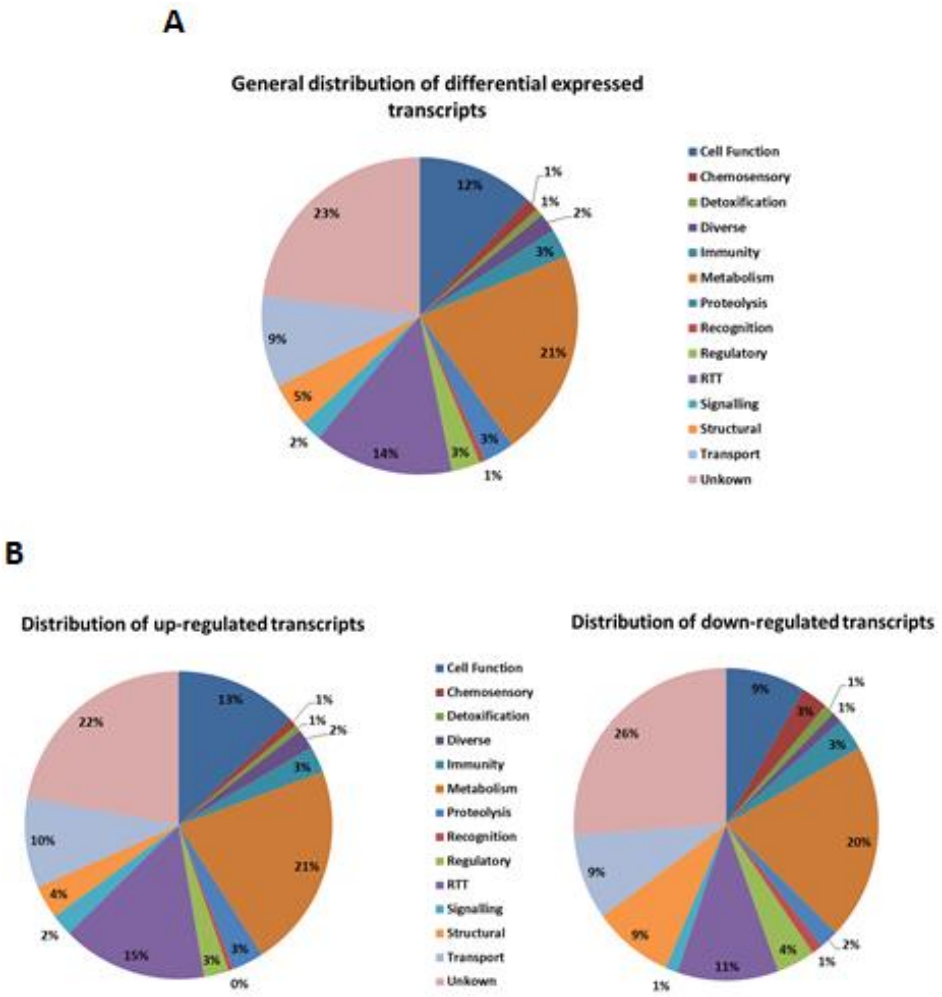



**Figure 22 - Results from functional grouping of *A. stephensi* transcripts from RNA-seq.** A) Results of general distribution of differential expressed transcripts by functional classes. B) Distribution results of up and down-regulated transcripts by functional classes.

The analysis of RNA sequencing data showed that in *Anopheles stephensi* salivary glands a total of 2536 genes were found to be significantly differential expressed in response to *P. berghei* infection for the conditions above mentioned. From these, 1996 (79%) were upregulated and 540 (21%) were found to be down regulated (data not shown).

Metabolism, RTT, cell function and transport were the functional classes with more transcripts considering a general distribution, while recognition, detoxification and chemosensory presented less number of transcripts in *A. stephensi* SG (Figure 22A). Metabolism and transport includes 539 and 238 transcripts, respectively; and chemosensory and recognition contains 33 and 16 transcripts (data not shown).

The most overexpressed transcripts were found in classes such as metabolism, replication-transcription-translation (RTT), cell function and transport and the same was observed for the most under expressed transcripts.

About 429 transcripts of metabolism class were up-regulated while the rest 110 transcripts were down-regulated (data not shown). This predominance of up-regulated transcripts is common between classes. Besides that, regulatory, immunity, proteolysis and chemosensory were classes with low percentage of transcripts.

A hypergeometric test was set by GOstas analysis, obtaining transition metal ion binding (GO:0046914) and cytoplasm (GO:0005737) as the over and under represented GO terms in this RNA-seq set, respectively. These GO terms were presented in metabolism, cell function and transport (data not shown) which are the most predominant functional classes of this RNA-seq result.

The aim, at this point, was to check the expression of genes related to various functional classes and with higher and lower expression. Genes associated with metabolism, RTT, cell function and transport functional classes were also selected due to their relevance in mosquito SG-*Plasmodium* interactions.

### 3.1.2. Confirmation of differential expression by qPCR

Mice were used to infect mosquitoes, when the parasitaemia reached 5-10% (Figure 23) and the presence of microgametocytes capable of exflagellation was observed.

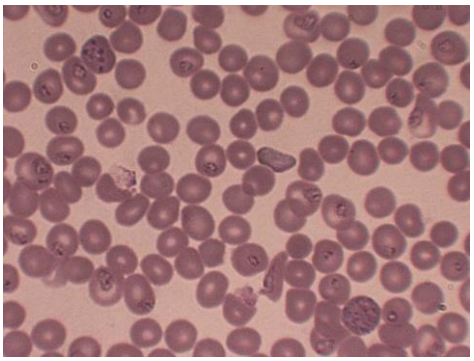


**Figure 23 - Parasitaemia of 5%.** Presence of several blood stages of *Plasmodium berghei*, examined with a 100x oil immersion objective of a light-microscope (Original from the author Joana Couto).

From the all mosquito groups used in different assays, a small fraction was picked-up from each cage and the infection confirmation carried on. As represented in Figure 24, 8-9 days post-blood meal (PBM) the presence of oocysts was confirmed by observing midgut tissues under a UV light microscope.

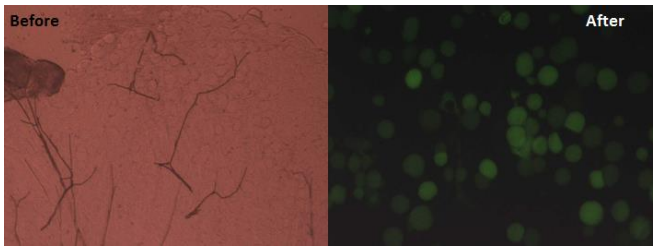



**Figure 24 - Presence of oocysts**. In left are represented a part of a midgut before exposition, and in right after exposed to UV light (Original from the author Joana Couto).

For groups showing a high level of infection (more than 60%), salivary glands (SG) were dissected (Figure 25) 18-19 days PBM to proceed to RNA extraction. Groups showing low infection rates and/or absence of exflagellation were rejected.

**Figure 25 - Salivary glands dissected** (Original from the author Joana Couto)**.**

RNA was quantified and analysed for purity on a ND-1000 *NanoDrop®* spectrophotometer (NanoDrop ND1000, Thermo Fisher Scientific, MA) (Appendix IV).

Integrity of RNA was evaluated by *Agilent 2100 Bioanalyser® system* (Agilent Technologies) (Appendix V). These data can be displayed as a gel-like image and/or as electropherogram. Considering all electropherograms, only one sample (RNA Total) presented

some degradation, which was not a concern once this sample was used to perform a standard curve for relative quantification.

The real-time polymerase chain reaction (qPCR) is an accurate technique used to evaluate gene expression in a defined biological system.

qPCR was used to measure and further validate results obtained from the transcriptomic analysis of the SG from *A. stephensi*. Based on the ontological analysis, a total of 13 genes identified as differentially expressed by RNA-seq, 10 upregulated and 3 down regulated in response to infection, were analysed by qPCR and results compared with those obtained from RNA-seq, as referred above (Appendix VI).

To evaluate gene expression level by qPCR, a first step concerned to the detection of the presence of primer dimers. In these experiments, as expected, no primer dimers were detected when inspecting the melting curves. Primer pairs were chosen when displaying amplification efficiency greater than 90%, except for AGAP004524, SCRB6 and INOSITOL genes that showed an efficiency of 89.5 ± 0.3%. The observation of melt curves also indicated that only one amplicon was produced for each gene amplified (Figure 26).

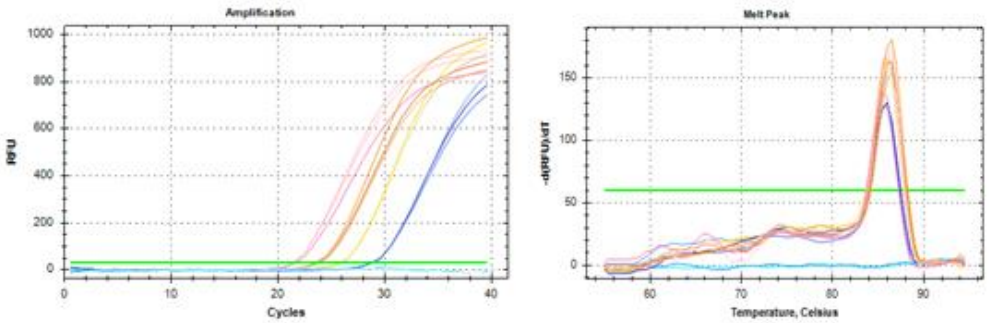



**Figure 26 – Amplification and melt peak chart of a standard curve.**

Reference genes are genes showing a constitutive expression across samples and were utilized to quantify changes in gene expression. Therefore, the relative expression results obtained were normalized against both *A. gambiae* ribosomal protein S7 (RPS7) and elongation factor (ELF) genes. A 2 delta Ct (ΔΔCt) method was performed to proceed an evaluation of the correlation between the expression values detected by RNA-seq and qPCR

for the 13 genes tested by calculating Pearson's correlation (Figure 27). Pearson value was slightly different between reference genes.

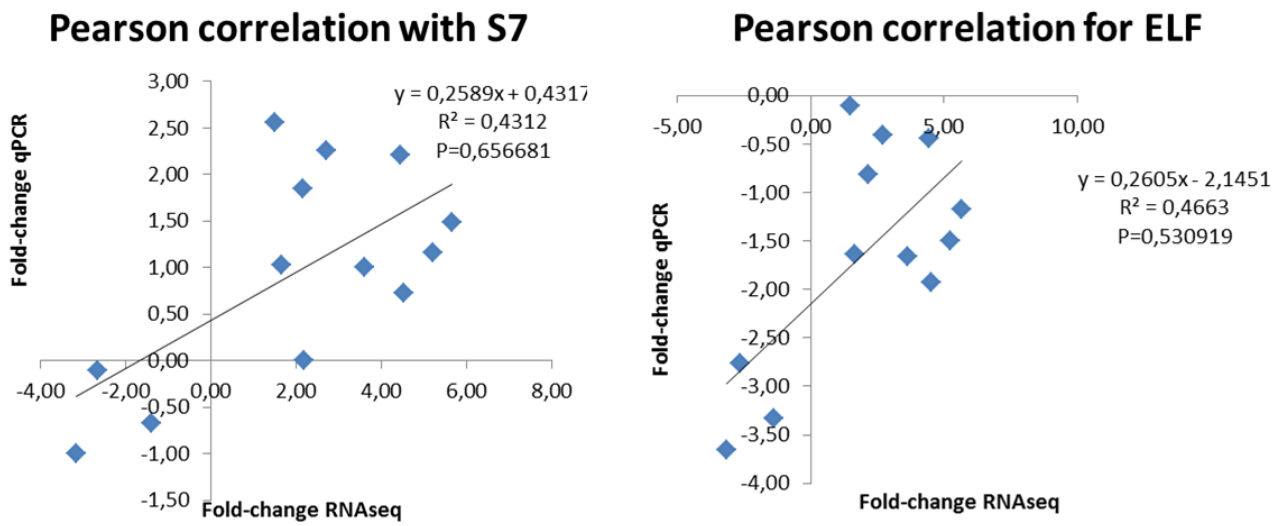



**Figure 27 - Correlation between qPCR and RNA-seq results.** The Pearson's correlation coefficient was calculated using the Excel software; a strong positive relationship correlation was observed (P = 0,656681 for S7 and *P* = 0,530919 for ELF).

Pearson correlation shows whether two variables are related or not and how strongly. *P* value was determined to find the strength between the variables between qPCR and RNA-seq. When RPS7 was used as a reference gene, the results have shown that the *P* value was 0.67; contrary, when ELF was used the *P* value was 0.53. Working with RPS7 as reference gene, fold-changes of qPCR have a strong correlation to those obtained by RNA-seq results than when assaying with ELF (Appendix VI).

## 3.2. *In vivo* gene knockdown in *Anopheles stephensi* by RNA interference

RNA interference methodology is widely used to knockdown genes that, by targeting mRNA, can provide a more information about gene function. After a deep analysis of transcripts, having in consideration the expression level and gene ontology classification, one transcript (ASTE009391) displaying a fold-change of 2.19 in RNA-seq results, showing a

secondary active sulfate transmembrane transporter activity in molecular function and known that as being related to SLC26 proteins was chosen for further studies. This transcript with 641bp is a transmembrane transporter of compounds essential to parasite survival and, mainly due to these characteristics; it was selected to further silencing assays.

In an attempt to identify physical interactions between *prestin* and other proteins, for a better understanding of intracellular signaling pathways and biochemical processes; following protein-protein interactions prediction (Figure 28) was performed.

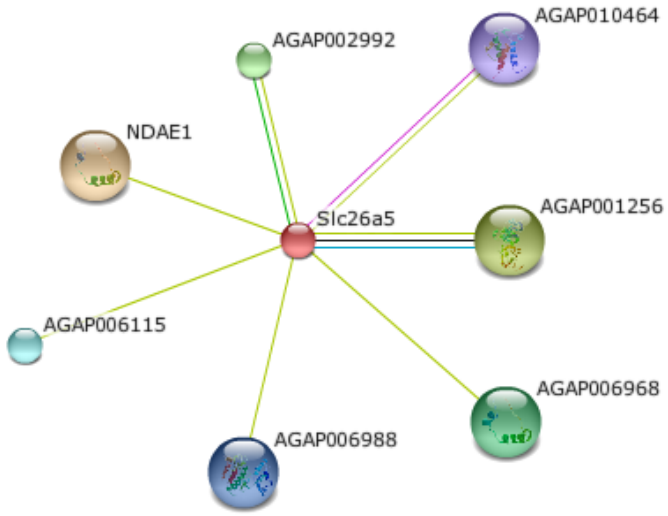



**Figure 28 – STRING analysis.** STRING database (http://string-db.org) can give a network of predicted associations for *prestin* (AGAP010725, an orthologue of ASTE009391). The network nodes are proteins: NDAE1 (AGAP009736), AGAP002992, AGAP006968, AGAP006115, AGAP001256, AGAP006988 and AGAP010464. The edges represent the predicted functional associations. A red line indicates the presence of fusion evidence; a green line - neighbourhood evidence; a blue line - co-occurrence evidence; a purple line - experimental evidence; a yellow line – text mining evidence; a light blue line - database evidence; a black line - co expression evidence.

STRING analysis showed proteins that interact with *prestin* are related to metabolism of sulfate, which it is an ion transported by *prestin*. Besides that, *NDAE1* protein that interact with *prestin*, share the same molecular function. Despite of this result, *prestin* may have an important role in the transport of compounds essential to parasite survival, as mentioned above, and to analyse this function is necessary to perform RNAi assays.

The ASTE009391 gene (*prestin*) was silenced using RNA interference technology as referred above. The effect of gene knockdown-RNAi mediated in *P. berghei* infection in *A. stephensi* mosquitoes was evaluated.

