UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

THE ROLE OF A NEWLY DISCOVERED FAMILY OF TRANSCRIPTION FACTORS, APIAP2, DURING THE DEVELOPMENT OF THE RODENT MALARIA PARASITE *PLASMODIUM BERGHEI*

Ana Rita Batista Gomes

Mestrado em Microbiologia Aplicada

2010

UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

THE ROLE OF A NEWLY DISCOVERED FAMILY OF TRANSCRIPTION FACTORS, APIAP2, DURING THE DEVELOPMENT OF THE RODENT MALARIA PARASITE *PLASMODIUM BERGHEI*

Ana Rita Batista Gomes

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
2010

Abbreviations

3' UTR	3' untranslated region
5' UTR	5' untranslated region
ApiAP2	Apicomplexan Apetala 2 family of transcription factors
cDNA	complementary desoxyribonucleic acid
CITH	CAR-I and fly TrailerHitch Homolog
DAPI	4',6'-diamidino-2-phenylindole
DNA	Desoxyribonucleic acid
DOZI	Development of Zygote Inhibited
DTT	Dithiothreitol
EEF	Exoerythrocytic form (liver stage)
EST	Expressed sequence tags
FBF	fem-3 binding factor
HIV	Human immunodeficiency virus
IFA	Immunofluorescence assay
iRBCs	Infected red blood cells
iv	Intra-venous
KO	Knock-out
mRNP	messenger ribonucleoprotein
NLS	Nuclear localization signal
PCR	Polymerase chain reaction
p.i.	post-infection
RNA	Ribonucleic acid
RT-PCR	Reverse-transcription polymerase chain reaction
SDS	Sodium lauryl sulphate
TF	Transcription factor
TBV	Transmission blocking vaccine
TR	Translational repression
uis4	up-regulated in infective sporozoites 4
WBC	White blood cell
WHO	World Health Organisation
WT	Wild type
PUF	pumilio protein

Table of contents

Acknowledgements

This project was developed at Instituto de Medicina Molecular in UPAMOL unit (Unidade de Parasitologia Molecular) under the direct supervision of Dr. Gunnar Mair and Dr Céline Carret and was funded by FCT Project PTDC/BIA-BCM/105610/2008, and GAPIC project 200913.

To our neighbours, Malaria Unit a special thanks to Fernanda and Ana for all the help with the mosquitoes experiments and for all the nice moments we spent together. To Kirsten, Vanessa and Sílvia thank you for all the help and suggestions; and to Patrícia, *the brain*, thanks for all the laughs when I wanted to change my title to "Troubleshooting in *Plasmodium* protocols". I am also grateful to Dr. Maria Mota e Dr. Miguel Prudêncio for the fruitful comments during labmeetings. To Ana Roberto thank you for all the help with the never-ending bureaucratic issues.

To UPAMOL team a sincere thank you for having me in during the Microbiology degree and now during the Master's degree. Ana e Jorge muita sorte para os vossos projectos de doutoramento!

To Gunnar and Céline, a "massive" thanks for all the training, support, friendship and patience during these two years. Thank you for keeping the standards high, Cambridge would not be an option without you. I wish you all the best in the world and I promise not disappoint you....

To Prof. Rita Zilhão a special thanks for getting me started in the Molecular Biology business.

To David, Nina, Vera and Cat, for everything, because it wouldn't make sense without you....

Finally I would like to express my gratitude to Prof. Rogério Tenreiro for his kindness and help.

These last two years have been great.

To all, thank you.

Publications

Gomes-Santos, C.B., J; Prudêncio, M; Carret, C; **Gomes, ARB**; Pain, A; Feltwell, T; Khan, S; Waters, A; Janse, C; Mair, G and Mota, M, Developmental transition of *Plasmodium* sporozoites into liver-stage forms is regulated by the RNA binding protein pumilio. *MPMXXI, 2010, 12-16 September*.

Best oral presentation Prize in "Tracking the role of translationally silenced transcription factors in Plasmodium berghei – the functions of AP2 factors during early sexual developmental mosquito stages" in the 12th annual Science Workshop organized by GAPIC (Medicine Faculty Office that allows undergraduate students to be involved in research projects) Medicine faculty, Lisbon, Portugal.

Abstract

Malaria is a vector borne disease that causes 1 million deaths annually. Gene expression control of *Plasmodium* parasites is poorly understood. A recently identified group of AP2/ERF-like proteins, ApiAP2 proteins, may represent the main transcriptional regulators. Gene expression in gametocytes (sexual precursor cells) is regulated by translational repression through the storage of certain mRNAs (for example transcripts encoding the surface proteins P25 and P28) in translationally silent P granule-like complexes. Only after fertilization are they activated, providing essential proteins for ookinete transformation during its early mosquito stage development. The main objective of this study was to understand whether the 3' untranslated region (UTR), widely believed to target mRNAs to translationally quiescent P bodies, are responsible for the silencing of certain ApiAP2 molecules in female gametocytes. Using a combination of bioinformatics, imaging, molecular and genetic tools we aimed to GFP-tag AP2 C-termini and at the same time replace the endogenous 3' UTR with a known one that allow translation in gametocytes. We successfully generated a stable transgenic *Plasmodium berghei* line, AP2-O::GFP (PBANKA_090590), that expresses AP2-O under the control of its endogenous promoter with an altered 3' UTR. Infection patterns (growth kinetics, parasite morphology and sporozoite numbers) of this mutant line were analysed and shown to be comparable to wild type parasites. Furthermore, AP2-O::GFP expression was restricted to early mosquito stages (i.e. female gamete and ookinete) when AP2-O is naturally expressed. These experiments indicate that the 3' UTR of AP2-O is not sufficient to trigger silencing in female gametocytes.

Key words: Malaria, ApiAP2 factors, gametocyte, translational repression, 3'-UTR

Resumo

A malária é uma doença causada por um parasita do género *Plasmodium* e transmitida pela picada de um mosquito do género *Anopheles*. De acordo com o último relatório da Organização Mundial de Saúde (OMS) esta doença ameaça sensivelmente metade da população mundial, ceifando um milhão de vidas anualmente. As crianças com idade inferior a cinco anos são as mais afectadas por este flagelo. Actualmente encontram-se descritas mais de 5000 espécies de *Plasmodium* que infectam não só o Homem mas também animais como macacos, pequenos roedores, aves e répteis. Entre as quatro espécies que infectam o Homem, a mais virulenta e prevalente é *Plasmodium falcifarum*. Nas últimas décadas tem-se verificado um aumento preocupante das resistências aos fármacos. Deste modo, actualmente vários programas de investigação científica dedicam-se ao estudo de potenciais alvos para vacinas. Consoante o estádio que a vacina ambiciona bloquear e o efeito surtido, assim a sua designação. Se o objectivo for bloquear a fase hepática, será uma vacina profiláctica, se o bloqueio for direccionado para a fase sanguínea será uma vacina "curativa" e se procurar interromper a infecção no mosquito é considerada uma vacina de bloqueio de transmissão. Neste último exemplo a vacina não tem como objectivo proteger o indivíduo imunizado mas sim desencadear uma resposta imunitária humoral que será capaz de bloquear o desenvolvimento do parasita num mosquito que pique, posteriormente, o indivíduo vacinado, reduzindo-se o número de parasitas com capacidade infecciosa na comunidade.No entanto, para desenvolver este tipo de tecnologia é necessário um conhecimento mais profundo da biologia molecular do parasita, nomeadamente ao nível do seu desenvolvimento sexual (fase do mosquito).

O genoma de *Plasmodium falciparum* codifica sensivelmente 5400 genes, contudo, mais de 60% apresenta uma reduzida ou mesmo nenhuma homologia relativamente aos genes dos outros eucariotas. Entre as espécies de *Plasmodium* que infectam o Homem e os roedores verifica-se uma elevada sintenia entre blocos de cromossomas e equivalente predição do número total de genes. Este facto permite que estudos efectuados no modelo animal possam ser extrapolados com relativa credibilidade para o caso da infecção em humanos. Assim, neste projecto foi utilizado o parasita de roedores, *P. berghei*. A utilização do modelo animal apresenta inúmeras vantagens, uma vez que não só as técnicas moleculares são mais eficientes para este modelo como todo o ciclo de vida pode ser mantido no laboratório.

Actualmente, o controlo transcricional dos parasitas da malária encontra-se fracamente descrito. A rápida expansão e adaptação que segue a transmissão entre hospedeiros (vertebrado e mosquito) sugerem uma complexa coordenação a nível molecular. Deste modo, devido ao reduzido número de activadores transcricionais identificados, foi sugerido que a regulação da expressão génica seja efectuada a nível pós-transcricional. De facto, foi demonstrado que este tipo de regulação ocorre no gametócito feminino e medeia o silenciamento de centenas de transcritos, mantendo-os num estado quiescente em grânulos citoplasmáticos. A tradução destes transcritos só decorre na fase do mosquito, após fertilização. Pensa-se que um motivo localizado na região não traduzida da extremidade 3' (3'UTR) seja responsável pelo recrutamento dos transcritos para os grânulos. Parasitas incapazes de produzir a proteína DOZI e/ou CITH funcionais, não geram zigotos viáveis. Estas proteínas estão presentes nestes complexos de repressão e pensa-se que por interagirem

fisicamente com os transcritos, quando ausentes destabilizam o complexo levando à desregulação de cerca de 300 transcritos. Entre esses transcritos encontram-se 5 factores de transcrição da família ApiAP2. As proteínas, ApiAP2, foram recentemente descritas e actualmente constituem o principal grupo de reguladores da transcrição destes parasitas; estas contêm domínios de ligação ao DNA, homólogos aos encontrados na família de factores de transcrição AP2/ERF, identificada em plantas, como *Arabidopsis*. O facto destes factores de transcrição estarem desregulados no mutante com deleção do gene *dozi* sugere que talvez tenham um papel preponderante na activação transcricional no estádio de zigoto.

Este projecto teve como objectivos a caracterização dos padrões de expressão genética de genes ApiAP2 e avaliar quais se encontravam regulados a nível pós-transcricional nos estádios de transmissão, nomeadamente no gametócito e no esporozoito. Para mais, este trabalho também visou determinar as funções moleculares de cinco factores ApiAP2 (PBANKA_090590 [também designado *ap2-o*], PBANKA_140370, PBANKA_093910, PBANKA_010290, PBANKA_131320), que são potencialmente regulados por mecanismos pós-transcricionais pelo facto de se encontrarem desregulados no mutante com delecção do gene *dozi.* Para tal tentou-se interferir com o seu padrão de expressão genética. Foram então utilizadas técnicas moleculares, genéticas, de microscopia e bioinformáticas para confirmar tanto os modelos genéticos dos cinco alvos propostos como o modelo proposto para a sua regulação (supressão da tradução). Vectores de transfecção foram preparados de modo gerar genes de fusão com o gene *gfp* (*green fluorescent protein*) para que a expressão das respectivas proteínas pudesse ser seguida por análise de fluorescência. A transfecção foi efectuada numa estirpe selvagem de *P. berghei* ANKA. Posteriormente, a correcta integração dos vectores no genoma dos parasitas foi avaliada por técnicas de PCR. Para induzir alterações no padrão de expressão destas moléculas, ou seja evitar o seu recrutamento para os grânulos citoplasmáticos dos gametócitos, os vectores de transfecção foram construídos de modo a que a sua integração no genoma do parasita deslocalizasse a região 3'UTR, substituindo-a por outra de um transcrito traduzido constitutivamente.

As populações de mutantes obtidas foram então clonadas seguindo-se a análise e detecção das proteínas de fusão fluorescentes. Tendo em conta que estas moléculas são factores de transcrição, ao tentar induzir precocemente a sua expressão esperava-se induzir também a expressão dos genes que controlam. Se esta indução resultasse na produção de proteínas de superfície (típicas da fase do mosquito) capazes de gerar uma resposta imunitária no hospedeiro vertebrado então essas proteínas seriam potenciais alvos para vacinas de bloqueio de transmissão.

Numa primeira análise os resultados da análise bioinformática sugeriram uma localização nuclear para estas proteínas. A análise por imunofluorescência, em combinação com técnicas moleculares de RT-PCR indicou ainda que as cinco moléculas ApiAP2 são reguladas pós-transcricionalmente no gametócito feminino.

A genotipagem dos parasitas transfectados revelou a presença de uma linha de parasitas mutantes para o gene *ap2-o (ap2-o::gfp*). Sendo naturalmente expresso nos estádios iniciais da fase do mosquito a presença da proteína AP2-O::GFP na fase sanguínea indicaria um papel para a região 3' UTR no silenciamento do respectivo transcrito. No entanto após análise de todos os estádios da

fase sanguínea por técnicas de fluorescência a proteína não foi detectada. Procedeu-se também à infecção de mosquitos e todos os indicadores de infecciosidade eram comparáveis aos da estirpe selvagem, nomeadamente, número de oocistos, número de esporozoitos gerados e capacidade de infectar novo hospedeiro vertebrado.

Tendo em conta os resultados obtidos procedeu-se à observação de oocinetos - o estádio onde foi demonstrada a presença da proteína AP2-O. Para tal efectuaram-se culturas *in vitro* de oocinetos. Utilizando técnicas de microscopia de amostras vivas foi detectada a presença da proteína de fusão AP2-O::GFP tanto em oocinetos como em gâmetas femininos não fertilizados, estádio onde expressão desta proteína nos gâmetas não tinha sido previamente descrita.

Estes resultados sugerem que no caso do gene *ap2-o* a repressão da tradução no gametócito ou é independente da região 3' UTR ou requer a interacção com um outro motivo, como por exemplo a região 5' UTR. Deste modo, idealizam-se como experiências futuras uma alteração concomitante de ambas as regiões não traduzidas (3' e 5') e/ou a sobre-expressão deste gene através da introdução de um promotor de expressão constitutiva, como por exemplo o promotor do gene *efl*α.

Palavras chave: Malária, factores ApiAP2, gametócito, repressão da tradução, 3'-UTR

1- Introduction

Infectious diseases are still in the top 10 of cause of death, with malaria, HIV and tuberculosis accounting for ~10% of total deaths worldwide[1]. Although the widespread use of vaccines and drugs has severely decreased mortality from infectious diseases in developed countries, they continue to be prevalent in the Third World; not only is this catastrophic for these countries, but some of these diseases are beginning to emerge or re-emerge in developed countries[2].

