

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL



**LOCALIZATION OF MRNA STORAGE
COMPLEXES IN *PLASMODIUM BERGHEI*
THROUGHOUT THE LIFE CYCLE**

Marcelo Luís Monteiro Pereira

Mestrado em Microbiologia Aplicada

2011

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL



**LOCALIZATION OF MRNA STORAGE
COMPLEXES IN *PLASMODIUM BERGHEI*
THROUGHOUT THE LIFE CYCLE**

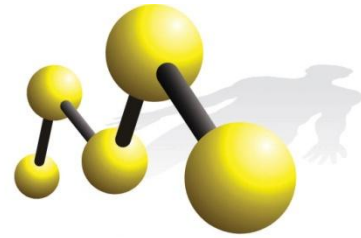
Marcelo Luís Monteiro Pereira

Dissertação orientada por:

Prof. Doutora Rita Zilhão, Faculdade de Ciências, Universidade de Lisboa

Doutor Gunnar R. Mair, Instituto de Medicina Molecular

**Mestrado em Microbiologia Aplicada
2011**



**INSTITUTO DE
MEDICINA MOLECULAR**

Faculdade de Medicina
Universidade de Lisboa

**LOCALIZATION OF MRNA STORAGE
COMPLEXES IN *PLASMODIUM BERGHEI*
THROUGHOUT THE LIFE CYCLE**

Marcelo Luís Monteiro Pereira

MASTER THESIS

2011

This thesis was fully performed at the Molecular Parasitology Unit in Instituto de Medicina Molecular under the direct supervision of Dr. Gunnar Mair.

Prof. Dr. Rita Zilhão was the internal designated supervisor in the scope of the *Master in Applied Microbiology* of the Faculty of Sciences of the University of Lisbon.

Abbreviations

3' UTR	3' untranslated region
4E	eIF4E
4G	eIF4G
5' UTR	5' untranslated region
BLAST	Basic local alignment search tool
cDNA	complementary deoxyribonucleic acid
CITH	<u>C</u> AR- <u>I</u> and fly <u>T</u> railer <u>H</u> itch <u>h</u> omolog
DNA	Deoxyribonucleic acid
DOZI	<u>D</u> evelopment of <u>z</u> ygote <u>i</u> nhibited
DTT	Dithiothreitol
EEF	Exo-erythrocytic form
eIF	Eukaryotic translation initiation factor
FELASA	Federation of laboratory animal science associations
GFP	Green fluorescent protein
HoBo	<u>H</u> omolog of <u>B</u> runo
HoMu	<u>H</u> omolog of <u>M</u> usashi
I.P.	Intraperitoneal injection
mRNP	messenger ribonucleoprotein
PABP	poly(A)-binding protein
PCR	Polymerase chain reaction
PhyML	Phylogenetic estimation using maximum likelihood
RBC	Red blood cell
RBP	RNA-binding proteins
RNA	Ribonucleic acid
RRM	RNA recognition motif
RT-PCR	Reverse-transcription polymerase chain reaction
SPF	Specific pathogen free
TIC	Translation initiation complex
WAG	Whelan and Goldman
WT	Wild type

Table of contents

Abbreviations.....	0
Acknowledgments.....	6
Abstract	7
Resumo	8
1 – Introduction.....	10
2 - Material and Methods	17
2.1 - Animal work.....	17
2.2 - Parasite Lines.	17
2.3 - Mice infections	17
2.4 - Mosquito infection.	17
2.5 - Determination of parasitemia and gametocytemia by Giemsa stain.....	18
2.6 - Ookinete cultures	18
2.7 - GFP fusion protein expression in blood stages.....	19
2.8 - GFP fusion gene expression in mosquito stage parasites.	19
2.9 - RNA extraction from gametocytes and sporozoites and RT-PCR.....	20
2.10 - HoMu and 4E tree and alignment.....	21
2.11 - HoMu and 4E Domain analysis	21
2.12 - Retrieval of <i>P. berghei</i> gene models and putative untranslated regions.....	21
3 – Results.....	22
3.1- HoMu and 4E are evolutionary highly conserved proteins	22
3.2 - HoMu is transcribed in gametocytes and sporozoites.....	27
3.3 - Strong GFP expression was detected in HoMu::GFP and 4E::GFP gametocytes	28
3.4 - HoMu::GFP and 4E::GFP are expressed in oocysts.....	30
3.5 - HoMu::GFP and 4E::GFP are expressed in mosquito salivary gland sporozoites	31
3.6 - HoMu::GFP and 4E::GFP parasites show normal blood stage development and morphology.	33
3.7 - Mosquito stage development of HoMu::GFP and 4E::GFP parasite lines is normal.....	35
3.8 - Re-infection of mice with mutant parasite line sporozoites – HoMu::GFP and 4E::GFP re-infections were unsuccessful	36
3.9 - More than 2000 <i>P. berghei</i> genes contain consensus RBS in their 3' untranslated regions	37

4 – Discussion	38
5 – Appendix.....	41
7 – References	46

Acknowledgments

This project was developed in Unidade de Parasitologia Molecular (UPAMOL) at Instituto de Medicina Molecular (IMM) under the supervision of Dr Gunnar Mair.

From the UPAMOL team, I would like to thank Gunnar for supervising my project; Natalia for helping me with molecular biology experiments: Celine, Ana and Jorge for giving me support with mice experiments.

From malaria unit team I would like to thank Fernanda Baptista and Ana Parreira for mosquitoes providing and supervising and for helping me with mosquitoes dissection; António Mendes for helping me with dissection and imaging of mosquito's salivary glands.

To José Rino and António Temudo from Biolmaging Unit, thank you for helping me with all imaging work.

Finally I would like to thank Prof. Rita Zilhão for supervising my project and Prof. Rogério Tenreiro for evaluating my work during the last two years.

Abstract

Plasmodium is the causative agent of malaria, a disease that caused 781 thousands deaths during the year of 2009. The apicomplexan responsible for this disease have shown to be very well adapted to both mosquito and Vertebrate hosts, regulating its gene expression often in a posttranscriptional manner. Recently the protein HoMu and the translation initiation factor eIF4E (4E) were found forming complexes with translationally arrested mRNAs (e.g. transcripts encoding the surface proteins p25 and p28). Storage of these mRNAs in the female gametocyte will permit the parasite to have them readily available to be translated once it enters into the mosquito. Using bioinformatics tools we showed that HoMu and 4E are conserved proteins. Their homologs are distributed in a wide range of species and participate in both translation initiation and arrest of transcripts containing specific sequences in their 3' untranslated regions (UTRs). The main purpose of this study was to understand when those proteins are expressed during the *Plasmodium berghei* life cycle and where they are localized. We also show that parasite lines expressing GFP fused to HoMu and 4E are able to establish infection with normal parasitological patterns compared to the wild type. By life cell imaging of these mutant lines we find that HoMu and 4E are strongly expressed in gametocytes, oocysts, and sporozoites suggesting that they are important during the transmission phases of the *P.berghei* life cycle.

Key words: Malaria, HoMu, eIF4E, gametocyte, post-transcriptional regulation

Resumo

A malária é uma doença causada pelo parasita unicelular *Plasmodium* e transmitida pela picada do mosquito *Anopheles*. Segundo os dados da organização mundial de saúde (OMS) esta doença vitimou em 2009 cerca de 780 mil pessoas, sendo a maioria crianças que habitavam na África subsariana. São quatro as espécies deste parasita que infectam o Homem: *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae*, e *Plasmodium falciparum*, sendo a última a causadora da forma mais severa da doença conhecida como malária cerebral.

Actualmente a investigação em parasitologia molecular procura desvendar quais os processos celulares que estão por de trás da interacção deste parasita com o hospedeiro. Como tal e visto ser ainda difícil reproduzir todo o ciclo de vida de *Plasmodium* em laboratório é necessário utilizar um hospedeiro passível de ser mantido em cativeiro e que sirva de modelo para o estudo da malária. A malária infecta também roedores e é sabido que as espécies do parasita que os infectam partilham elevado número de genes com as espécies da malária humana. Como tal é muitas vezes utilizado em investigação o parasita de roedores *Plasmodium berghei*.

O parasita da malária caracteriza-se pela sua rápida adaptação ao hospedeiro. Durante a fase sanguínea o parasita invade os eritrócitos e diferencia-se em células precursoras dos gametas, os gametócitos. Estas células caracterizam-se por possuir complexos contendo mRNA e proteínas específicos no seu citoplasma onde a tradução é inibida. Pensa-se que estes complexos estão relacionados com o controlo da expressão génica no parasita da malária. De facto poucos factores de transcrição são ainda conhecidos em *Plasmodium*, por outro lado, as proteínas de ligação ao RNA são particularmente abundantes neste parasita. Estas observações reforçam a hipótese de que, em *Plasmodium* a regulação génica ser feita preferencialmente ao nível pós-transcricional.

Em gametócitos femininos de *Plasmodium* foram identificadas ribonucleoproteínas (mRNP) como DOZI e CITH formando complexos onde o mRNA é armazenado. Ao gerar-se linhas celulares mutantes do parasita sem estas proteínas codificadas no seu genoma verificou-se que este não se diferenciava em oocinetos após a formação do zigoto. Fazendo parte dos mesmos complexos foram encontradas outras proteínas, entre elas uma proteína homóloga de Musashi (HoMu) e um factor de iniciação da tradução eIF4E (4E). Estas proteínas são o alvo deste estudo e possuem homologia com proteínas de diferentes espécies de seres vivos. HoMu por exemplo possui

homologia com a proteína Musashi do sapo *Xenopus*. Neste animal a proteína Musashi liga-se ao mRNA *mos* de oócitos activando a sua tradução. Noutras espécies, no entanto, Musashi caracteriza-se por inibir a tradução de certos mRNAs, armazenando-os no citoplasma em complexos de mRNP. Já 4E por outro lado é um factor de iniciação da tradução em células eucariotas que juntamente com outros factores permite activar o mRNA para ser traduzido. No entanto esta proteína participa também em processos de inibição da tradução, integrando complexos de mRNP onde o mRNA é armazenado. Em *Drosophila* por exemplo a proteína Cup liga-se a 4E impedindo a formação do complexo de iniciação da tradução.

Este projecto teve como objectivo principal identificar o padrão de expressão das proteínas HoMu e 4E durante o ciclo de vida do parasita *P. berghei*. Para isso foram utilizadas duas linhas celulares contendo os genes *homu* e *eif4e* de fusão com o gene *gfp* (*green fluorescent protein*) de forma a possibilitar a identificação da expressão de HoMu e 4E por análise da fluorescência. As duas proteínas revelaram ser expressas na fase sanguínea bem como no mosquito, sugerindo a importância de ambas no ciclo de vida de *Plasmodium*.

Foi demonstrado neste projecto que as linhas celulares GFP possuem um desenvolvimento normal quando comparado com a linha selvagem. A cinética de crescimento do parasita mutante durante a fase sanguínea foi idêntica à do selvagem e os números de esporozoítos nas glândulas salivares do mosquito foram normais. Isto comprova que as linhas celulares GFP utilizadas neste projecto são viáveis para o estudo das proteínas HoMu e 4E.

Finalmente, a análise bioinformática das regiões 3' UTR (*untranslated regions*) dos quase 5000 genes codificantes de proteínas identificados até então em *P. berghei* revelou que muitos deles possuem sequências *consensus* nas quais as proteínas Musashi de outros seres vivos se ligam. Isto leva-nos a sugerir que o homólogo de Musashi em *Plasmodium* possa também ligar-se às regiões 3' UTR de diversos mRNAs inibindo ou não a tradução dos mesmos.

Palavras-chave: Malária, HoMu, eIF4E, gametócito, regulação pós-transcricional

1 – Introduction

Malaria is a mosquito-borne disease that caused 781 thousand deaths during the year of 2009 (WHO 2010), the majority of them young children in sub-Saharan Africa (Figure 1). The disease is caused by apicomplexan parasites of the genus *Plasmodium* which are transmitted during the bite of the primary host, the *Anopheles* mosquito. Four species of *Plasmodium* infect humans: *Plasmodium vivax*, *Plasmodium ovale*, *Plasmodium malariae*, and *Plasmodium falciparum* which cause the most severe form of the disease.

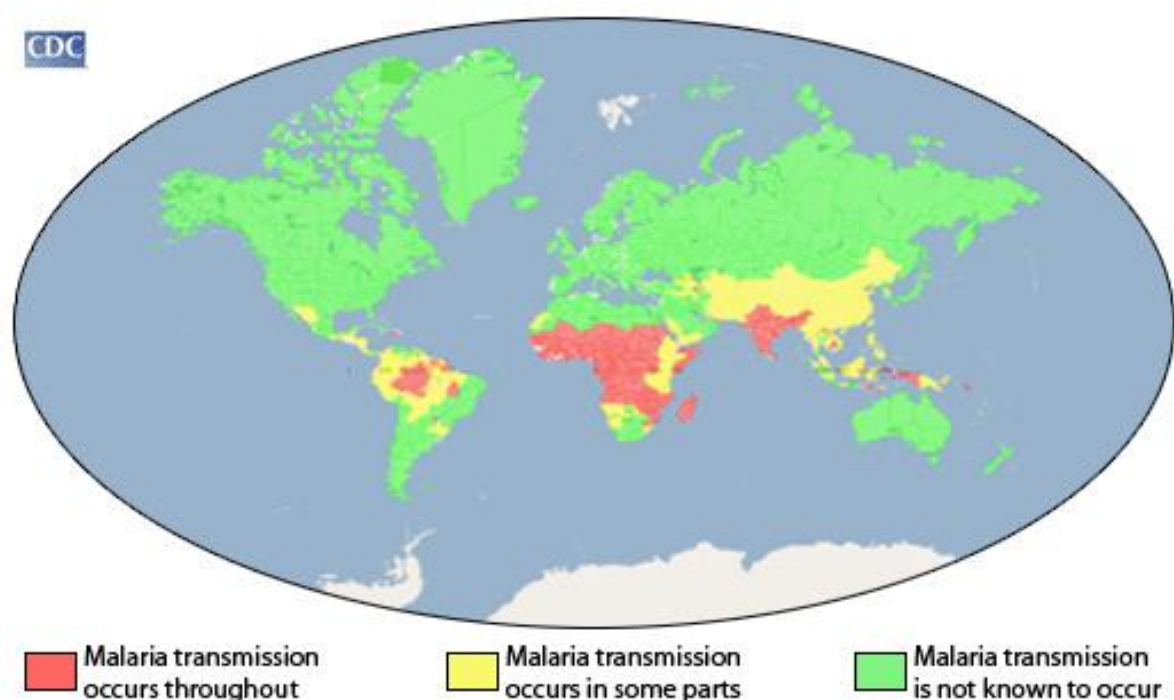


Figure 1| Extent of malaria. Map of the approximated parts of the world where malaria transmission occurs. Adapted from CDC website (<http://www.cdc.gov/malaria/>).