The *prestin* gene was amplified by PCR using the primers with T7 promoter (Forward–5' TAATACGACTCACTATAGGGAGAGGAAGGGCATGAGAGTGGTA-3' and Reverse–5' TAATACGACTCACTATAGGGAGAAGTACACCAGCACCGGAAAG-3') and amplicons analysed on a 0.5X TBE, 1.5% (w/v) agarose gel (Figure 29A). Amplification at optimal conditions resulted in different PCR products. To assure which band was the one related with *prestin*, all bands were purified and sequenced. Sequences from both primers forward and reverse were aligned using BLAST and ClustalW tools from Vectorbase resource (having *prestin* (ASTE009391) sequence from *A. stephensi*). Only the first purified band (Figure 29 A and B) showed a high identity of 99.8% (reverse primer) and 96.3% (forward primer) to *prestin* sequence (Appendix VII) and therefore, DNA corresponding to this amplification was selected to use as template to synthetize *ds*RNA. The resulting *ds*RNA was purified, quantified and checked on a 0.5X TBE, 1.5% (w/v) agarose gel (Figure 29C). Concentration of purified *ds*RNA was approximately 3.5 mg/µL.

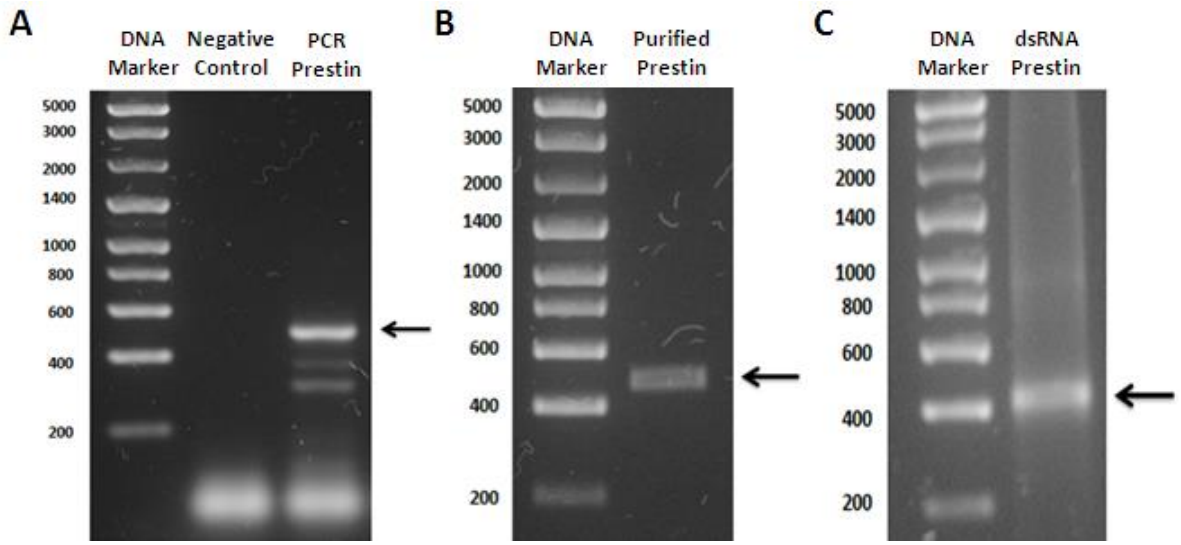



**Figure 29 - Electrophoresis gel of *prestin* PCR, purified PCR product and *prestin ds*RNA.**

### 3.2.1. *Double stranded* RNA injection in *Anopheles stephensi*

To infer about the function of *prestin* and its eventual relation with *P. berghei* infection, knockdown experiments were carried on *A. stephensi* mosquitoes. For knockdown, 69 nl (4µg/ml) of *ds*RNA were injected in each female mosquito previously infected with *P. berghei.* Controls were inoculated in the same conditions with the reference gene *β2M*. Three independent experiments were performed using for each group (*prestin* and *β2M*) 350 mosquitoes. After injection, the viable mosquitos were counted every day until the 18st day PBM (Figure 30).

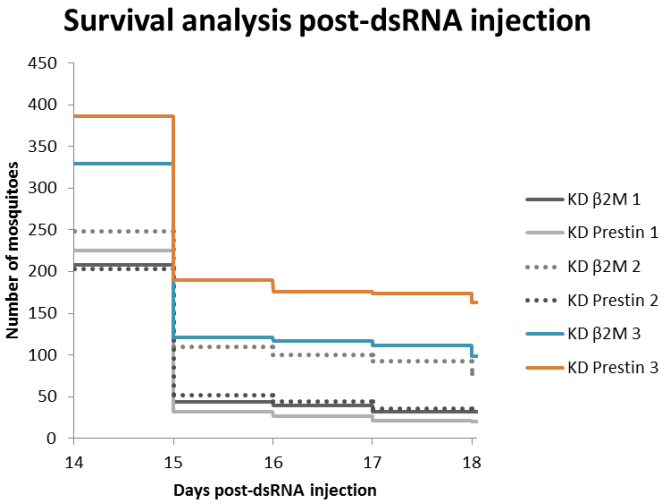



**Figure 30 - Survival analysis post-*ds*RNA injection.** Day 14 represents the time of *ds*RNA injection and the other days represent days after injection. Survival analysis is represented for each condition of knockdown (KD*β2M* and KD*prestin*) of each pool (1, 2 and 3). Significant differences were present when comparing *prestin* to *β2M,* in pool 2 and 3 (*p* value of 0.0001 and 0.00001, respectively). In contrast with pool 1 that not shows significant differences between conditions (*p* value of 0.345). Significant differences were presented when comparing number of mosquitos before injection (Day 14) and the last day post-*ds*RNA injection (Day 18) (*p* value of 0.00001).

### 3.2.2. Gene knockdown assessment

The effect of gene knockdown on the infection level was assessed by evaluating the number of sporozoites. For this, SG were extracted on day 18 and homogenized in PBS, to further quantification in a haemocytometer. Figure 31 shows that no significant reduction of

sporozoites was obtained after gene knockdown, when analysing the all three experiments. However, the first pool of mosquitoes showed a reduction of 92% (reduction of $80.4x10^4$ to $6.2x10^4$ sporozoites/mL) (data not shown).

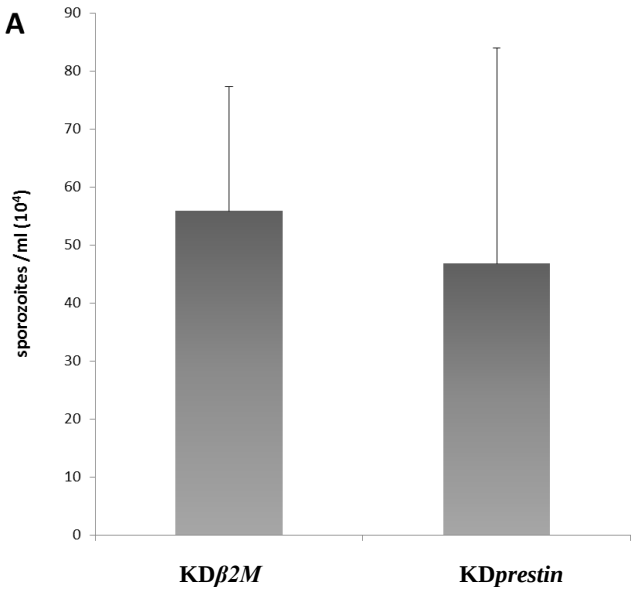



**Figure 31 - Sporozoites quantification for gene knockdown.** Number of sporozoites *per* mL is represented for each condition of knockdown (KD*β2M* and KD*prestin*) of each pool (1, 2 and 3). No significant difference was observed.

# Chapter 4: Discussion

Treatment and control of malaria disease have been the focus of several research projects aiming the improvement in antimalarial medicines through focused research[196].

The understanding of the interactions between the malarial parasite and the mosquito vector has recently been augmented by different methods including the culture sporogonic stages of some parasite species, new technologies, and genetic transformation of both parasite and vector, that will allow the development of new intervention strategies[197] and then elucidate the mechanisms of insect-vector specificity and monitor the spread and infection-route of the pathogen.

RNA-seq is a new approaches that have been increasingly used in many model organisms including mosquitoes [198,199]. In this work, the analysis of the transcriptome of *A. stephensi* female SG was performed to identify genes with differential expression in response to infection with *P. berghei*. In *A. stephensi*, a total of 2536 genes were found to be significantly differential expressed in response to *P. berghei* infection, from which 1996 (79%) were upregulated and 540 (21%) were found to be down regulated. This indicates an enrichment of upregulated genes in *A. stephensi* SG in response to malaria parasite, and this was also present in others experiences [200,201]. Metabolism, RTT, cell function and transport were the functional classes with more number of transcripts considering general distribution (confirmed by GOstat analysis); while recognition, detoxification, chemosensory and diverse presented less transcripts in *A. stephensi* (Figure 22A). The most overexpressed transcripts were found in classes such as metabolism, RTT, cell function and transport and the same was observed in the most under expressed transcripts. This distribution can be justified by the importance of such transcripts to infection or development of *Plasmodium*. In general, transcripts grouped in functional class of metabolism can be related to products involved in process to limit parasite development or production of metabolites needed to sporozoites progress[202,203]. RTT and cell function are related to cellular changes, such as cellular structural components that are modified during *Plasmodium* infection[204,205]. Transport is a class that include transcripts that can be induced under *Plasmodium* infection[202,206,207]. As some detox components, in *A.*

*gambiae* for example, are suppressed under *Plasmodium* infection, the same might happen in *A. stephensi* [208]. *A. stephensi* is more susceptible to *P. berghei* than *A. gambiae*, and that can be justified by minor number of transcripts belonged to immunity functional class[209] leading to a weak immune response of *A. stephensi*.

Looking closer within each functional class, it was possible to identify some transcripts having a significant expression level and an eventual relevant role in the infection process. In cell function class, the transcript AGAP010035 (Fold-change of 2.34), which is implicated with the *TOR* pathway, is related with the initiation of nutrient transport proteins synthesis in the mosquito fat body[210], essential for mosquito survival[211]. In chemosensory class, CPIJ017114 (Fold-change of 3.64) and AGAP002079 (Fold-change of 1.93) belong to *arrestin* family, and some authors [212,213] have suggested that altering olfactory genes could lead to a reduction of malaria transmission and other globally important medical and veterinary diseases. CYP6AA3 (Fold-change of 1.08) and other P450 cytochromes genes are presented in RNA-seq, included in detoxification class, once it is a response to malaria infection by *Anopheles* and are implicated with insecticide resistance[214]. In Immunity class, there are several groups of transcripts that interact each other and play a central role in immune response of *Anopheles* mosquito. For example, AGAP006974 (*TOLL9*; Fold-change of 2.69) belongs to *Toll* pathway, which is regulated by AGAP009515 (*REL1*; Fold-change of 2.59) and AGAP006747 (*REL2*; Fold-change of 2.12) that are NFκB transcripton factors-like[215]. Although the *Toll* pathway limits the development of *P. berghei*, *P. falciparum* and/or *P. vivax* [216], the immune deficiency (*Imd*) signalling pathway is the most effective pathway against the human malaria parasite and that is why are attractive candidates for genetic modification to create a mosquito with an immune response that overwhelms the parasite[217]. Also influenced by *REL1* and *REL2, TOR* pathway, which e.g. includes transcript AGAP010035 (Fold-change of 2.34) can be affected by signals arising from the parasite's invasion resulting in an immune response[211]. Other groups of transcripts are also related with Immunity class such as *leucine-rich repeat* (*LRR*) domains family, that also influences the survival of *Plasmodium* parasites[218], *clip domain serine proteases* (*CLIPs*), implicated in extracellular proteolytic cascades that can promote *Plasmodium* ookinetes elimination in *Anopheles* midgut[219] and *thioester-containing protein 1* (*TEP1*) that after cleaved, binds to the surfaces of bacteria and

later stages of parasites through the conserved thioester bond influencing pathogens clearance by phagocytosis, *in vitro* [220]. Related to this class, the proteolysis class was found upregulated in mosquitoes SG infected with *P. berghei*, associated to serine-type endopeptidases, that may be involved in proteolytic events during blood-feeding on vertebrate hosts or in digestion of extracellular matrix components [221,222]. Metabolism class includes 21.25% transcripts (539 transcripts, 429 upregulated and 110 down regulated) with a documented crucial role in malaria infection. Some examples are: AGAP004175 (fold-change of 4.17) that has *orotate phosphoribosyltransferases* (*OPRT*) activity, allowing *de novo* biosynthesis of pyrimidines required by *P. falciparum*[223]; and AGAP003184 (fold-change 2.67) that are related to heme-biosynthetic pathway, augment *Plasmodium* survival[224,225]. Of 238 transcripts included in transport class, ion transport transcripts were predominant, up-regulated transcripts related to *solute carrier family* (*SLC*) proteins, such as AGAP007054 (*SLC13*), AGAP007743 (*SLC16*), AGAP005405 (*SLC 39*), AGAP000097 (*SLC 25*), AGAP005537 (*SLC 35*), AGAP010725 (*SLC26*) and *NDAE1* (*SLC4*) were also present showing a high expression level.

Real time PCR enables both detection and quantification of gene expression being used to confirm RNA-seq results. Quantification can be either an absolute number of copies or a relative amount when normalized to DNA input or additional normalizing genes [226]. For qPCR assays, two reference genes candidates were selected, the *A. gambiae* ribosomal protein S7 (RPS7) and elongation factor (ELF) genes. During this study, Pearson correlation was determined to show whether qPCR and RNA-seq results were related or not and how which was the most suitable reference gene amongst RPS7 and ELF, to use in qPCR normalization, i.e. genes showing a constitutive expression across samples. Comparing the results of the two genes, RPS7 was considered to be a better choice once the *P* value for this gene showed a stronger correlation than ELF (Figure 27) and with a constitutive expression.

Moreover, as can be seen in Appendix VI, fold-changes from qPCR normalized against RPS7 demonstrated more homogenous results than ELF, confirming RPS7 as a better candidate as a reference gene in future studies.

As referred above, a list of 20 genes showing to be differentially expressed was chosen to confirm RNA-seq results using the qPCR approach. These genes were selected based on the

ontological analysis and expression level in response to infection (Appendix VI). From the 13 genes with confirmed expression, ASTE009422 having an arrestin domain showed to be highly downregulated when assayed by both RNAseq and qPCR techniques, agreeing to the fact that this gene is implicated with sensory pathways[227] being present in high levels in male mosquitoes to regulate olfactory-driven behaviours, and less expressed in females[228].

Among the upregulated genes those chosen included paxillin (ASTE007758), a gene related to metabolism (ASTE004032), basic helix-loop-helix (ASTE011434), chondroitin 4-sulfotransferase (ASTE007038),clip domain serine proteases ASTE008188 (*CLIPA4*), serpin ASTE001475 (*SRPN6*) and ASTE009391 (*prestin*).The first gene is thought to encode a cytoskeletal actin binding protein that allows focal adhesion [222229,230] and might be related with P. berghei adhesion in the salivary glands and invasion. ASTE004032 that belong to the metabolism functional class was expected to be up-regulated, once metabolism is crucial to development of malaria parasite and vector as well. ASTE011434 (bHLH), that is devoid in apicomplexan organism [231], is upregulated in *Anopheles* and, interesting, is highly expressed in *Artemisia annua L.* and implicated in the production of artemisin[232].

In this RNA-seq data we found, as well, the chondroitin 4-sulfotransferase (ASTE007038) transcript upregulated in SG. The transport of chondroitin glycosaminoglycans in the apical midgut microvilli of *A. gambiae* is mediated by chondroitin 4-sulfotransferase, leading to a better adhesion for parasite in to the midgut[233,234]. ASTE007038 might also be related with *P. berghei* adhesion in the SG once chondroitin glycosaminoglycans are transported through these tissues. Regarding the CLIP family that includes *CLIPA4228*, it was previously described[235] that they play an essential role of activation or suppression during a specific immune process defined as melanization[236], consisting in a deposition of melanin allowing lyse of parasites[237]. An upregulated gene ASTE001475[238], also called *SRPN6*, has been involved in insect immunity facilitating invasion[239] or stimulating anti-parasitic responses[240].