Parasitic protozoans are responsible for some of the most devastating infectious diseases. These unicellular, eukaryotic organisms are responsible for human diseases like amoebiasis, Chagas' disease, African sleeping sickness, malaria, leishmaniosis and toxoplasmosis. Malaria is a vector borne disease caused by Apicomplexan parasites from the genus *Plasmodium*. More than 5000 species of *Plasmodium* have been described[3]. They infect humans but also animals including monkeys, rodents, birds, and reptiles. Out of the four species that are human pathogens, *P. falciparum* is the most virulent and widespread one. According to the WHO 2009 report, malaria endangers half of the world's population [Fig.1], and is responsible for 3-500 millions of clinical cases claiming 1 million lives every year[4].

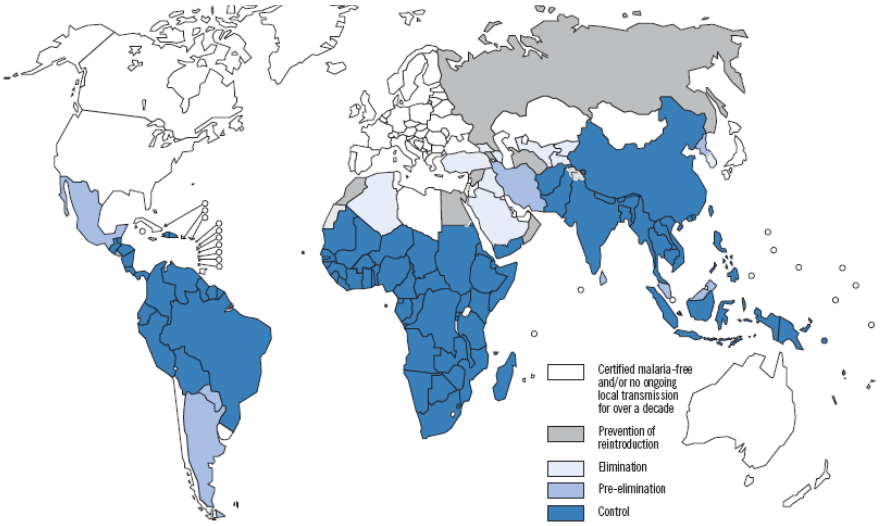

Fig. 1| Malaria is endangering half of the world's population. Malaria-free and malaria-endemic countries in phases of control, pre-elimination and elimination in the end of 2008. In "Pre-elimination" countries malaria test positivity rate is less than 5% during the malaria season (>5% "Control" phase); "Elimination" countries have zero incidence of locally transmitted infections; "Malaria free" countries have no locally transmitted infections for over a decade. (Adapted from WHO report 2009)

Despite intensive efforts during the 20th century the impact of this disease on human health is increasing as financial constraints and emerging resistance to current drug therapies hamper malaria control programs. Owing to the growing resistance to drug therapies an anti-malarial vaccine is considered the holy grail of malaria research. Vaccines can approach three different steps of the lifecycle: the pre-erythrocytic or liver stage (preventive), erythrocytic or blood stage (curative) and sexual stage parasites (transmission blocking; TBV) [5]. Anti-sexual-stage or TBVs intend to reduce parasite transmission between individuals. Unlike liver-stage or asexual-blood-stage vaccines that reduce human infection rates or lower the disease severity, transmission blocking vaccines do not protect an individual directly from acquiring malaria[6], instead, by inducing the production of antibodies

in the host that will inhibit parasite development in the mosquito midgut, they altruistically could provide community protection through decreased overall parasite load and transmission[7].

Recently the AMA1-based malaria vaccine FMP2.1/AS02A (blood stage) showed a good safety profile and tolerability, leading to more than 100-fold increase of anti-AMA1 antibody levels. AMA1 is an apical membrane protein that plays a critical role during merozoite invasion of erythrocytes. This malaria vaccine is now being evaluated in a Phase 2 efficacy trial [8]. Concerning TBV, to date studies have focused on two major gametocyte-specific surface antigens — Pfs48/45 and Pfs230 — and the two ookinete-specific surface coat proteins P25 and P28. Antisera from mice immunized with recombinant Pvs25 (P25 from *P. vivax* parasites) completely prevent the appearance of oocysts in mosquitoes that had ingested the antisera with *P. vivax* parasites[9]. This vaccine is now in phase I trials to determine safety and immunogenicity. However, preliminary tests prompted measurable but weak reactions in healthy volunteers; when combined with an adjuvant to boost the body's immune response severe reactions in two participants deferred the trial.

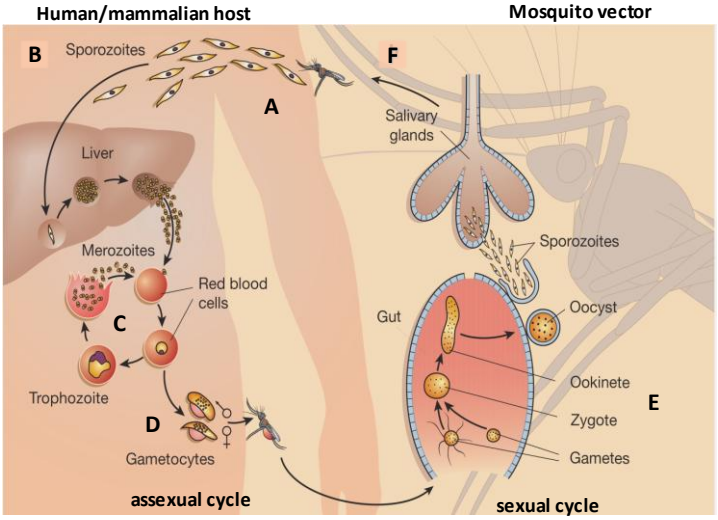

Fig. 2|Malaria life cycle. The parasites moves between two different hosts – the vertebrate and the mosquito vector. The asexual cycle is initiated when the vertebrate host is bitten by an infected *Anopheles* mosquito. The injected sporozoites migrate to the liver where they generate thousands of merozoites; when they leave the hepathocyte they invade erythrocytes. Once gametocytes, the sexual precursors, are taken by another mosquito, the sexual stage (i.e. mosquito phase) begins. (Adapted from Dyann Wirth, 2002).

The malaria parasite progresses through a remarkable series of developmental and maturational stages during its complex life cycle [Fig.2][10]. A human infection is initiated when *Plasmodium* sporozoites are injected during the blood meal of an infected, female *Anopheles* mosquito [Fig.2a]. During the bite an average of 123 (Ana Rita: é mesmo 123?) sporozoites[11] are transferred into the blood stream and quickly migrate to the liver [Fig.2b]; there they multiply giving rise to thousands of merozoites in 2-16 days, depending on the *Plasmodium* species. Eventually these merozoites bud off from the hepatocyte through so-called merosomes into the liver sinusoids[12] capable of infecting red blood cells (RBCs), thus starting the blood stage of the infection. While the initial stages of the infection (i.e. liver stage) are asymptomatic, repeated infection of erythrocytes by merozoites causes the symptoms and pathologies of malaria (fever, anaemia, respiratory problems,

sequestration of parasites in the brain can lead to cerebral malaria and death). Inside the erythrocyte, a single merozoite replicates and undergoes successive differentiations from ring through trophozoite stage eventually generating 16-32 schizonts; finally the RBC ruptures and releases new merozoites, which in turn infect new erythrocytes [Fig.2c]. During the asexual cycle in the blood a subset of parasites bypasses asexual multiplication and differentiates into sexually committed cells: the female and male gametocytes (or sexual precursor cells) [Fig. 2d]. These forms are arrested in the G0 phase and only re-enter the cell cycle to produce gametes after having been ingested by a mosquito. In the mosquito midgut fertilisation of a female by a male gamete results in the formation of the zygote that develops into a motile, invasive ookinete. The ookinete crosses the midgut wall and forms an oocyst on the basolateral lamina [Fig. 2e] where it will generate thousands of oocyst-derived sporozoites. When mature, these sporozoites migrate through the hemocoel to the mosquito's salivary glands, hence closing the cycle [Fig 2f] [13].

Although life cycle progression is well established in several *Plasmodium* species and is studied in the lab, the molecular cues that drive *Plasmodium* development and differentiation are only recently being unravelled. The *P. falciparum* genome encodes around 5400 genes[14] but more than 60% of those encode proteins with weak or no homology to other eukaryotes, and have not been assigned a function experimentally. Human and rodent malaria parasites share highly syntenic chromosome blocks with similar number of predicted genes[15-17]. As maintaining the complete *P. falciparum* life cycle in the laboratory is still a difficult task, the conserved synteny and clear linkage groups between human and rodent parasites (*P. berghei* is used in this project) strengthens the credibility of using the rodent model in many studies[17]. The transcriptome and proteome of the different malaria parasites is highly regulated with up to 30% specificity for certain life cycle stages[18,19]; how this is accomplished on a molecular level is largely unknown. Until recently, comparative genomics analyses of Apicomplexans suggested a scarcity of specific transcription factors (TF) despite comparable numbers of genes and signalling pathways found in other unicellular eukaryotes[20,21]. The ratio of the total number of genes in the genome to the total number of TFs in yeast is in the range of 25–30, while for *Plasmodium* it is in the range of 350–800[22,23]. The parasitic lifestyle could explain part of this discrepancy, as parasites may require less intricate regulation during much of their life cycle compared to free-living organisms who face many homeostatic challenges. However, Apicomplexans have an extensive set of structural and regulatory chromosomal proteins, as well as cytoplasmic signalling proteins, such as kinases and GTPases, that are found in comparable numbers to other eukaryotes [23-25]. The most severe and rapid changes in environment are experienced by the parasite during transmission between vertebrate and invertebrate host and require a fast adaptation and modulation of gene expression. Transmission from one host to the next results in a reduction of the parasite population; each transmission to and from the mosquito vector is therefore followed by the generation of a newly adapted parasite form and the expansion of the remaining population; for example the fertilized female gamete transforms within 18 hours into the motile ookinete and gives rise to a single oocyst that can generate thousands of sporozoites; a single sporozoite, once established within a human liver cell, will expand into thousands of liver stage merozoites.

The absence of clearly recognizable TF's led to the speculation that gene expression in *Plasmodium* may preferentially be regulated post-transcriptionally through a combination of mRNA degradation and translational mechanisms; or rely on epigenetics[26-28]. Post-transcriptional gene regulation has been shown or suggested to play an essential role during parasite transmission to and from the mosquito vector. Silencing of mRNAs (crucial during sexual differentiation, gametogenesis and embryogenesis of multicellular eukaryotes) regulates the translation of specific transcripts during gametocytogenesis and transformation of fertilized female gametes into ookinetes, the mosquito-infective form [29]. In metazoans (*Drosophila* for example) targeting mRNAs for silencing relies on distinct interactions of RNA-binding proteins with specific, short motifs commonly situated within the 3' untranslated regions (UTR) of mRNAs thereby influencing mRNA stability and accessibility to the translational apparatus [29-31]. In *P. berghei* it has been shown that in the female gametocyte a pool of translationally repressed mRNAs (containing more than 300 transcripts) is stored in cytoplasmic P-body like complexes (P granules) protecting them from degradation; after fertilization, these mRNAs are activated and translated generating essential proteins for ookinete development and mosquito infection [27]. Two proteins, DOZI and CITH, define a stable,16-protein strong cytoplasmic mRNP that includes the cap-binding protein eIF4E, poly(A) binding protein and Alba-domain proteins. Immunoprecipitation experiments in gametocytes showed that well-characterised silenced transcripts such as *p25* and *p28* co-elute with DOZI::GFP and CITH::GFP indicating a physical association with this mRNP [Fig. 3].

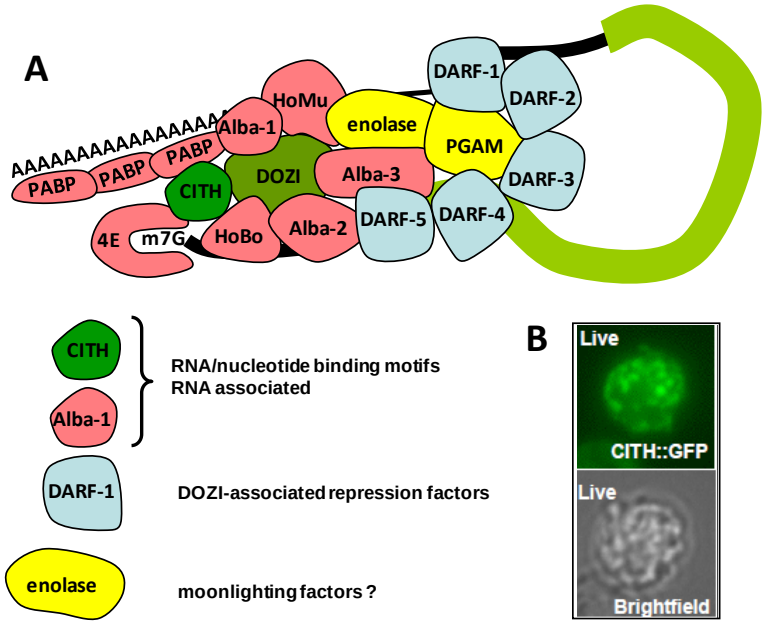

Fig. 3|DOZI and CITH defined P-granules of female gametocytes. A Protein composition of DOZI and CITH-containing granule as defined by immunoprecipitation and mass-spectrometric analysis of female gametocyte P granules. **B** Live microscopy of CITH::GFP in female gametocyte. (Adapted from Mair et al. 2010)

DOZI, a DDX6 group RNA helicase, is predominantly expressed in the female gametocyte with a speckled appearance [16], and acts as an essential stabiliser of translationally repressed transcripts.

DOZI knockout parasites abort zygote to ookinete transformation, probably due to the observed de-stabilization of more than 300 mRNAs [27,29].