The *Plasmodium* life cycle starts when a mosquito with sporozoites in its salivary glands bites the intermediate host (a human or an experimental mouse host) to take a blood meal and injects infected saliva. Once in the circulatory system, sporozoites head for the liver where they establish an asymptomatic primary infection. Inside hepatocytes they develop and change into non motile and replicating exo-erythrocytic forms (EEFs); this change was induced in vitro by a shift in temperature (Doi, Shinzawa *et al.* 2011). In the next 8- 30 days the parasite differentiates into thousands of merozoites which are released to liver sinusoids after hepatocyte rupture (Prudencio, Rodriguez *et al.* 2006). In

the blood, these merozoites infect red blood cells causing the symptomatic phase of the disease. Malaria is characterized by severe anemia, fever, vomiting and fatigue.

Inside red blood cells (RBC) the parasite differentiates first into ring forms, followed by trophozoites and mature schizonts. When the infected RBC finally ruptures it releases new merozoites that will infect new RBCs. Alternatively some parasites, once inside RBCs, differentiate into sexual erythrocytic stages (gametocytes). Only these forms, when taken up by a mosquito, are capable of infecting it. In the mosquito midgut, changes in pH and temperature will trigger gametocytes differentiation into gametes (Billker, Shaw *et al.* 1997). Sexual reproduction takes place with the fusion of male and female gametes, and formation of a zygote that develops into ookinete. This motile form can cross the midgut epithelium and establish an oocyst to generate sporozoites that will be released into the hemocoel after 10-15 days. Free sporozoites eventually invade salivary glands of the mosquito (Figure 2)(Wirth 2002).

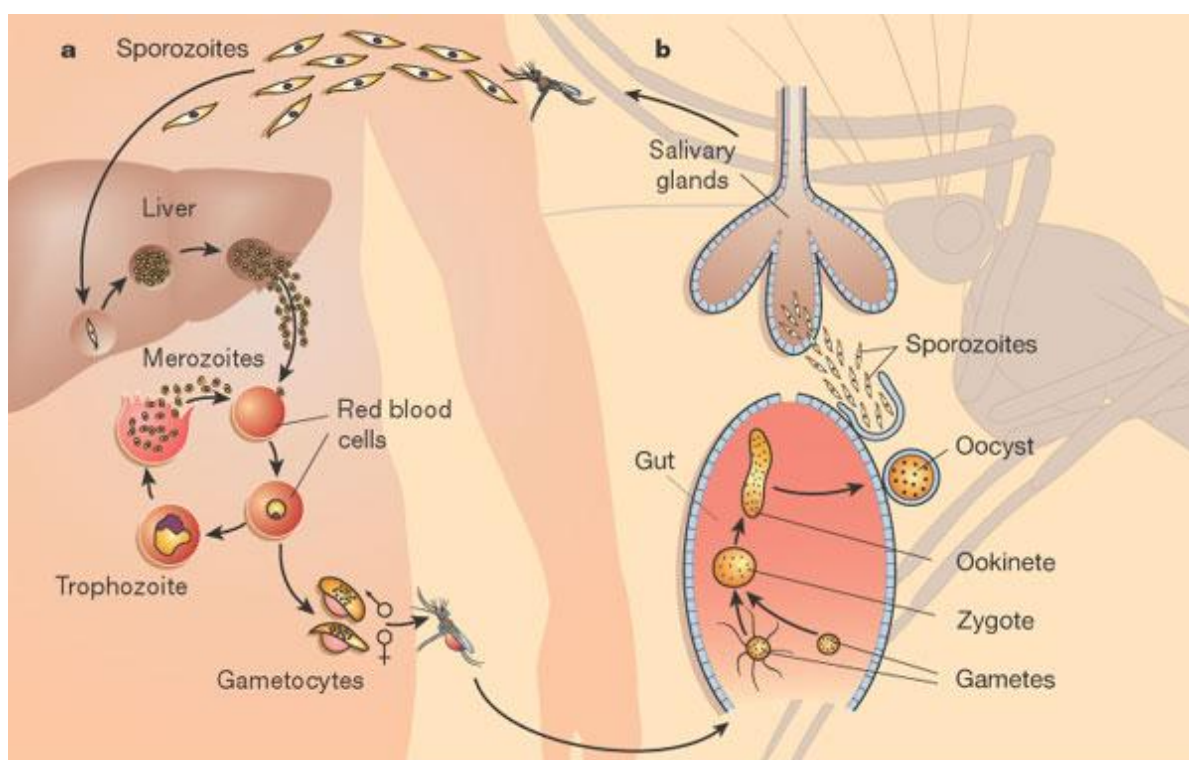


Figure 2 | Malaria life cycle. The model represents the different stages of *Plasmodium* in both human (on the left) and mosquito (on the right) hosts. Adapted from (Wirth 2002).

The *Plasmodium falciparum* genome encodes around 5300 genes. The majority of proteins encoded in the genome is of unknown function as 60% of them lack sufficient similarity to proteins encoded in other organisms. Not surprisingly, compared to free-living eukaryotic genomes, the *P. falciparum* genome possesses a large number of

genes that are important in host-parasite interactions and host immune system evasion (Gardner, Hall et al. 2002).

It is still difficult to reproduce the entire life cycle of the human malaria parasite *P. falciparum* in a laboratory setting, therefore an animal model in which *Plasmodium*'s infection is known to occur and which can be kept and more easily manipulated is central. Malaria is known in rodents; importantly human and rodent malaria parasites share a high number of genes (Carlton, Vinkenoog et al. 1998; Carlton, Angiuoli et al. 2002; Hall, Karras et al. 2005). This is the reason why rodent malaria parasites like *Plasmodium yoelii* or *Plasmodium berghei*, which is used in this study, have been developed as models to study human malaria.

Understanding how *Plasmodium* gene expression – from transcription to protein translation – is regulated is fundamental for developing new drugs and vaccines against this parasite. The discrepancy between mRNA and protein levels found during the *P. falciparum* life cycle (Le Roch, Johnson et al. 2004) reinforced the possibility that this parasite widely regulates gene expression post-transcriptionally. This hypothesis is also supported by the fact that the *Plasmodium* genome contains only $\frac{1}{3}$ of the known transcription associated factors commonly found in the genome of free-living eukaryotes. Instead, RNA-binding proteins (RBP) that can regulate translation and RNA decay are abundant (Coulson, Hall et al. 2004). Understanding the mechanisms how these factors regulate gene expression is clearly important.

In eukaryotes, unlike prokaryotes, transcription and translation occur in different compartments of the cell. Transcription and pre-mRNA processing (capping, cis-splicing and polyadenylation) occur in the nucleus and involves countless RBPs that assemble into large complexes. After pre-mRNA processing most processing proteins disassemble from the mature mRNA. Those RBPs related with nuclear export however remain attached and assemble to form a cargo-carrier complex. Once mRNA is exported to cytoplasm it can be directed to translation initiation complexes (TIC).

Messenger RNAs possess well-characterized, general sequence elements that are fundamental to translational regulation. These include the poly(A) tail and the m⁷G cap nucleotide; both support mRNA stability and act as determinants for translation efficiency. For translation initiation 5' UTR (untranslated regions) elements like the 5' cap, are central. Translation initiation starts with the binding of the small (40S) ribosomal subunit at or in proximity of cap. Aided by translation initiation factors such as eIF3 and eIF4E (4E) the ribosomal subunit scans the mRNA until it reaches the start codon

(AUG). Initiation factors are then released, the large (60S) ribosomal subunit on the other hand binds to form the 80S ribosome and elongation begins (Wilkie, Dickson et al. 2003).

Often additional, RNA-specific regulatory motifs are located within a transcript's UTRs. These mostly short motifs, function as binding sites for RBPs that can control translation initiation. Homologs of such RBPs were identified in *P. berghei* (Mair, Lasonder et al. 2010) and include: DOZI (development of zygote inhibited; an homolog of yeast DEAD box RNA helicase Dhh1); poly(A)-binding protein (PABP; which contain RNA recognition motifs - RRM); an homolog of Bruno, a protein that mediates translational repression of *oskar* mRNA in the ovary of *Drosophila* (Kim-Ha, Kerr et al. 1995); CITH, a Trailer Hitch protein (homolog of CAR-I protein in *Caenorhabditis elegans*). An homolog of CITH was found as part of a cytoplasmic RNA-protein complex of *Drosophila* embryo. This complex also includes the translation repression factor Me31b (an homolog of DOZI) and Cup, a protein also known to translationally repress *oskar* mRNA in *Drosophila* (Wilhelm, Buszczak et al. 2005). *oskar* mRNA contains a 3' UTR response element for silencing where Bruno binds (Kim-Ha, Kerr et al. 1995), showing that translational repression can occur by RBP binding to a 3' UTR in mRNA.

All the above RBPs were found associated with translationally quiescent mRNAs in cytoplasmic mRNA-protein complexes (P-granules) in female *P. berghei* blood stage gametocytes (Mair, Lasonder et al. 2010). Genetic disruption of *dozi* and *cith* led to the targeting of more than 350 different transcripts for mRNA degradation pathways clearly identifying a stabilizing effect of these factors on certain mRNAs. In addition, gene deletion mutants showed post-fertilization development defects; this demonstrates that mRNAs translationally silenced, are necessary for ookinete development of the fertilized zygote. A similar process is also found in *C. elegans*, where the RNA helicase CGH-1 (the DOZI homolog) is important in translational repression and stability of stored maternal mRNAs (Rajyaguru and Parker 2009).

DOZI acts together with CITH (homolog of CAR-I and fly Trailer Hitch) and 14 additional factors in *Plasmodium* gametocyte mRNPs (messenger ribonucleoprotein) complexes; apart from the ones mentioned above, they include Alba domain DNA/RNA binding proteins, translation initiation factor 4E and a Musashi homolog (Mair, Lasonder et al. 2010) (Figure 3). Together they stabilize mRNAs like *p25* and *p28* that encode proteins essential for zygote development and mosquito midgut invasion. In this study, HoMu (homolog of Musashi with two RNA Recognition Motifs) and 4E will be central.

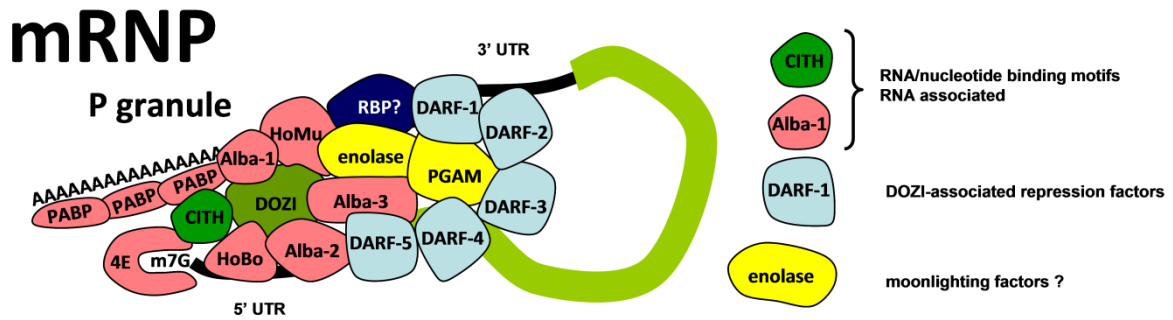


Figure 3 | *Plasmodium berghei* P granule. Model of gametocyte mRNP complex proteins that were isolated by immunoprecipitation (Mair, Lasonder et al. 2010). DOZI and CITH clearly affect the abundance of quiescent mRNA in female gametocytes; all additional proteins were identified by reciprocal immunoprecipitation of DOZI::GFP and CITH::GFP and are implicated in maintaining mRNAs translationally repressed until ookinete development.

Homologs of HoMu can be found in a wide range of species. Its function is unknown in most of them, but in mice, *Drosophila* and *Xenopus* this protein is well studied (Table 1). Musashi-1 for example, competes with PABP for eIF4G (4G) binding (Kawahara, Imai et al. 2008). It is known that PABP binds to 4G to promote translation initiation (Kahvejian, Svitkin et al. 2005). Therefore, competition of Musashi-1 with PABP1 for 4G binding may cause an arrest in translation. Indeed, examples of this are known: Musashi-1 prevents *numb* mRNA from being translated by binding to a 3' UTR sequence in mouse cells (Table 1) (Imai, Tokunaga et al. 2001); in *Drosophila* Musashi repress translation of *ttk69* by binding to a specific sequence in its 3'UTR (Table 1)(Okano, Imai et al. 2002). In the above cases Musashi regulates gene expression by mRNA binding and translation inhibition. In *Xenopus* (Table 1) on the other hand Musashi is responsible for maternal binding and translational activation of *mos* mRNA inside oocytes (Charlesworth, Wilczynska et al. 2006). Consequently it would be important to know if and how HoMu can activate or repress translation of certain *Plasmodium* mRNAs in blood stage gametocytes or other life cycle stages.

Species	Target mRNA	Consensus sequence	Function	Effect
Mouse	numb	(G/A)U(1-3)AGU GUUAGUUAGUUAGUU	Notch signaling	Repression
<i>Xenopus</i>	mos	AUAGU	Activate MAP permitting meiosis progression	Activation
<i>Drosophila</i>	ttk69	GU ₃₋₆ (G/AG)	Notch signaling	Repression

Table 1 | Musashi function in metazoans. Table of Musashi targets, consensus binding sequences in 3' UTRs of target mRNAs, and their role in three metazoan species where this RNA binding protein is well characterized.