Within the transport functional class, SG transport proteins has shown to be important for parasite development and for sporozoite-SG interactions as parasite factors involved in the interaction with insect cells[241]. Membrane transport proteins are likely to be located on the outer surface of the parasite cell and mediate the passage across the membrane bilayer of

specific molecules and/or ions serving a diverse range of physiological roles, mediating the uptake of nutrients into cells, the removal of metabolic wastes and xenobiotics (including drugs), and the generation and maintenance of transmembrane electrochemical gradients [242]. In some cases the native capacities of the host cell membrane probably suffice to meet the parasite needs [243].

For this reason some transmembrane transporters can contribute to the permanence of the parasite in the host. Malaria transmission requires *Plasmodium* invasion of the *Anopheles* mosquito midgut and then to salivary glands. In midgut phase is involved the adhesion of *Plasmodium* to chondroitin sulfate ($ChSO_4$) [244], midgut epithelial ligand. Sulfate ($SO_4^{2-}$) for $ChSO_4$ biosynthesis is supplied by $SO_4^{2-}$ uptake. *Slc26a5*, also known as Dipteran *prestin*, have a function as $Cl^-$ exchange systems for $HCO^{3-}$, oxalate$^{2-}$, $SO_4^{2-}$ and formate [245]. However, *prestin* can be found also in the salivary glands and gastric caeca[246]. Poorly understood at molecular level, even in Diptera (Drosophila and *Anopheles*), can be hypothesized that control of *Slc26a5* activity could influence parasite (*Plasmodium*) invasion[246].

Meeting the main purpose of this research work, which is to find new genes or proteins capable of decreasing mosquito vector capacity and considering the importance of membrane proteins of *Anopheles spp.* SG for parasite invasion and transmission and the prevalence of *SLC* proteins in transport class, the AGAP010725 transport gene was selected for further characterization. The choice of this gene was based on its fold-change, molecular function, pathway and length. ASTE009391 (*A. stephensi* ortholog of AGAP10725) was selected, including a fold-change of 2.19 and a secondary active sulfate transmembrane transporter activity (as confirmed by *TMHMM Server* v. 2.0 [168]), related to *SLC26* proteins. This transcript with 641bp was found up regulated in RNA-seq with a fold-change of 2.19, however in qPCR experiments show a fold-change of 0.01. Looking through qPCR results from the three pools analysed, the fold-change variation was between -0.26 and 0.31 (data not shown). These differences in the expression of *prestin* can be related to heterogeneity within and between the pools, to a short biological half-life of *prestin* mRNA in the SGs, to extraction procedure which was performed at slight different time points (hours), or to alternative splicing of the gene products that may have importance in tissue translation selectivity [247]. Although, the

results of qPCR were less clear regarding to the upregulation of the gene *prestin* in comparison with the RNA-seq, this study proceeded to RNAi assays.

During silencing assays, statistical analysis showed that the number of deaths in mosquitoes was significantly affected by the knockdown of *prestin* (P>0.001) in pool 2 and 3 (*p* value of 0.0001 and 0.00001, respectively), in comparison with the *β2M* controls. Contrarily, pool 1 did not show significant differences between both conditions (*p* value of 0.345). Significant differences were presented when comparing number of mosquitos before injection (Day 14) and the last day post-*ds*RNA injection (Day 18) (*p* value of 0.00001) (Figure 30).

Considering the sporozoites in the SG evaluated by microscopy, statistical analysis have shown that in the overall results for pools 1, 2 and 3, the presence and number of this parasite stage is not significantly influenced by the injection of *prestin ds*RNA, when compared with control. However, analysing the number of sporozoites of each pool separately it was observed that only pool 1 had a considerable reduction of this number ($6.2 \times 10^4$ sporozoites/ mL) in comparison with the control ($80.4 \times 10^4$ sporozoites/ mL).

To further elucidate about all results obtained during this study, namely the influence of *prestin* knockdown in the survival of the mosquitoes and the presence and number of sporozoites in the SGs, it would be fundamental to conduct qPCR assays to determine the differential expression of this gene. Also, it is important to further analyse the differential expression of the *off-target* gene (ASTE006714) once the similarity with *prestin* is high (Appendix VIII).

# Chapter 6: Conclusions and future perspectives

It is well known and largely disseminated by various research institutions and by researchers themselves that exploring new targets to develop transmission blocking vaccines to be used together with other control measures is crucial to eradicate and control malaria disease.

A comparison of the transcriptome of *A. stephensi* salivary glands non-infected and *P. berghei* infected was performed and a catalogue of differentially expressed genes assembled from transcriptomic data analysis constituting a source of fundamental information that can be used by all the scientific community to select potential targets for malaria candidate vaccines.

Among the differentially expressed genes analysed by transcriptomics and qPCR, an up regulated $Cl^-$ transmembrane transporter gene, *prestin*, was investigated after *P. berghei* infection. This interesting gene encodes for a transmembrane protein that might be related to the $Cl^-$ and $HCO_3^{2-}$ exchange in the mosquito SG, allowing the optimal conditions for parasite development.

Future experiments should focus in the confirmation of the expression of *prestin* after gene knockdown by qPCR. This approach will contribute for a better understanding about this gene function in the mosquito SG during infection by *P. berghei*, elucidating the results here obtained, such as the influence of *prestin* knockdown in the survival of the mosquitoes and the presence and number of sporozoites in the SGs.

In an attempt for a better understanding of intracellular signaling pathways and biochemical processes related to *prestin*, NDAE1 function must be evaluated once it interact with *prestin* and share the same molecular function. This can be accomplished by evaluating NDAE1 protein function *via* knockdown and the effect of silencing of both genes on *Anopheles* infected by the *Plasmodium* parasite. Immunolocalisation assays of these proteins in the SG allied with RNAi data would contribute for a better knowledge regarding their function and interaction with parasites.

# Bibliography


1. Sharakhova, M. V *et al.* A physical map for an Asian malaria mosquito, *Anopheles stephensi*. *American Society of Tropical Medicine and Hygiene*.**83,** 1023–7 (2010).

2. WHO. WHO | Malaria. (2014). at <http://www.who.int/mediacentre/factsheets/fs094/en/>

3. CDC. CDC - Malaria. (2014). at <http://www.cdc.gov/malaria/>

4. Holt, R. A. *et al.* The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298,** 129–49 (2002).

5. Faulde, M. K., Rueda, L. M. & Khaireh, B. A. First record of the Asian malaria vector *Anopheles stephensi* and its possible role in the resurgence of malaria in Djibouti, Horn of Africa. *Acta Tropica.* **139,** 39–43 (2014).

6. Kappe, S., Kaiser, K. & Matuschewski, K. The Plasmodium sporozoite journey: a rite of passage. *Trends Parasitology.* **19**(3)**,** 135–143 (2003). at <http://www.sciencedirect.com/science/article/pii/S1471492203000072>

7. Wang, S. & Jacobs-Lorena, M. Genetic approaches to interfere with malaria transmission by vector mosquitoes. *Trends Biotechnology.* **31,** 185–93 (2013).

8. Summerton, J. E. Morpholino. siRNA, and S-DNA compared: impact of structure and mechanism of action on off-target effects and sequence specificity. *Current Topics in Medicinal Chemistry.* **7,** 651–60 (2007).

9. Pedersen, A. B. & Fenton, A. Emphasizing the ecology in parasite community ecology. *Trends in Ecology & Evolution.* **22,** 133–9 (2007).

10. World Health Organization. WHO | Malaria. (2014). at <http://www.who.int/topics/malaria/en/>

11. Calouste Gulbenkian Foundation. Medical Microbiology. *Calouste Gulbenkian Foundation*. (Vol1, No. 5, pp. 1064-1097) Tipografia António Coelho Dias, 1988.

12. Canali, S. Researches on thalassemia and malaria in Italy and the origins of the 'Haldane hypothesis'. *Medicina nei secoli.* **20,** 827–46 (2008).

13. Bairwa, M. *et al.* Malaria vaccine: a bright prospect for elimination of malaria. *Human Vaccines & Immunotherapeutics.* **8,** 819–22 (2012).

14. Cox, F. E. History of the discovery of the malaria parasites and their vectors. *Parasites & Vectors.* **3,** 5 (2010).

15. Camargo, E. P. Malária, maleita, paludismo. *Ciência e Cultura.* **55,** 26–29

16. CDC. CDC - Malaria - About Malaria - Where Malaria Occurs. (2010). at <http://www.cdc.gov/malaria/about/distribution.html>

17. WHO. World Malaria Report 2014. (2014). at <http://www.who.int/malaria/publications/world_malaria_report_2014/en/>

18. CDC. CDC - Malaria - About Malaria. (2012). at <http://www.cdc.gov/malaria/about/index.html>

19. Thera, M. A. & Plowe, C. V. Vaccines for malaria: how close are we? *Annual Review of Medicine.* **63,** 345–57 (2012).

20. Adl, S. M. *et al.* The revised classification of eukaryotes. *Journal of Eukaryotic Microbiology.* **59,** 429–93 (2012).

21. Singh, B. & Daneshvar, C. Human infections and detection of *Plasmodium knowlesi*. *Clinical Microbiology Reviews.* **26,** 165–84 (2013).

22. White, N. J. *Plasmodium knowlesi*: the fifth human malaria parasite. *Clinical Infectious Diseases.* **46,** 172–3 (2008).

23. Wang, R., Smith, J. D. & Kappe, S. H. I. Advances and challenges in malaria vaccine development. *Expert Reviews in Molecular Medicine.* **11,** e39 (2009).

24. Goodman, A. L. *et al.* The utility of *Plasmodium berghei* as a rodent model for anti-merozoite malaria vaccine assessment. *Scientific Reports.* **3,** 1706 (2013).

25. Harbach, R. Culicidae | Mosquito Taxonomic Inventory. (2008). at <http://mosquito-taxonomic-inventory.info/simpletaxonomy/term/6045>

26. Sinka, M. E. *et al.* A global map of dominant malaria vectors. *Parasites & Vectors.* **5,** 69 (2012).

27. Sokhna, C., Ndiath, M. O. & Rogier, C. The changes in mosquito vector behaviour and the emerging resistance to insecticides will challenge the decline of malaria. *Clinical Microbiology and Infection.* **19,** 902–7 (2013).

28. Sinka, M. E. *et al.* The dominant Anopheles vectors of human malaria in the Asia-Pacific region: occurrence data, distribution maps and bionomic précis. *Parasites & Vectors.* **4,** 89 (2011).

29. Sinka, M. E. *Anopheles mosquitoes - New insights into malaria vectors*. (InTech, 2013). doi:10.5772/3392

30. CDC. CDC - Malaria - About Malaria - Biology - Mosquitoes - *Anopheles* Mosquitoes. (2012). at <http://www.cdc.gov/malaria/about/biology/mosquitoes/>

31. World Health Organization. Malaria in tropics disease research. 13th programme report. *WHO, Geneva* (1997). at <https://scholar.google.pt/scholar?hl=en&as_sdt=0,5&as_ylo=1997&q=author%3Awho+malaria+1997#5>

32. Thielman, A. & Hunter, F. Photographic key to the adult female mosquitoes (Diptera: Culicidae) of Canada. *Canadian Journal of Arthropod Identification.* (2007). at <http://www.biology.ualberta.ca/bsc/ejournal/th_04/th_04.html>

33. Lamb, T. J., Brown, D. E., Potocnik, A. J. & Langhorne, J. Insights into the immunopathogenesis of malaria using mouse models. *Expert Reviews in Molecular Medicine.* **8,** 1–22 (2006).

34. Calouste Gulbenkian Foundation. Parasitologia Veterinária. *Calouste Gulbenkian Foundation* (Vol 2, No. 3, pp. 302-713). Sociedade Tipografica Lda., 1988.

35. Lea, A. O., Dimond, J. B. & Delong, D. M. Role of Diet in Egg Development by Mosquitoes (*Aedes aegypti*). *Science.* **123,** 890–891 (1956).

36. Hoffman, S. L. Adaptive clinical trials of three *Pf*SPZ products for development of a whole sporozoite vaccine that prevents *Plasmodium falciparum* infection, disease and transmission. *Malaria Journal.* **11,** O48 (2012).

37. Sauerwein, R. W., Roestenberg, M. & Moorthy, V. S. Experimental human challenge infections can accelerate clinical malaria vaccine development. *Nature Reviews Immunology*. **11,** 57–64 (2011).

38. Miller, L. H., Baruch, D. I., Marsh, K. & Doumbo, O. K. The pathogenic basis of malaria. *Nature.* **415,** 673–9 (2002).

39. Hoffman, S. L. *et al.* Protection of humans against malaria by immunization with radiation-attenuated *Plasmodium falciparum* sporozoites. *The Journal of Infectious Diseases.* **185,** 1155–64 (2002).

40. Wang, S. & Jacobs-Lorena, M. Genetic approaches to interfere with malaria transmission by vector mosquitoes. *Trends Biotechnology.* **31,** 185–93 (2013).

41. Ferrolho, J., Domingues, N. & Domingos, A. The Role of Regulatory CD4 CD25 T Cell Subset in Host Homeostasis during Protozoan Infection: An Overview. *researchgate.net* at <http://www.researchgate.net/profile/Joana_Ferrolho3/publication/275021971_The_Role_of_Regulatory_CD4CD25_T_Cell_Subset_in_Host_Homeostasis_during_Protozoan_Infection_An_Overview/links/552ea51d0cf2acd38cbbd45e.pdf>

42. Basilico, N., Spaccapelo, R. & D'Alessandro, S. Malaria Diagnosis, Therapy, Vaccines, and Vector Control. *Human and Mosquito Lysozymes.* 19-43 (2015). at <http://link.springer.com/chapter/10.1007/978-3-319-09432-8_2>

43. WHO. World Malaria Report 2013. 286 (2013). doi:10.1038/nature.2013.13535

44. Wilson, M. L. Malaria rapid diagnostic tests. *Clinical Infectious Diseases.* **54,** 1637–41 (2012).

45. Bisoffi, Z., Gobbi, F., Buonfrate, D. & Ende, J. Van den. Diagnosis of malaria infection with or without disease. *Mediterranean Journal of Hematology and Infectious Diseases* **4,** 2012036 (2012).

46. Mouatcho, J. C. & Goldring, J. P. D. Malaria rapid diagnostic tests: challenges and prospects. *Journal of Medical Microbiology.* **62,** 1491–505 (2013).