Although the precise mechanism is still unknown, gene deletion mutants of the RNA binding protein pumilio-2 causes a de-regulation of life cycle progression independent from environmental factors triggering early EEF differentiation while still inside the mosquito salivary glands. The PUF family of proteins is named after the two founding members, *Drosophila* Pumilio protein and *Caenorhabditis elegans* FBF (fem-3 binding factor). Apart from the Pumilio homology domain, these proteins are highly diverse and bind to mRNA 3' UTRs to modulate translation in a variety of eukaryotic species; they do so either by enhancing turn-over or repressing translation, and act combinatorially with other regulatory proteins[32]. For example, together with Nanos Pumilio regulates the translation of *hunchback* mRNA in *Drosophila* embryos. All malaria parasites have two conserved Puf proteins that are differentially expressed in gametocytes[33]. In human malaria the protein Puf2 has its highest level of protein expression in salivary gland sporozoites [34] and was found in cytoplasmic speckles in *P. berghei* sporozoites[35] [Fig. 4A-D]; the pattern resembles somewhat DOZI::GFP distribution in the female gametocyte[36] and the sporozoite [Fig. 4E, F].

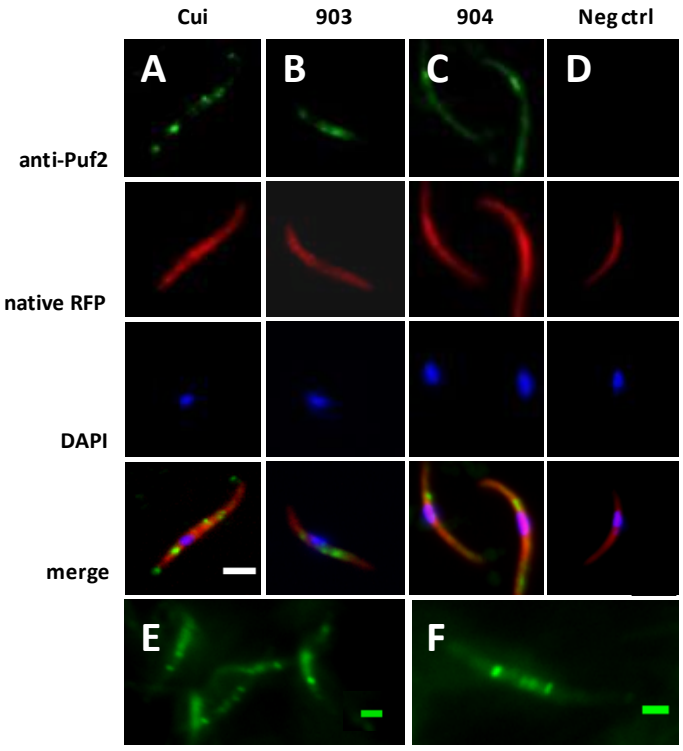

Fig. 4|Puf-2 immunofluorescence (A-D) and live DOZI::GFP localization (E, F) in wild type salivary gland sporozoites. Note the few speckles in the cytoplasm of Puf-2 staining and DOZI::GFP localization reminiscent of P-bodies identified in female gametocytes. Cui, 903 and 904 denote 3 polyclonal anti-Puf-2 antisera used in IFA. Scale bars: A-D = 3 μm; E-F = 2 μm

Both examples show that post-transcriptional regulation of gene expression is essential for rapid life cycle progression in the newly parasitized host. Among the transcripts destabilized and therefore probably translated after fertilization in DOZI but also CITH (an Sm-like protein that co-purifies with

DOZI [27]) gene deletion parasites are several ApiAP2 transcription factors [see Table 1]. One of these mRNAs (AP2-O, PBANKA_090590) was recently shown to co-elute with DOZI and CITH::GFP in pull-down experiments from gametocytes, while the protein appears to be expressed only in developing ookinetes. AP2-O is responsible for stage-specific transcription of a small subset of genes known to be important for midgut invasion in the ookinete, hence its designation as AP2-O; its KO impairs healthy ookinete generation and consequently mosquito infection[37]. The down-regulation of 4 additional factors in DOZI gene deletion mutants strongly indicates that they play an important – however unknown – role during transcriptional activation in the newly-formed zygote and developing ookinete, possibly extending into the early oocyst stage. Gene deletion of the Puf-2 RNA binding protein resulted in transcriptional adaptations characteristic of early EEF stages possibly driven by the de-regulation of 3 AP2 transcripts (PBANKA_052170, PBANKA_010950, PBANKA_083520) [Table 1], again supporting translational control as an important mechanism of gene expression regulation for the AP2 proteins during malaria parasite transmission.

Table 1|AP2's identified in mutant parasite lines with defects in transmission. logFc values are log2-fold changes of ratio between mutant versus wild type parasite as measured by microarray profiling. DOZI values are from RNA from Nycodenz gradient purified mixed gametocytes, while pumilio values are from RNA isolated from hand-dissected salivary gland, 27 day old sporozoites.

	genedb	mutant	logFc	reference
	PBANKA_140370	DOZI	-2.3	27,36
	PBANKA_131320	DOZI	-2.3	27,36
	PBANKA_093910	DOZI	-3.5	27,36
	PBANKA_010290	DOZI	-3.3	27,36
ap2-o	PBANKA_090590	DOZI	-6.2	27,36,37
	PBANKA_052170	Puf2	-2.3	35
	PBANKA_010950	Puf2	1.6	35
	PBANKA_083520	Puf2	1.4	35

AP2 factors were only very recently identified in Apicomplexans by sensitive sequence profile searches. They were shown to be encoded in *Plasmodium* as a large, lineage-specific family of DNA binding proteins that belong to the ApiAP2 family. These proteins contain one or more copies of a conserved, ~60-residue AP2 domain[38] which are homologous to the well-characterized and widespread plant AP2 family of transcription factors [20]. In *P. falciparum*, 27 genes encoding AP2

proteins have been identified [14] with clear orthologs for 26 in the rodent malaria parasite *P. berghei*. In plants, these factors are involved in key developmental steps, such as flower organogenesis and seed development, and regulate gene expression during stress responses[39]. Sometimes also called the AP2/ERF (ethylene-responsive factor) domain, AP2/ERF DNA-binding proteins comprise the second largest class of regulators in *Arabidopsis*, with 145 members[40]. Through heteronuclear multidimensional NMR analysis its 3-D solution structure revealed a 3-stranded anti-parallel β-sheet and an α-helix packed approximately parallel to the β-sheet[41] [Fig.5]. In addition, the β-sheet contacts 8 out of 9 consecutive base pairs in the major groove of DNA, demonstrating that the β-sheet is crucial for the formation of the protein-DNA complex. This new kind of DNA interaction pattern differs from the one usually seen for a similar arrangement of secondary structural elements based on α and β strands

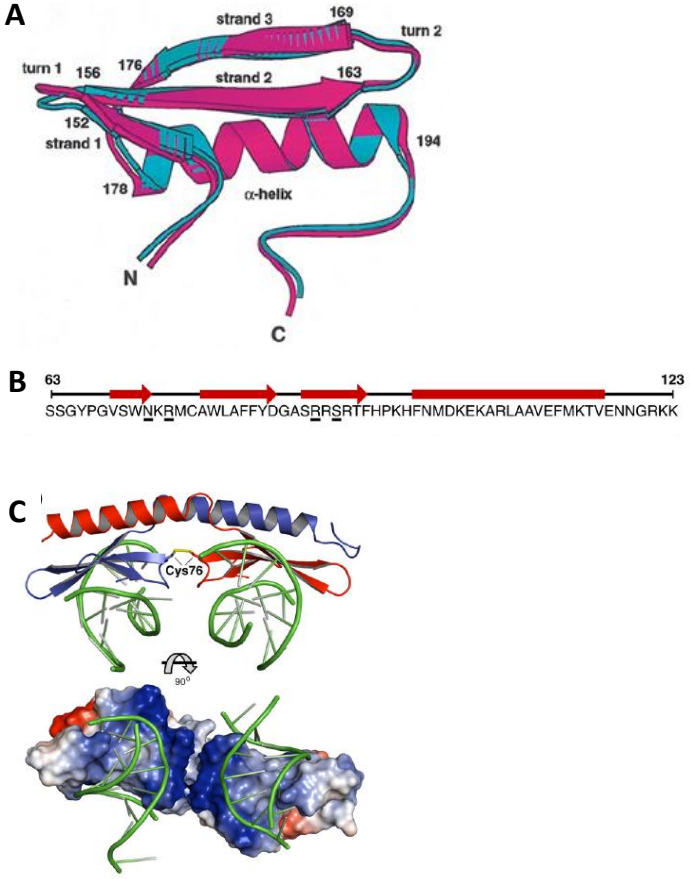

Fig. 5|The structure of an AP2 domain from plants (A) and *Plasmodium* (B, C). A The secondary structural elements of the plant AP2 domain, e.g. the α-helix, are labeled with the amino acid positions at the ends, e.g. 178. The N- and C-termini are labeled (Adapted from Allen et al. 1998). **B** Secondary structural elements of the AP2 domain of PF14_0633, the *P. berghei* AP2-Sp ortholog, are highlighted above the primary sequence as arrows (β-sheet) and a rectangles (α-helix). Residues that make base-specific contacts to double-stranded DNA are underlined. **C** (Top) Ribbon diagram of the crystal structure of a domain-swapped dimer of the AP2 domain (one monomer shaded blue, another red) bound to independent double-stamded DNA molecules (shaded green). Disulfide-bonded Cys^{76} residues are indicated and shaded yellow. (Bottom) Orthogonal view with protein electrostatic surface potentials highlighted (blue, electropositive; red, electronegative) (Adapted from Lindner et al. 2009)

where α helices bind to DNA (e.g. zinc finger of SWI5) [41,42]. Recently, the ApiAP2 domain of the gene PF14_0633 (AP2-Sp ortholog PBANKA_132980), responsible for the transcription of sporozoite stage specific genes[43] was analysed and the core structure shown to be maintained. However, two striking characteristics of the structure of the PF14_0633 AP2 domain during interaction with double-stranded (ds) DNA are that it dimerizes through a three-dimensional domain-swapping mechanism in which the α-helix of one protomer is packed against the β-sheet of its dimer mate and that the dimeric structure aligns the single native cysteine residues within proximity to form a disulfide bond [Fig 5][38]. The plant AP2/ERF domain is found either as a single module or in tandem arrangements: single-domain AP2 genes are usually involved in environmental stress responses including thermotolerance, dehydration and ethylene response, whereas proteins with tandem domains are implicated in development[44]. The Apicomplexan ApiAP2 subfamily contains usually a single domain, but proteins with 2–4 AP2 domains have also been described[20]. Nevertheless, the biological meaning of this architecture is unknown. Outside of the conserved AP2 domain, sequence homology between ApiAP2 proteins from different *Plasmodium* species is low, making the identification of additional, functionally relevant domains (e.g. transcriptional activation domains) difficult. Indeed, the average pairwise distance within the ApiAP2 family is much greater than the average pairwise distance within the plant AP2 domain family (2.6 versus 1.2; measured using the JTT score matrix)[45]. The size of ApiAP2 proteins is also highly variable ranging from hundreds to thousands of amino acids (e.g. PFI1665w – 200aa; PF13_0235 – 3858aa [Table 3 – Appendix 2]. A global yeast two-hybrid study suggested that ApiAP2 proteins in *Plasmodium* interact with each other and with the chromatin remodelling factor histone acetyltransferase GCN5[46]. Binding to chromatin remodelling factors may serve to recruit these complexes to specific chromosomal locations and facilitate interaction with the core transcription machinery[47]. In addition, Apicomplexan AP2-domains possess affinities for a much more diverse set of DNA sequences (i.e TAGCTA for PBANKA_090590[37] and TGCATGCA for PF14_0633[38]) than in plants where an overwhelming majority of AP2/ERF proteins bind the canonical GCC box (A/GCCGCC)[48].

In the following experiments I aimed to define the specific expression patterns of AP2 factors in the *P. berghei* rodent malaria model and identify which factors are prone to post-transcriptional gene regulation at the time of transmission. These AP2's were chosen based on published microarray data obtained from DOZI KO parasites indicating that these molecules are regulated through translational repression. In order to understand the molecular functions of putatively, translationally silenced ApiAP2 TFs during parasite transmission, mis-timed expression was attempted to be induced by genetic modification. Using a combination of bioinformatics, molecular and genetic tools the first goal was to confirm the gene models of 5 selected AP2 genes that were GFP-tagged. In addition, this phenomenon was checked by imaging techniques (FISH) and RT-PCR using cDNA prepared from RNA that co-precipitated with CITH::GFP. Once confirmed by sequencing the vectors were transfected into a *P. berghei ANKA* wild type strain and mutants drug selected and genotyped by PCR. Note that the constructs contained the 3' UTR from the *dhfr* gene, which is not regulated by translational repression. Hence, the integration by single cross-over, replaced the endogenous 3' UTR of the target gene thus supposedly interfering with the silencing phenomenon. Mutant lines obtained

were cloned by limiting dilution. The activation of the mutant factors was followed through the analysis of the GFP expression.

In short by forcing early expression of these factors (without the endogenous 3' UTR the silencing may be bypassed) we hoped to induce precocious expression of the downstream targets of these transcription factors. Should some of these downstream targets be translated during the blood stage of the infection and able to elicit an immune response from the host (i.e. surface proteins), we would have a new candidate transmission blocking vaccine.

2- Materials and Methods

2.1 – Bioinformatics. Genedb (www.genedb.org) gene accession numbers for AP2-1 (a.k.a. AP2-O), AP2-2, AP2-3, AP2-4 and AP2-5 are PBANKA_090590, PBANKA_140370, PBANKA_093910, PBANKA_010290, PBANKA_131320, respectively. The reference numbers of ESTs used to check the AP2-3 gene model are DC219303, DC233365 and DC223368 and were from www.ncbi.nlm.nih.gov.

2.2 – Predicted subcellular localization and NLS predictions. All gene models were analysed for subcellular localization using PSORT (Prediction of Protein Localization Sites; version 6.4). Nuclear Localization Signals (NLS) were identified using software from CUBIC (Columbia University Bioinformatics Center[49]). ClustalW2 software, provided by EMBL-EBI, was used to perform multiple sequence alignments, and to identify aligned basic stretches of amino acids near their N-terminus (putative NLS sequences).