The protein 4E is a conserved m⁷G cap binding protein that promotes cap-dependent translation initiation. 4E is part of the eIF4F complex. This complex contains factors that are responsible for unwinding the cap-proximal region of mRNA to make ribosomal attachment possible (eIF4A, eIF4B and eIF4H). Circularization of the mRNA is achieved by 4G interaction with PABP which in turn bind to the mRNA poly(A) tail. Simultaneously 4E interacts with 4G closing the mRNA in a circle. Unwinding and circularization permits the 43S pre-initiation complex to identify the mRNA start codon and to initiate translation elongation (Sonenberg and Hinnebusch 2009; Jackson, Hellen et al. 2010). These steps can be controlled by several means: for example, in *Drosophila* translation is driven by proteins that compete with 4G for 4E binding. In *Drosophila* oocytes, maternal unlocalized *nanos* mRNA translation is repressed by binding of the Cup protein to 4E, avoiding the recruitment of 4G (Nelson, Leidal et al. 2004). Cup is also required during *oskar* translational repression and was shown interacting with Bruno and with 4E (Nakamura, Sato et al. 2004). In *P. berghei* an homolog of Bruno (HoBo) was found in mRNP complexes of gametocytes also containing DOZI, CITH, HoMu and 4E, by mass-spectroscopic analysis (Mair, Lasonder et al. 2010). So far, no Cup homolog protein was found in *P. berghei*. It is therefore unknown whether a similar mechanism occurs in this parasite. Translation repression of *oskar* mRNA can occur without interaction of Cup with 4E (Chekulaeva, Hentze et al. 2006). Maybe a similar process can occur with HoBo containing mRNP complexes in *P. berghei*.

Translational control of mRNAs plays an important role during transmission of the malaria parasite from and to the mosquito. Only 3 proteins have been characterized in detail: they are DOZI and CITH, as well as Puf2 (identified in small speckles of salivary gland sporozoites (Gomes-Santos, Braks et al. 2011)). DOZI and CITH are core components of gametocyte mRNP complexes necessary to ookinete development (Mair, Lasonder et al. 2010). In *P. berghei* Puf2 regulates translation of early liver stage mRNAs (Gomes-Santos, Braks et al. 2011; Muller, Matuschewski et al. 2011). In *P.*

falciparum Puf2 also plays a role during gametocytogenesis (formation of gametocytes); differentiation of male gametocytes is increased when Puf2 is disrupted (Miao, Li et al. 2010).

The aim of this study is to phenotypically characterize parasite forms that express C-terminally GFP-tagged (green fluorescent protein tagged) forms of HoMu and 4E (HoMu::GFP and 4E::GFP), identify temporal and spatial protein expression patterns, and thereby evaluate their potential to be used in biochemical characterization of larger protein complexes involved in regulation of gene expression at a post-transcriptional level during parasite transmission.

2 - Material and Methods

2.1 - Animal work. All mice used in this work were Balb/c from Charles River|France. They were kept in an SPF (Specific Pathogen Free) Rodent Facility at the IMM. All animal experiments were in accordance with FELASA (Federation of Laboratory Animal Science Associations) recommendations. I am holder of a FELASA B animal license for laboratory animal experimentation given at the IMM following FELASA recommendations for animal laboratory experimentation.

2.2 - Parasite Lines. The following *P. berghei* ANKA parasite lines were used in this work: Pb683 (DOZI::GFP); Pb909 (CITH::GFP); 1750cl3 (Puf2::GFP); Pb910 (HoMu::GFP); Pb964 (4E::GFP) and 259cl2 (wild type GFP). All The parasite lines used in this work express GFP. In the wild type line GFP is expressed constitutively. The lines DOZI::GFP, CITH::GFP, Puf2::GFP, HoMu::GFP and 4E::GFP express the respective proteins DOZI, CITH, Puf2, HoMu and 4E fused with GFP. All lines except the wild type GFP (WT::GFP) contain a pyrimethamine resistance marker (dhfr/ts) in their genomes.

2.3 - Mice infections. Mice infections with DOZI::GFP, CITH::GFP, Puf2::GFP, HoMu::GFP, 4E::GFP and WT::GFP parasite lines were carried out by intraperitoneal injection (I.P.) of up to 200 μ L of thawed glycerol stocks cryopreserved at -80 °C (glycerol/PBS solution with 30% glycerol v/v). The parasitemias of the cryopreserved vials were 4% for DOZI::GFP, 3% for CITH::GFP, 6% for WT::GFP, and for Puf2::GFP, 250 μ L contained 10^7 parasites. Parasitemia levels were allowed to reach a maximum of 10% whereupon mice were sacrificed in a carbon dioxide chamber. When necessary blood was taken by cardiac puncture using an insulin syringe containing 50 μ L Heparin from a stock solution of 200 I.U./mL of heparin dissolved in RPMI-1640 medium.

2.4 - Mosquito infection. For mosquito infection 2 mice per line were infected via I.P. from cryopreserved vials. At day 4 exflagellation was checked; if positive, mice were anesthetized with a mixture of anesthetics: 12% Imalgene 1000 plus 8% Rompum (2%) (V/V); 25 female mosquitoes were allowed to blood feed for ½ hour. Midguts were checked for oocysts 13 days after the blood feeding. Sporozoites were observed and counted 21 days after infection.

To observe exflagellation of male gametocytes, slides were made at 4th day after mice infection. Exflagellation was induced by a decrease in temperature (from 37°C to 21°C) and a rise in pH from 7.3 to 8.0. Rise of pH in a blood droplet was achieved by blowing onto it. The lowering in temperature was done by leaving the slide at room temperature

for 10 minutes. Slides were done by taking a drop of blood from the mice tail into a slide after piercing it with a G26 syringe needle and squeezing the blood from it. Microscopic observation was performed in Leica DM2500 brightfield microscope at x1000 magnification.

2.5 - Determination of parasitemia and gametocytemia by Giemsa stain. The percentage of parasitemia was measured in blood droplets obtained from infected mice. Briefly, tails were pierced with a 26G syringe needle; the blood squeezed gently onto a glass slide and smeared using a second slide; the smear was air-dried, fixed with 100% methanol for 10 seconds, and stained with Giemsa solution for 10 minutes; finally the smear was rinsed with tap water and air-dried for 10 minutes. Number of infected red blood cells was then counted using a Leica DM2500 brightfield microscope. Parasitemia is the average percentage of infected red blood cells with any parasite development stage in all counted red blood cells and was determined by counting all infected versus non infected erythrocytes in 3 microscopic fields at x1000 magnification. Gametocytemia is the average percentage of infected red blood cells with gametocytes and was determined by the same way as parasitemia.

2.6 - Ookinete cultures. Ookinete cultures were developed expressing DOZI::GFP, CITH::GFP, Puf2::GFP, HoMu::GFP, 4E::GFP and WT::GFP. Mice infected with those lines were sacrificed once parasitemias reached 10% and the whole blood was taken by heart puncture. The total blood of each mouse was mixed with 10 mL of 1x RBC lysis buffer (0.15M NH₄CL, 0.01M KHCO₃, 1mM Na₂EDTA; pH 7.4). The mixture was then incubated on ice for 4 minutes and spin for 8 minutes 2000 RPM in a falcon centrifuge at room temperature. The resulting pellet was mixed with ookinete culture medium: RPMI-1640 Medium HEPES Modification, with L-glutamine and 25mM HEPES, without sodium bicarbonate (Sigma R4130); 92µM hypoxanthine (sigma H9377); 50000 units/L of Pen/Strep (Jena Bioscience ML-105); 100µM xanturenic acid (sigma D120804) in 6mM NaHCO₃; filter-sterilized; final pH 7.4. 20% FBS was added before mixing with the pellet. The cultures were incubated at 19°C for 16 hours with shaking at 50 RPM. A sample was taken from each culture after 15 minutes incubation to check exflagellation. Giemsa stain was done after 16 hours incubation by taking 100 µL from each culture, pellet them by centrifugation at 600g for 4 minutes at room temperature and re-suspend them in a small volume of supernatant (up to 5 µL) in order to make smears.

2.7 - GFP fusion protein expression in blood stages. To test GFP fusion protein expression in blood stages, mice were I.P. injected with DOZI::GFP, CITH::GFP, Puf2::GFP, HoMu::GFP, 4E::GFP and WT::GFP cryopreserved infected blood (3 mice per each line).

For live microscopy of all lines, slides were done at day 4 post I.P. infection by mixing a drop of blood from mice tail in 100 μ L of 1x PBS, pellet it by brief (9 seconds) centrifugation at 16100 g and by resuspend it in 10 μ L of 1 in 100 dilution of Hoechst 33342 stock (Sigma). Slides were then mounted with a cover slip. Observation of the slides was done in Zeiss Axiovert 200M (for HoMu::GFP, 4E::GFP and WT::GFP) and in Zeiss LSM 510 Meta confocal microscope (for DOZI::GFP, CITH::GFP and Puf2::GFP).

2.8 - GFP fusion gene expression in mosquito stage parasites. To test GFP expression in sporozoites inside salivary glands of the mosquitoes and inside midgut oocysts, *Anopheles stephensi* were infected by blood feeding on anaesthetized mice from all parasite lines. On day 13 post infection 5 to 10 mosquitoes of each GFP line were taken out of the incubator, placed at -20°C for 5 minutes to induce dormancy and fluorescence were checked under fluorescence stereomicroscope (Leica MZ10F) with x6.3 magnification. Four mosquitoes per each line were dissected under a Leica scope in order to extract their midguts. At day 21 post infection five DOZI::GFP mosquitoes were dissected and salivary glands were taken. A slide with the glands was mounted with a cover slip and observed at Axiovert 200M Microscope at x100 magnification. The other lines were also dissected at day 21 post infection. The removed salivary glands were smashed with a homogenizer and passed through a mesh filter in order to isolate the sporozoites and eluting them in PBS 1x. A spin was performed at 16100g during 1 minute at room temperature, supernatant was removed and pellet was re-suspended in 20 μ L of 1/100 dilution of Hoechst stock. Slides of salivary glands and midguts were done to observe in vivo the fluorescence in a widefield fluorescent microscope Leica DM5000B at x10 magnification (midguts) and at Zeiss Axiovert 200M at x100 magnification (salivary glands).

Filtered sporozoites taken from 5 dissected mosquitoes of each HoMu::GFP and 4E::GFP lines at day 21 post feeding on infected mice were diluted 1 to 5 in PBS1x. Sporozoites were count in a Neubauer chamber.

2.9 - RNA extraction from gametocytes and sporozoites and RT-PCR. In order to retain gametocytes but eliminate asexually replicating forms, sulfadiazine was given in water at a concentration of 20mg/L during 48h to a mouse infected with *P. berghei* ANKA PBA1019 parasite line with a parasitemia of 3%. Three days later the mouse was sacrificed and the blood was taken by heart puncture and transferred to pre-warmed PBS (37°C). The mixture was centrifuged 8 min at 450g at room temperature; the resulting pellet was re-suspended in 5 mL of PBS containing Nycodenz (49%) and centrifuged 20 min at 450g without brakes at room temperature with a swing out rotor. The grayish interface containing gametocytes was collected and centrifuged again 8 min 450g, at room temperature, this time with brakes. The resulting pellet (gametocytes) was resuspended in 1mL of TRIzol; 200 µL of chloroform were added, the mixture was vortexed for 10 seconds, left on ice for 5 minutes and centrifuged for 10 minutes at 12000g at 4°C. The aqueous phase was mixed with 500 µL of isopropanol and centrifuged under the same conditions; the supernatant was discarded and the pellet washed with 70% ethanol. Finally the pellet was dissolved in 51 µL of DEPC RNase free water and treated with DNase-I (Invitrogen) for 15 minutes at room temperature in a total volume of 60 µL followed by 10 min inactivation at 65°C with 6 µL EDTA.

After DNase-I treatment, 1µL of the total RNA solution was used to measure the RNA concentration in a Nanodrop 1000 spectrophotometer. 1 µg of the total RNA was then reverse transcribed with 1 µg of Oligo(dT) primers. Reaction contained: 1µL of Oligo(dT) (i.e. 1 µg), 1µL of 10mM dNTP mix and RNase-free water up to 12 µL. The mixture was heated to 65°C for 5 minutes and then quickly chilled on ice followed by adding these reagents: 4 µL of 5x first strand buffer; 2 µL of 0.1M DTT and 1 µL of RNaseOUT (40 units/µL). The contents were mixed and incubated at 42°C for 2 minutes; 1 µL (200 units) of SuperScript II (Invitrogen) was added and the mixture was incubated at 42°C during 50 minutes. The reaction was inactivated by heating at 70°C for 15 minutes and the remaining RNA removed from cDNA by adding 5 units (1µL) of RNase H (New England Biolabs) and incubating it 37°C for 20 minutes.

Sporozoite cDNA was previously generated in the lab by Dr Gunnar Mair. For Reverse Transcription PCR, cDNAs were diluted to a concentration of 10 ng/µL and 2 µL (20ng) of these dilutions were used to perform RT-PCR. The conditions used in this RT-PCR are listed in Table S3.

All primers used and their melting temperatures are shown in Table S4. Melting temperatures of the primers were determined at

www.finnzymes.fi/tm_determination.html. The parameters of the calculator were kept standard.

2.10 - HoMu and 4E tree and alignment. BlastP (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to identify *Plasmodium* Musashi and 4E homologs from 23 and 14 species respectively (accession numbers shown in tables S1 and S2). Alignments of their conserved motifs were performed with MUSCLE software (Tables S5 and S6). Phylogenetic trees of the retrieved alignments were generated by PhyML - Phylogenetic estimation using Maximum Likelihood (Bootstrap: 100 replicates; Model of substitution: Whelan and Goldman - WAG).

2.11 - HoMu and 4E Domain analysis. The analysis of Musashi RNA binding domains and 4E IF4E domain was performed by Pfam (<http://pfam.sanger.ac.uk>) in 9 and 8 species respectively. The information about interactions between Musashi and other proteins in the species *Mus musculus*, *Xenopus tropicalis* and *Drosophila melanogaster* was retrieved by String 8.1 (<http://string81.embl.de/>) and is shown in Table 1.

2.12 - Retrieval of *P. berghei* gene models and putative untranslated regions.

Searching in Plasmodb (<http://plasmodb.org/plasmo/>) with “pbanka*” gave rise to 4902 genes; genomic sequences of 400 bps downstream from stop codons of all *Plasmodium berghei* genes were downloaded and converted into an excel sheet (data not shown) in which consensus Musashi binding sequences of *Xenopus*, mouse and *Drosophila* were searched.

3 – Results

Translational repression plays an important role in a developmental context; mRNAs transcribed but stored for translation at a later time point are needed to direct post-fertilization growth and development. Such mechanisms are not only essential for higher, multicellular eukaryotes but were recently found to be essential during the development and transmission of the malaria parasite *Plasmodium berghei* to and from the mosquito vector. When the parasite crosses this species barrier, it needs to adapt quickly to the new environment; changes in gene expression patterns coincide with the generation of new parasite life cycle forms. DOZI, CITH and Puf2 were shown to play key roles in blood stage gametocytes and during the transition of the parasite from salivary gland sporozoites to early liver stage parasites (Mair, Braks *et al.* 2006; Mair, Lasonder *et al.* 2010; Gomes-Santos, Braks *et al.* 2011; Muller, Matuschewski *et al.* 2011).