47. Sharma, Y. Structure and possible function of heat-shock proteins in Falciparum malaria. *Comparative Biochemistry and Physiology - Part B.* **102.3**, 437-444 (1992). at <http://www.sciencedirect.com/science/article/pii/030504919290033N>

48. Minota, S. & Cameron, B. Autoantibodies to the constitutive 73-kD member of the hsp70 family of heat shock proteins in systemic lupus erythematosus. *Journal of experimental medicine.* **168.4**, 1475-1480 (1988). at <http://jem.rupress.org/content/168/4/1475.abstract>

49. Zhang, M., Hisaeda, H. & Kano, S. Antibodies specific for heat shock proteins in human and murine malaria. *Microbes and infection.* **3.5**, 363-367 (2001). at <http://www.sciencedirect.com/science/article/pii/S1286457901013910>

50. Na, B., Park, J., Lee, H. & Lin, K. Characterization of *Plasmodium vivax* heat shock protein 70 and evaluation of its value for serodiagnosis of tertian malaria. *Clinical and Vaccine Immunology.* **14.3**, 320-322 (2007). at <http://cvi.asm.org/content/14/3/320.short>

51. Jani, D., Nagarkatti, R., Beatty, W. & Angel, R. HDP—a novel heme detoxification protein from the malaria parasite. *PLoS Pathogens.* **4.4** (2008). at <http://dx.plos.org/10.1371/journal.ppat.1000053.g007>

52. Vinayak, S., Rathore, D. & Kariuki, S. Limited genetic variation in the *Plasmodium falciparum* heme detoxification protein (HDP). *Infection, Genetics and Evolution.* **9.2**, 286-289 (2009). at <http://www.sciencedirect.com/science/article/pii/S1567134808002566>

53. Sherman, I. Biochemistry of *Plasmodium* (malarial parasites). *Microbiological Reviews.* **43.4**, 453(1979). at <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC281489/>

54. Yuvaniyama, J. Insights into antifolate resistance from malarial DHFR-TS structures. *Nature Structural & Molecular Biology.* **10.5**, 357-365(2003). at <http://www.nature.com/nsmb/journal/v10/n5/abs/nsb921.html>

55. Kattenberg, J., Versteeg, I. & Migchelsen, S. New developments in malaria diagnostics: monoclonal antibodies against *Plasmodium* dihydrofolate reductase–thymidylate synthase, heme detoxification protein and glutamate rich proteins. *MAbs* (Vol 4, No. 1, pp. 120-126). Taylor & Francis (2012). at <http://www.tandfonline.com/doi/abs/10.4161/mabs.4.1.18529>

56. Jepsen, M. & Jogdand, P. The malaria vaccine candidate GMZ2 elicits functional antibodies in individuals from malaria endemic and non-endemic areas. *Journal of Infection.* **208,3,** 479-488 (2013). at <http://jid.oxfordjournals.org/content/208/3/479.short>

57. Higgins, S., Xing, K., Kim, H. & Kain, D. Systemic release of high mobility group box 1 (HMGB1) protein is associated with severe and fatal *Plasmodium falciparum* malaria. *Malaria Journal.* **12,** 105 (2013). at <http://www.biomedcentral.com/content/pdf/1475-2875-12-105.pdf>

58. Jr, R. B. A simple method to detect *Plasmodium falciparum* directly from blood samples using the polymerase chain reaction. *American Society of Tropical Medicine and Hygiene.* **46.4,** 416-426 (1992). at <http://europepmc.org/abstract/med/1575288>

59. Snounou, G., Viriyakosol, S. & Jarra, W. Identification of the four human malaria parasite species in field samples by the polymerase chain reaction and detection of a high prevalence of mixed infections. *Molecular and biochemical parasitology.* **58.2**, 283-292 (1993). at <http://www.sciencedirect.com/science/article/pii/0166685193900508>

60. Han, E. & Watanabe, R. Detection of four *Plasmodium* species by genus-and species-specific loop-mediated isothermal amplification for clinical diagnosis. *Journal of*

*Clinical Microbiology.* **45.8**, 2521-2528 (2007). at <http://jcm.asm.org/content/45/8/2521.short>

61. Iseki, H., Kawai, S. & Takahashi, N. Evaluation of a loop-mediated isothermal amplification method as a tool for diagnosis of infection by the zoonotic simian malaria parasite *Plasmodium knowlesi. Journal of Clinical Microbiology.* **48.7**, 2509-2514 (2010). at <http://jcm.asm.org/content/48/7/2509.short>

62. Lau, Y., Fong, M., Mahmud, R. & Chang, P. Specific, sensitive and rapid detection of human *Plasmodium knowlesi* infection by loop-mediated isothermal amplification (LAMP) in blood samples. *Malaria Journal.* **10.1**, 1-6 (2011). at <http://link.springer.com/article/10.1186/1475-2875-10-197>

63. Amina, K., Giuliana, G. & Prato, M. From control to eradication of malaria: the end of being stuck in second gear? *Asian Pacific Journal of Tropical Medicine.* **3,5**, 412-420 (2010). at <http://www.sciencedirect.com/science/article/pii/S1995764510601010>

64. Ghebreyesus, T., Lynch, M. & Coll-Seck, A. The Global Malaria Action Plan for a Malaria-Free World. *Geneva, Switzerland: Roll Back Malaria Partnership. 274* (2008). at <http://scholar.google.pt/scholar?q=The+Global+Malaria+Action+Plan—for+a+malaria+free+world.&btnG=&hl=en&as_sdt=0,5#0>

65. WHO. WHO | Country antimalarial drug policies: by region. (2012). at <http://www.who.int/malaria/am_drug_policies_by_region_afro/en/>

66. Antinori, S., Galimberti, L., Milazzo, L. & Corbellino, M. *Plasmodium knowlesi*: the emerging zoonotic malaria parasite. *Acta Tropica.* **125,2,**191-201 (2013). at <http://www.sciencedirect.com/science/article/pii/S0001706X12003440>

67. Price, R., Douglas, N. & Anstey, N. New developments in *Plasmodium vivax* malaria: severe disease and the rise of chloroquine resistance. *Current Opinion in Infectious Diseases.* **22,5,** 430-435 (2009). at <http://journals.lww.com/co-infectiousdiseases/abstract/2009/10000/new_developments_in_Plasmodium_vivax_malaria_.3.aspx>

68. Barber, B. & William, T. A prospective comparative study of knowlesi, falciparum, and vivax malaria in Sabah, Malaysia: high proportion with severe disease from *Plasmodium. Clinical Infectious Diseases.* **56,3**, 383-397 (2013). at <http://cid.oxfordjournals.org/content/56/3/383.short>

69. White, N. J. *et al.* Malaria. *Lancet*. **383,** 723–35 (2014).

70. Faucon, Benoît; Murphy, Colum; Whalen, J. Africa's Malaria Battle: Fake Drug Pipeline Undercuts Progress. *The Wall Street Journal* (2013). at <http://www.wsj.com/articles/SB10001424127887324474004578444942841728204>

71. Mackey, T. K. & Liang, B. A. The global counterfeit drug trade: patient safety and public health risks. *Journal of Pharmaceutical Sciences.* **100,** 4571–9 (2011).

72. Greenwood, B. Malaria vaccines: Evaluation and implementation. *Acta Tropica.* **95,3**, 298-304 (2005). at <http://www.sciencedirect.com/science/article/pii/S0001706X05001063>

73. Nussenzweig, R., Vanderberg, J., Most, H. & Orton, C. Protective immunity produced by the injection of x-irradiated sporozoites of *Plasmodium berghei*. *Nature*. **216**, 160-162 (1967). at <http://www.nature.com/articles/216160a0>

74. CLYDE, D. & MOST, H. Immunization of man against sporozite-induced falciparum malaria. *American Society of Tropical Medicine and Hygiene.* **266,3**, 169-177 (1973). at <http://journals.lww.com/amjmedsci/Abstract/1973/09000/Immunization_of_man_against_sporozite_induced.2.aspx>

75. Marsh, K. & Kinyanjui, S. Immune effector mechanisms in malaria. *Parasite Immunology*. **28,** 51–60

76. Carter, R. & Mendis, K. N. Evolutionary and Historical Aspects of the Burden of Malaria. *Clinical Microbiology Reviews.* **15,** 564–594 (2002).

77. Doolan, D. L., Dobaño, C. & Baird, J. K. Acquired immunity to malaria. *Clinical Microbiology Reviews.* **22,** 13–36, Table of Contents (2009).

78. Drakeley, C., Sutherland, C., Bousema, J. T., Sauerwein, R. W. & Targett, G. A. T. The epidemiology of *Plasmodium falciparum* gametocytes: weapons of mass dispersion. *Trends Parasitology.* **22,** 424–30 (2006).

79. Tiono, A. B. *et al.* Dynamics of malaria transmission and susceptibility to clinical malaria episodes following treatment of *Plasmodium falciparum* asymptomatic carriers: results of a cluster-randomized study of community-wide screening and treatment, and a parallel entomology. *BMC Infectious Diseases.* **13,** 535 (2013).

80. Pinkevych, M. *et al.* The dynamics of naturally acquired immunity to *Plasmodium falciparum* infection. *PLoS Computational Biology.* **8,** e1002729 (2012).

81. Epstein, J., Tewari, K., Lyke, K. & Sim, B. Live attenuated malaria vaccine designed to protect through hepatic CD8+ T cell immunity. *Science.* **334**, **6055**, 475-480 (2011). at <http://www.sciencemag.org/content/334/6055/475.short>

82. Seder, R., Chang, L. & Enama, M. Protection against malaria by intravenous immunization with a nonreplicating sporozoite vaccine. *Science.* **341, 6152**, 1359-1365 (2013). at <http://www.sciencemag.org/content/341/6152/1359.short>

83. Barry, A. E., Schultz, L., Buckee, C. O. & Reeder, J. C. Contrasting population structures of the genes encoding ten leading vaccine-candidate antigens of the human malaria parasite, *Plasmodium falciparum*. *PLoS One* **4,** e8497 (2009).

84. Daily, J. P. Malaria vaccine trials--beyond efficacy end points. *The New England Journal of Medicine.* **367,** 2349–51 (2012).

85. Seder, R. A. *et al.* Protection against malaria by intravenous immunization with a nonreplicating sporozoite vaccine. *Science* **341,** 1359–65 (2013).

86. Ménard, R. *et al.* Looking under the skin: the first steps in malarial infection and immunity. *Nature Reviews Microbiology.* **11,** 701–12 (2013).

87. VanBuskirk, K. Preerythrocytic, live-attenuated *Plasmodium falciparum* vaccine candidates by design. *Proceedings of the National Academy of Sciences.* **106, 31**, 13004-13009 (2009). at <http://www.pnas.org/content/106/31/13004.short>

88. Nardin, E. & Nussenzweig, R. T cell responses to pre-erythrocytic stages of malaria: role in protection and vaccine development against pre-erythrocytic stages. *Annual Review of Immunology.* **11,1**, 687-727 (1993). at <http://www.annualreviews.org/doi/pdf/10.1146/annurev.iy.11.040193.003351>

89. Rieckmann, K. Sporozoite induced immunity in man against an Ethiopian strain of *Plasmodium falciparum*. *Transactions of the Royal society of Tropical Medicine and Hygiene.* **68,3**, 258-259 (1974). at <https://scholar.google.pt/scholar?hl=en&q=Trans+R+Soc+Trop+Med+Hyg+68%3A+258%E2%80%93259+&btnG=&as_sdt=1%2C5&as_sdtp=#0>

90. WHO. Malaria vaccine development. (2014). at <http://www.who.int/malaria/areas/vaccine/en/>

91. Agnandji, S. First results of phase 3 trial of RTS, S/AS01 malaria vaccine in African children. *The New England Journal of Medicine.* **365, 20**, 1863-75 (2011). at <http://researchonline.lshtm.ac.uk/39130/>

92. Agnandji, S. A phase 3 trial of RTS, S/AS01 malaria vaccine in African infants. *The New England Journal of Medicine.* **367, 24**, 2284-2295 (2012). at <http://researchonline.lshtm.ac.uk/427472/>

93. Tran, T. M., Portugal, S., Draper, S. J. & Crompton, P. D. Malaria Vaccines: Moving Forward After Encouraging First Steps. *Current Tropical Medicine Reports.* **2,** 1–3 (2015).

94. Heppner, D. G. *et al.* Towards an RTS,S-based, multi-stage, multi-antigen vaccine against falciparum malaria: progress at the Walter Reed Army Institute of Research. *Vaccine* **23,** 2243–50 (2005).

95. Thompson, F. M. *et al.* Evidence of blood stage efficacy with a virosomal malaria vaccine in a phase IIa clinical trial. *PLoS One* **3,** e1493 (2008).

96. Douglas, A. D. *et al.* Comparison of modeling methods to determine liver-to-blood inocula and parasite multiplication rates during controlled human malaria infection. *The Journal of Infectious Diseases.* **208,** 340–5 (2013).

97. Fowkes, F., Richards, J., Simpson, J. & Beeson, J. The relationship between anti-merozoite antibodies and incidence of *Plasmodium falciparum* malaria: A systematic review and meta-analysis. *PLoS Medicine.* **7, 1**, 110 (2010). at <http://dx.plos.org/10.1371/journal.pmed.1000218>

98. Goodman, A. & Draper, S. Blood-stage malaria vaccines—recent progress and future challenges. *Annals of Tropical Medicine and Parasitology.* **104,3**, 189-211 (2010). at <http://www.maneyonline.com/doi/abs/10.1179/136485910X12647085215534>

99. Richards, J. & Beeson, J. The future for blood-stage vaccines against malaria. *Immunology & Cell Biology.* **87,5**, 377-390 (2009). at <http://www.nature.com/articles/doi:10.1038%2Ficb.2009.27>

100. Rogerson, S., Hviid, L. & Duffy, P. Malaria in pregnancy: pathogenesis and immunity. *The Lancet Infectious Diseases.* **7,2**, 105-117 (2007). at <http://www.sciencedirect.com/science/article/pii/S1473309907700221>

101. Hviid, L. The role of *Plasmodium falciparum* variant surface antigens in protective immunity and vaccine development. *Human Vaccines & Immunotherapeutics.* **6,1**, 84-89 (2010). at <http://www.tandfonline.com/doi/abs/10.4161/hv.6.1.9602>

102. Wright, G. J. & Rayner, J. C. *Plasmodium falciparum* erythrocyte invasion: combining function with immune evasion. *PLoS Pathogens.* **10,** e1003943 (2014).

103. Tran, T. M. *et al.* Naturally acquired antibodies specific for *Plasmodium falciparum* reticulocyte-binding protein homologue 5 inhibit parasite growth and predict protection from malaria. *The Journal of Infectious Diseases.* **209,** 789–98 (2014).

104. Douglas, A. D. *et al.* Neutralization of *Plasmodium falciparum* merozoites by antibodies against *Pf*RH5. *The Journal of Immunology.* **192,** 245–58 (2014).

105. Srinivasan, P. *et al.* Immunization with a functional protein complex required for erythrocyte invasion protects against lethal malaria. *Proceedings of the National Academy of Sciences.* **111,** 10311–6 (2014).

106. Raj, D. K. *et al.* Antibodies to *Pf*SEA-1 block parasite egress from RBCs and protect against malaria infection. *Science* **344,** 871–7 (2014).

107. Sinden, R. A biologist's perspective on malaria vaccine development. *Human Vaccines & Immunotherapeutics.* **6,1**, 3-11 (2010). at <http://www.tandfonline.com/doi/abs/10.4161/hv.6.1.9604>

108. Sutherland, C. Surface antigens of *Plasmodium falciparum* gametocytes—A new class of transmission-blocking vaccine targets? *Molecular and Biochemical Parasitology*. **166,2**, 93-98 (2009). at <http://www.sciencedirect.com/science/article/pii/S0166685109000917>

109. Greenwood, B. & Fidock, D. Malaria: progress, perils, and prospects for eradication. *The Journal of Clinical Investigation.* **118, 4**, 1266 (2008). at <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2276780/>

110. Pradel, G. Proteins of the malaria parasite sexual stages: expression, function and potential for transmission blocking strategies. *Parasitology*. **134,14**, 1911-1929 (2007). at <http://journals.cambridge.org/abstract_S0031182007003381>

111. Sattabongkot, J. & Tsuboi, T. Blocking of transmission to mosquitoes by antibody to *Plasmodium vivax* malaria vaccine candidates *Pvs*25 and *Pvs*28 despite antigenic polymorphism in field. *American Society of Tropical Medicine and Hygiene.* **69,5**, 536-541 (2003). at <http://www.ajtmh.org/content/69/5/536.short>

112. Tachibana, M., Sato, C. & Otsuki, H. *Plasmodium vivax* gametocyte protein *Pvs*230 is a transmission-blocking vaccine candidate. *Vaccine.* **30,10**, 1807-1812 (2012). at <http://www.sciencedirect.com/science/article/pii/S0264410X12000047>

113. Wu, Y., Ellis, R., Shaffer, D. & Fontes, E. Phase 1 trial of malaria transmission blocking vaccine candidates *Pfs*25 and *Pvs*25 formulated with montanide ISA 51. *PLoS One.* **3,7**, e2636 (2008). at <http://dx.plos.org/10.1371/journal.pone.0002636>

114. Dinglasan, R. & Jacobs-Lorena, M. Flipping the paradigm on malaria transmission-blocking vaccines. *Trends Parasitology.* **24,8**, 364-370 (2008). at <http://www.sciencedirect.com/science/article/pii/S1471492208001542>

115. Ranson, H., N'Guessan, R. & Lines, J. Pyrethroid resistance in African anopheline mosquitoes: what are the implications for malaria control? *Trends Parasitology.* **27.2,** 91-98 (2011). at <http://www.sciencedirect.com/science/article/pii/S1471492210001753>

116. Davies, T. G. E., Field, L. M., Usherwood, P. N. R. & Williamson, M. S. DDT, pyrethrins, pyrethroids and insect sodium channels. *IUBMB Life* **59,** 151–62 (2007).

117. Donnelly, M. J. *et al.* Does kdr genotype predict insecticide-resistance phenotype in mosquitoes? *Trends Parasitology.* **25,** 213–9 (2009).