2.3 – Mice infection and parasite growth. Females BALB/c mice breed in Charles River Company with 6-8 weeks of age were infected with *P. berghei* parasites (WT or mutant) from -80ºC frozen vials by intra-peritoneal injection of 10^7 iRBCs. Two days post-infection and during 4-6 days, parasitaemias were checked in a 10% Giemsa stained thin blood smear made from 1 drop (~2-3μL) of tail blood (%= nº of iRBC/total nº RBC, per field).

2.4 – Parasite lines. The following parasite lines were used in these experiments: Wild type GFP expressor line: *P.berghei ANKA GFPcon* 259cl2 [50] referred as WT-GFP in this manuscript; wild type *P.berghei ANKA* cl15cy[16] (non-fluorescent line) and CITH::GFP parasites *P. berghei ANKA* 909cl1[36].

2.5 – Confirmation of AP2 gene models by RT-PCR. All five gene models (AP2-1 to AP2-5) were analysed by RT-PCR using gametocyte cDNA. Gametocytes were purified from infected blood through a 49% Nycodenz/PBS-solution (v/v) gradient. cDNA was generated using 1 μg of total RNA (extracted with the TRIzol method, see Appendix I) with SuperScript II according to Appendix I indications. Oligonucleotides used for PCR amplification were designed based on Expressed Sequence Tags (EST) and predicted gene models from www.genedb.org and www.plasmodb.org. PCR targeted especially the 3' ends of the predicted ORF's beyond the stop codon, as it is the site for C-terminal GFP-tag fusion. As far as AP2-4 (PBANKA_010290) is concerned, there were no ESTs available for *P. berghei* and PlasmoDB's prediction for this gene model was incomplete by the time this experiment was performed. Therefore, this model was analyzed based on a contig assembled based on an EST's multi-alignment of different orthologs for this gene found in *P. falciparum*, *P. chabaudi* and *P. yoelii*. RT-PCR primers used are listed in Table 4, Appendix II). In each 50 μL reaction 50ng of cDNA was used with a PCR mix described in Appendix I. The cycling program used was 94°C **-** 5min initial denaturation, followed by 40 cycles of 94°C **-** 30s, 43°C **-** 1min, 62°C **-** 90s. A final round of 62°C **-**

10min allowed complete elongation of PCR products. The extension temperature used was 62°C due to the low GC-content of the *P. berghei* genome. RT(-) and internal controls (PCR across introns) were used to ensure that cDNA samples were free of genomic DNA. Clone Manager 9 (Professional Edition) was used as a bioinformatic tool for all sequence manipulations as well as primer design.

2.6 – Comparative *ap2* expression profile analyses by RT-PCR. In order to define the transcription profiles of each of the 24 *P. berghei* AP2 transcripts in transmission stage parasites, RT-PCR's were performed on stage-specific cDNA generated from gametocyte and salivary gland sporozoite cDNA. cDNA was generated as described in section 2.5 from RNA isolated through the TRIzol method (see Appendix I). The PCR mix used is described in Appendix I and the oligonucleotides are listed in Table 5, Appendix II. The cycling program used to amplify 50ng of cDNA was 94°C - 5 min for initial denaturation, followed by 40 rounds of 94°C - 30s, 55°C – 30s, 62°C - 30s. A final round of 62°C - 10 min allowed complete elongation of PCR products. The extension temperature used was 62°C due to the low GC-content of *P. berghei* DNA. RT (-) and internal controls (PCR across introns) were used to ensure that cDNA samples were free of genomic DNA. Clone Manager 9 (Professional Edition) was used as a bioinformatic tool for all sequence manipulations as well as primer design.

2.7 – Translational repression of *ap2's* transcribed in blood stage gametocytes. To test whether specific AP2's are maintained in CITH::GFP mRNP's, all AP2 genes shown to be actively transcribed in gametocyte stage (i.e. positive hits in the experiment described in section 2.6) were amplified by PCR from cDNA prepared from CITH::GFP immunoprecipitation experiments: input, α-GFP (specific for the CITH::GFP fusion protein), α-c-myc (control IP), protein G beads, available in the lab from previous experiments from Dr. Gunnar Mair. RT-PCR conditions followed the same scheme as described in section 2.6.

2.8 - Fluorescent in situ hybridisation (FISH) and DOZI::GFP immunofluorescence. Blood smears were prepared from BALB/c mice infected with *P. berghei* ANKA, air-dried and fixed for 15 minutes in a 4% PFA, 5% acetic acid solution followed by 1x PBS washing 3x 5 min and dehydration through a graded ethanol series (70-90-100%, 3min each); these slides were stored at -20°C. DNA probes for all the five AP2 targets (PBANKA_090590, PBANKA_140370, PBANKA_093910, PBANKA_010290, PBANKA_131320) were prepared by nick translation in the presence of the fluorofore Atto590, dUTP-rhodamine equivalent, according to [27] procedures described by Jena Bioscience for Atto590 NT Labelling (PP-305L-590). 1 μg of purified PCR product was used as template (primers used are listed in table 6, Appendix II) in the mix and incubated for 90 min at 15ºC in an Eppendorf mixer. To block unspecific binding 50 μg of fish sperm DNA (D7290 Sigma Aldrich) and 50 μg of yeast tRNA (R8759 - Sigma Aldrich) was added to the probe before purification by precipitation in 100% ethanol/ 3M sodium acetate (pH=5.6). Before hybridization, slides were treated 5 minutes with 0.1% pepsin in 10mM HCl at 37°C, washed with PBS and dehydrated through an ethanol series. Both slides and probes (2 ng/μl) were denaturated for 7 minutes at 80°C. Hybridization was performed overnight using 20 ng of probe, upside down, in a moist chamber at 37°C. Subsequently,

slides were quickly rinsed in 0.1% TWEEN-20 in 2XSSC to remove the cover slip, and washed twice in pre-warmed 50% formamide in 2X SSC (pH=7) for 15 minutes at 37°C and finally washed in TNT buffer (0.1M Tris, 0.15M NaCl, 0.05% TWEEN-20 [pH7.5]) twice for 3 minutes. For DOZI::GFP detection, slides were then incubated with ROCHE monoclonal mouse anti-GFP (Cat.No. 11814460001), diluted 1/400, in 1% BSA, for 1 hour at 37°C. This was followed by 3x15min washes in 1xPBS and incubation with secondary Jackson immune research donkey α mouse 488 (715-486-150), diluted 1/1000 in 1% BSA, for 1 hour at 37°C. Again, slides were washed 3x15min in 1XPBS. Finally, slides were incubated with 5 μg/mL of DAPI nuclear stain, washed in 1x PBS for 5 min and mounted in Fluoromount G medium. Slides were analyzed on a Zeiss LSM 510 META microscope, a confocal point-scanning microscope and images were processed using ImageJ software v.1.38a.

2.9 – Generation of transfection plasmids. Primers were designed based on the results of the confirmed gene models and included restriction sites for cloning into the tagging plasmid; PCR products for each ORF were approximately 1kb upstream (but not including) the stop codon and designed to contain a single restriction enzyme cutter within the PCR amplicon and flanking restriction sites to be used in the cloning steps. All primers are listed in table 7, Appendix II; restriction sites are highlighted. Single cutters for AP-1 and AP-2 were *Eco*RI (Fermentas-ER0271), *Bsm*I (NEB-R0134), respectively and *Cla*I (NEB-R0197) for AP-3, AP-4 and AP-5. In each reaction, 10 ng of *P. bergh*ei genomic DNA (isolated according to phenol-chloroform extraction method, further described in Appendix I) was amplified using the PCR mix described in Appendix I, according to the following cycling program: 94°C - 5min for initial denaturation, followed by 25 rounds of 94°C - 30s, 43°C - 1min, 62°C – 1min. The cycling program ended with a round of 62°C - 10min to allow complete elongation of the PCR products.

AP2-5 lacks a single cutter within the target sequence needed for linearization of the final targeting plasmid. Therefore, a *Cla*I-site was inserted by a overlapping PCR mutagenesis protocol. Four primers were designed: primers 140 and 142 as forward primers; and 141 and 143 as reverse ones. Primers 142 and 143 contain an overlapping region with *Cla*I restriction sequence. Primer sequences are listed in Table 7 (Appendix 2). First, PCR 1 [140|141] and PCR 2 [142|143] were performed using the following cycling program 94°C - 5min, followed by 25 rounds of 94°C - 30s, 43°C – 30s, 62°C – 30s, ending with a round of 62°C - 10min. Then a third PCR was run using 50 ng of PCRs 1 and 2 (apart from the 10ng of gDNA) that worked as primers during the first five cycles; after cycle no. 5 outside primers 140 and 143 were added. In this PCR extension time lasted for 150 seconds instead of the 30s used for PCRs 1 and 2.

All PCR products were purified using the ROCHE PCR purification kit. PCR products (inserts) and generic GFP tagging plasmid (backbone) were digested with the same or compatible restriction enzymes. For AP2-O elimination of the *Eco*RI-site was needed in order to create a single cutter within the insert and permit linearization of the tagging vector; this is a required feature for homologous recombination. Thus, this was achieved by cutting the plasmid and the insert with enzymes with different restriction enzymes that produce compatible ends; in this case *Mun*I for the insert and *Eco*RI for the plasmid. After ligation both restriction sites were lost. The AP2-5 insert was cut with *SwaI* and

*Not*I, while the vector was cut with *Eco*RV and *Not*I. *Eco*RV and *Swa*I generate blunt ends which after ligation are lost.

Apart from the GFP gene, all tagging plasmids contain the ampicillin resistance marker to allow bacterial amplification of the construct and the *Toxoplasma dhfr* gene that allows selection of mutant *P. berghei* parasites with pyrimethamine (7 mg/ml in drinking water). PCR products were cloned in frame with eGFPM3 to produce C-terminal AP2-fusion genes.

Vector-insert ligations were carried out using 1 unit of Fermentas T4 ligase, 1x Fermentas ligase buffer and 300ng of linearized plasmid and 120ng of insert in a molar ratio insert to backbone of 3:1; reactions were carried out for 2 hours at room temperature. Chemically-competent cells used were *Escherichia coli* DH5α prepared according to procedures described in appendix I. After ligation, 100μL of freshly thawed cells, previously stored at -80ºC, were added and incubated on ice for 20 minutes. Heat shock was performed for 45 seconds at 42°C followed by two minutes on ice. Subsequently, cells were grown in Sigma LB medium without ampicillin at 37° for 45 minutes on a shaker and plated on Sigma LB agar plates supplemented with ampicillin (100 μg/mL). Plasmid DNA was extracted from single colonies grown overnight by boiling mini-prep method (see Appendix I, for details); correct ligation of the insert into the plasmid was confirmed by Restriction Map analysis using a combination of gene-specific single cutters and plasmid backbone enzymes, as well as sequencing (Stabvida) of final plasmids with oligonucleotide 1753 (5'-CCGTATGTTGCATCACCTTCACCC-3').

Genomic DNA, plasmid and insert concentrations were determined with a Nanodrop®-1000 Spectrophotometer and confirmed by running 50 ng on a 1% agarose gel.

2.10 – *Plasmodium berghei* transfection. Once linearized using the single cutter, plasmids were transfected into *P. berghei* parasites; mutant parasites were amplified in inbred mice. Transfection procedures were performed according to published data[51]. Briefly, parasitaemia of infected BALB/c mice was followed to reach 1-3% at which point the mouse was sacrificed by CO_2 inhalation. Blood was collected through cardiac puncture using a heparinised syringe and an overnight culture of schizonts (the blood-stage replicative form) was prepared. On the following day mature schizonts were purified on a 55% Nycodenz/PBS-solution (v/v) gradient and 10^6 parasites electroporated with 10 μg of linearized plasmid using AMAXA technology (program U33). 5 independent sets of electroporated parasites were injected into 5 different naïve mice. 24h post infection mice were treated with pyrimethamine (7 mg/mL in drinking water) to select mutant parasites. Parasitaemia was monitored daily for 7 days by counting of GIEMSA-stained thin, blood smears prepared according to section 2.3. Parasitaemias were counted and parasite morphology was analyzed.

2.11 – Genotyping of the transfected parasite lines and cloning by limiting dilution. In successful mice infections (with parasitaemias reaching ~5-10%), mice were sacrificed by CO_2 inhalation and blood was collected through cardiac puncture using a heparinised syringe. RBCs were lysed using RBC lysis buffer (0.15M NH4Cl, 0.01M KHCO3, 1mM Na_2EDTA; pH 7.4). Parasite genomic DNA was prepared by proteinase K digestion followed by purification with phenol-chloroform (see genomic DNA preparation in Appendix I). Parasites were genotyped to check integration of the tagged gene into the

parasite's genome. Five PCRs were performed in each parasite line that targeted: the 5' integration site [58|182], the 3' integration site [87|408], the *Toxoplasma dhfr* resistance marker [250|251], the RNA polymerase II gene as a positive control PCR for both wild type and mutant parasite populations [84|85], and wild type gene (negative control) [182|100]. Primers are listed in table 8, in Appendix II and PCR reaction mixes were prepared as described in Appendix I. The cycling program was: 94°C - 5min for initial denaturation, followed by 25 rounds of 94°C - 30s, 43°C – 30s, 62°C – 90s. The cycling program ended with a round of 62°C - 10min to allow complete elongation of the PCR products.

Limiting dilution was finally used to clone transgenic parasite populations. To this end, a BALB/c mouse was infected with the mutant line. When parasitaemia reached 0.3-1%, the mouse was sacrificed by CO_2 inhalation and the blood was collected by cardiac puncture. Then, using a Neubauer chamber to count the total number of RBC, blood was diluted in 1XPBS by successive dilution and 0.7 parasites were injected i.v. into 10 naïve mice. One week p.i. parasitemia was checked for 5 days by counting GIEMSA-stained blood smears. In order to select the mutant clones, the successful infections parasites were genotyped by PCR using three different sets of primers: 423|444 as a positive control [uis4 gene - PBANKA_050120], 182|100 to amplify the WT gene and 87|408 to amplify the mutated gene. PCR conditions followed the above-mentioned indications.

2.12 – Mutant parasite growth kinetic analyses. In order to compare the growth kinetics of mutant AP2::GFP with wild type parasites two sets of 6 BALBc mice each were infected with 10^7 parasites. Parasitaemias were followed for 5 days by counting GIEMSA-stained blood smears; in each smear 10 fields with total RBC numbers of 500-900 were counted. The resulting kinetics were analyzed in an un-paired T-test (H0: mutant= WT). Blood stage morphologies were compared and imaged using an upright brightfield microscope Leica DM2500 equipped with a colour camera. Pictures were treated using ImageJ software v.1.38a.