3.1- HoMu and 4E are evolutionary highly conserved proteins. A frequently used method to identify the potential function of a protein is to evaluate its conservation in close and distant phylogenetic groups. We wanted to know if HoMu and 4E are conserved in species from different Kingdoms. Therefore we used the protein BLAST (Basic Local Alignment Search Tool) tool to identify proteins similar to *P. berghei* HoMu and 4E. For HoMu we retrieved proteins from 23 different species including bacteria and archaea (Table S1). For 4E, proteins from 14 species were used (Table S2). We generated two alignments of the retrieved proteins with MUSCLE software. For HoMu the resulting tree (Figure 4A) shows that the apicomplexan group is composed by two *Plasmodium* species and one from the related apicomplexan *Cryptosporidium*. The closest relatives are proteins from three non-animal groups, all containing plants species: one containing species of archaea, bacteria and green algae the other containing brown algae and the last with two species of higher plants only. In animals, the *Anopheles gambiae* mosquito, which is a vector of *Plasmodium* is contained along with the fly *Drosophila*, the nematode *C. elegans*, the platyhelminth parasite *Schistosoma* and 4 species of vertebrates. In vertebrates it is possible to distinguish two groups corresponding to two different isoforms: Musashi-I (*Mus musculus* and *Bos Taurus*) and Musashi-II (*Homo sapiens* and *Xenopus laevis*).

The 4E tree differs from the HoMu one (Figure 5A). The groups are distributed differently with the apicomplexan group also containing the brown alga *Ectocarpus siliculosus*. Like in the HoMu tree here is also a green algae containing group but this one also includes *Arabidopsis thaliana* which in HoMu tree was included in brown algae group instead.

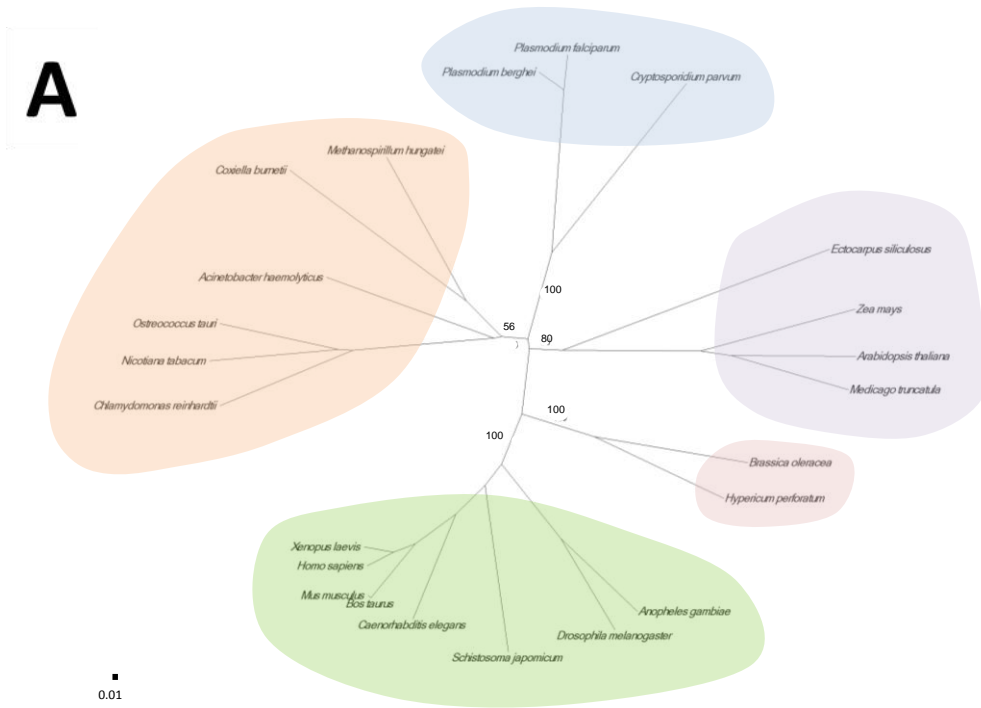
Animal species distribution in 4E tree is substantially different: vertebrates form one group with the plant *Zea mays* and *A. gambiae* mosquito; two other invertebrate species (*D. melanogaster* and *C. elegans*) form another group.

In summary we found that HoMu is a very well conserved protein with homologs present in such diverse groups as bacteria, archaea, plants, algae and animals. Similarly 4E is also highly conserved in these groups but is absent from bacteria and archaea.

Domain analysis of all HoMu homologs showed that they contain the same domain structure with two central RRM1 (RNA Recognition Motif 1) domains, with exception of the bacterium *Acinetobacter haemolyticus* which only possess one RRM1 domain (Figure 4B). 4E is highly conserved in all of the 8 species shown (Figure 5B) and got higher scores than HoMu in the Pfam alignment, meaning that amino acid sequence of the 4Es IF4E domain is more conserved than the one in RRM1 domains of HoMu.

We also aimed to find how deeply are RRM1 and IF4E domains containing proteins distributed among several well-known parasite species. A search in EUPATHDB for RRM1 and IF4E domains retrieved a total of 2075 (RRM1) and 95 (IF4E) proteins from several species. For RRM1 55 of the 2075 proteins are from *P. berghei* (Figure 4C). For IF4E only one (4E) is from *P. berghei* (Figure 5C). The *Plasmodium* genus possesses a total of 445 proteins with RRM1 domain and is followed by *Toxoplasma gondii* with 202 RRM1 domain containing proteins, *Trypanosoma* genus with 163, and *Trichomonas vaginalis* with 131. On the other hand IF4E domain distribution is different from the RRM1, with the genus *Trypanosoma* and the specie *T. vaginalis* having 13 IF4E domain containing proteins each. The *Plasmodium* genus has 7 proteins, and finally the species *Entamoeba histolytica* and *T. gondii* have 6 proteins with IF4E domain.

In summary we found that HoMu RRM 1 domains are present in a wide range of life species and that RRM1 containing proteins are abundantly found in different genus of parasites, particularly in *Plasmodium* and in *Toxoplasma*. The IF4E domain, although in less numbers, is also present in proteins of several parasites, particularly in *T. vaginalis* and in *Trypanosoma* genus.



B

Species	Proteins	Accession numbers	Domain(s) length (aas)
<i>A. haemolyticus</i>		ZP_06726647.1	70
<i>A. gambiae</i>		XP_321067.4	69
<i>A. thaliana</i>		NP_181639.1	65; 70
<i>C. elegans</i>		NP_497799.1	69; 67
<i>C. parvum</i>		XP_628599.1	68; 67
<i>D. melanogaster</i>		NP_733108.2	69
<i>H. sapiens</i>		NP_733839.1	69; 67
<i>P. berghei</i>		PBANKA_081770	68; 69
<i>X. laevis</i>		NP_001084027.1	69; 67

C

RRM 1 domain in several EuPathDB species

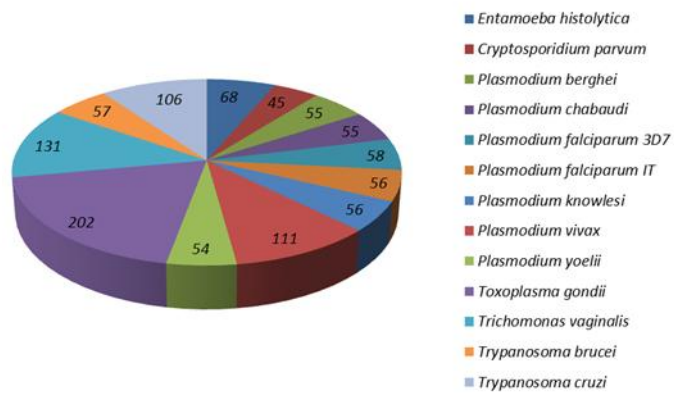
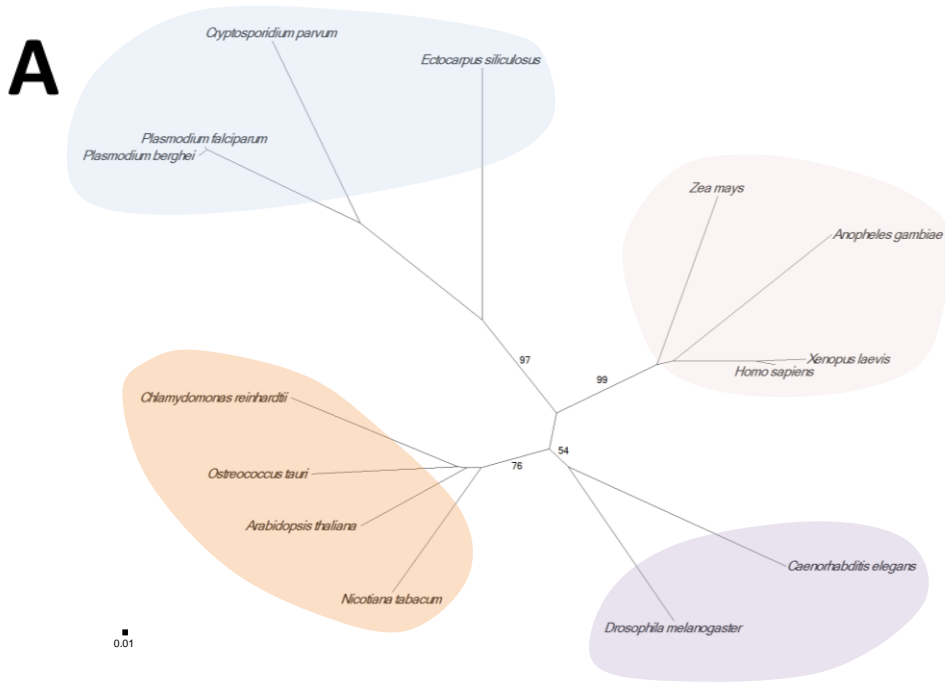


Figure 4 | Bioinformatic analyses of Plasmodium berghei HoMu. A - Phylogenetic tree from multiple sequence alignment of the conserved region of RRM proteins from different species generated by PhyML (Bootstrap: 100 replicates; Model of substitution: WAG). **B** - Graphic representation of RRM 1 domains and their lengths in different species generated by Pfam. Pfam bit scores of the alignments between query sequences and consensus sequence of the RRM 1 domain (PF00076) and which calculation is based on hidden Markov model (HMM) are: *Acinetobacter haemolyticus*: 75.2; *Anopheles gambiae*: 60.5 and 60.4; *Arabidopsis thaliana*: 54.4 and 47.0; *Caenorhabditis elegans*: 65.2 and 55.8; *Cryptosporidium parvum*: 48.7 and 64.4; *Drosophila melanogaster*: 57.3; *Homo sapiens*: 69.8 and 60.4; *Plasmodium berghei*: 54.1 and 55.7; and *Xenopus laevis*: 69.3 and 54.7. **C** - Representation of the number of proteins with RRM1 domain in several parasite species (Data from EuPathDB).



B

Species	Proteins	Accession numbers	Domain(s) length (aas)
<i>A. gambiae</i>		XP_311951.4	159
<i>A. thaliana</i>		NP_193538.1	152
<i>C. elegans</i>		NP_499751.2	155
<i>C. parvum</i>		XP_627474.1	176
<i>D. melanogaster</i>		NP_651654.1	130
<i>H. sapiens</i>		AAC39871.1	162
<i>P. berghei</i>		PBANKA_041280	175
<i>X. laevis</i>		NP_001089549.1	162

C

IF4E domain in several EuPathDB species

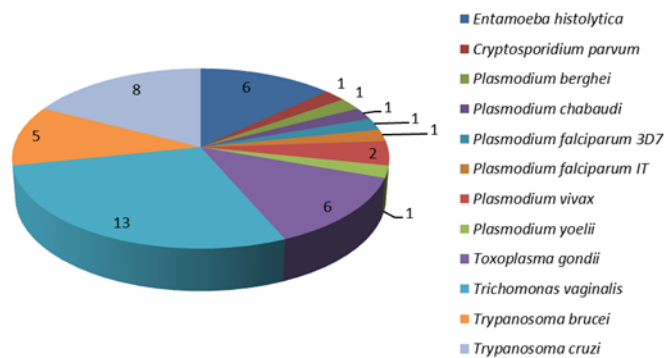


Figure 5 | Bioinformatic analysis of *Plasmodium berghei* eIF4E. **A** – Phylogenetic tree from multiple sequence alignment of the conserved region of IF4E domain containing proteins from different species generated by PhyML (Bootstrap: 100 replicates; Model of substitution: WAG). **B** - Graphic representation of IF4E domains and their lengths in different species generated by Pfam. Pfam bit scores of the alignments between query sequences and consensus sequence of the IF4E domain (PF01652) and which calculation is based on hidden Markov model (HMM) are: *Anopheles gambiae*: 172.5; *Arabidopsis thaliana*:189.6; *Caenorhabditis elegans*:181.1; *Cryptosporidium parvum*:136.9; *Drosophila melanogaster*:166.7; *Homo sapiens*:203.2; *Plasmodium berghei*: 138.2; and *Xenopus laevis*: 192.4. **C** - Representation of the number of proteins with IF4E domain in several parasite species (Data from EuPathDB).

3.2 - HoMu is transcribed in gametocytes and sporozoites. It was previously shown that *dozi* and *puf2* (the proteins are clearly involved in translational regulation) are transcribed in blood stage gametocytes and sporozoites (Gomes-Santos, Braks *et al.* 2011; Muller, Matuschewski *et al.* 2011). Here, by using reverse transcription PCR of gametocyte and sporozoites total RNA we show that *cith* and *homu* are transcribed in both transmission life cycle stages of *P. berghei* (Figure 6). As expected, expression of *spect2* and *trap*, which were used as controls, was only detected in sporozoites.