118. Malaria Vectors. Global Plan for Insecticide Resistance Management. (2012). at <http://whqlibdoc.who.int/Publications/2012/9789241564472_eng.pdf>

119. Ghosh, A. Targeting *Plasmodium* ligands on mosquito salivary glands and midgut with a phage display peptide library. *Proceedings of the National Academy of Sciences.* **98.23,** 13278-13281 (2001). at <http://www.pnas.org/content/98/23/13278.short>

120. Gwadz, R., Kaslow, D. & Lee, J. Effects of magainins and cecropins on the sporogonic development of malaria parasites in mosquitoes. *Infection and Immunity*.**57,9**, 2628-2633 (1989). at <http://iai.asm.org/content/57/9/2628.short>

121. Kokoza, V. & Ahmed, A. Blocking of *Plasmodium* transmission by cooperative action of Cecropin A and Defensin A in transgenic *Aedes aegypti* mosquitoes. *Proceedings of the National Academy of Sciences.* **107,18,** 8111-8116 (2010). at <http://www.pnas.org/content/107/18/8111.short>

122. *Human and Mosquito Lysozymes*. (Springer International Publishing, 2015). doi:10.1007/978-3-319-09432-8

123. Hurwitz, I., Fieck, A. & Read, A. Paratransgenic control of vector borne diseases. *International Journal of Biological Sciences.* **7,9,** 1334 (2011). at <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3221369/>

124. Benedict, M. & Robinson, A. The first releases of transgenic mosquitoes: an argument for the sterile insect technique. *Trends Parasitology*.**19,8**, 349-355 (2003). at <http://www.sciencedirect.com/science/article/pii/S1471492203001442>

125. Catteruccia, F., Crisanti, A. & Wimmer, E. Transgenic technologies to induce sterility. *Malaria Journal*.**8 (**Suppl 2), S7 (2009). at <http://www.biomedcentral.com/content/pdf/1475-2875-8-S2-S7.pdf>

126. Crawford, J. E. *et al. De novo* transcriptome sequencing in *Anopheles funestus* using Illumina RNA-seq technology. *PLoS One* **5,** e14202 (2010).

127. Illumina. An Introduction to Next-Generation Sequencing Technology. (2011). at <http://www.illumina.com/content/dam/illumina-marketing/documents/products/illumina_sequencing_introduction.pdf>

128. Maher, C. A. *et al.* Transcriptome sequencing to detect gene fusions in cancer. *Nature* **458,** 97–101 (2009).

129. Hornett, E. A. & Wheat, C. W. Quantitative RNA-Seq analysis in non-model species: assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species. *BMC Genomics* **13,** 361 (2012).

130. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456,** 470–6 (2008).

131. Sultan, M. *et al.* A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321,** 956–60 (2008).

132. Wang, X. *et al.* Transcriptome-wide identification of novel imprinted genes in neonatal mouse brain. *PLoS One* **3,** e3839 (2008).

133. Fontanillas, P. *et al.* Key considerations for measuring allelic expression on a genomic scale using high-throughput sequencing. *Molecular Ecology.* **19 Suppl 1,** 212–27 (2010).

134. Creighton, C. J., Reid, J. G. & Gunaratne, P. H. Expression profiling of microRNAs by deep sequencing. *Briefings in Bioinformatics.* **10,** 490–7 (2009).

135. Haas, B. J. & Zody, M. C. Advancing RNA-Seq analysis. *Nature Biotechnology.* **28,** 421–3 (2010).

136. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics.* **10,** 57–63 (2009).

137. Sboner, A., Mu, X. J., Greenbaum, D., Auerbach, R. K. & Gerstein, M. B. The real cost of sequencing: higher than you think! *Genome Biology.* **12,** 125 (2011).

138. Cheung, F. *et al.* Sequencing Medicago truncatula expressed sequenced tags using 454 Life Sciences technology. *BMC Genomics* **7,** 272 (2006).

139. Jeukens, J., Renaut, S., St-Cyr, J., Nolte, A. W. & Bernatchez, L. The transcriptomics of sympatric dwarf and normal lake whitefish (*Coregonus clupeaformis* spp., Salmonidae) divergence as revealed by next-generation sequencing. *Molecular Ecology.* **19,** 5389–403 (2010).

140. Fang, Z., Martin, J. & Wang, Z. Statistical methods for identifying differentially expressed genes in RNA-Seq experiments. *Cell & Bioscience.* **2,** 26 (2012).

141. Guttman, M. *et al. Ab initio* reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nature Biotechnology.* **28,** 503–10 (2010).

142. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology.* **29,** 644–52 (2011).

143. Birol, I. *et al. De novo* transcriptome assembly with ABySS. *Bioinformatics* **25,** 2872–7 (2009).

144. Toth, A. L. *et al.* Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science* **318,** 441–4 (2007).

145. Collins, L. J., Biggs, P. J., Voelckel, C. & Joly, S. An approach to transcriptome analysis of non-model organisms using short-read sequences. *Genome Informatics* **21,** 3–14 (2008).

146. Adiconis, X. *et al.* Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nature Methods.* **10,** 623–9 (2013).

147. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* **5,** 621–8 (2008).

148. Corney, D. C. RNA-seq Using Next Generation Sequencing. *Materials and Methods* (2015). at <http://www.labome.com/method/RNA-seq-Using-Next-Generation-Sequencing.html>

149. Corney, D. C. & Jefferson, T. RNA-seq Using Next Generation Sequencing. *Materials and Methods* **3,** (2013).

150. Fuller, C. W. *et al.* The challenges of sequencing by synthesis. *Nature Biotechnology* **27,** 1013–23 (2009).

151. Schulz, M., Zerbino, D., Vingron, M. & Birney, E. Oases: robust *de novo* RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics.* **28,8**, 1086-1092 (2012). at <http://bioinformatics.oxfordjournals.org/content/28/8/1086.short>

152. Haas, B., Papanicolaou, A. & Yassour, M. *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature Protocols.* **8,8**, 1494-1512 (2013). at <http://www.nature.com/articles/doi:10.1038/nprot.2013.084>

153. Schulz, M. H., Zerbino, D. R., Vingron, M. & Birney, E. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics* **28,** 1086–92 (2012).

154. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biology.* **10,** R25 (2009).

155. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25,** 2078–9 (2009).

156. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25,** 1105–11 (2009).

157. Trapnell, C. *et al.* Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature. Biotechnology.* **28,** 511–5 (2010).

158. Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22,** 1658–9 (2006).

159. Li, H. *et al.* Determination of tag density required for digital transcriptome analysis: application to an androgen-sensitive prostate cancer model. *Proceedings of the National Academy of Sciences* **105,** 20179–84 (2008).

160. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology.* **5,** R80 (2004).

161. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biology.* **11,** R106 (2010).

162. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–40 (2010).

163. Conesa, A. *et al.* Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21,** 3674–6 (2005).

164. Kanehisa, M., Goto, S. & Sato, Y. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic acids research.* **42, D1**, D199-D205 (2014). at <http://nar.oxfordjournals.org/content/42/D1/D199.short>

165. Franceschini, A. & Szklarczyk, D. STRING v9. 1: protein-protein interaction networks, with increased coverage and integration. *Nucleic acids research.* **42, D1**, D808-D815 (2013). at <http://nar.oxfordjournals.org/content/41/D1/D808.short>

166. Horton, P. *et al.* WoLF PSORT: protein localization predictor. *Nucleic acids research.* **35,** W585–7 (2007).

167. Blum, T., Briesemeister, S. & Kohlbacher, O. MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. *BMC Bioinformatics* **10,** 274 (2009).

168. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *Journal of Molecular Biology.* **305,** 567–80 (2001).

169. Reimand, J., Kull, M., Peterson, H., Hansen, J. & Vilo, J. g:Profiler--a web-based toolset for functional profiling of gene lists from large-scale experiments. *Nucleic acids research.* **35,** W193–200 (2007).

170. Reimand, J., Arak, T. & Vilo, J. g:Profiler--a web server for functional interpretation of gene lists (2011 update). *Nucleic acids research.* **39,** W307–15 (2011).

171. Beissbarth, T. & Speed, T. P. GOstat: find statistically overrepresented Gene Ontologies within a group of genes. *Bioinformatics* **20,** 1464–5 (2004).

172. Metzker, M. Sequencing technologies—the next generation. *Nature Reviews Genetics.* **11,1**, 31-46 (2010). at <http://www.nature.com/articles/doi:10.1038/nrg2626>

173. Auer, P. L. & Doerge, R. W. Statistical design and analysis of RNA sequencing data. *Genetics* **185,** 405–16 (2010).

174. McIntyre, L. M. *et al.* RNA-seq: technical variability and sampling. *BMC Genomics* **12,** 293 (2011).

175. Marioni, J. C., Mason, C. E., Mane, S. M., Stephens, M. & Gilad, Y. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research.* **18,** 1509–17 (2008).

176. Nagalakshmi, U. *et al.* The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320,** 1344–9 (2008).

177. Lee, C., Kim, J., Shin, S. G. & Hwang, S. Absolute and relative QPCR quantification of plasmid copy number in *Escherichia coli*. *Journal of Biotechnology.* **123,** 273–80 (2006).

178. Schmittgen, T. D. & Livak, K. J. Analyzing real-time PCR data by the comparative CT method. *Nature Protocols.* **3,** 1101–1108 (2008).

179. Bogdanove, A. J. & Voytas, D. F. TAL effectors: customizable proteins for DNA targeting. *Science* **333,** 1843–6 (2011).

180. Mittal, V. Improving the efficiency of RNA interference in mammals. *Nature Reviews Genetics.* **5,** 355–65 (2004).

181. Ketting, R. F. The many faces of RNAi. *Developmental Cell.* **20,** 148–61 (2011).

182. Mittal, V. Improving the efficiency of RNA interference in mammals. *Nature Reviews Genetics.* **5,** 355–65 (2004).

183. Robinson, M., McCarthy, D., Chen, Y. & Smyth, G. edgeR: differential expression analysis of digital gene expression data. (2010). at <ftp://ftp2.uib.no/pub/bioconductor/2.6/bioc/vignettes/edgeR/inst/doc/edgeR.pdf>

184. Nogueira, F. & Rosário, V. E. do. Methods for assessment of antimalarial activity in the different phases of the Plasmodium life cycle. *Revista Pan-Amazônica de Saúde.* **1,** 109–124 (2010).

185. Wright, K. A. The anatomy of salivary glands of *Anopheles stephensi Liston*. *Canadian Journal of Zoology*. **47,** 579–587 (1969).

186. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research.* **29,** 45e–45 (2001).

187. Bustin, S. A. *et al.* The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry.* **55,** 611–22 (2009).

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

189. Lawrence, I. & Lin, K. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* (1989). at <http://www.jstor.org/stable/2532051>

190. Milligan, J. F. & Uhlenbeck, O. C. Synthesis of small RNAs using T7 RNA polymerase. *Methods in Enzymology.* **180,** 51–62 (1989).

191. Pinto, S. B., Kafatos, F. C. & Michel, K. The parasite invasion marker *SRPN6* reduces sporozoite numbers in salivary glands of *Anopheles gambiae*. *Cellular Microbiology.* **10,** 891–8 (2008).

192. Wilhelm, B. T. *et al.* Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature* **453,** 1239–43 (2008).

193. Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods.* **5,** 621–8 (2008).

194. Cloonan, N. *et al.* Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nature Methods.* **5,** 613–9 (2008).

195. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics.* **10,** 57–63 (2009).

196. Ridley, R. G. Medical need, scientific opportunity and the drive for antimalarial drugs. *Nature* **415,** 686–93 (2002).

197. Sinden, R. E. Molecular interactions between *Plasmodium* and its insect vectors. *Cellular Microbiology.* **4,** 713–24 (2002).

198. Aguilar, R. *et al.* Global gene expression analysis of *Anopheles gambiae* responses to microbial challenge. *Insect Biochemistry and Molecular Biology.* **35,** 709–19 (2005).

199. Jiang, X. *et al.* Genome analysis of a major urban malaria vector mosquito, *Anopheles stephensi*. *Genome Biology.* **15,** 459 (2014).

200. Xu, X. C., Qu, F. Y., Song, G. H. & Xu, J. N. [Enrichment and screening of up-regulated genes of the mosquito *Anopheles stephensi* in response to malaria parasite]. *Chinese journal of parasitology & parasitic diseases.* **19,** 325–9 (2001).

201. Oduol, F., Xu, J., Niare, O., Natarajan, R. & Vernick, K. D. Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria. *Proceedings of the National Academy of Sciences.* **97,** 11397–402 (2000).

202. Rosinski-Chupin, I. *et al.* SAGE analysis of mosquito salivary gland transcriptomes during *Plasmodium* invasion. *Cellular Microbiology.* **9,** 708–24 (2007).

203. Peterson, T. M. L., Gow, A. J. & Luckhart, S. Nitric oxide metabolites induced in *Anopheles stephensi* control malaria parasite infection. *Free Radical Biology & Medicine.* **42,** 132–42 (2007).

204. Dimopoulos, G. *et al.* Genome expression analysis of *Anopheles gambiae*: responses to injury, bacterial challenge, and malaria infection. *Proceedings of the National Academy of Sciences.* **99,** 8814–9 (2002).

205. Vlachou, D., Schlegelmilch, T., Christophides, G. K. & Kafatos, F. C. Functional genomic analysis of midgut epithelial responses in *Anopheles* during *Plasmodium* invasion. *Current Biology.* **15,** 1185–95 (2005).

206. Slavic, K. *et al.* Life cycle studies of the hexose transporter of *Plasmodium* species and genetic validation of their essentiality. *Molecular Microbiology.* **75,** 1402–13 (2010).

207. Liu, K., Dong, Y., Huang, Y., Rasgon, J. L. & Agre, P. Impact of trehalose transporter knockdown on *Anopheles gambiae* stress adaptation and susceptibility to *Plasmodium falciparum* infection. *Proceedings of the National Academy of Sciences.* **110,** 17504–9 (2013).

208. Kumar, S. *et al.* The role of reactive oxygen species on *Plasmodium* melanotic encapsulation in *Anopheles gambiae*. *Proceedings of the National Academy of Sciences.* **100,** 14139–44 (2003).

209. Xu, X. *et al.* Transcriptome analysis of *Anopheles stephensi-Plasmodium berghei* interactions. *Molecular and Biochemical Parasitology.* **142,** 76–87 (2005).

210. Hansen, I. A., Attardo, G. M., Park, J.-H., Peng, Q. & Raikhel, A. S. Target of rapamycin-mediated amino acid signaling in mosquito anautogeny. *Proceedings of the National Academy of Sciences.* **101,** 10626–31 (2004).

211. Rono, M. K., Whitten, M. M. A., Oulad-Abdelghani, M., Levashina, E. A. & Marois, E. The major yolk protein vitellogenin interferes with the anti-*plasmodium* response in the malaria mosquito *Anopheles gambiae*. *PLoS Biology.* **8,** e1000434 (2010).

212. Merrill, C. E. *et al.* Visual arrestins in olfactory pathways of *Drosophila* and the malaria vector mosquito *Anopheles gambiae*. *Proceedings of the National Academy of Sciences.* **99,** 1633–8 (2002).

213. Walker, W. B., Smith, E. M., Jan, T. & Zwiebel, L. J. A functional role for *Anopheles gambiae* Arrestin1 in olfactory signal transduction. *Journal of Insect Physiology.* **54,** 680–90 (2008).

214. Félix, Rute; Silveira, H. in *ISBN* 978–953 (INTECH Open Access Publisher, 2012).

215. Meister, S. *et al.* Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito *Anopheles gambiae*. *Proceedings of the National Academy of Sciences.* **102,** 11420–5 (2005).

216. Frolet, C., Thoma, M., Blandin, S., Hoffmann, J. A. & Levashina, E. A. Boosting NF-kappaB-dependent basal immunity of *Anopheles gambiae* aborts development of *Plasmodium berghei*. *Immunity* **25,** 677–85 (2006).

217. Garver, L. S. *et al. Anopheles* Imd pathway factors and effectors in infection intensity-dependent anti-*Plasmodium* action. *PLoS Pathogens.* **8,** e1002737 (2012).