2.13 – Transcription of the *ap2-o::gfp* transgene. After experiments described in section 2.11 the infected blood from three mice of each set was collected and RNA was extracted on the same day in order to obtain equivalent cDNAs from both mutant and WT mixed blood stage parasites. RNA and cDNA were prepared according to procedures described in appendix I. The mutant transcript's presence was checked in RT-PCR using 50 ng of cDNA and primers listed in table 9, Appendix II. The cycling program was: 94°C - 5min for initial denaturation, followed by 35 rounds of 94°C - 30s, 43°C – 30s, 62°C – 35s. The cycling program ended with a round of 62°C - 10min to allow complete elongation of the PCR products. An internal control (PCR across introns) was used to ensure that cDNA samples were free of genomic DNA.

2.14 – GFP detection in immunofluorescence assay (IFA) in the mutant AP2-O::GFP cloned line. Mutant parasites were analysed live and by IFA in PFA–fixed slides. Thin blood smears for IFAs were prepared from infected blood with mutant AP2-O::GFP parasites and fixed for 15 minutes in a 4% PFA solution. Fixation was followed by 1XPBS washing (2x10min). Subsequently, slides were washed in 0.1M Glycine/PBS solution for 10 minutes and permeabilized for 10 minutes in 0.1% Triton-X. After

that, slides were blocked in 1% BSA/PBS (fresh), for 20 minutes at room temperature. Then, for AP2-O::GFP detection, slides were incubated with ROCHE monoclonal mouse anti-GFP (cat. no. 11814460001), diluted $^{1}/_{400}$, in 1% BSA/PBS (fresh), for 1 hour at 37°C. This was followed by 3x 15min washes in 1XPBS and incubation with secondary Jackson immune research donkey α-mouse 488 (715-486-150), diluted $^{1}/_{1000}$ in 1% BSA/PBS (fresh), for 1 hour at 37ºC. Again slides were washed 3x 15min in 1XPBS. Finally, slides were incubated with 5 μg/mL of D9564 DAPI nuclear stain, washed in 1XPBS for 5 min and mounted in Fluoromount-G medium. During incubations slides were kept upside down, in the dark. Slides were analyzed on a Zeiss Axiovert 200M microscope, a widefield microscope.

Fresh blood was also checked for GFP fluorescence. Few drops of blood (less than 20μL) were collected from the mouse tail, incubated for 2min with 0.4% Hoechst (bisBenzimide H 33342 trihydrochloride) at room temperature and then washed in 1mL of 1x PBS. Mutant parasites were checked immediately for GFP fluorescence on a Zeiss Axiovert 200M microscope. All images were processed using ImageJ software v.1.38a.

2.15 – Live GFP imaging in synchronised culture of schizonts. Experiments from section 2.14 suggested the presence of fluorescence in the schizont stage. Hence an overnight culture of schizonts was prepared in order to image specifically this stage. Synchronization was performed according to published data[51]. Briefly, a BALB/c mouse was infected with the AP2-O::GFP mutant line and when a parasitaemia of 1-3% was reached the mouse was sacrificed by CO2 inhalation and blood was collected by cardiac puncture between 1 to 3 pm. Then blood was washed and resuspended in complete RPMI1640 medium, flushed with a gas mixture of 5% CO2, 5% O2, 90% N2 using a 0.2-μm filter unit connected to the gas hose and incubated overnight at 36.5ºC with shaking (50rpm).

On the following morning the culture was sampled; sample (~100μL) was incubated for 2minutes with 0.4% Hoechst (bisBenzimide H 33342 trihydrochloride), at room temperature and then washed in 1mL of 1x PBS. Parasites were checked immediately for native GFP fluorescence on a Zeiss LSM 510 META microscope, a confocal point-scanning microscope and images were processed using ImageJ software v.1.38a.

2.16 – Western blot analysis of AP2-O::GFP blood stage protein extract. In order to detect the presence of AP2-O::GFP protein, by a second approach, an αGFP Western blott was performed. Protein extract was prepared from blood-stage parasites after mouse sacrifice with parasitaemia ~15%. After blood collection by cardiac puncture, blood was lysed in RBC lysis buffer (0.15M NH_4Cl, 0.01M $KHCO_3$, 1 mM Na_2EDTA; pH 7.4) and resuspended in 2x SDS buffer (10% w/v SDS, 20 % v/v Glycerol 0.2 M Tris-HCl, pH 6.8 0.05% w/v Bromophenol blue). Prior to the run, 100mM of DTT was added to the samples (WT::GFP, WT non GFP, AP2-O::GFP), followed by boiling for 10 minutes. Samples were run in a 4% stacking/10% resolving NF-acrylamide/bis- acrylamide solution 40 % (29:1) gel in TGS buffer (25 mM Tris, 192 mM glycine and 0.1% SDS, pH ~ 8.6.) for 150 minutes at 80V. Next, the gel was washed in transfer buffer (25 mM Tris, 192 mM glycine and 20% methanol) for 10 minutes and then transferred onto a nitrocellulose membrane at 200mA (constant amperage) for 110

minutes. Then a quick wash in Ponceau Red confirmed the transfer was done properly. The membrane was then blocked in 5% milk/(Tween-PBS(0,05%v/v)) for one hour at room temperature and incubated for another hour with the primary antibody, ROCHE monoclonal rabbit anti-GFP (Cat. No. 11814460001), diluted 1/3000 in blocking solution. After, the membrane was washed twice in 0.05% Tween-PBS(v/v) for 15 minutes and then incubated one hour with the secondary antibody anti-rabbit HRP (Jackson immune research, 211-032-171) diluted 1/5000 in blocking solution. A final wash in 0.05% Tween-PBS(v/v) 2x 15 minutes was done before incubation with Millipore ECL system (cat no WBKLS0050) for 5 minutes. Fuji medical x-ray film was exposed for one hour and revealed.

2.17 – Gametocyte fitness and mosquito infection with wild type and AP2-O::GFP parasites.

Male gametocyte fitness of both ap2-o::GFP mutant and wild type parasites was checked through an ex-flagellation assay. On day 4-5 p.i. a small drop of blood was collected from the tail of an infected mouse onto a slide, and exflagellation triggered by warm blow of air. Release of motile, male gametes from the male gametocyte was quantified 8-12 minutes later.

After observing exflagellation for both wild type and ap2-o::GFP transgenic parasites, *Anopheles stephensi* female mosquitos were allowed to blood feed for 30 minutes on BALB/c mice infected with these parasites. These mosquitoes were provided by Fernanda Baptista and Ana Parreira from Dr. Maria Mota's group and starved for 24h before the blood meal. Mice were anesthetized with ketamine/xylazine (12% ketamine, 0.16% xylazine, in PBS; intraperitoneal injection, 10mL/Kg) and then allowed to recover before returning to the mice house. Then mosquitoes were starved for 48h in order to select only the ones that blood fed on the mice and kept at 21ºC for 18 days.

On day 10 p.i. mosquito infection was checked by quantifying the number of oocysts present in the mosquitos' midguts. In total 15 mosquitoes from each batch (WT and ap2-o::GFP) were dissected; 9 in 2% mercury-chrome that stains the oocysts in orange hence allowing direct counting, and 6 in 1x PBS to check the number of oocysts by looking at the presence of native GFP (if present).

Midguts were imaged in a Leica DM2500B widefield microscope and images were processed using ImageJ software v.1.38a.

2.18 – Completion of the malaria life cycle – analysis of sporozoite infectivity. Sporozoite infectivity was analysed by passing the infection from mosquitoes back to BALB/c mice. In this experiment both the ability of ap2-o::GFP parasites to leave the salivary gland as well as to infect the mouse was checked independently. To this end two sets of mice were used; the first (n=3) was anesthetized with ketamine/xylazine as described in section 2.17 and 10 infected mosquitoes (both ap2-o::GFP and WT) were allowed to blood feed on them. The second set of mice (n=4) was infected by intravenous injection in the ocular vein with 20,000 hand dissected sporozoites. Salivary glands dissection was done after anesthetizing mosquitoes by cold (10 minutes at -20ºC) using a Leica scope. On day 2 p.i. and for 1 week parasitaemias were monitored.

2.19 – Live imaging of ookinete stage – ookinete culture. In order to check the ookinete stage for morphologic alterations and native GFP fluorescence an overnight *in vitro* culture of ookinetes was prepared for both the WT line and the mutant ap2-o::GFP. Exflagellation was checked as described in section 2.16 in a young BALB/c mouse infection (i.e. no more than 5 days old) to evaluate gametocyte fitness. Provided that at least one exflagellation centre is observed per field under the microscope (1000x magnification), the mouse was sacrificed by CO_2 inhalation and blood collected by cardiac puncture. Infected blood was washed in 1x PBS and WBC depleted using a Whatman CF11 cellulose column, previously equilibrated in 1xPBS. The collected blood was centrifuged for 10 minutes at 350 rcf, 4ºC and resuspended in ookinete 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 $NaHCO_3$; filter-sterilized; final pH 7.4. 20% FBS (sigma F2442) was added just before to add the purified blood. This stage is highly sensitive to pH/temperature/medium components; therefore the culture medium is critical, if necessary temperature and pH can be corrected before adding the blood. The culture was incubated for 16h at 19ºC shaking at 50 rpm. On the following day cultures were checked for the presence of ookinetes using live microscopy, as described for schizonts live imaging on section 2.14., and in GIEMSA stained smears.

3- Results

3.1- Subcellular localization and NLS predictions – AP2-2 contains six putatives NLS

PSort software was used to predict the most likely sub-cellular localization for each AP2 molecule. Nuclear localization were returned for AP2-1 (at 30% certainty), AP2-3 (30%), and AP2-4 (54%), while high values were predicted for AP2-2 and AP2-5, 94.8% and 88%, respectively. These 2 molecules were additionally screened for potential nuclear localization signals using CUBIC software: only AP2-2 was predicted to have 6 motives [Table 2]. Regardless, all 5 molecules possess short stretches of basic amino acids near the N-terminus which may function as non-canonical NLS's.

Table 2| Results from CUBIC software analysis for AP2-2. Six motives were found in AP2-2; the number gives position of the motif; all NLS's are highlighted in blue.

CUBIC software analysis	
Input Sequence	MIKEGDIKQKEQINNDQFKHKCCYNYALSNNYFIYPKDNSLFINGIFNNINNNV GPYYGNIFYKDKMFFSGRSGGL**KRKKKKKDERVINTCSAKRLEFFYPKKKR** **R**QRVGLIQNSRKNIVYDNVLKRFLVYYYKQGIQVFRSFSCK**KKKKFESARNK** **AIILSKQYNKKYTRK**RNLEKDNSKDHLIINNSTNLVVRHDYDIRKKVKIVPDKN KSGYRGVFYDSSYHAYICTYNEAGIRKYQIFKIKNNDYLEAYNLAVMCRRYK LFKNYQFVSQRNRVRSGRIHLK
NLS's found	KRKKKKK - 76 RKKKKKD - 77 RKKKKKD - 77 RKKKKKDERVINTCSAKRLEFFYPKKK - 77 KKKRR - 101 KKKKFESARNKAIILSKQYNKKYTRK -146

3.2 - Confirmation of *Plasmodium* gene models by RT-PCR–An extra intron was found in AP2-3

All gene models predicted at PlasmoDB (version 4.4) were verified by RT-PCR and compared to PCR amplicons generated using genomic DNA. Gene models for AP2-1, AP2-2, AP2-4 and AP2-5 were in agreement with RT-PCR results and neither of them revealed additional introns. As for AP2-3, an alignment generated between the annotated genomic locus, the corresponding cDNA sequence (both from Plasmodb) and a contig manually assembled based on ESTs (retrieved from www.ncbi.nlm.nih.gov) demonstrated clear differences [Fig.6A]. RT-PCR results validated our EST contig data [Fig. 6B], revealing the presence of a second intron at the 3' end of the open reading frame that was omitted in the PlasmoDB gene. Our conclusion was based on the PCR product amplified from cDNA using primer combination 117|120 which is smaller in size when compared to the genomic one. In addition, reverse primer 119, was only able to amplify the cDNA as it spans exons 2 and 3. On the other hand, primer 118 was only able to amplify genomic DNA since it extends from exon 3 up to intron 2. Sequencing results from TOPO-TA cloning of the amplicon 117|120 also confirmed the presence of the second intron in the AP2-3 gene [Fig.6c].

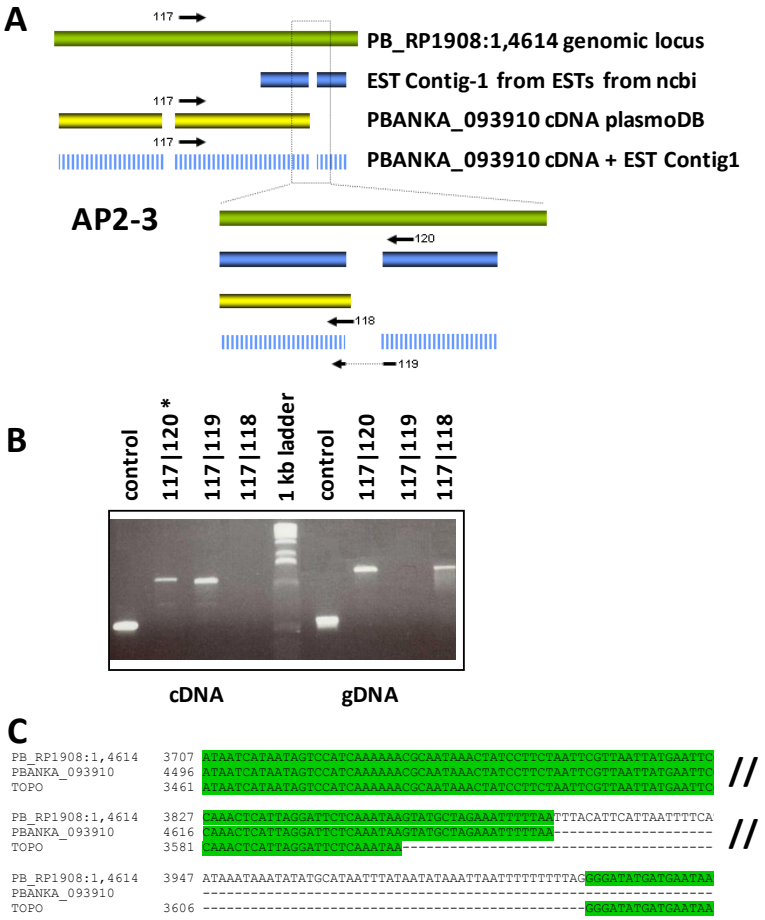

Fig. 6|Predicted and actual exon-intron organization of *ap2-3*. A Schematic representation of available, conflicting sequence data for the AP2-3 gene indicating the position of primers used for PCR. **B** PCR on cDNA and genomic DNA show that an additional intron is present within the 3' end of the predicted ORF. ***** denotes PCR product 117|120 using cDNA as template; this was TOPO-cloned and sequence analyzed. **C** Sequence alignment of *P. berghei* contig PB_RP1908:1,4614, predicted ORF from PBANKA_093910 and sequenced PCR product 117|120.