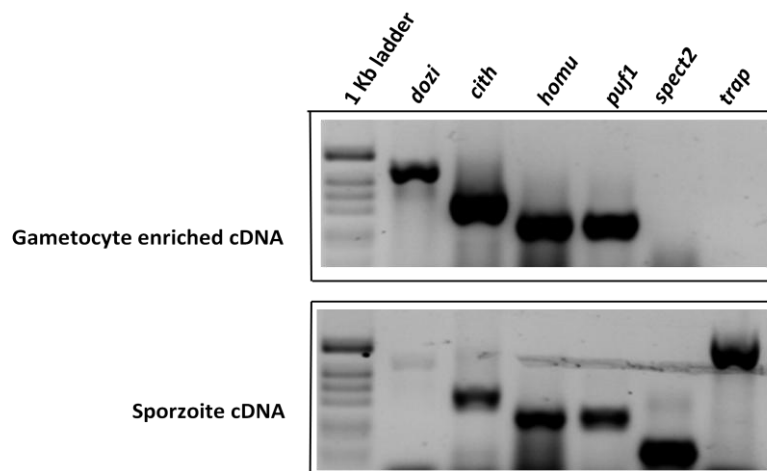


Figure 6 | Expression of genes involved in translational repression. Reverse Transcriptase PCR showing *homu* expression compared to *dozi*, *cith* and *puf1*. *spect2* and *trap* are known to be expressed only in sporozoites.

3.3 - Strong GFP expression was detected in HoMu::GFP and 4E::GFP

gametocytes. Since RT-PCR results showed that *homu* gene is transcribed strongly in gametocytes, native HoMu::GFP expression analysis in live parasites was performed in order to check if *homu* is also translated in gametocytes with particular comparison to *dozi* and *cith*. 18 mice were infected with DOZI, CITH, Puf2, HoMu, 4E and wild type GFP lines and a drop of blood was taken from each mouse at day 4 post I.P. infection. Fluorescence was analyzed under fluorescence and confocal microscopes. GFP positive gametocytes were found in all lines including Homu::GFP and 4E::GFP. GFP expression was particularly strong in gametocytes of all lines except Puf2::GFP (Figure 7). GFP signal was also detected in other blood stages of HoMu::GFP and 4E::GFP. This means that *homu* and *4e* are being translated not only during the gametocyte stage of *P. berghei* but also in other blood stages, as trophozoites (both lines) and schizonts (4E::GFP). The cytoplasmic GFP distribution pattern inside parasite gametocytes is similar between HoMu::GFP and 4E::GFP, but is slightly different from previously observed DOZI and CITH expression patterns (Mair, Braks et al. 2006; Mair, Lasonder et al. 2010).

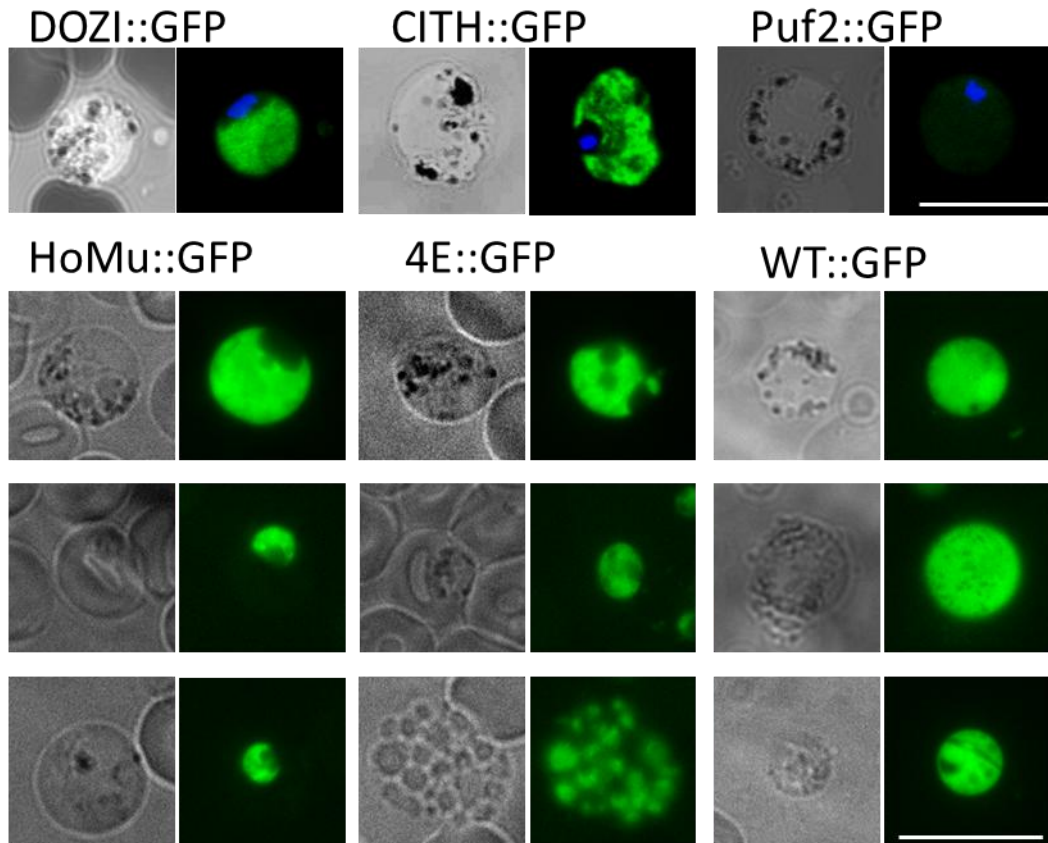


Figure 7 | Transgene expression in blood stage parasites. Live cell pictures of the fluorescent live stages of the cell lines DOZI::GFP, CITH::GFP and Puf2::GFP (above) taken by confocal microscope at x63 magnification. Pictures of HoMu::GFP, 4E::GFP and WT::GFP are shown below and were taken by fluorescent microscope at x100 magnification. In HoMu::GFP and 4E::GFP, the pictures in the middle are both Trophozoites. For the pictures below: the HoMu::GFP one (left) is a trophozoite and the 4E::GFP (middle) is a schizont. Green represents the GFP fluorescence, blue is Hoechst nucleus staining and gray is bright field. Scale bars=10 μ m.

In summary we found that HoMu::GFP and 4E::GFP are abundantly expressed in gametocytes. While DOZI::GFP and CITH::GFP are predominantly observed in female gametocytes (Mair, Braks et al. 2006; Mair, Lasonder et al. 2010), HoMu::GFP and 4E::GFP are also abundant in asexual stage parasites.

3.4 - HoMu::GFP and 4E::GFP are expressed in oocysts. One of the aims of this work was to identify at which life cycle stage(s) HoMu and 4E are expressed. In order to check if HoMu::GFP and 4E::GFP are expressed in oocysts, 12 mice were infected with these lines and also with DOZI::GFP, CITH::GFP, Puf2::GFP and WT::GFP. At day 4 post I.P. infection, exflagellation was checked in all lines and confirmed. Infected mice were therefore anaesthetized and 25 female mosquitoes allowed to feed. On day 13 post feeding, live observations of mosquito abdomens were made using a fluorescent stereomicroscope (Figure 8A). Midguts were dissected from the abdomens and observed at fluorescent microscope (Figure 8B). No fluorescence was observed in GFP-tagged DOZI and CITH lines. Puf2::GFP was seen in a few oocysts, while HoMu::GFP and 4E::GFP were strong (Figure 8A, B).

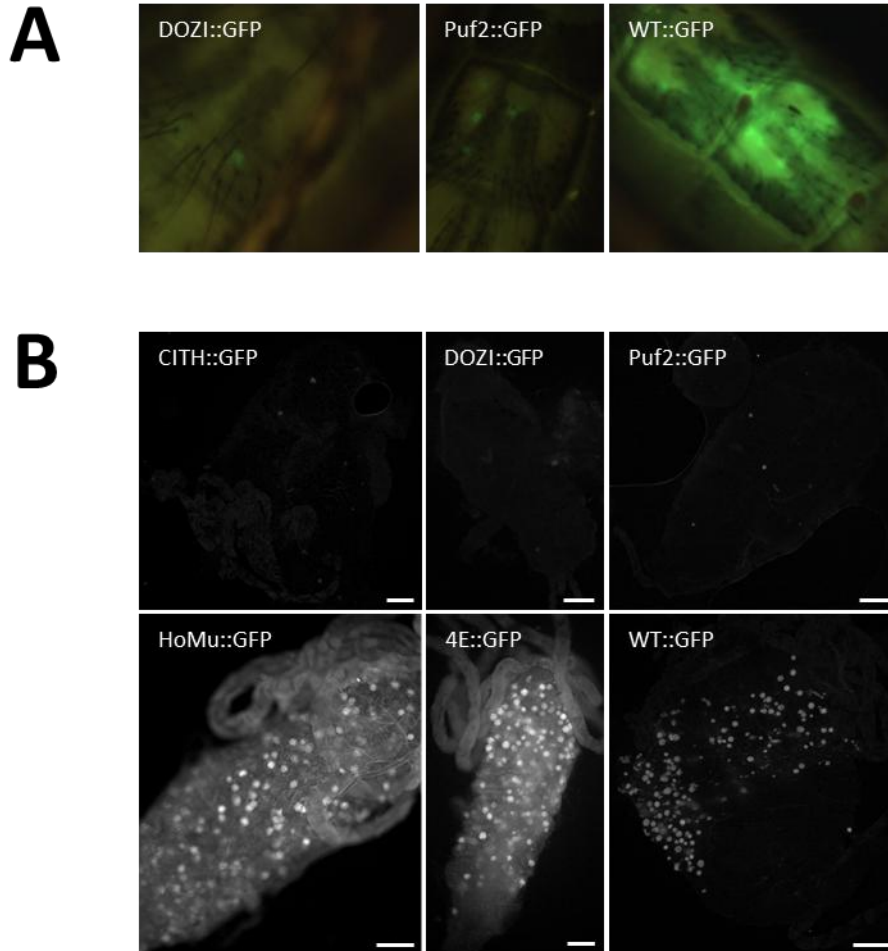


Figure 8 | Transgene expression during mosquito midgut development.

A - Fluorescence microscopy pictures of live mosquitoes abdomens taken with fluorescence stereomicroscope at x6.3 magnification. **B**- Dissected midguts pictures taken by fluorescence microscope at x10 magnification. Fluorescent oocysts are white dots Scale bars: 100 µm.

In summary we showed that only Puf2::GFP, HoMu::GFP and 4E::GFP are expressed in oocyst parasite life cycle stage.

3.5 - HoMu::GFP and 4E::GFP are expressed in mosquito salivary gland

sporozoites. Data from RT-PCR have shown that *homu* is transcribed abundantly in sporozoites. Therefore, in order to know if it is also translated in this life cycle stage, the expression of HoMu::GFP was compared with the expression of DOZI, CITH, Puf2, 4E and wild type GFP lines.

On day 21 post feeding on DOZI::GFP, CITH::GFP, Puf2::GFP, WT::GFP as well as HoMu::GFP and 4E::GFP lines, 5 mosquitoes per line were hand-dissected in order to remove the salivary glands. Fluorescent and bright field microscopic analysis of salivary glands revealed a strong GFP signal in all lines (Figure 9). The GFP distribution pattern of Puf2::GFP is speckled and similar to the one shown in previous data (Gomes-Santos, Braks et al. 2011). DOZI::GFP and CITH::GFP sporozoites also have a speckled distribution of the GFP but less abundantly than Puf2::GFP. The speckled distribution of GFP in these lines suggests that the expression of the proteins DOZI, CITH and Puf2 in cytoplasm is localized and not uniformly distributed. Such localized expression patterns are typical for mRNPs that could be involved in translational regulation. HoMu::GFP and 4E::GFP parasites have a more uniform distribution of GFP than the previous ones meaning that HoMu and 4E expression in cytoplasm is not localized. Surprisingly, sporozoites were not detected in mosquitoes that fed on WT::GFP infected mice. The reason why needs to be determined by repeating the experiment.

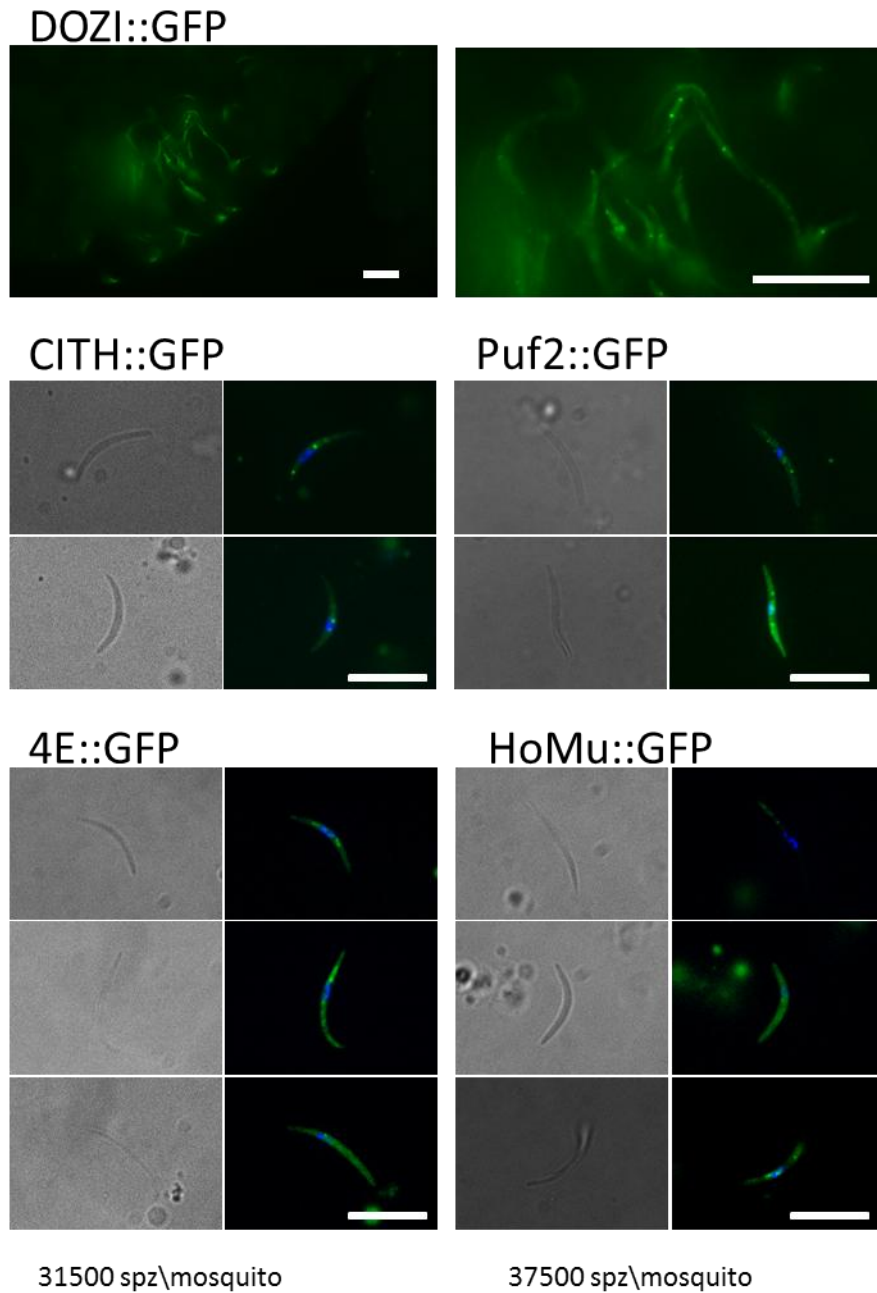


Figure 9 | Transgene expression salivary gland sporozoites. Live sporozoites of DOZI::GFP (above), CITH::GFP (middle left), Puf2::GFP (middle right), 4E::GFP (below left) and HoMu::GFP (below right) lines taken from dissected salivary glands of mosquitoes. Green is GFP, blue is Hoechst nucleus staining and gray is bright field. Pictures were taken with a fluorescent microscope at x100 magnification. Sporozoite mean numbers in salivary glands of 4E::GFP and HoMu::GFP mosquitoes are shown below. Scale bars: 10 μ m.