218. Mitri, C. *et al.* Fine pathogen discrimination within the APL1 gene family protects *Anopheles gambiae* against human and rodent malaria species. *PLoS Pathogens.* **5,** e1000576 (2009).

219. Volz, J., Osta, M. A., Kafatos, F. C. & Müller, H.-M. The roles of two clip domain serine proteases in innate immune responses of the malaria vector *Anopheles gambiae*. *The Journal of Biological Chemistry.* **280,** 40161–8 (2005).

220. Blandin, S. *et al.* Complement-Like Protein TEP1 Is a Determinant of Vectorial Capacity in the Malaria Vector *Anopheles gambiae*. *Cell* **116,** 661–670 (2004).

221. Das, S. *et al.* Transcriptomic and functional analysis of the *Anopheles gambiae* salivary gland in relation to blood feeding. *BMC Genomics* **11,** 566 (2010).

222. Baker, D. A. *et al.* A comprehensive gene expression atlas of sex- and tissue-specificity in the malaria vector, *Anopheles gambiae*. *BMC Genomics* **12,** 296 (2011).

223. Zhang, Y., Luo, M. & Schramm, V. L. Transition states of *Plasmodium falciparum* and human orotate phosphoribosyltransferases. *Journal of the American Chemical Society.* **131,** 4685–94 (2009).

224. Nagaraj, V. A. *et al.* Malaria parasite-synthesized heme is essential in the mosquito and liver stages and complements host heme in the blood stages of infection. *PLoS Pathogens.* **9,** e1003522 (2013).

225. Ke, H. *et al.* The heme biosynthesis pathway is essential for *Plasmodium falciparum* development in mosquito stage but not in blood stages. *The Journal of Biological Chemistry.* **289,** 34827–37 (2014).

226. Heid, C. A., Stevens, J., Livak, K. J. & Williams, P. M. Real time quantitative PCR. *Genome Research.* **6,** 986–94 (1996).

227. Fox, A. N., Pitts, R. J., Robertson, H. M., Carlson, J. R. & Zwiebel, L. J. Candidate odorant receptors from the malaria vector mosquito *Anopheles gambiae* and evidence of down-regulation in response to blood feeding. *Proceedings of the National Academy of Sciences.* **98,** 14693–7 (2001).

228. Merrill, C. E., Pitts, R. J. & Zwiebel, L. J. Molecular characterization of arrestin family members in the malaria vector mosquito, *Anopheles gambiae*. *Insect Molecular Biology.* **12,** 641–650 (2003).

229. Vlachou, D., Schlegelmilch, T., Christophides, G. K. & Kafatos, F. C. Functional genomic analysis of midgut epithelial responses in *Anopheles* during *Plasmodium* invasion. *Current Biology.* **15,** 1185–95 (2005).

230. Pruijssers, A. J. & Strand, M. R. PTP-H2 and PTP-H3 from *Microplitis demolitor Bracovirus* localize to focal adhesions and are antiphagocytic in insect immune cells. *Journal of Virology.* **81,** 1209–19 (2007).

231. *Non-coding RNAs and Epigenetic Regulation of Gene Expression: Drivers of Natural Selection*. (Horizon Scientific Press, 2012). at <https://books.google.com/books?hl=en&lr=&id=HvWuosBg2tcC&pgis=1>

232. Ji, Y. *et al.* Cloning and characterization of AabHLH1, a bHLH transcription factor that positively regulates artemisinin biosynthesis in *Artemisia annua*. *Plant and Cell Physiology.* **55,** 1592–604 (2014).

233. Dinglasan, R. R. *et al. Plasmodium falciparum* ookinetes require mosquito midgut chondroitin sulfate proteoglycans for cell invasion. *Proceedings of the National Academy of Sciences.* **104,** 15882–7 (2007).

234. Sinnis, P. *et al.* Mosquito heparan sulfate and its potential role in malaria infection and transmission. *The Journal of Biological Chemistry.* **282,** 25376–84 (2007).

235. Simões, M. L., Gonçalves, L. & Silveira, H. Hemozoin activates the innate immune system and reduces *Plasmodium berghei* infection in *Anopheles gambiae*. *Parasites & Vectors.* **8,** 12 (2015).

236. Christophides, G. K. *et al.* Immunity-related genes and gene families in *Anopheles gambiae*. *Science* **298,** 159–65 (2002).

237. Barillas-Mury, C. CLIP proteases and *Plasmodium* melanization in *Anopheles gambiae*. *Trends Parasitology.* **23,** 297–9 (2007).

238. Abraham, E. G. *et al.* Analysis of the *Plasmodium* and *Anopheles* transcriptional repertoire during ookinete development and midgut invasion. *The Journal of Biological Chemistry.* **279,** 5573–80 (2004).

239. Michel, K., Budd, A., Pinto, S., Gibson, T. J. & Kafatos, F. C. *Anopheles gambiae* SRPN2 facilitates midgut invasion by the malaria parasite *Plasmodium berghei*. *EMBO Reports.* **6,** 891–7 (2005).

240. Abraham, E. G. *et al.* An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites. *Proceedings of the National Academy of Sciences.* **102,** 16327–32 (2005).

241. Suzuki, S. *et al.* Interaction between the membrane protein of a pathogen and insect microfilament complex determines insect-vector specificity. *Proceedings of the National Academy of Sciences.* **103,** 4252–7 (2006).

242. Kirk, K., Martin, R. E., Bröer, S., Howitt, S. M. & Saliba, K. J. *Plasmodium* permeomics: membrane transport proteins in the malaria parasite. *Current Topics in Microbiology and Immunology.* **295,** 325–56 (2005).

243. Ginsburg, H. Transport pathways in the malaria-infected erythrocyte: characterization and their use as potential targets for chemotherapy. *Memórias do Instituto Oswaldo Cruz* **89,** 99–109 (1994).

244. Dinglasan, R. R. *et al.* Disruption of *Plasmodium* falciparum development by antibodies against a conserved mosquito midgut antigen. *Proceedings of the National Academy of Sciences.* **104,** 13461–6 (2007).

245. Hirata, T., Czapar, A. & Brin, L. Ion and solute transport by Prestin in *Drosophila* and *Anopheles*. *Journal of Insect Physiology.* **58,4**, 563-569 (2012). at <http://www.sciencedirect.com/science/article/pii/S0022191012000108>

246. Brin, L. R. *et al.* Characterization of *Anopheles gambiae* Slc26a5 and potential role in malaria. *FASEB J* **24,** 1002.25– (2010).

247. Black, D. L. Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry.* **72,** 291–336 (2003).

# Appendix I

**Table 1 - Biological characteristics of *Plasmodium* species.**

| **Species** | *P. vivax* | *P. malariae* | *P. falciparum* | *P. ovale* |
|---|---|---|---|---|
| **Incubation period (days)** | 8-27 | 15-30 | 8-25 | 9-17 |
| **Duration of the erythrocyte cycle (hours)** | 48 | 72 | 48 | 48 |
| **Number of merozoites *per* schizont** | 10000 | 2000 | 40000 | 15000 |

# Appendix II

**Table 2 - List of primers used for validation by qPCR.**

| Gene | Vectorbase code | Primer Forward (5'-3') | Primer Reverse (5'-3') | Length (bp) | Annealing conditions (ºC/s) |
|---|---|---|---|---|---|
| ***SRPN6*** | ASTE001475 | CAGCATGAACTGGAAGCGTA | GTTCGGGACATCGTCGTAGT | 142 | 55 |
| ***AGAP004524*** | ASTE004032 | ACTATCAAGCGCACAGATGC | GATGCCAAACGGTTTTCTGT | 103 | 55 |
| ***CLIPA4*** | ASTE008188 | CTGGACGCTATCACACCAAC | ACTTGTTCTGCTCGGCAAAT | 121 | 55.5 |
| ***TOLL9*** | ASTE010442 | GCCGGACTTCAGTTGCTTAC | CACTAGCAGCAGGATGACGA | 116 | - |
| ***AGAP009577*** | ASTE010763 | CATCCGGGTTCGATAGTACG | TGTTAGCACTGGGACGTCTG | 105 | 56 |
| ***Paxillin (PAX)*** | ASTE007758 | GGTGTGAACACAACGCAGAA | GCAGTGGCTGCAGGTAAAGT | 117 | 57 |
| ***AGAP003844*** | ASTE011434 | GAACAGTATCGGCTCGGGTA | TTGCAATGTTCCTTGAGGTG | 116 | 55 |
| ***AGAP005721*** | ASTE007038 | CGCCAGCAGTTGCTACAGT | GCGCTGGTAGTGGTAGTGGT | 132 | 58 |
| ***AGAP005796*** | ASTE009772 | GTTCCGGTTGGCAATCATAC | GTGCAGCTTGGTGAAGGTTT | 103 | 55 |
| ***AGAP004170*** | ASTE004117 | GGCGAGGAGATGTACCAGTG | TTACAGATCGTGGCGGTGTA | 107 | 57 |
| ***AGAP007702*** | ASTE001733 | TCAATCATCCGGACACGATA | CGAGACATCGTTGAAAGCAA | 101 | 53 |
| ***Inositol oxygenase (INO)*** | ASTE005290 | GATCTGGATCTGCCGAACAT | GCTCACCGTAGAAAGCCATC | 135 | 53 |
| ***helicase DDX24/MAK5 (HELI)*** | ASTE010887 | TTGAAACAAGATGCGACTGC | CGCTTAGTGATCCTGGCTTC | 118 | 55 |
| ***Nep1 (AGAP002808)*** | ASTE007528 | AGCTGGCGATGTCTTTCAGT | GATCCCTCGGTGTAGTCCAA | 133 | 56 |
| ***ARRDC2 (ARRE)*** | ASTE009422 | GACAACCAAGGTCCAACGAT | GTACTGCACCCAACCGTACC | 114 | 57 |
| ***SCRB6*** | ASTE000417 | CCTATCGCATGTCCACCTTT | ATCTTCATAGCCCCAGAGCA | 113 | 55 |
| ***AGAP006353*** | ASTE000811 | CTCGTAATTCCTCGCAAACC | AGCCTTCCTCCAGTCCTTGT | 124 | 56 |
| ***AGAP003844*** | ASTE011434 | GAACAGTATCGGCTCGGGTA | TTGCAATGTTCCTTGAGGTG | 116 | 55 |
| ***AGAP010794*** | ASTE008166 | TCGTACTTGCCCTTTCATCC | CAGTGAGGGCAATGATGTTG | 120 | 54 |
| ***RPS7*** | ASTE004816 | TCCTGGAGGATCTGGTGTTC | GATGGTGGTCTGCTGGTTCT | 113 | 60 |
| ***AGAP005128(ELF)*** | ASTE005097 | TGCGGATTACGTGAAGAACA | ACGATYTTGCTRACGCCAAC | 141 | 60 |
| ***AGAP010725 (PRESTIN)*** | ASTE009391 | ATTGCTGTCACGGTACCTTC | GGCTGAAATCTGGCAATGTT | 101 | 59.2ºC/s |

# Appendix III

**Table 3 - List of primers used for double-stranded RNA synthesis.**

| Gene | Vectorbase code or GenBank code | Primer Forward (5'-3')* | Primer Reverse (5'-3')* | Length (bp) | Annealing conditions (ºC/s) |
|---|---|---|---|---|---|
| **beta-2 microglobulin (*β2M*)**** | NM_009735 | CACCCCCACTGAGACTGATACA | CACCCCCACTGAGACTGATACA | 450 | 64ºC/s |
| **ASTE009391 (*prestin*)***** | ASTE009391 | GGAAGGGCATGAGAGTGGTA | AGTACACCAGCACCGGAAAG | 471 | 64ºC/s |

*All primers forward contained T7 promoter sequences (5' – TAATACGACTCACTATAGGGAGA – 3') at the 5'end.

**An exogenous gene, mouse beta-2microglobulin (*β2M*) (GenBank: NM_009735) was used as control for the knockdown experiments.

*** For ASTE009391 were tested three pairs of primers from exon 1, 5 and 6. But only primers from exon 1 amplified to high yields.

# Appendix IV

**Table 4 - RNA samples for qPCR. Concentrations and absorbance of samples from SG extraction of control (C) and infected (I) experiments.**

| RNA sample ID | Concentration (ng/µL) | Abs 260/280 | Abs 260/230 |
|---|---|---|---|
| **RNA Total** | 875.88 | 1.95 | 0.61 |
| **RNA SG C 1** | 30.58 | 1.67 | 0.10 |
| **RNA SG I 1** | 69.62 | 1.71 | 0.30 |
| **RNA SG C 2** | 83.01 | 1.65 | 0.42 |
| **RNA SG I 2** | 259.11 | 1.86 | 0.66 |
| **RNA SG C 3** | 34.08 | 1.52 | 0.67 |
| **RNA SG I 3** | 36.48 | 1.66 | 0.66 |

# Appendix V

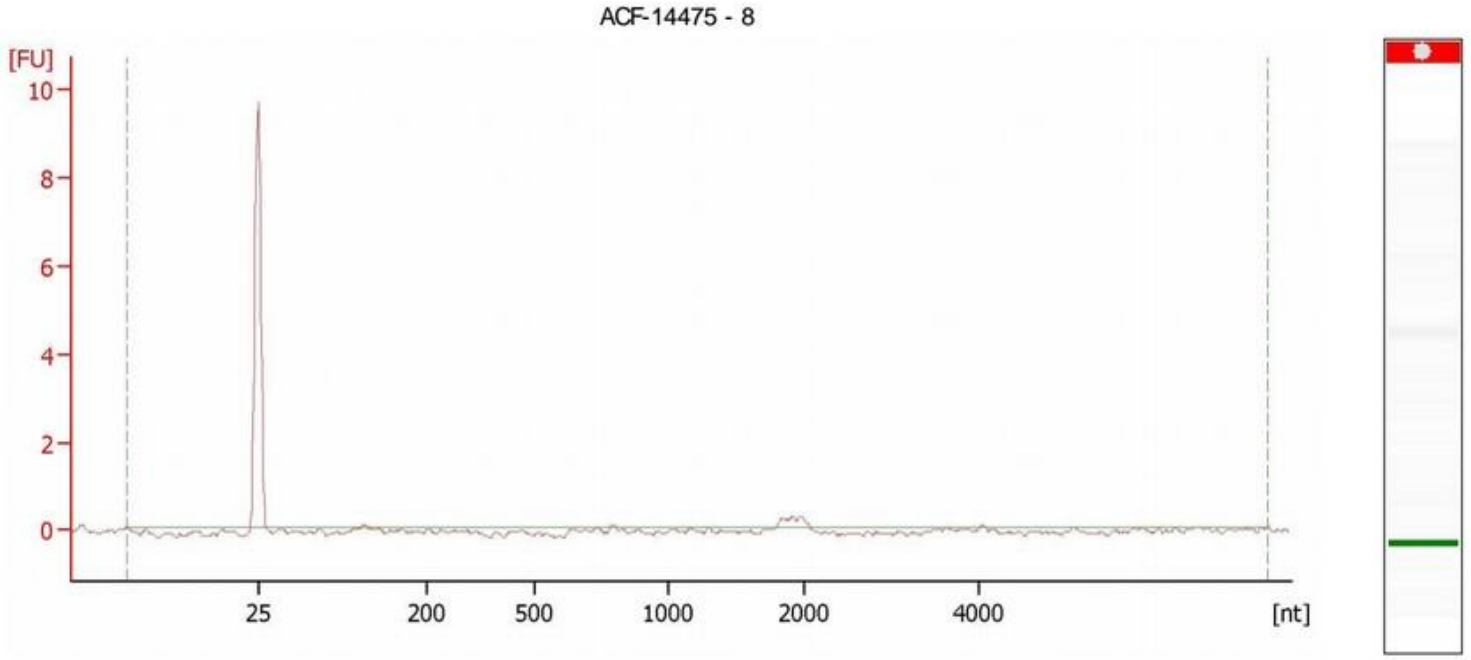
RNA Total



Assay Class: Eukaryote Total RNA Nano
Data Path: C:\...Eukaryote Total RNA Nano_DE24802219_2015-04-13_10-47-25.xad

Created: 4/13/2015 10:47:25 AM
Modified: 4/21/2015 3:50:56 PM

Electropherogram Summary Continued ...