3.3 – Comparative AP2 expression profile analysis by RT-PCR – 15 out of 24 AP2 molecules are highly expressed in gametocyte stage

Primers were designed for 24 *P. berghei* AP2 genes and the presence of the corresponding transcripts in transmission stages was tested by RT-PCR using cDNA generated from mixed gametocytes and 19 to 24-day old sporozoites. According to the RT-PCR results, 15 out of 24 molecules are highly expressed in gametocytes [Fig. 7a], while in sporozoites 11 AP2 transcripts were detected [Fig. 7b]. 5 AP2 molecules were transcribed in neither stage. Interestingly only 7 genes have overlapping expression patterns, while 11 are clearly transcribed in a stage-specific fashion. 9 are exclusively detected in gametocytes and 4 are specific for sporozoites. Together, these results suggest that AP2 transcription is tightly regulated itself, in a developmental, stage-specific manner. All 5 AP2's found to be down-regulated in DOZI gene deletion mutants were strongly expressed in gametocyte cDNA. As expected, *ap2-sp* (PBANKA_132980)[43] is highly expressed in sporozoites but also in gametocytes. Interestingly, out of the 3 AP2 genes that were shown to be differentially expressed in mutant salivary gland sporozoites (i.e. Puf2 gene deletion parasites), only the down-regulated gene (PBANKA_052170) is present in the sporozoite cDNA sample, while the two up-

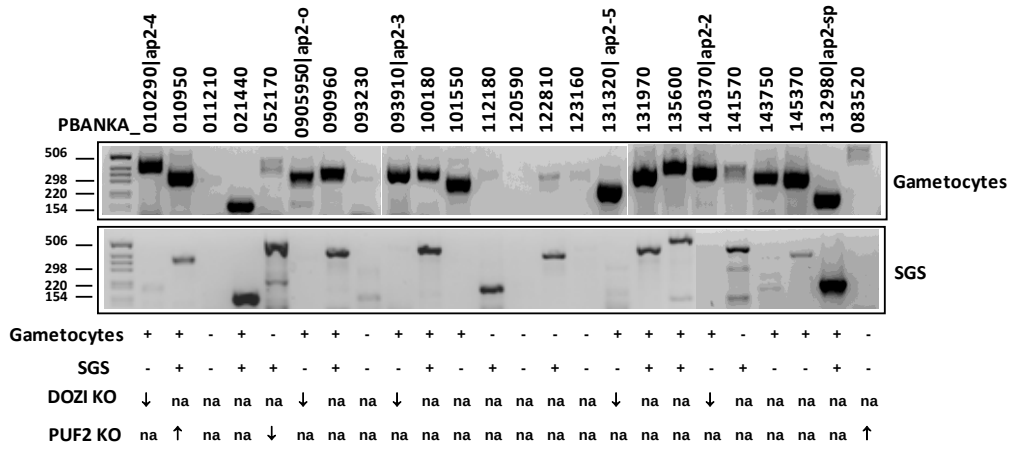

Fig. 7|Comparative RT-PCR analysis of *ap2* expression in malaria transmission stages. RT-PCR was performed using cDNA generated from Nycodenz-gradient purified mixed (male/female) gametocytes from the CITH::GFP line (upper panel), or from hand-dissected, wild type salivary gland sporozoites (SGS; lower panel). Below: top 2 lines depict positive and negative expression hits schematically. The DOZI KO and PUF2 KO lines indicate with arrows the observed de-regulation profile in DOZI and PUF2 KO mutants. **↑-**upregulated in the KO**;↓ -** downregulated in the KO**; na-** not altered in the KO.

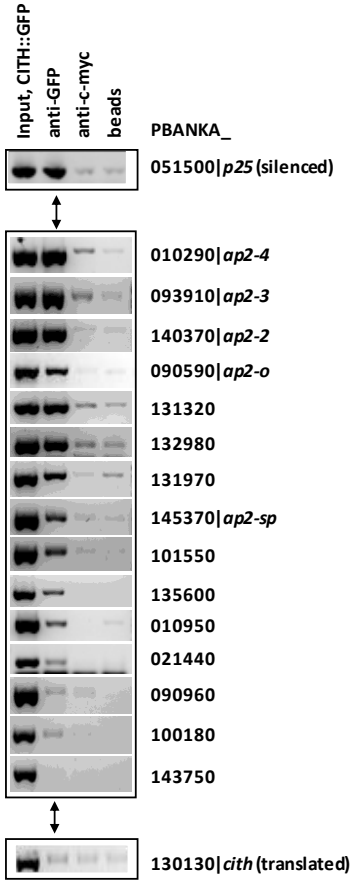

Fig. 8|RT-PCR analysis of *ap2* transcript co-immunoprecipitation with CITH::GFP. RT-PCR was done with cDNA generated from Nycodenz-gradient purified mixed (male/female) gametocytes from the CITH::GFP clonal line; 15 out of 24 tested molecules are potentially associated with P-granules. *ap2* enrichment levels are compared to known silenced (*p25*) and translated (*cith*) mRNAs. PBANKA_ together with number (e.g. 051500) denotes genedb accession number.

regulated genes (PBANKA_010950 and PBANKA_083520) are absent [Table 1]. This suggests that the transcription of these 2 genes is indeed initiated during the transformation of sporozoites into early EEF stages and may contribute to the global transcriptional changes observed in these 2 mutants.

3.4 – RT-PCR results suggest that 9 AP2 molecules are translationally repressed

Prompted by the large number of transcribed AP2 genes in the gametocyte we wanted to test which are candidates for translational silencing. AP2-O was recently found to be important for the transcription of ookinete-specific genes; the small number of mRNAs identified in KO mutants[37] suggested that additional AP2's contribute to the transcription of ookinete-specific genes. We therefore analysed which of the 15 gametocyte-transcribed AP2 genes associated with CITH::GFP (a component of the female gametocyte silencing complex); such an association would indicate that these transcripts are translationally regulated in gametocytes awaiting translation in the newly formed zygote in a manner similar to *p25* and *p28*. RT-PCR was used with template cDNA reverse transcribed from mRNA extracted from CITH::GFP immunoprecipitation experiments. Transcript enrichment in specific α-GFP (CITH::GFP) IP's was compared to input material, control IP's were performed with α-c-myc antibodies and beads only. Out of the 15 genes tested 9 to 11 AP2 genes were enirched, suggesting their presence in a CITH-defined mRNP. These results suggest that the mRNAs are regulated by translational silencing in the female gametocyte. Interestingly, the 5 targeted AP2 molecules identified previously to be down regulated in the DOZI KO parasite were strongly enriched. As a positive control we used *p25*, a known silenced gene that co-IPs with DOZI and CITH; as negative control the known translated gene *cith* was used [Fig.8].

3.5 – FISH; AP2-O co-localizes with DOZI

Microarray data [27] and above RT-PCR's suggested that several AP2's were being kept in female-specific P-granules. Therefore thin blood smears were analysed by FISH to identify whether co-localisation of specific mRNAs could be found together with P granule components DOZI. Rhodamine-labelled AP2 DNA probes were hybridized to DOZI::GFP blood smears; this line allows easy detection of gametocyte P-granules [Fig. 9]. Two controls were included (panels A and C); in the top panel a TAMRA poly d(T) 50mer was used as probe to label total poly-adenylated mRNA. The probe's signal was found in all parasites, clearly punctuated in gametocytes. On the lower panel a labelled probe for *p28* (known to be silenced in P-granules) was used to show widespread co-localization with the P granule component CITH. Out of 5 AP2 probes prepared, only *ap2-o* probe generated the pattern given by the controls (panel B). Altogether, evidence supports that *ap2-o* transcript is targeted to P-granules in the *P. berghei* female gametocyte.

3.6 – 5 AP2 transfection plasmids were generated

PCR products for each ORF of approximately 1 kb immediately upstream of the stop codon were generated based on the confirmation results described previously (see chapter 3.2). PCR primers were designed to include restriction sites flanking the amplicons to allow ligation into the GFP-tagging vector plasmid; a single restriction enzyme cutter was ensured to be present for plasmid

linearization prior to transfection. From the 5 genes amplified, only AP2-5 lacked the single cutter in its sequence; a *Cla*I restriction site was therefore inserted by overlapping PCR and confirmed by digestion of the final PCR product with *Cla*I. The single cutter enzyme was *Eco*RI for AP2-1, *Bsm*I for AP-2 while, *Cla*I was used for AP2-3, AP2-4 and AP2-5. All PCR products were cloned into a generic GFP-tagging plasmid and final constructs confirmed by restriction map analysis and sequencing of two independent clones. Prior to transfection clones were linearized with the appropriate single cutter and stored in a 0.3M sodium acetate/70% ethanol solution at -20°C.

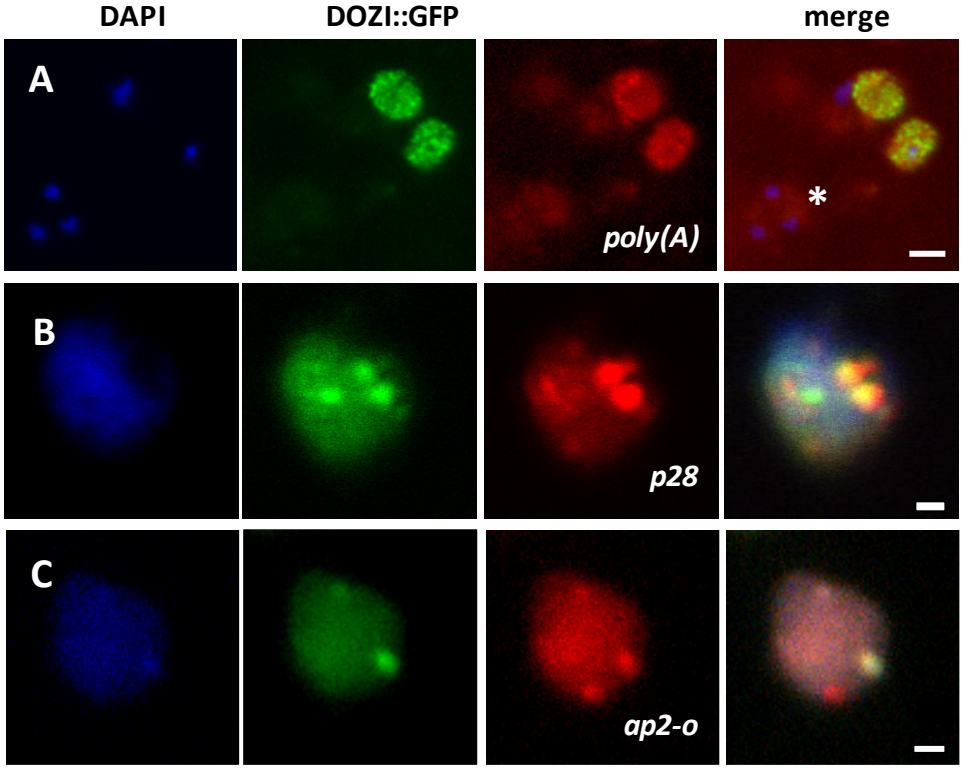

Fig. 9|FISH analyses of gametocytes. A Mixed blood stage parasites were hybridized to a 50mer, 3'-labeled TAMRA poly d(T) to visualize total polyadenylated mRNA. Note the strong staining of gametocytes, and the weak staining in 3 schizonts (*). Scale bar = 6 μm. **B** *p28* FISH combined with DOZI::GFP staining. Scale bar = 2 μm. **C** *ap2-o* FISH combined with DOZI::GFP staining. Scale bar = 2 μm.

3.7 – AP2-O::GFP was successfully transfected

All five constructs were transfected into wild type *P. berghei*. Following electroporation, parasites were injected i.v. into wild type BALB/c mice and pyrimethamine drug treatment in drinking water ensured the selection of mutant parasites [Fig. 10]. Out of the 5 transfections, 3 generated positive infections in the rodent host, namely AP2-O::GFP, AP2-2::GFP and AP2-4::GFP. However, following diagnostic PCR, only AP2-O::GFP proved to be correctly integrated in the genomic locus [Fig. 11A], the other two infections were not integrated. A second, independent, AP2-O::GFP construct was successfully transfected later. Following these results, AP2-O::GFP transgenic parasites were cloned by limiting dilution. Drug selection initially results in a mixed population of both mutant and wild type parasites, although the later ones are present in a very small percentage. The AP2-O::GFP mutant line was successfully cloned by limiting dilution [Fig 11B]. In addition, total RNA from mixed

blood stages was extracted and transcription of the *ap2-o::gfp* transgene was verified by RT-PCR. Several PCRs were performed targeting both the wild type and mutant transcripts as well as the *gfp* tag. As shown in [Fig 11C] the transgene is transcribed in mixed blood stages with no traces of transcription of the remaining, truncated wild type *ap2-o* gene.