In summary we found that all GFP-tagged proteins are expressed in the mosquito salivary gland sporozoites.

3.6 - HoMu::GFP and 4E::GFP parasites show normal blood stage development and morphology. Parasite lines with GFP tagged proteins are not only important tools for protein localization studies and detection of protein levels throughout the life cycle of malaria parasite (see above). If they behave as wild type protein functionally, such mutant parasites can be used as a starting point for the isolation of larger protein complexes (mRNPs in our case). It is therefore essential to identify whether GFP-tagging and a pyrimethamine resistance marker (dhfr/ts) introduction into the parasite genome has resulted in a cost to parasite fitness. Therefore we checked whether these parasite lines grow as wild type forms throughout the life cycle or show morphological abnormalities. Sets of 3 BALB/c mice for each parasite line were injected in parallel with the same amount of cryopreserved blood infected with similar parasitemias. Parasitemias of 4E::GFP, HoMu::GFP and WT::GFP lines were determined on GIEMSA stained thin blood smears during 10 days from the day of I.P. infection and were plotted in a growth curve (Figure 10A above). Gametocytemias of the same lines were determined at day 4 post infection and are shown in Figure 10A (below). A comparison of the parasitemias and gametocytemias in wild type GFP line with the experimental lines in both graphs reveals no apparent difference therefore it can be suggested that neither HoMu nor 4E tagged with GFP results in slower growth of asexual blood stage parasites or impaired development of gametocytes. Pictures from giemsas used in parasitemia and gametocytemia determination were taken with the brightfield Leica DM2500 microscope. All of the development blood stages shown reveal no morphological differences between mutant HoMu and 4E GFP lines and the wild type GFP line (Figure 10B).

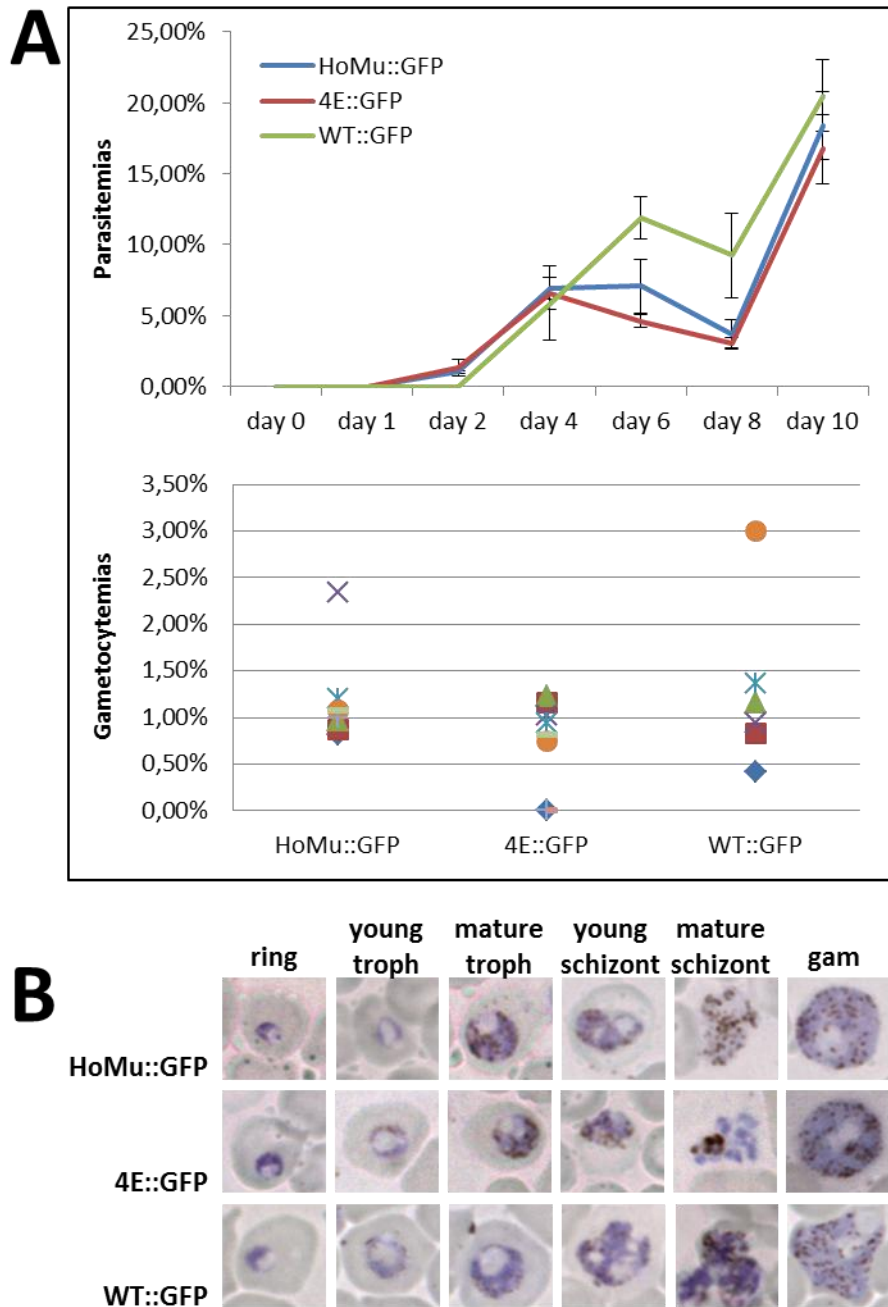


Figure 10 | Growth characteristics of all mutant parasite lines in Balb/C mice. **A** –Parasitemias growth curves (above) of HoMu and 4E and wild type GFP lines were determined during 10 days from I.P. infection day; Gametocytemias (below) of HoMu::GFP, 4E::GFP and WT::GFP were determined at day 4 post I.P. infection; each dot represents the gametocyte percentage value in a microscopic field. **B** – Light microscopic analysis of Giemsa stained blood smears of WT::GFP compared to HoMu::GFP and 4E::GFP lines. Blood stage forms names from left to right: ring, young trophozoite, mature trophozoite, young schizont, mature schizont and gametocyte.

It can be summarized that *P. berghei* parasites from the mutant lines HoMu::GFP and 4E::GFP show normal development and morphology during the blood stage.

3.7 - Mosquito stage development of HoMu::GFP and 4E::GFP parasite lines is

normal. Translation inhibition of certain mRNAs is important for the *P. berghei* parasite to progress from the zygote to the ookinete stage. DOZI and CITH were shown to be essential in this process (Mair, Braks et al. 2006; Mair, Lasonder et al. 2010). As both 4E and HoMu were detected in DOZI-containing mRNPs, we wanted to verify that GFP-tagged versions of these proteins still allow ookinete formation and mosquito infection in wild type numbers. We therefore checked exflagellation on day 4 post I.P. infection of mice with all experimental lines; it was confirmed that all lines were exflagellating. As was already shown above, mosquitoes that fed on the same mice at the same day became infected showing that gametocytes from all mutant and wild type GFP lines were able to develop normally. In order to test if gametocytes of the DOZI, CITH, Puf2, HoMu and 4E GFP-tagged lines have the same fitness as wild type parasites in ookinete development, we also performed in vitro ookinete cultures prepared from the mice used in growth curves. When mice had reached 10% or more of parasitemia three ookinete cultures per each line were prepared and the gametocyte exflagellation was checked before overnight incubation of the cultures. Exflagellation, a sign that male gametocytes are developing into male gametes, could not be confirmed in ookinete cultures of DOZI, CITH, Puf2, HoMu, 4E and wild type GFP lines after 15 minutes incubation at 19°C. In order to check if there is a difference between sporozoite load in salivary glands of mosquitoes infected with HoMu::GFP and 4E::GFP parasite lines and the WT::GFP, mosquito's salivary glands were used to determine the average sporozoite load. An average of 37500 and 31500 sporozoites per mosquito were found after dissection of 5 mosquitoes infected with HoMu::GFP and 4E::GFP, respectively. Previous data shown that the numbers of salivary gland WT::GFP sporozoites were in an average of 31 600 (Muller, Matuschewski et al. 2011). My data from the infected mosquitoes with mutant lines HoMu::GFP and 4E::GFP is in accordance with these observations; hence there is no impairment in the infection of mosquitoes salivary glands with these mutant parasite lines.

In summary, although no ookinetes were recovered from in vitro culture, gametocytes from blood that was used at day 4 post I.P. infection were exflagellating and mosquitoes that fed on the same mice at the same day became infected suggesting that gamete

generation, fertilization and ookinete formation in mutant lines was not impaired. Also the movement of mature sporozoites from oocysts to the mosquito's salivary glands was not impaired.

3.8 - Re-infection of mice with mutant parasite line sporozoites – HoMu::GFP and 4E::GFP re-infections were unsuccessful. The mosquito infection experiment suggested that GFP tagged proteins are expressed in a localized fashion during sporozoite stage; also the numbers of salivary gland sporozoites appeared in a normal range of wild type parasites. In order to check whether these lines transmit normally to the subsequent rodent host, I set up a series of mosquito feeding experiments. 2 mice per each line were used. All lines except HoMu::GFP and 4E::GFP were able to transmit from mosquitoes to at least one mouse (Figure 11). The reason for the failure to transmit is unclear, and will need to be determined in a repeat of the experiment.

Parasite line	Infected / total mice
DOZI::GFP	1/2
CITH::GFP	2/2
Puf2::GFP	2/2
HoMu::GFP	0/2
4E::GFP	0/2
WT::GFP	1/2

Figure 11| Re-infection of mice with mutant parasite line sporozoites. Number of mice that became infected after mosquitoes feeding. 25 Mosquitoes per each line were previously infected by feeding on parasitized mice (at 4th day post I.P. infection); Re-infections were performed 21 days post feeding.

3.9 - More than 2000 *P. berghei* genes contain consensus RBS in their 3'

untranslated regions. Having established that the mutant lines show no growth defects, indicating that HoMu::GFP function like the untagged wild type protein, we wanted to determine if the extent of known Musashi-binding elements is present in the *P. berghei* genome. Genes that possess consensus RNA binding sequences (RBS) in their 3' untranslated regions are likely to be translationally regulated, therefore it's important to know if *P. berghei* genes also contain these RBS.

With the aim to find RBS in 3' untranslated regions of the *P. berghei* genes a search in Plasmodb with "pbanka*" was performed; this returned 4902 protein coding genes. From all of these we retrieved 400 bps downstream from stop codons and searched for known Musashi binding elements; 2294 genes had the *Xenopus* consensus sequence 3'-AUAGU and 24 genes had the mouse partial consensus sequence 3'-GUUAGUUA. The entire consensus Musashi binding sequence in mouse is 3'-GUUAGUUAGUUAGUU and was not found. Finally for the *Drosophila* consensus sequence 3'-GUUUUUAG, 59 *P. berghei* genes were found.

In summary we found that more than 2000 *P. berghei* genes contain the consensus RBS of the *Xenopus* frog *mos* RNA. Other species consensus sequences were found in less number. Whether those are real targets for HoMu binding in *P. berghei* needs to be evaluated experimentally. The wild type behavior of the HoMu::GFP line will aid in this.

4 – Discussion

Abrupt changes in the environment obligate the malaria parasite to adapt its gene expression profile when its gamete precursor cells, the gametocytes, are taken up by a female *Anopheles* mosquito. During this transmission the parasites experience a variation in pH and temperature. The pH increases from 7.3 in mouse blood to 8.0 in the mosquito's midgut and the temperature decreases from 37 °C to 20°C (Billker, Shaw *et al.* 1997). The parasite therefore has to adapt quickly, changing its gene expression profile to these new conditions when it enters the mosquito. Translational repression has been shown to be a major process of gene expression control in *Plasmodium* during this process. Proteins involved in translational silencing of the transcripts *p25* and *p28* were identified in cytoplasmic mRNP complexes. These complexes include DOZI, CITH, HoMu and 4E. DOZI and CITH have been shown experimentally to be involved in translation arrest and storage of female gametocyte mRNAs (Mair, Braks *et al.* 2006; Mair, Lasonder *et al.* 2010). Here we show that GFP-tagged versions of HoMu and 4E are expressed in a similar fashion in gametocytes. The GFP signal of the *P. berghei* parasite lines HoMu::GFP and 4E::GFP was strong compared to constitutively expressed GFP of the wild type line, with a mildly localized distribution pattern.

It was also demonstrated here through protein alignments, phylogenetic trees and domain analyses that HoMu and 4E are highly conserved with homologs in a wide range of species. HoMu homologs in mice and *Drosophila* are known to repress translation (Imai, Tokunaga *et al.* 2001; Okano, Imai *et al.* 2002), however the *Xenopus* homolog is responsible for translation activation (Charlesworth, Wilczynska *et al.* 2006). Homologs of the *P. berghei* 4E also play an important role in translation initiation and arrest. The dual role of these proteins could therefore explain the less localized distribution of GFP signal of the lines HoMu::GFP and 4E::GFP in gametocyte cytoplasm. 4E certainly can participate in mRNPs at translation initiation, as well as arrest processes during the storage of mRNAs for translation at a later time. Both TIC complexes and translationally quiescent mRNP complexes are widely distributed in the cytoplasm; it is therefore impossible to distinguish these mRNPs based on GFP fluorescence alone. The proteins DOZI and CITH, on the other hand, are unknown to participate in TIC and are hence more specific, inhibiting translation of a smaller number of transcripts that are more localized to certain regions of the cytoplasm, what possibly makes the distribution pattern of fused with GFP mutant lines of these proteins appear spotty.