Overall Results for sample 4 : ACF-14475 - 8

| | | | |
|---|---|---|---|
| RNA Area: | 0.7 | RNA Integrity Number (RIN): | N/A (B.02.08) |
| RNA Concentration: | 0 ng/µl | Result Flagging Color: | |
| rRNA Ratio [28s / 18s]: | 0.0 | Result Flagging Label: | RIN N/A |

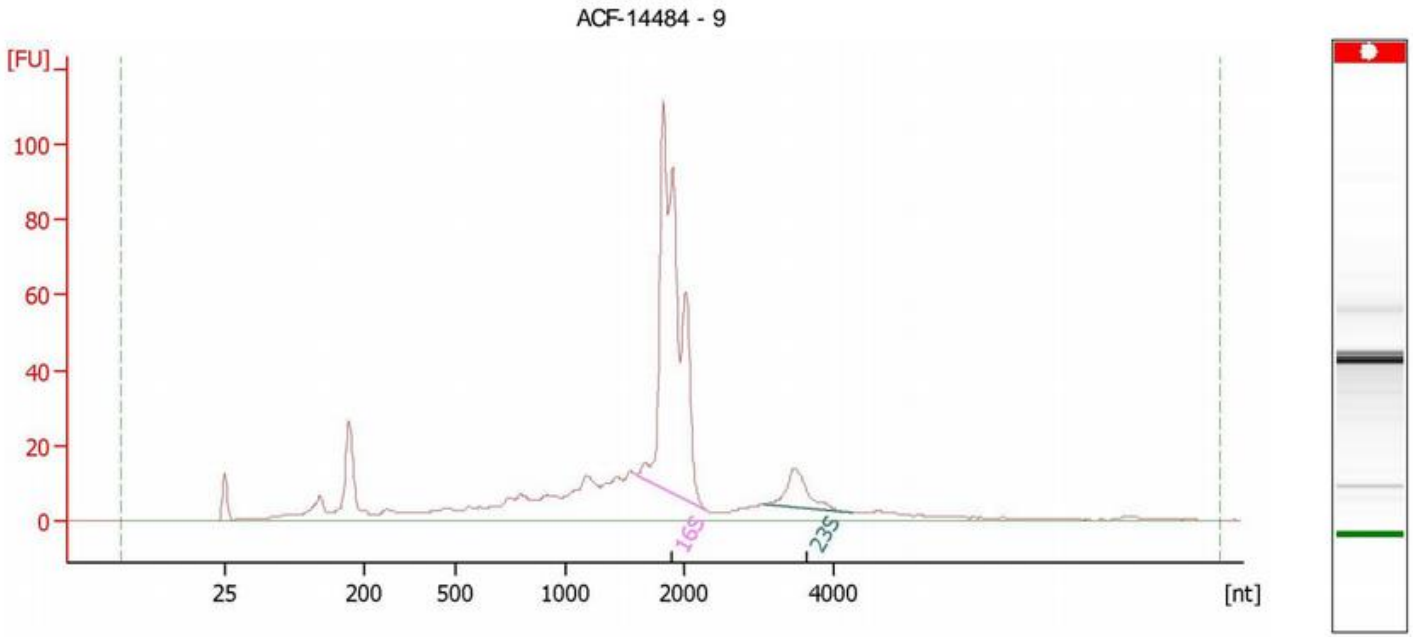
RNA SG C1



Assay Class: Prokaryote Total RNA Pico
Data Path: C:\...rokaryote Total RNA Pico_DE24802219_2015-04-17_15-28-48.xad

Created: 4/17/2015 3:28:48 PM
Modified: 4/17/2015 5:51:22 PM

Electropherogram Summary Continued ...




Overall Results for sample 1 : ACF-14484 - 9

| | | | |
|---|---|---|---|
| RNA Area: | 612.4 | rRNA Ratio [23s / 16s]: | 0.1 |
| RNA Concentration: | 18,393 pg/µl | RNA Integrity Number (RIN): | N/A (B.02.08) |

Fragment table for sample 1 : ACF-14484 - 9

| Name | Start Size [nt] | End Size [nt] | Area | % of total Area |
|---|---|---|---|---|
| 16S | 1,590 | 2,287 | 232.6 | 38.0 |
| 23S | 3,052 | 4,268 | 24.1 | 3.9 |

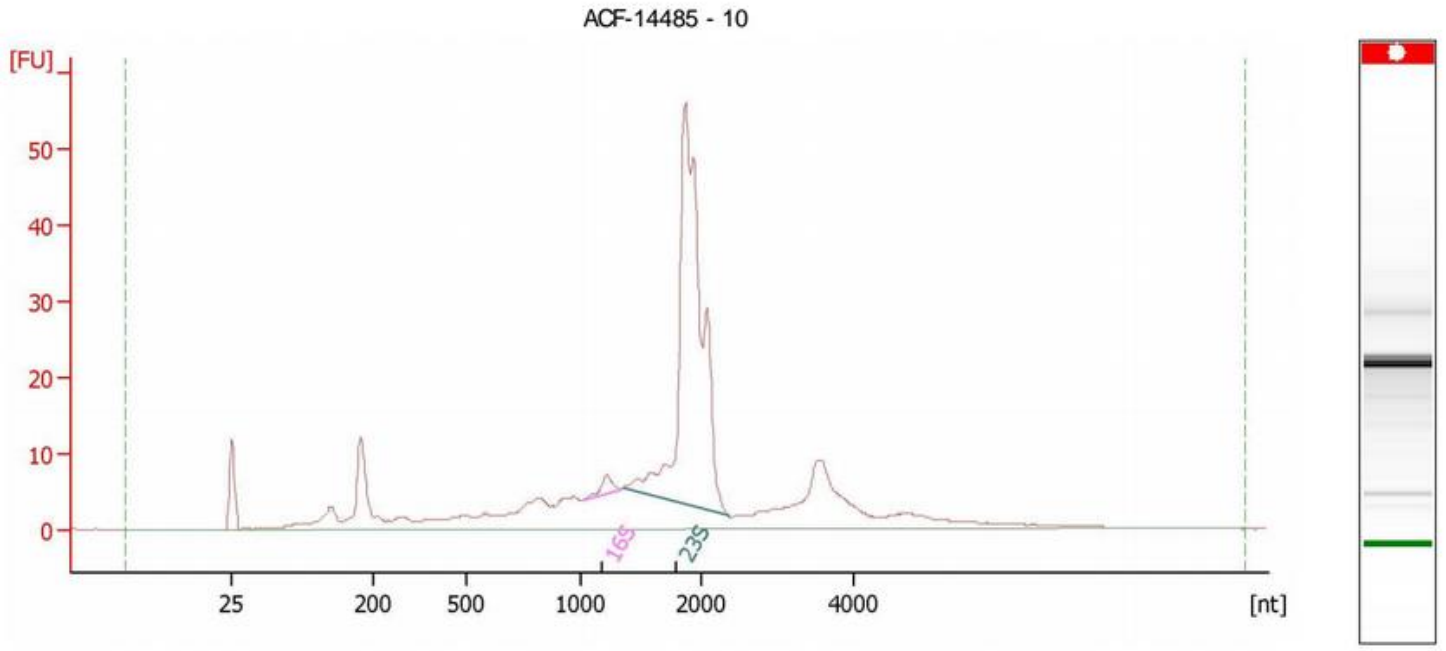
RNA SG I1



Assay Class: Prokaryote Total RNA Pico
Data Path: C:\...rokaryote Total RNA Pico_DE24802219_2015-04-17_15-28-48.xad

Created: 4/17/2015 3:28:48 PM
Modified: 4/17/2015 5:51:22 PM

Electropherogram Summary Continued ...




Overall Results for sample 2 : ACF-14485 - 10

| | | | |
|---|---|---|---|
| RNA Area: | 352.9 | rRNA Ratio [23s / 16s]: | 41.9 |
| RNA Concentration: | 10,600 pg/µl | RNA Integrity Number (RIN): | N/A (B.02.08) |

Fragment table for sample 2 : ACF-14485 - 10

| Name | Start Size [nt] | End Size [nt] | Area | % of total Area |
|---|---|---|---|---|
| 16S | 1,025 | 1,332 | 3.2 | 0.9 |
| 23S | 1,351 | 2,348 | 135.2 | 38.3 |

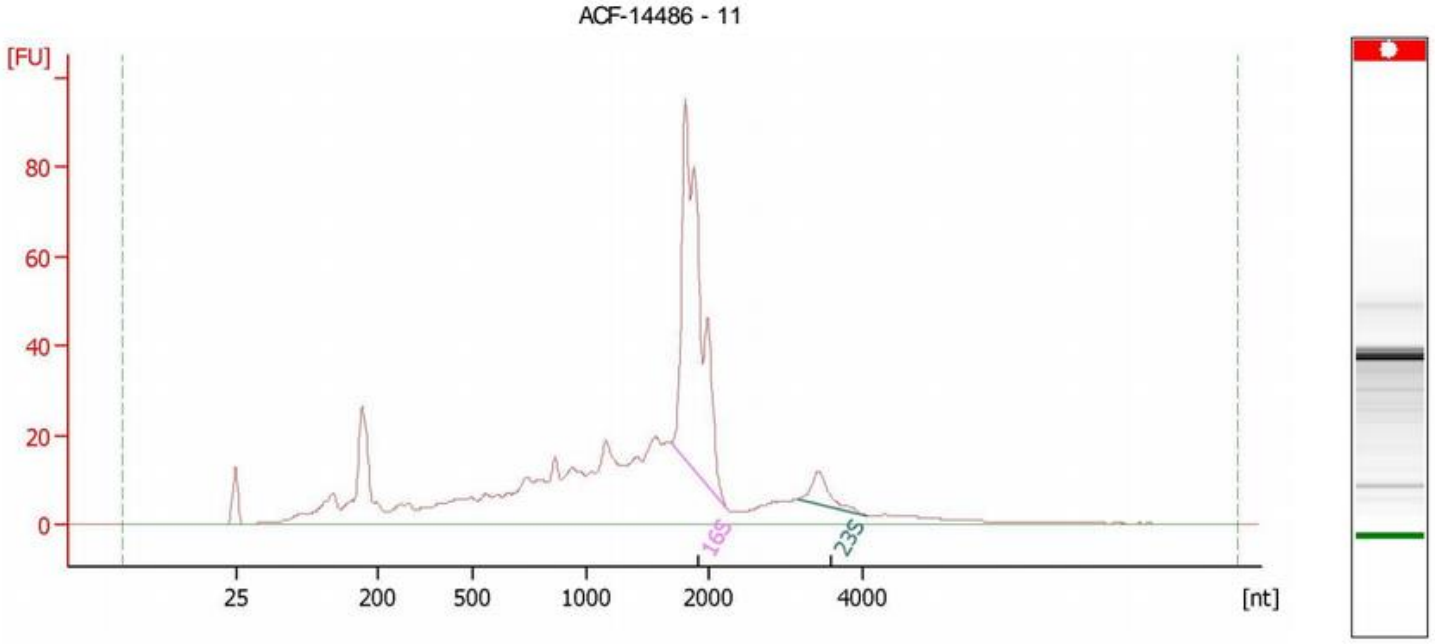
RNA SG C2



Assay Class: Prokaryote Total RNA Pico
Data Path: C:\...rokaryote Total RNA Pico_DE24802219_2015-04-17_15-28-48.xad

Created: 4/17/2015 3:28:48 PM
Modified: 4/17/2015 5:51:22 PM

Electropherogram Summary Continued ...




Overall Results for sample 3 : ACF-14486 - 11

| | | | |
|---|---|---|---|
| RNA Area: | 712.7 | rRNA Ratio [23s / 16s]: | 0.1 |
| RNA Concentration: | 21,405 pg/µl | RNA Integrity Number (RIN): | N/A (B.02.08) |

Fragment table for sample 3 : ACF-14486 - 11

| Name | Start Size [nt] | End Size [nt] | Area | % of total Area |
|---|---|---|---|---|
| 16S | 1,679 | 2,253 | 174.6 | 24.5 |
| 23S | 3,140 | 4,072 | 13.8 | 1.9 |

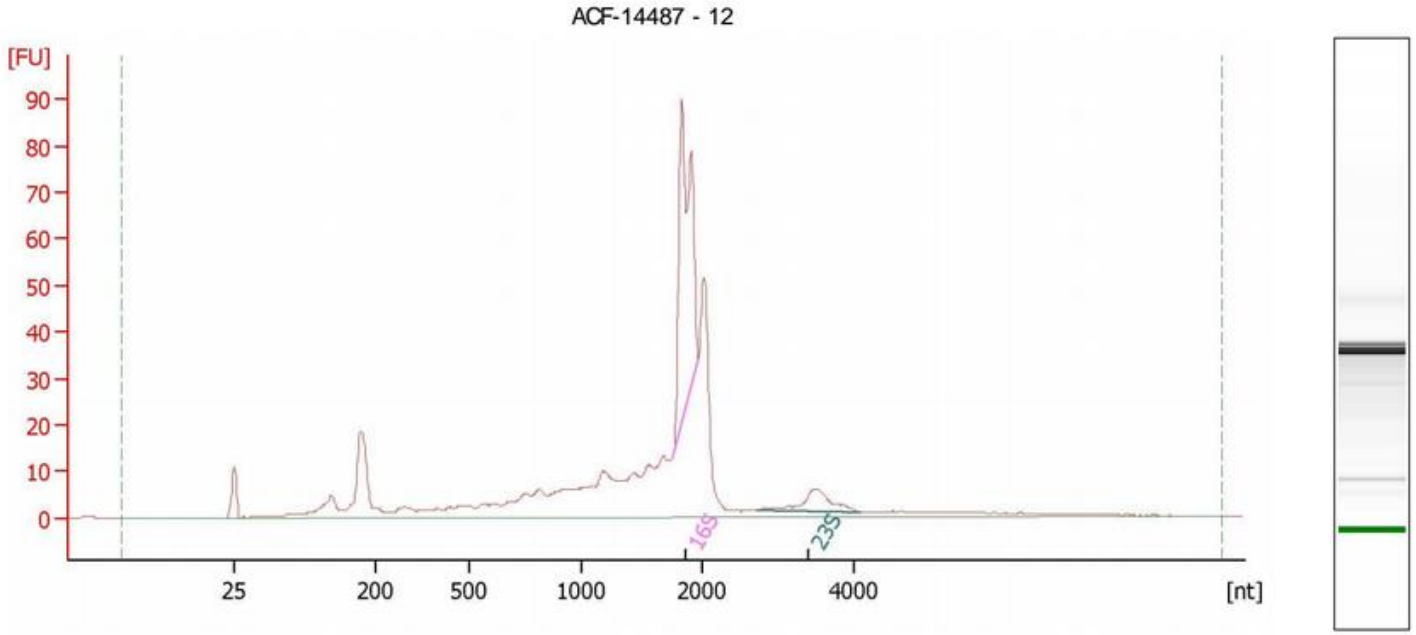
RNA SG I2



Assay Class: Prokaryote Total RNA Pico
Data Path: C:\...rokaryote Total RNA Pico_DE24802219_2015-04-17_15-28-48.xad

Created: 4/17/2015 3:28:48 PM
Modified: 4/17/2015 5:51:22 PM

Electropherogram Summary Continued ...