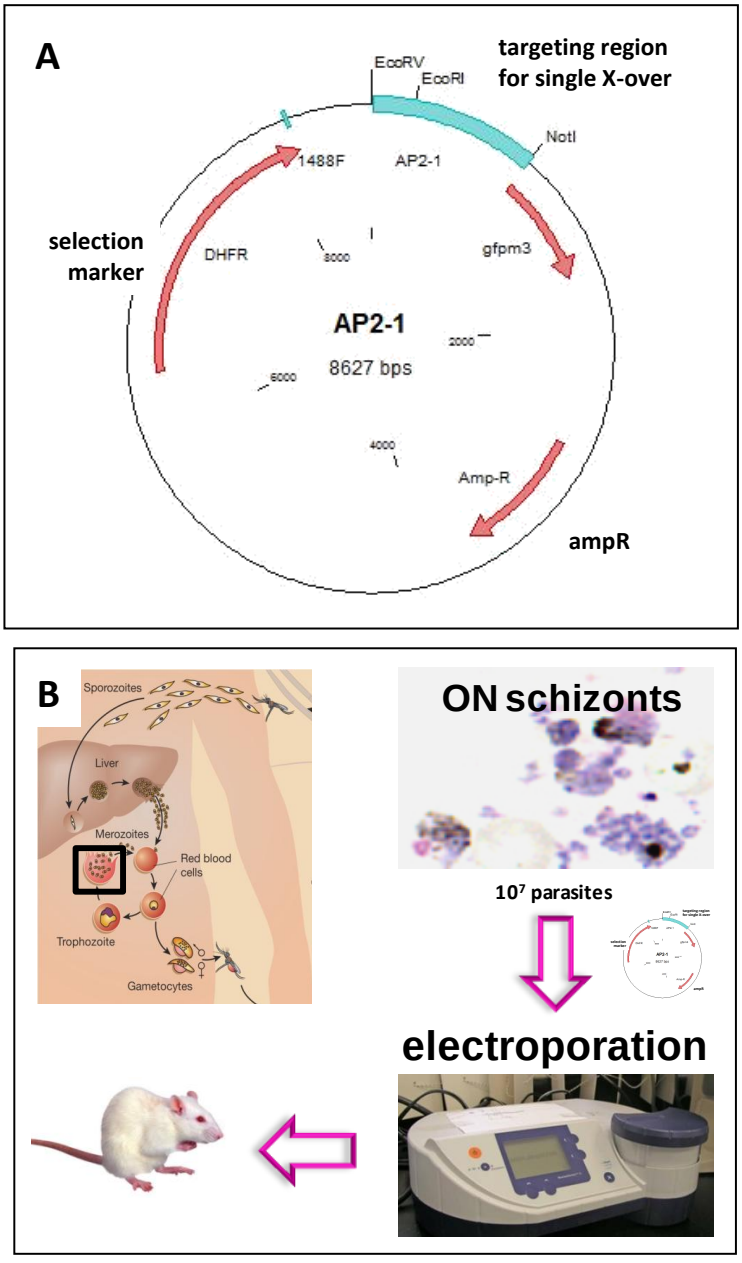

Fig. 10|Transfection of tagging constructs into schizonts. A Tagging plasmid for *ap2-o* C-terminal GFP-tagging. Note the single restriction site *Eco*RI within the targeting region which is used for linearization of the plasmid prior to transfection and homologous recombination by single cross-over. Plasmids for AP2-2, 3, 4, and 5 were generated accordingly (not shown) and sequence analyzed. Each plasmid contains the pyrimethamine-resistant marker *Toxoplasma gondii* DHFR gene. **B** Schematic *Plasmodium berghei* transfection protocol.

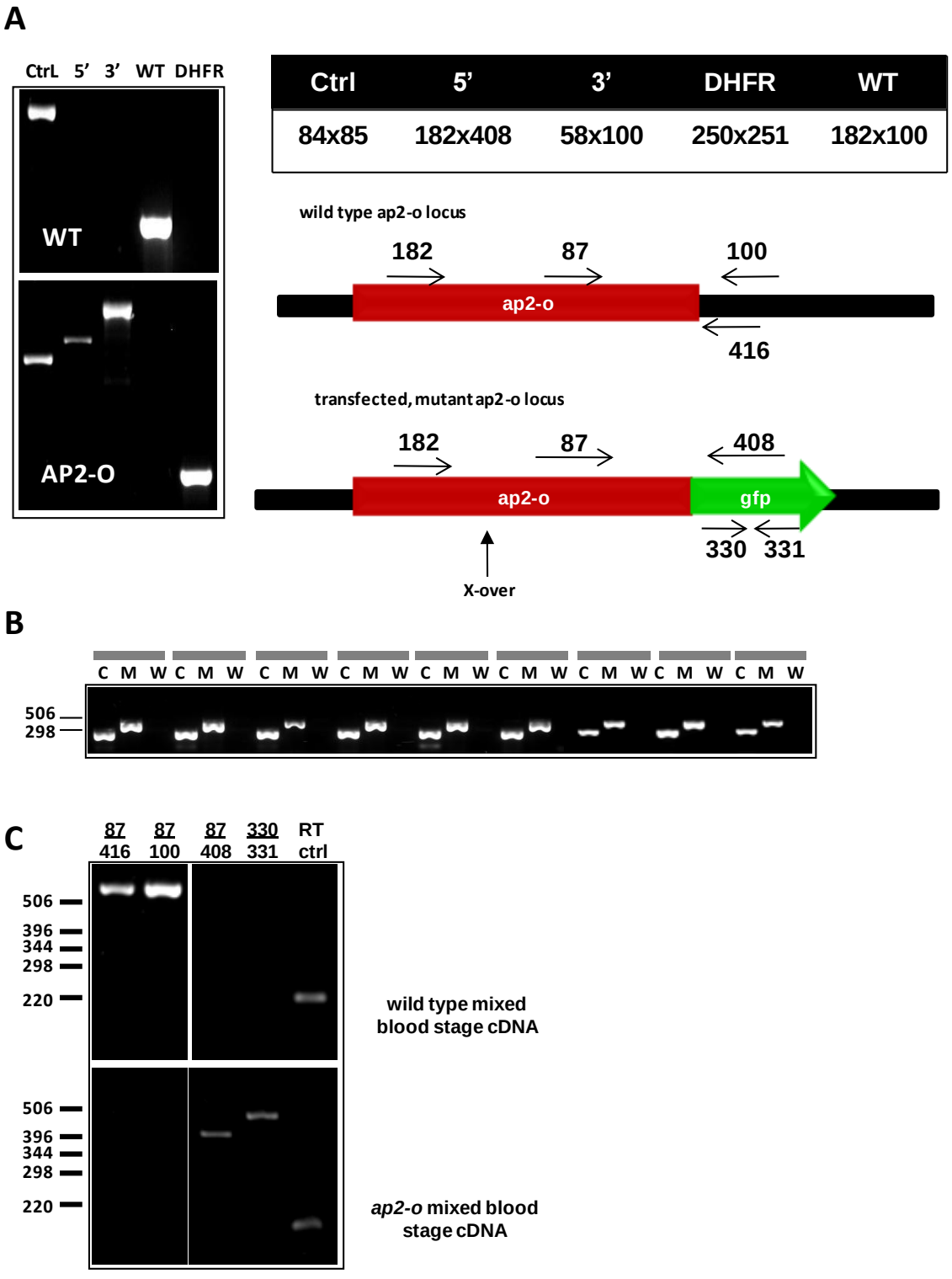

Fig. 11| Genotyping of cloned AP2::GFP transgenic parasites. A PCR analyses of AP2-O mutant and wild type parasite populations prior to cloning. Primers 84|85 were used to amplify RNApol II (+ control); Primers 250|251 were used to amplify the Toxoplasma *dhfr* gene, present in the construct (not shown); **B** Genotyping of 9 independent AP2-O::GFP clones obtained by limiting dilution. 3 reactions were performed in each clone: C = positive control 423|444, M = mutant 87/408, W = wild type 182|100. No wild type parasite contamination was detected. **C** RT-PCR analyses of wild type and mutant ap2-o::gfp transcription. No expression of the truncated wt gene was detected. 87|416, 87x100 wild type ap2-o; 87x408 ap2-o::gfp; 330|331 gfp; ctrl = RT control, pcr across introns, cDNA size =188bp, gDNA size 331bp.

3. 8 – No GFP protein expression was detected in AP2-O::GFP blood stages

Since RT-PCR results showed that the *ap2-o::gfp* transgene was being transcribed in blood stage parasites, immunofluorescence assays and Western blots were performed to check whether *ap2-o::gfp* was being translated. Thin smears of mutant mixed blood stages were fixed and immunostained with α-GFP antibodies; these experiments revealed no clear GFP signal. Furthermore, protein extract from wild type (both GFP+ and GFP- reference lines) and mutant parasites was prepared and analysed by Western blot. Once again, no GFP expression was detected in the mutant lines (data not shown). To check native AP2::GFP expression in live parasites a drop of blood was collected from infected mice, incubated with Hoechst and viewed under the microscope. All parasite stages (i.e rings, trophozoites, schizonts and gametocytes) were GFP negative (data not shown).

3.9 – AP2-O::GFP and wild type parasites grow equally well in mice

Introducing a tag and a resistance marker into the parasite genome can result in a cost to parasite fitness. In order to check potential altered growth kinetics of the mutant AP2::GFP parasite, two sets of 6 BALB/c mice were infected in parallel with the wild type and the mutant lines. Development of parasitemias (10 fields/smear with 500-900 RBCs each) were followed for five days and plotted [Fig. 12A]. This kinetic analysis of wild type and mutant parasite lines during asexual development revealed no significant differences (unpaired t-test, p-value= 0.33; H0: mutant =wt). Morphology of the different stages was also analysed [Fig.12B] revealing no apparent life stage differences between mutant parasites and wild type as shown from ring to young and mature trophozoite, young and mature schizont, as well as the sexual precursor cells (gametocytes).

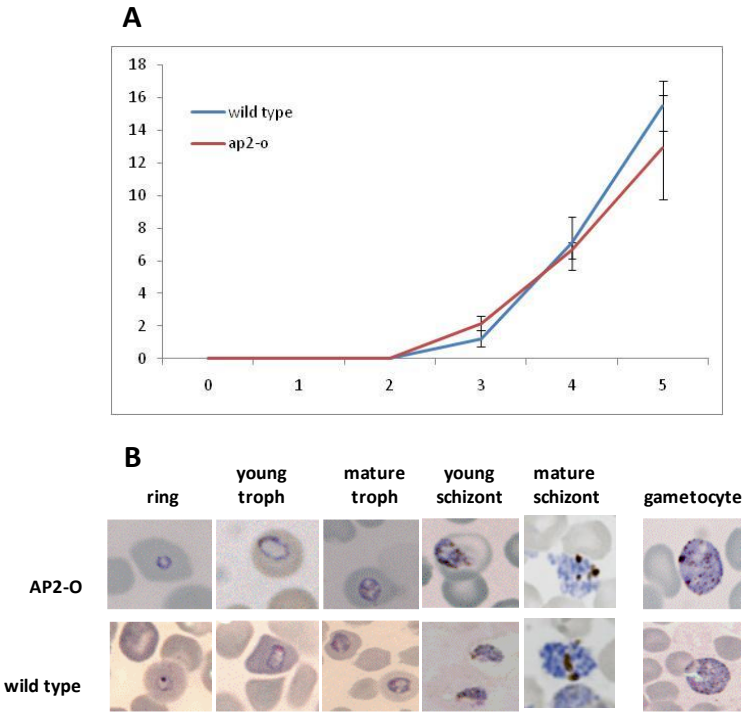

Fig. 12| AP2::GFP transgenic parasites have wild type, asexual growth characteristics. **A** Kinetic analysis of wild type and mutant parasite lines during asexual development revealed no significant differences as given by unpaired t-test (p-value= 0.33). **B** Light microscopic analyses of wild type compared to mutant AP2::GFP GIEMSA-stained blood smears, including gametocytes. No morphological differences were detected.

3.10 – AP2-O::GFP mutant gametocytes established a productive mosquito infection

Infection of the *A. stephensi* mosquito vector with *P. berghei* parasites [Fig. 13A, B] depends on the rapid generation of male (exflagellation) and female gametes, fertilization and transformation within 24 hours of the zygote into the motile ookinete. It is the ookinete that eventually escapes the hostile environment of the mosquito midgut and establishes the proper mosquito infection through the establishment of the oocyst stage. Exflagellation of male gametes is triggered when the microgametocyte encounters the mosquito midgut environment which is characterized by higher pH and a temperature change to 21C. These conditions can be induced *in vitro*, and 8-12 minutes after induction exflagellating males were counted under a 100X bright-field microscope; the AP2-O::GFP transgenic line generated male gametes in a comparable fashion to wild type, GFP+ *P. berghei,* with an average number of exflagellation centres of 1.25 and 1.5, respectively [Fig. 13C-1]. Successful oocyst formation was also very similar to wild type parasites (average number: wt=136,6; AP2-O::GFP mutant= 95,6; Fig. 13C-2). No GFP fluorescence was detected in AP2-O::GFP-infected mosquitoes [Fig.13 F and G]. Mercury-chrome stained oocysts are shown magnified [Fig. 13H]. On day 18th p.i. salivary glands were hand dissected and the number of sporozoites counted in a Neubauer chamber. Although variable [Fig 13C-3], the average number of sporozoites per salivary gland was similar between the mutant and the wild type line, with 21850 and 23726 parasites per mosquito, respectively.

3.11 – AP2-O::GFP sporozoites were able to re-infect the rodent host

Only female *Anopheles* blood feed on the vertebrate host. Although blood is not vital for their survival (both male and female mosquitoes feed on glucose solution), females need supplemental nutrients including proteins and iron for egg development. If infected with malaria during a previous blood meal, a female *Anopheles* can inoculate sporozoites to a new, subsequent host. AP2-O KO parasites produce morphologically somewhat degenerated ookinetes that are unable to establish oocysts and consequently sporozoites[37]. In this project the endogenous 3' UTR of *ap2-o* was exchanged for the translation competent *dhfr/ts* 3'-UTR with the aim to promote early translation of transcript in blood stage gametocytes. As the expression patterns of most AP2's are largely unknown, we analysed whether salivary gland sporozoites were affected by this mutation. Thus, we checked whether AP2-O::GFP mutant sporozoites were able to leave the mosquitoes' salivary glands and infect the rodent host: a set of mice were infected i.v. with 20 000 hand-dissected sporozoites, and another set by mosquito bite of 10 starved mosquitoes. After infection, the occurrence of blood stage parasites was examined at 3 time points: 18, 22 and 30 days p.i. AP2-O::GFP sporozoites proved to be infective in both cases (direct i.v. injection and mosquito bite).