The analysis of the 3' untranslated regions of all genes revealed that more than 2000 of them contain the consensus sequence for *Xenopus* Musashi binding. This reinforces the possibility that HoMu can actually be a general inhibitor or activator of translation. However to confirm this hypothesis a HoMu consensus binding sequence should be determined experimentally.

It was shown in published data that translation arrest also plays a role during mosquito stage (Gomes-Santos, Braks *et al.* 2011; Muller, Matuschewski *et al.* 2011). When a *Plasmodium* infected mosquito feeds on a mammalian host, sporozoites inside their salivary glands will enter into the host connective tissues and subsequently into the host blood circulation and will experiment environment changes, like a shift in temperature that proved to trigger sporozoite conversion to EEFs *in vitro* (Doi, Shinzawa *et al.* 2011). Gene expression in salivary gland sporozoites is regulated by translational repression of liver specific transcripts preventing them to turn into EEFs before having invaded a hepatocyte (Gomes-Santos, Braks *et al.* 2011; Muller, Matuschewski *et al.* 2011). The pumilio protein Puf2 plays a role in this process, however it is unclear whether those transcripts are arrested and stored in sporozoite P granules.

Here we show for the first time that DOZI, CITH, HoMu and 4E proteins fused with GFP are expressed in sporozoites. The fluorescent signal in these lines is reminiscent of cytoplasmic speckles suggesting that sporozoite DOZI, CITH, HoMu and 4E can play a role in translation inhibition and/or activation of a specific set of transcripts. Fluorescent microscopic analysis of dissected midguts revealed that HoMu and 4E proteins are also expressed in the oocyst phase of *P. berghei* life cycle. The function that they could play in these phase need to be elucidated in further experiments on oocysts with for example *homu* and *4e* disrupted gene expression *P.berghei* lines.

No differences were found between HoMu::GFP, 4E::GFP lines and wild type GFP lines at levels of parasitemias and gametocytemias during blood stage development. Although exflagellation was confirmed, showing that at least male gametocytes develop into male gametes, *in vitro* ookinete cultures failed. GFP lines infected and developed normally in the mosquito. Their sporozoite numbers are in accordance with previous data determined in wild type references lines that constitutively express GFP during the whole *P. berghei* life cycle (Muller, Matuschewski *et al.* 2011). This data strongly suggest that HoMu::GFP and 4E::GFP mutant lines do not have impairment during blood stage and mosquito development. Reinfection of mice with infected mosquitoes was successful

only with DOZI, CITH and Puf2 GFP lines as well as a wild type GFP line. HoMu::GFP and 4E::GFP line parasites were unable to re-infect naive mice by unknown reasons.

Data from RT-PCR shown that like *dozi* and *puf1*, *cith* and *homu* are also transcribed in both gametocytes and sporozoites. Now that mRNA expression of *homu* is proved, it will be important to determine whether disruption of this gene affects gametocyte and/or sporozoite development. To achieve this, a mutant line without *homu* in its genome will be generated. The mRNA expression of *4e* was not tested; however it's likely that this gene is transcribed throughout the *Plasmodium* life cycle since it's an eukaryotic translation initiation factor. Nevertheless RT-PCR of *4e* as well as real time PCR of *homu* will be considered in the future to get more detailed data about their expression in both mosquito and blood stages.

In conclusion, we were able to show that mutant parasite lines HoMu::GFP and 4E::GFP develop normally in mosquito and in asexual life cycle stages (blood stages). It will now be possible to identify if there is co-localization of these proteins with transcripts translationally regulated during these stages. Immunoprecipitation and mass spectrometry analyses of protein complexes isolated from HoMu::GFP and 4E::GFP lines using anti-GFP antibodies can be used in future studies to identify proteins forming complexes with 4E and HoMu at different life cycle stages.

5 – Appendix

Species	Protein ID
<i>Acinetobacter haemolyticus</i>	ZP_06726647.1
<i>Anopheles gambiae</i>	XP_321067.4
<i>Arabidopsis thaliana</i>	NP_181639.1
<i>Bos taurus</i>	DAA20633.1
<i>Brassica oleracea</i>	ABD64922.1
<i>Caenorhabditis elegans</i>	NP_497799.1
<i>Chlamydomonas reinhardtii</i>	XP_001689671.1
<i>Coxiella burnetii</i>	NP_819948.1
<i>Cryptosporidium parvum</i>	XP_628599.1
<i>Drosophila melanogaster</i>	NP_733108.2
<i>Ectocarpus siliculosus</i>	CBJ30102.1
<i>Homo sapiens</i>	NP_733839.1
<i>Hypericum perforatum</i>	ADK92872.1
<i>Medicago truncatula</i>	ABD28657.1
<i>Methanospirillum hungatei</i>	YP_503023.1
<i>Mus musculus</i>	NP_032655.1
<i>Nicotiana tabacum</i>	AAF66823.1
<i>Ostreococcus tauri</i>	CAL54641.1
<i>Plasmodium berghei</i>	PBANKA_081770
<i>Plasmodium falciparum</i>	PFI0820c
<i>Schistosoma japonicum</i>	CAX72901.1
<i>Xenopus laevis</i>	NP_001084027.1
<i>Zea mays</i>	NP_001149939.1

Table S1 | ncbi (www.ncbi.nlm.nih.gov) and EUPATHDB (www.eupathdb.org) accession numbers of proteins used in the HoMu bioinformatic analyses

Species	Protein ID
<i>Anopheles gambiae</i>	XP_311951.4
<i>Arabidopsis thaliana</i>	NP_193538.1
<i>Caenorhabditis elegans</i>	NP_499751.2
<i>Chlamydomonas reinhardtii</i>	XP_001693235.1
<i>Cryptosporidium parvum</i>	XP_627474.1
<i>Drosophila melanogaster</i>	NP_651654.1
<i>Ectocarpus siliculosus</i>	CBN73980.1
<i>Homo sapiens</i>	AAC39871.1
<i>Nicotiana tabacum</i>	AAU06579.1
<i>Ostreococcus tauri</i>	XP_003075025.1
<i>Plasmodium berghei</i>	XP_674865.1
<i>Plasmodium falciparum</i>	XP_001351220.1
<i>Xenopus laevis</i>	NP_001089549.1
<i>Zea mays</i>	ACG31883.1

Table S2 | ncbi (www.ncbi.nlm.nih.gov) and EUPATHDB (www.eupathdb.org) accession numbers of proteins used in the 4E bioinformatic analyses

Temperature	Time	# Cycles
95°C	3 min	1
95°C	10 sec	35
43°C\50°C(<i>spect 2</i> and <i>trap</i>)	30 sec	
62°C	30 sec	
62°C	10 min	1

Table S3 | PCR cycling program

Gene	FW	mT (°C)	Reverse	mT (°C)	gDNA	cDNA	sequence forward primer	sequence reverse primer
DOZI	546	59.30	548	49.59	641	439	TAATTGTGTCGCTTCAAATG	TAATTCITTTATCATAGCAG
CITH	549	48.04	550	60.25	313	313	GAAAAAAGCAAAGATGTATTATCTG	ATAGGCTGGGTATCTGTAAATG
Musashi	822	61.80	823	62.38	243	243	CGTCTAATAGCGCCTTTGTC	ACCTTTAGATCGGCCTTCC
Puf1	479	59.00	480	65.00	415	239	ATAAGTGTTTCATGGAACCCG	TTACGCAGCACCCATGCC
TRAP	433	56.00	432	57.00	559	559	CATGTTATTCCAATGCTCAC	AACATTCACCTCCATTCTCC
SPECT2	440	59.00	441	58.00	289	146	AAGGAGTTTCAGCTATGCAC	CAGTTCATTTATGCCTGACC

Table S4 | Oligonucleotide primers used for RT-PCR (Fig 5)

Localization of mRNA storage complexes in *Plasmodium berghei* throughout the life cycle /2011

C.parvum	83RSPVSR...	RVMV	VRN	IS	Q	T	S	D	K	V	F	L	N	L	E	K	F	G	E	I	E	D	S	T	I	R								
P.Berghei	92SSPSTR...	RLM	VRN	IP	S	T	K	D	Q	F	L	K	Y	E	T	F	G	E	I	E	D	G	L	I	V	R								
P.falciparum	136SSPSTR...	RLM	VRN	IP	S	T	R	D	Q	F	L	K	Y	E	S	F	G	E	I	E	D	G	L	I	V	R								
C.reinhardtii	109RKSGVG...	N	I	F	K	N	L	D	K	A	I	D	A	K	A	L	H	D	T	F	S	A	F	G	K	L	S	C	K	V	A	T		
N.tabacum	110RRSGNG...	N	I	F	K	N	L	D	K	A	I	D	A	K	A	L	H	D	T	F	S	A	F	G	N	L	S	C	K	V	A	V		
O.tauri	115RRSGVG...	N	I	F	K	N	L	D	K	A	I	D	N	K	A	L	L	D	T	F	A	Q	F	G	T	I	S	A	K	V	A	M		
B.oleracea	1MESSC...	K	L	F	G	G	L	S	M	E	T	T	E	D	R	L	R	E	Y	F	Q	S	F	G	E	V	L	E	A	V	I	M	K	
H.perforatum	1MQSDNG..	K	L	F	V	G	G	L	S	M	E	T	T	E	D	R	L	K	E	Y	F	G	S	F	G	D	V	E	A	V	I	M	K	
A.gambiae	68HDPAPG...	K	L	F	V	G	G	L	S	M	E	T	T	E	D	R	L	K	E	Y	F	G	M	F	G	T	V	T	D	V	L	I	M	K
D.melanogaster	180	NKSGSSGRSTP	K	L	F	V	G	G	L	S	M	E	T	T	E	D	R	L	K	E	Y	F	N	M	F	G	T	V	T	D	V	L	I	M	K
S.japonicum	73DSNDPG..	K	M	F	G	G	L	S	P	T	T	S	E	V	L	R	D	Y	F	Q	K	Y	G	E	L	R	E	Y	M	I	M	R		
C.elegans	40GSQDPG...	K	M	F	G	G	L	S	M	O	T	T	A	N	L	R	D	Y	F	G	R	F	G	E	V	N	E	C	M	V	M	R		
B.taurus	15SPHDPG...	K	M	F	G	G	L	S	M	O	T	T	Q	E	L	R	E	Y	F	Q	F	G	E	V	K	E	C	L	V	M	R			
M.musculus	15SPHDPG...	K	M	F	G	G	L	S	M	O	T	T	Q	E	L	R	E	Y	F	Q	F	G	E	V	K	E	C	L	V	M	R			
H.sapiens	13FSPSS...	K	M	F	G	G	L	S	M	O	T	S	P	S	L	R	D	Y	F	S	K	F	G	E	T	R	E	C	M	V	M	R		
X.laevis	16SQHDPG...	K	M	F	G	G	L	S	M	O	T	S	P	S	L	R	D	Y	F	N	K	F	G	E	T	R	E	C	M	V	M	R		
E.siliculosus	158DIPVNK...	K	I	F	V	R	G	L	P	M	E	T	N	D	Q	S	L	R	A	V	E	Q	Y	G	E	L	A	B	A	T	V	V	M	
Z.mays	121ADPAQR...	K	I	F	V	H	C	L	G	M	D	T	T	T	T	T	L	T	E	A	F	G	P	Y	G	E	L	D	L	K	V	T		
A.thaliana	123EDLVHR...	K	I	F	V	H	C	L	G	M	D	T	K	A	S	L	I	D	A	F	K	Q	Y	G	E	L	E	D	C	K	V			
M.truncatula	122GDASHR...	K	I	F	V	H	C	L	G	M	D	T	S	A	T	L	I	N	A	F	S	Q	Y	G	E	L	E	D	C	K	A	V	T	
C.burnetii	39KSDNND...	Q	I	V	G	N	L	P	H	V	V	E	N	L	H	Q	Y	F	S	R	F	G	A	T	E	S	V	K	I	V	R			
A.haemolyticus	1MNHFLGFSM	K	I	L	V	R	N	I	D	R	S	V	I	D	A	E	V	L	E	L	K	Q	Y	G	R	V	E	S	C	V	V	V	K	
M.hungatei	1MEGK...	R	L	Y	V	G	N	L	P	H	S	A	N	E	T	Q	I	R	E	L	A	P	F	G	T	E	N	V	K	L	E			