Overall Results for sample 4 : ACF-14487 - 12

| | | | |
|---|---|---|---|
| RNA Area: | 476.2 | rRNA Ratio [23s / 16s]: | 0.2 |
| RNA Concentration: | 14,301 pg/µl | RNA Integrity Number (RIN): | 6.5 (B.02.08) |

Fragment table for sample 4 : ACF-14487 - 12

| Name | Start Size [nt] | End Size [nt] | Area | % of total Area |
|---|---|---|---|---|
| 16S | 1,754 | 1,965 | 91.1 | 19.1 |
| 23S | 2,697 | 4,085 | 14.6 | 3.1 |

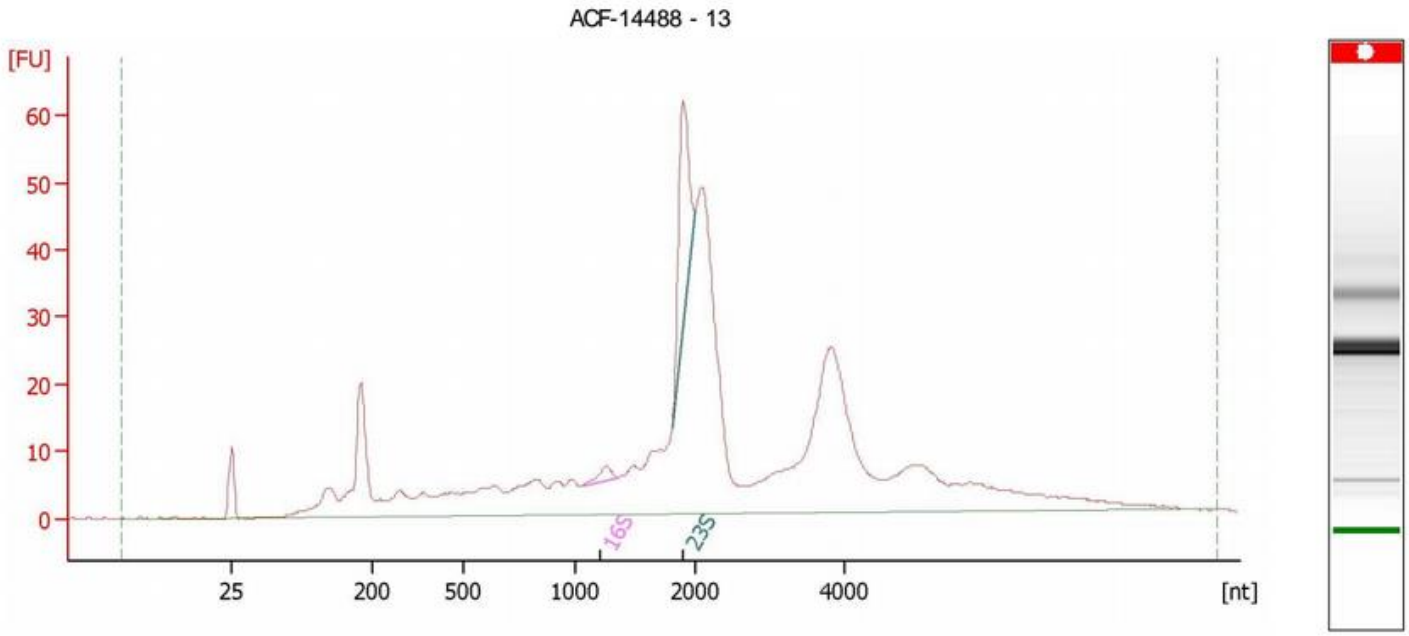
RNA SG C3



Assay Class: Prokaryote Total RNA Pico
Data Path: C:\...rokaryote Total RNA Pico_DE24802219_2015-04-17_15-28-48.xad

Created: 4/17/2015 3:28:48 PM
Modified: 4/17/2015 5:51:22 PM

Electropherogram Summary Continued ...




Overall Results for sample 5 : ACF-14488 - 13

| | | | |
|---|---|---|---|
| RNA Area: | 620.2 | rRNA Ratio [23s / 16s]: | 10.4 |
| RNA Concentration: | 18,627 pg/µl | RNA Integrity Number (RIN): | N/A (B.02.08) |

Fragment table for sample 5 : ACF-14488 - 13

| Name | Start Size [nt] | End Size [nt] | Area | % of total Area |
|---|---|---|---|---|
| 16S | 1,060 | 1,355 | 3.5 | 0.6 |
| 23S | 1,797 | 1,990 | 35.8 | 5.8 |

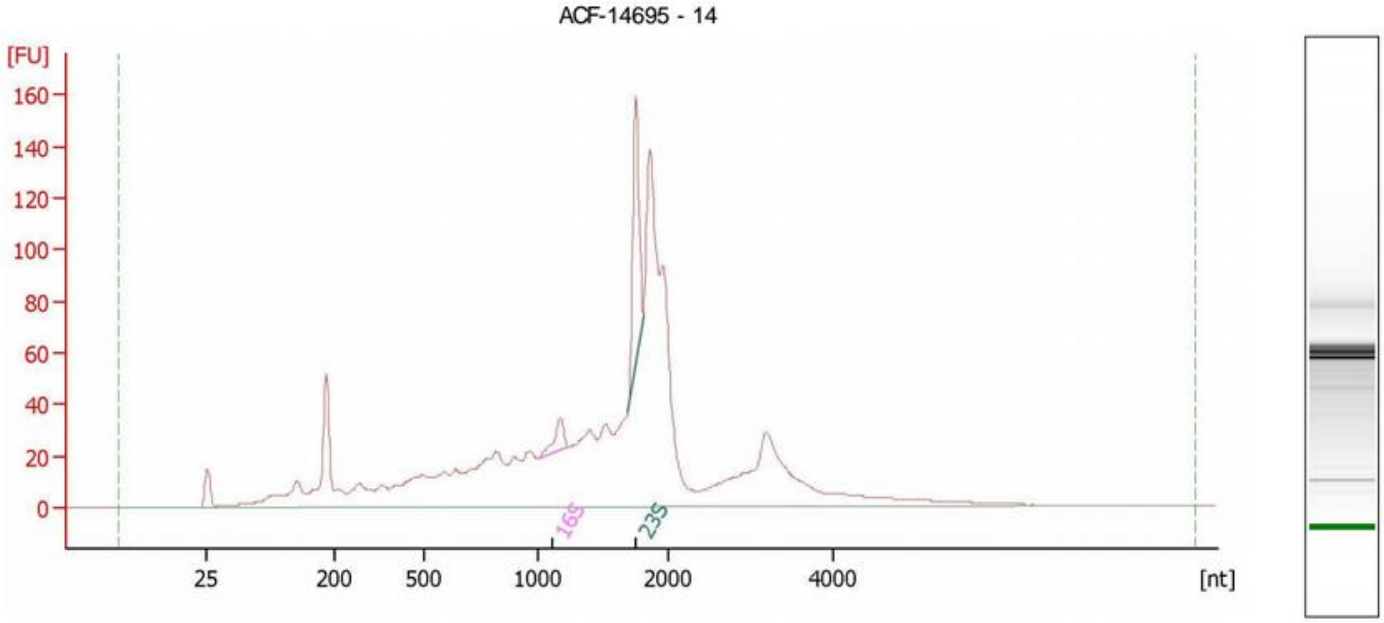
RNA SG I3



Assay Class: Prokaryote Total RNA Pico
Data Path: C:\...rokaryote Total RNA Pico_DE24802219_2015-05-15_16-14-40.xad

Created: 5/15/2015 4:14:40 PM
Modified: 5/18/2015 2:18:17 PM

Electropherogram Summary Continued ...




Overall Results for sample 4 : ACF-14695 - 14

| | | | |
|---|---|---|---|
| RNA Area: | 1,434.6 | rRNA Ratio [23s / 16s]: | 4.9 |
| RNA Concentration: | 5,261 pg/µl | RNA Integrity Number (RIN): | 5.2 (B.02.08) |

Fragment table for sample 4 : ACF-14695 - 14

| Name | Start Size [nt] | End Size [nt] | Area | % of total Area |
|---|---|---|---|---|
| 16S | 1,003 | 1,237 | 15.4 | 1.1 |
| 23S | 1,688 | 1,805 | 76.3 | 5.3 |

# Appendix VI

**Table 5- Differential gene expression in *A. stephensi* infected with *P. berghei.***

| Code | qPCR S7 | qPCR ELF | RNAseq |
|---|---|---|---|
| **ASTE004117** | 1,02 | -1,64 | 1,66 |
| **ASTE001733** | -0,11 | -2,76 | -2,65 |
| **ASTE009422** | -0,68 | -3,33 | -1,39 |
| **ASTE007758** | 0,73 | -1,93 | 4,54 |
| **ASTE004032** | 1,16 | -1,50 | 5,22 |
| **ASTE010763** | 1,49 | -1,17 | 5,67 |
| **ASTE011434** | 2,21 | -0,45 | 4,45 |
| **ASTE007038** | 1,85 | -0,81 | 2,17 |
| **ASTE009772** | 2,25 | -0,41 | 2,71 |
| **ASTE000811** | -1,00 | -3,65 | -3,16 |
| **ASTE008188** | 1,00 | -1,66 | 3,62 |
| **ASTE001475** | 2,55 | -0,10 | 1,50 |
| **ASTE009391** | 0.01 | -4.67 | 2.19 |

For each transcript code is represented the mean of the triplicates fold-changes from qPCR normalized against RPS7 and ELF using ΔΔCt. Is also represented the fold-change from each transcript from RNA-seq results.

# Appendix VII

>(Exon 1) primer forward sequenced
TTAATCATCCTACCTTCTACCAGGAGCTGAACGAAGAACAACCGAGCTCGGAAACGGTGGCGGCACGTGATGGCGC
ACTGCCGCCGATCGCTACCGTGTCCCATCCGATGTACCAGCAGCACCAGCTAAACGATGCGTATCAATATCGGAAAC
CGAAACGGGCACTGCATCGGGAGCTGGTGGCCAGCGTACGGCGGATGGATGCAAAGACGTGCTGCAGTACCGTCT
TTCCGCTGACGACCTGGTTGCCGGAGTACTCGTGGAGCAAGGATTTGGTGCGCGACTTGATTAGTGGGTGTACGGT
GGCTGTGATGCACATACCGCATGGTATCGGGTACGCTCTGTTGGCCAACGTGCCTCCGCTCGCTCGCATCCCTATGG
CTCTCCTTCCGGCGCTGGCGTTCTTCCTCATATACTGACTCCTATTAACCTCTATACCTCAATCTTACATTCACCTTTAC
CTTCTCCCCTACCCTATCCCCCCCCCTTTAATCTCTTCTTTTCTTATTCT

>(Exon 1) primer reverse sequenced
AAAATAAAAGCCATATAGATGCCGACGATCGGTGGCACGTTGGCCAACAGCGCGTACCCGATACCCTGCGGTATGTG
CATCACAGCCACCGTACACCCACTAATCAAGTCGCGCACCAAATCCTTGCTCCACGAGTACTCCGGCAACCAGGTC
GTCAGCGGAAAGACGGTACTGCAGCACGTCTTTGCATCCATCCGCCGTACGCTGGCCACCAGCTCCCGATGCAGTG
CCCGTTTCGGTTTCCGATATTGATACGCATCGTTTAGCTGGTGCTGCTGGTACATCGGATGGGACACGGTAGCGATCG
GCGGCAGTGCGCCATCACGTGCCGCCACCGTTTCCGAGCTCGGTTGTTCTTCGTTCAGCTCCTGGTAGAAGGTAGG
ATTGACAATACCACTCTCATGCCCTTCCTCTCCCTATAGTGAGTCGTATTAACCCCCCCTCAATTCCTCCTCCTCAAAC
CAACACATTATCATTCCTCCCTTACCTCTATCACAACACCCTTTTTTCCACTCTCCTTTACACTCTCTAAATC

# Appendix VIII

Length=511

| Sequences producing significant alignments: | | Score (Bits) | E Value |
|---|---|---|---|
| ASTE009391-RA | hypothetical protein\|protein_coding\|KB664609:126... | 654 | 0.0 |
| ASTE006714-RA | hypothetical protein\|protein_coding\|KB664766:801... | 64.4 | 5e-10 |
| ASTE008320-RA | hypothetical protein\|protein_coding\|KB664477:608... | 35.6 | 0.24 |
| ASTE009021-RA | hypothetical protein\|protein_coding\|KB665132:533... | 33.7 | 0.84 |
| ASTE008308-RA | hypothetical protein\|protein_coding\|KB664477:531... | 33.7 | 0.84 |
| ASTE002695-RA | hypothetical protein\|protein_coding\|KB664543:687... | 33.7 | 0.84 |
| ASTE014913-RA | hypothetical protein\|protein_coding\|KB664461:188... | 31.9 | 2.9 |
| ASTE011007-RA | hypothetical protein\|protein_coding\|KB664508:309... | 31.9 | 2.9 |
| ASTE010270-RA | hypothetical protein\|protein_coding\|KB664655:551... | 31.9 | 2.9 |
| ASTE009441-RA | hypothetical protein\|protein_coding\|KB664457:337... | 31.9 | 2.9 |

```
> ASTE009391-RA  hypothetical protein|protein_coding|KB664609:12628-
15846:-1|gene:ASTE009391
Length=2082

 Score =   654 bits (724),  Expect = 0.0
 Identities = 385/400 (96%), Gaps = 0/400 (0%)
 Strand=Plus/Plus

Query  7    ATCCTACCTTCTACCAGGAGCTGAACGAAGAACAACCGAGCTCGGAAACGGTGGCGGCAC  66
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  41   ATCCTACCTTCTACCAGGAGCTGAACGAAGAACAACCGAGCTCGGAAACGGTGGCGGCAC  100

Query  67   GTGATGGCGCACTGCCGCCGATCGCTACCGTGTCCCATCCGATGTACCAGCAGCACCAGC  126
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  101  GTGATGGCGCACTGCCGCCGATCGCTACCGTGTCCCATCCGATGTACCAGCAGCACCAGC  160

Query  127  TAAACGATGCGTATCAATATCGGAAACCGAAACGGGCACTGCATCGGGAGCTGGTGGCCA  186
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  161  TAAACGATGCGTATCAATATCGGAAACCGAAACGGGCACTGCATCGGGAGCTGGTGGCCA  220

Query  187  GCGTACGGCGGATGGATGCAAAGACGTGCTGCAGTACCGTCTTTCCGCTGACGACCTGGT  246
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  221  GCGTACGGCGGATGGATGCAAAGACGTGCTGCAGTACCGTCTTTCCGCTGACGACCTGGT  280

Query  247  TGCCGGAGTACTCGTGGAGCAAGGATTTGGTGCGCGACTTGATTAGTGGGTGTACGGTGG  306
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  281  TGCCGGAGTACTCGTGGAGCAAGGATTTGGTGCGCGACTTGATTAGTGGGTGTACGGTGG  340

Query  307  CTGTGATGCACATACCGCATGGTATCGGGTACGCTCTGTTGGCCAACGTGCCTCCGCTCG  366
            |||||||||||||||||| ||||||||||||| |||||||||||||||| ||| |||
Sbjct  341  CTGTGATGCACATACCGCAGGGTATCGGGTACGCGCTGTTGGCCAACGTGCCACCGATCG  400

Query  367  CTCGCATCCCTATGGCTCTCCTTCCGGCGCTGGCGTTCTT  406
               |||||  ||||||  || ||||||  |||||  || |||
Sbjct  401  TCGGCATCTATATGGCGTTCTTTCCGGTGCTGGTGTACTT  440
```

```
> ASTE006714-RA  hypothetical protein|protein_coding|KB664766:801820-
804774:-1|gene:ASTE006714
Length=2034

 Score = 64.4 bits (70),  Expect = 5e-10
 Identities = 104/150 (69%), Gaps = 0/150 (0%)
 Strand=Plus/Plus

Query  242  CTGGTTGCCGGAGTACTCGTGGAGCAAGGATTTGGTGCGCGACTTGATTAGTGGGTGTAC  301
            |||| ||||||| |||||||    ||  |||||  |  ||||  |||| |||||    ||
Sbjct  285  CTGGCTGCCGGAATACTCGTTTCCCAGCGATTTTATTGGCGATCTGATCAGTGGCCTAAC  344

Query  302  GGTGGCTGTGATGCACATACCGCATGGTATCGGGTACGCTCTGTTGGCCAACGTGCCTCC  361
             || | ||||||||||||| ||||| || || || || || ||  | || ||| |||| ||
Sbjct  345  CGTCGGTGTGATGCACATCCCGCAGGGGATGGGCTATGCGCTAATCGCTAACATGCCACC  404

Query  362  GCTCGCTCGCATCCCTATGGCTCTCCTTCC  391
            | || |  | |||  || |||  || ||||
Sbjct  405  GATCACCGGTATCTATACGGCGTTCTTTCC  434
```