C.parvum	122	EK.NGRSRG	G	F	V	T	K	S	S	E	S	I	K	L	F	N	E	N	L	I	N	G	R	O	L	V	K	L	A	A	D	P	F	S	E	F	T	C	G	P	I	D	C	N	K	Q	V							
P.Berghei	131	EK.EGRSKG	G	F	V	T	K	Y	I	E	S	V	Q	K	L	K	G	S	H	L	D	N	K	E	L	Q	V	R	L	V	A	D	P	F	T																			
P.falciparum	175	EK.EGRSKG	G	F	V	T	K	Y	V	E	S	V	Q	K	L	K	S	S	H	L	D	N	K	E	L	Q	V	R	L	V	A	D	P	F	T																			
C.reinhardtii	148	DA.NEVS	K	G	G	F	V	H	E	D	Q	A	A	D	R	A	I	Q	T																																			
N.tabacum	149	DS.SGQS	K	G	G	F	V	Q	D	S	E	E	A	Q	K	A	E	K																																				
O.tauri	154	DS.AGNS	K	G	G	F	V	Q	E	T	A	E	A	Q	A	A	D	N																																				
B.oleracea	39	DRATGR	A	R	G	F	G	F	L	V	E	A	D	P	T	V	A	E	R	V	L	I	R	H	V	I	D	G	K	I	V	E	A	K	R	A	V	P	R	D	D	H	V	L	N	R	S	S	S	L	Q			
H.perforatum	40	DRTTGR	A	R	G	F	G	F	L	V	E	A	D	P	S	V	A	D	R	V	L	I	K	E	K	H	N	I	D	G	R	M	V	E	A	K	R	A	I	P	R	D	D	Q	N	I	P	S	R	N	S	V	S	
A.gambiae	107	DPVTQR	S	R	G	F	G	F	I	T	E	Q	E	P	S	V	D	K	V	L	Q	V	P	I	H	L	D	G	K	I	D	P	K	H	A	T	P	K	N	R	P	K	T	Q	S									
D.melanogaster	237	DPVTQR	S	R	G	F	G	F	I	T	E	Q	E	P	S	V	D	K	V	L	Q	V	P	I	H	L	D	G	K	I	D	P	K	H	A	T	P	K	N	R	P	R	Q											
S.japonicum	112	DPLTKR	S	R	G	F	G	F	V	T	E	S	D	P	M	C	V	E	K	V	L	E	A	A	P	H	L	D	F	K	I	D	P	K	L	A	V	P	R	K	S	G	O	N	T	K	I	S						
C.elegans	79	DPATKR	A	R	G	F	G	F	I	T	E	V	D	P	S	S	V	D	K	V	L	N	N	R	E	H	L	D	G	K	I	D	P	K	V	A	F	P	K	R	T	Q	A	K	L	V								
B.taurus	54	DPLTKR	S	R	G	F	G	F	V	T	E	M	D	Q	A	G	V	D	K	V	L	A	Q	S	R	H	E	L	D	S	K	T	D	P	K	V	A	F	P	R	R	A	Q	P	K	M	V							
M.musculus	54	DPLTKR	S	R	G	F	G	F	V	T	E	M	D	Q	A	G	V	D	K	V	L	A	Q	S	R	H	E	L	D	S	K	T	D	P	K	V	A	F	P	R	R	A	Q	P	K	M	V							
H.sapiens	51	DPTTKR	S	R	G	F	G	F	V	T	E	A	D	P	A	S	V	D	K	V	L	G	Q	P	H	E	L	D	S	K	T	D	P	K	V	A	F	P	R	R	A	Q	P	K	M	V								
X.laevis	55	DPTTKR	S	R	G	F	G	F	V	T	E	A	D	P	A	S	V	D	K	V	L	A	Q	P	H	E	L	D	S	K	T	D	P	K	V	A	F	P	R	R	A	Q	P	K	M	V								
E.siliculosus	197	DKMTQK	S	K	G	G	F	V	T	E	K	S	M	D	G	A	H	A	L	E	N	P	E	K	M	I	D	G	R	V	S	L	C	N	L	A	A	L	R	S	S	Q	P	L	P	A	R	A	G	G	M	Q		
Z.mays	160	DRNTGR	C	K	G	G	F	I	L	E	R	H	R	S	G	A	R	A	L	R	E	P	Q	K	I	G	N	R	S	T	A	C	O	L	A	S	V	G	P	V	P	G	A	P	N	S	A	A	P	A	P	V		
A.thaliana	162	DKVSG	S	K	G	G	F	I	L	E	K	S	R	S	G	A	R	N	A	L	K	Q	P	Q	K	I	G	T	R	M	T	A	C	O	L	A	S	I	G	P	V	Q	G	N	P	V	A	P	A	Q	H	F		
M.truncatula	161	DKVSG	S	K	G	G	F	I	L	E	K	R	R	S	G	A	R	N	A	L	K	E	P	Q	K	I	G	N	R	M	T	A	C	O	L	A	S	I	G	P	V	Q	T	P	V	A	A	A	P	A	L			
C.burnetii	78	NFR	T	G	R	S	K	G	A	F	V	T	V	T	P	K	Q																																					
A.haemolyticus	43	DEGTGR	S	K	G	G	F	V	E	M	T	N	P	R	E																																							
M.hungatei	38	Q.....	K	G	F	G	F	V	E	M	S	S	E																																									

Localization of mRNA storage complexes in *Plasmodium berghei* throughout the life cycle /2011

C.parvum	41	IPLSHEWIVWEQLNV--ETRKDLDYSNATKPVARFSSVQQFWWLWHNIPOPSSELLKGRM
P.berghei	31	LLLQYNWVIVWEQVSD-NKIKQSNNYKDYTRPLAKFNSVQKFWQLNRLPQPSDLLAQRSM
P.falciparum	31	LLLQYNWVIVWEQVSD-NKIKQSNNYKDYTRPLAKFNSVQKFWQLNRLPQPSDLLAQRSM
Z.mays	51	HPLRHKLVLWVTRRT--XGARSQSVEDNKKKIDFSTVESFWVCYCHLARPSSLPSPT--
A.gambiae	39	HKLQYTYCLWFGKK---GSHRAAEYDKSLHFVGRCASVEQWWSLYCHLIKPLSLKPYR--
X.laevis	32	HPLQYKTYTFWMSRRTPSRPASTHNYEQNIRPFGTVASVEQFWRITYSHIVRFGDLSGYS--
E.siliculosus	10	HPLQNTWCLWEHK---VAKSSADWGDMSQSLAEFAFVVEDFWKVYNNVPKPSQVMFDGSM
C.reinhardtii	21	HPLENRWTLWFDNPN--QKQTKQKQGGSLRSVYTFDTVEDFWCLYNNRTPSQVNVGA--
A.thaliana	60	HPLEHSWTFWFDNPA--VKSQTSWGSLLRPVFTFSTVEEFWWSLYNNMKHPSKLAHGA--
O.tauri	1	-----MGT--
N.tabacum	26	HKLERRWTFWFDNQS--KPKQGAWGSLLRKAYTFETVVEEFWWSLYDQIFKPSKLTANA--
C.elegans	12	YPLKRNWTFWMLNDE----RNKSWEDRLKRYTFNTVSEFWALYDARTPPSGLNALC--
D.melanogaster	33	HRLQNTWTLWGV----KYDPEISWEDMLKEIDSENTVEDFWNLYFRIDTPSKLNRCG--
C.parvum	99	IRESSDGSKSVDVAVLFLFKGCIQPMWEDPMNATGGHIFHRAWQSSVVPGEIDTMWNNLVL
P.berghei	90	TRFSDDGIFRIVDALMIFRDNIQPMWEDPANAGGGHFEYKILPKDFPYSQIDDEFWNNLVL
P.falciparum	90	TRFSEDDGIFRIVDALMIFRDNIQPMWEDPANSGGGHFEYKILPKDYFYSQIDDEFWNNLVL
Z.mays	107	-----DLHLFKGTIRPLWEDPANQGGKWIIRFKKAVS----GRFWEDLVL
A.gambiae	94	-----RLHLFKSGIKPMWEDPSNSKGGKRWIRIKKSKI----DRAWENVCM
X.laevis	90	-----DLHLFKGTIKPMWEDEANKNGGKWIIRLRKGLA----SRSWENLIL
E.siliculosus	66	RKKVGD---RVIDSWSIFFKGKIKPWEAPEENSHGGEWQ---SRSKMTPQALDLFWENMVL
C.reinhardtii	77	-----TYLFLFKGTIEPKWEDPKNLKGGSWTAPVFNKGDSSKKTLDAWWNLTVL
A.thaliana	116	-----DFYCFKHITIEPKWEDPICANGGKWTMTFPKEKS----DKSWLYTLL
O.tauri	4	-----DLHLFKGCIQPKWEDETCAGGKWTYYVPKQKD-GGELDECWLNLLL
N.tabacum	82	-----DLHLFKGCIQPKWEDPECASGGKWTVTSSRKAN----LETMWLETLM
C.elegans	65	-----DYNVFRDITQPMWEVPEPNSNGGRWLIVTDKGTK-PEMVDAILWLETLM
D.melanogaster	86	-----DYMFLFKGIRPMWEDPPNKGGGRWYKVDKRST--AELDKTWLDVLL
C.parvum	159	AVIGGSL-ENSSIVNGTIRLVDKLGGNK-GNIRVEIWFSDFNQSAHQALLKEHETLM--S
P.berghei	150	AVIGCSL-KHYDLITGIRLVDKLSTTRYGYIRIEIWTYTTITDENVRNHLRKLDEEHM--C
P.falciparum	150	AVIGCSL-KHYDLITGIRLVDKLSTTRYGYIRIEIWTYTTITDESVKNYLRKLDEEHM--C
Z.mays	149	VVVGDLQ- EYSDDVCGVVLVSRFNED----ILSVVNRNASDHQAVMALRDSIKRHL--K
A.gambiae	136	AMIGEQF-LVGSEICGVVLCTQYPED----VLSVWNRATLDTVSTNRIRDLRRIL--N
X.laevis	132	AMIGEQF-MVGSEICGVVLSIRFQED----ILSTWTKTANDQFSTVIRIRDTLRRVL--N
E.siliculosus	120	GLVGETV-DVGDEICGARVVDKSKNNHNVY-RFELWLRN-NDGNTAARILLERLACL--A
C.reinhardtii	124	ACIGEQF-TDGDEVNGLAVVIRARGD----RIELWSRTASNEAAQTMIGKQLKQYL--D
A.thaliana	158	ALIGEQF-DHGDEICGAVVNRGKQE----RISITWTKNASNEAAQVSIQKQWKEFL--D
O.tauri	50	CMIGEQF-SESTEICGAVVSVRQKQH----RIALWTKTASSEPEQTAIGRHFKDIL--E
N.tabacum	125	ALVGEQF-DESEICGVVASVRRSQD----KLSLWTKTASNEAIQMSIGKRWKEIL--E
C.elegans	111	ALVGEQFGKDMESICGLVCNVRGKGS----KISVWTKDCNDDETNRIGVVLKEKIMAA
D.melanogaster	131	CMIGETAC-DHCDQICGAFVRIIRKKNIN----KISVWTKADAGDECNEQC-----

Table S6 | eIF4E domain alignment by MUSCLE
(<http://www.ebi.ac.uk/Tools/msa/muscle/>)

7 – References

Billker, O., M. K. Shaw, G. Margos and R. E. Sinden (1997). "The roles of temperature, pH and mosquito factors as triggers of male and female gametogenesis of *Plasmodium berghei* in vitro." Parasitology **115** (Pt 1): 1-7.

Carlton, J. M., S. V. Angiuoli, et al. (2002). "Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*." Nature **419**(6906): 512-519.

Carlton, J. M., R. Vinkenoog, A. P. Waters and D. Walliker (1998). "Gene synteny in species of *Plasmodium*." Mol Biochem Parasitol **93**(2): 285-294.

Charlesworth, A., A. Wilczynska, P. Thampi, L. L. Cox and A. M. MacNicol (2006). "Musashi regulates the temporal order of mRNA translation during *Xenopus* oocyte maturation." EMBO J **25**(12): 2792-2801.

Chekulaeva, M., M. W. Hentze and A. Ephrussi (2006). "Bruno acts as a dual repressor of oskar translation, promoting mRNA oligomerization and formation of silencing particles." Cell **124**(3): 521-533.

Coulson, R. M., N. Hall and C. A. Ouzounis (2004). "Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*." Genome Res **14**(8): 1548-1554.

Doi, Y., N. Shinzawa, S. Fukumoto, H. Okano and H. Kanuka (2011). "Calcium signal regulates temperature-dependent transformation of sporozoites in malaria parasite development." Exp Parasitol **128**(2): 176-180.

Gardner, M. J., N. Hall, et al. (2002). "Genome sequence of the human malaria parasite *Plasmodium falciparum*." Nature **419**(6906): 498-511.

Gomes-Santos, C. S., J. Braks, et al. (2011). "Transition of *Plasmodium* sporozoites into liver stage-like forms is regulated by the RNA binding protein Pumilio." PLoS Pathog **7**(5): e1002046.

Hall, N., M. Karras, et al. (2005). "A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses." Science **307**(5706): 82-86.

Imai, T., A. Tokunaga, T. Yoshida, M. Hashimoto, K. Mikoshiba, G. Weinmaster, M. Nakafuku and H. Okano (2001). "The neural RNA-binding protein Musashi1 translationally regulates mammalian numb gene expression by interacting with its mRNA." Mol Cell Biol **21**(12): 3888-3900.

Jackson, R. J., C. U. Hellen and T. V. Pestova (2010). "The mechanism of eukaryotic translation initiation and principles of its regulation." Nat Rev Mol Cell Biol **11**(2): 113-127.

Kahvejian, A., Y. V. Svitkin, R. Sukarieh, M. N. M'Boutchou and N. Sonenberg (2005). "Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms." Genes Dev **19**(1): 104-113.

Kawahara, H., T. Imai, H. Imataka, M. Tsujimoto, K. Matsumoto and H. Okano (2008). "Neural RNA-binding protein Musashi1 inhibits translation initiation by competing with eIF4G for PABP." J Cell Biol **181**(4): 639-653.

Kim-Ha, J., K. Kerr and P. M. Macdonald (1995). "Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential." Cell **81**(3): 403-412.

Le Roch, K. G., J. R. Johnson, et al. (2004). "Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle." Genome Res **14**(11): 2308-2318.

Mair, G. R., J. A. Braks, et al. (2006). "Regulation of sexual development of Plasmodium by translational repression." Science **313**(5787): 667-669.

Mair, G. R., E. Lasonder, et al. (2010). "Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development." PLoS Pathog **6**(2): e1000767.

Miao, J., J. Li, Q. Fan, X. Li and L. Cui (2010). "The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite Plasmodium falciparum." J Cell Sci **123**(Pt 7): 1039-1049.

Muller, K., K. Matuschewski and O. Silvie (2011). "The Puf-Family RNA-Binding Protein Puf2 Controls Sporozoite Conversion to Liver Stages in the Malaria Parasite." PLoS One **6**(5): e19860.

Nakamura, A., K. Sato and K. Hanyu-Nakamura (2004). "Drosophila cup is an eIF4E binding protein that associates with Bruno and regulates oskar mRNA translation in oogenesis." Dev Cell **6**(1): 69-78.

Nelson, M. R., A. M. Leidal and C. A. Smibert (2004). "Drosophila Cup is an eIF4E-binding protein that functions in Smaug-mediated translational repression." EMBO J **23**(1): 150-159.

Okano, H., T. Imai and M. Okabe (2002). "Musashi: a translational regulator of cell fate." J Cell Sci **115**(Pt 7): 1355-1359.

Prudencio, M., A. Rodriguez and M. M. Mota (2006). "The silent path to thousands of merozoites: the Plasmodium liver stage." Nat Rev Microbiol **4**(11): 849-856.

Rajyaguru, P. and R. Parker (2009). "CGH-1 and the control of maternal mRNAs." Trends Cell Biol **19**(1): 24-28.

Sonenberg, N. and A. G. Hinnebusch (2009). "Regulation of translation initiation in eukaryotes: mechanisms and biological targets." Cell **136**(4): 731-745.

WHO (2010). "World Malaria Report 2010."

Wilhelm, J. E., M. Buszczak and S. Sayles (2005). "Efficient protein trafficking requires trailer hitch, a component of a ribonucleoprotein complex localized to the ER in Drosophila." Dev Cell **9**(5): 675-685.

Wilkie, G. S., K. S. Dickson and N. K. Gray (2003). "Regulation of mRNA translation by 5'- and 3'-UTR-binding factors." Trends Biochem Sci **28**(4): 182-188.

Wirth, D. F. (2002). "Biological revelations." Nature **419**(6906): 495-496.