3.12 – Fluorescence was detected in *in vitro* ookinete culture

Fertilization occurs 1-2 hours after a blood meal in the mosquito; 18h later the zygote has transformed into a motile ookinete. The absence of the AP2-O TF prevents proper ookinete development thus preventing definite mosquito infection [37]. The fact that our AP2-O::GFP mutant parasites were able to infect mosquitoes strongly suggested that this mutant line was able to express AP2-O::GFP protein during zygote to ookinete transformation. Therefore we decided to look at protein

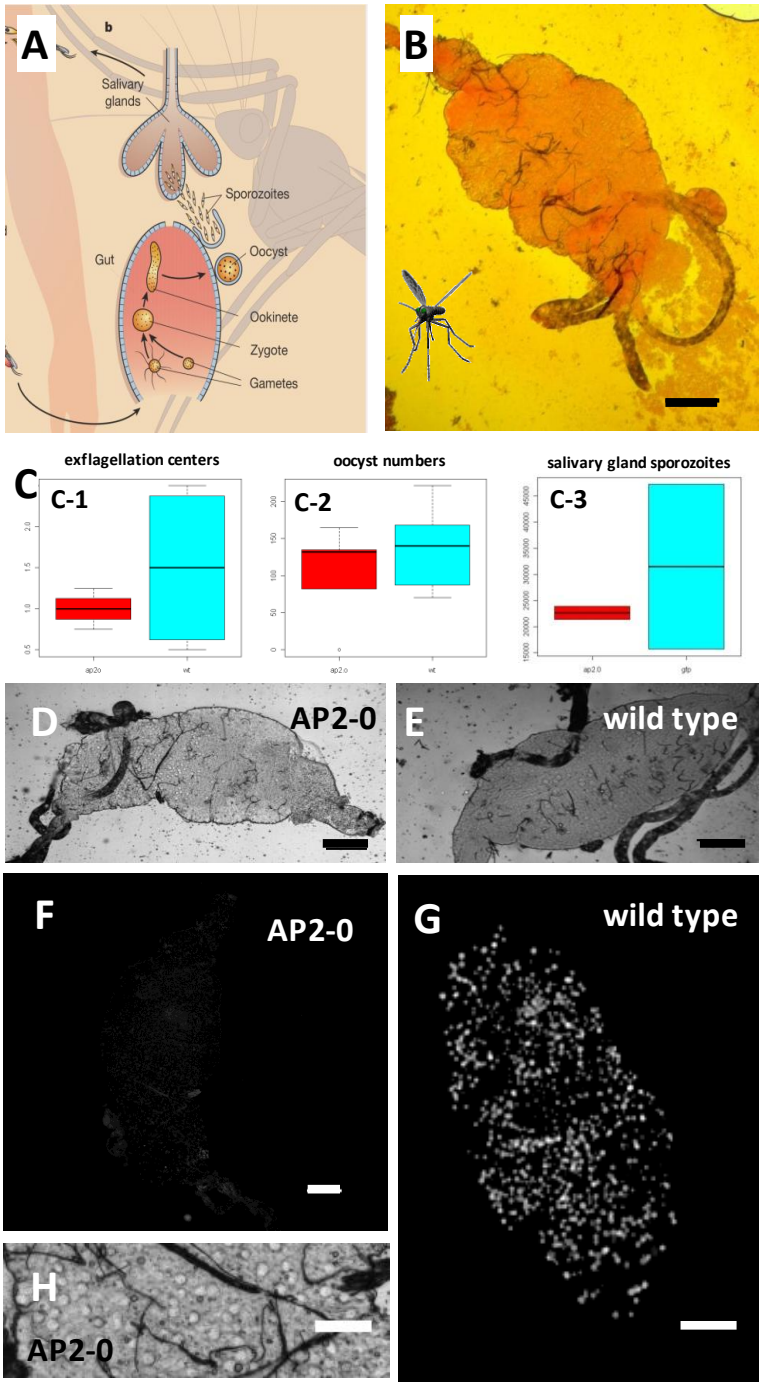

Fig. 13|AP2-O::GFP parasites develop normally in the mosquito.
A Schematic representation of the malaria life cycle in the mosquito. **B** Dissected mosquito midgut stained with mercury-chrome (scale bar, 50 μm). **C** Box plots comparing Exflagellation, oocysts and sporozoites from ap2-o::gfp (red) and wild type (blue). Exflagellation centers (n=4 independent infections), oocysts (n=9 mosquitoes each) and sporozoite (n=3 batch dissections) numbers per salivary glands. **D & E** Whole mosquito midgut mercury-chrome staining reveals the presence of wild type numbers of oocysts in the mutant(scale bar ,50μm). **F & G** Only wild type reference line is GFP+; fluorescence was detected in whole mosquito midguts by fluorescence microscopy(scale bar = 50μm) ; AP2-O parasites are GFP-, but are infective as revealed by mercury-chrome staining shown in **H** (scale bar 20 μm).

expression at the stage where AP2-O is naturally expressed: the ookinete. Overnight, *in vitro* ookinete cultures were prepared from 3 different parasite lines: wild type, CITH::GFP and AP2-O::GFP. Conversion from zygote to ookinete was not efficient for wild type parasites; only non-fertilized female gametes were found. Although with different patterns, both CITH::GFP and AP2-O::GFP lines generated green ookinetes [Fig14]. CITH::GFP fluorescence had had an apical, cytoplasmic distribution in ookinetes. Fluorescence for AP2-O revealed two patterns: restricted to the nucleus [Fig. 14, central panel] and dispersed in the cytoplasm [Fig. 14 right panel]. Interestingly, live non-fertilized female AP2-O::GFP gametes also showed a clear nuclear GFP expression pattern [Fig.15]. In summary, no AP2-O::GFP stage apart from the female gamete and ookinete was found to be GFP+.

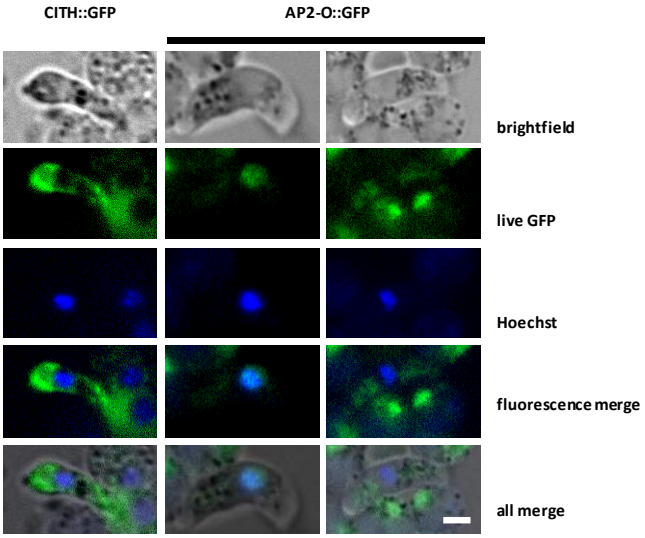

Fig. 14| GFP expression was detected in AP2-O::GFP parasites after incubation in ookinete medium. Live microscopy of ookinetes. Left panel: CITH::GFP ookinete, central and right panel: AP2-O::GFP ookinetes. Note that GFP expression in AP2-O ookinetes differ, being restricted to the nucleus in the central panel and present in the cytoplasm on the right one. (scale bar 2μm)

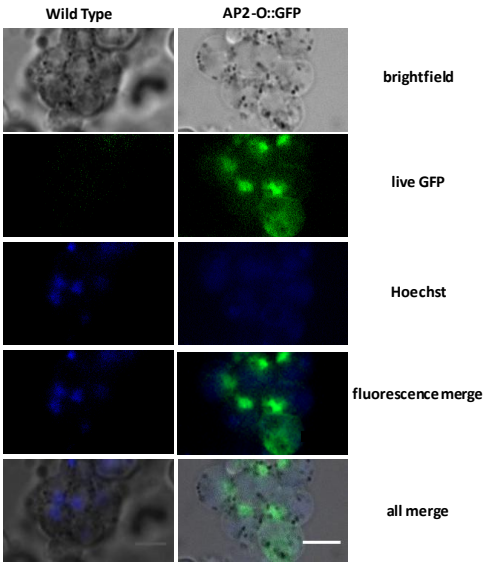

Fig. 15| GFP expression was detected in AP2-O::GFP gametes. Live microscopy of female gametes in ookinete culture. Left panel: wild type non fertilized female gametes, right panel: AP2-O::GFP female non-fertilized gametes. Note the presence of GFP expression only in AP2-O gametes. (scale bar 5μm)

4- Discussion

Future progress in the control of malaria parasite transmission will likely fall into three categories: (i) vector control to reduce transmission, (ii) antimalarial drugs that kill gametocyte stage parasites and (iii) genetic approaches to reduce or block transmission[52]. Transmission between the mammalian host and the vector imposes severe bottlenecks in the parasite's life cycle and is followed by a large expansion of the surviving population. The variety of host cell types and the parasite's morphological transitions during development suggest the utilization of fine-tuned regulation of gene expression at all stages. Indeed, in gametocytes and sporozoites (the distinct transmission stages between hosts) post-transcriptional regulation of gene expression has been shown to play a key role in the rapid adaptation in the newly parasitized host.

Among the transcripts shown to be regulated through translational repression mechanisms in the gametocyte (i.e. storage in P granule-like complexes) are *p25*, *p28* and the AP2 transcription factor *ap2-o*. Together with *ap2-o* four additional AP2 mRNAs are significantly down regulated in DOZI KO parasites suggesting a key role in the rapid adaptation to the new host environment in the mosquito midgut. Here we analysed the expression profiles in transmission stages of 24 known AP2 genes in *P. berghei;* they were tested by RT-PCR in gametocyte and sporozoite cDNA. Not only were the five down-regulated AP2's highly expressed in gametocytes but included 10 additional AP2 genes. Furthermore, 9 out of the 15 positive hits in gametocyte cDNA were found to be physically associated with CITH::GFP defined P-granules. In addition, FISH of *ap2-o* showed widespread co-localization with cytoplasmic P granules, reinforcing translational repression as regulatory mechanism for this particular gene.

Following published data[29-31], the motifs within the 3' UTR of silenced transcripts has been identified as the major target for RNA-binding proteins that target the transcript for translational repression thereby modulating mRNA stability, accessibility to the translational apparatus and therefore translational capacity. Our primary goal was to understand whether the 3' UTR region alone was sufficient to recruit the five targeted transcripts into P granules. Therefore, by exchanging the endogenous with the *dhfr* 3'UTR (known to allow translation of transgenes like GFP and RFP in gametocytes[29]) we expected to promote mis-timed expression by preventing targeting of the mRNA to translationally silent mRNPs. Upon gene model confirmation GFP-tagged vectors for 5 AP2 molecules were constructed, and confirmed by sequencing. Following transfection into *P. berghei* wild type schizonts, to incorporate the proposed alterations in the parasite's genome, one mutant line was obtained: AP2-O::GFP. AP2::GFP fluorescence was imaged in several stages by live microscopy and immunofluorescence. In addition, Western blot analysis was used to detect the presence of the protein. However, no signal was observed in blood stage parasites including female gametocytes. These results suggest that exchange of the endogenous AP2-O 3' UTR with the translation competent DHFR 3' UTR does not promote translation of the mutant gene; as in *p25* (but not *p28*) additional motifs in the 5' UTR may contribute to translational silencing. In future experiments we will therefore generate mutant parasites that contain – in addition to the changed 3' UTR – also a different 5' UTR under the control of the constitutively active *ef1*α promoter.

Kinetic and morphological analysis of the transgenic AP2::GFP mutant parasites during blood stage progression revealed no major differences. Infection of mosquitoes with the AP2-O::GFP parasite clone proceeded normally with wild type numbers of oocysts and salivary gland sporozoites. Sporozoite infectivity also revealed to be efficient at several time points of sporozoite maturation (18, 22 and 30 days p.i.) using either i.v. injection or mosquito bite experiments.

However, when we performed *in vitro* ookinete conversion assays with the same parasite clone we found no ookinetes. Although the number of CITH::GFP gametocyte transformation was also very low, the AP2-O::GFP culture showed the presence of phagocytosed female gametes; this was never observed in the CITH::GFP nor wild type experiments. Only after WBCs removal from the infected blood prior to incubation were AP2-O::GFP ookinetes observed. Although intra-erythrocytic gametocytes are refractory to phagocytosis (levels of gametocyte phagocytosis are comparable to uninfected RBC uptake[53]), free gametes are highly vulnerable to IgG opsonisation, agglutination and hence phagocytosis, as well as complement-mediated lysis[54]. In addition, shown by J. Healer et al. a malaria immune serum is more efficient in activating the phagocytosis cascade than a non-imune serum[53]. Interestingly, not only were the mutant ookinetes GFP positive (AP2-O::GFP expression was observed both in cytoplasmic and nuclear localization) but non fertilised female gametes were also clearly GFP positive, with staining restricted to the nucleus. In previously published experiments[37] where AP2-O was GFP tagged (N.B. only the tag was added, no other alteration was induced) expression prior to ookinete stage was not documented.

Our data suggest that translational silencing of *ap2-o* is 3' UTR-independent, or alternatively relies not only on motifs within the 3' UTR but depends also on additional motifs that may be localized to the 5' UTR; such a mechanism has also been proposed for *p25*. Thus, mutant *ap2-o::gfp* has followed the standard repression pathway, being stored in a translationally quiescent state in female gametocyte P granules; the transcript is only translated following development in ookinete culture medium, but is also translated in the absence of fertilization. Translation of silenced *p28* in the absence of fertilisation was previously reported in infertile CDPK4 KO mutants; P28 expression on the surface of unfertilised female gametes is clearly visible [Fig.16].

In conclusion, we were able to successfully generate an AP2-O::GFP mutant line where the endogenous 3'-UTR was replaced by the PbDHFR 3'-UTR known to be constitutively translated. Infection patterns (growth kinetics, parasite morphology, sporozoite numbers) were comparable to the wild type and no AP2-O::GFP expression was detected in blood stage parasites. AP2-O::GFP expression was only observed during early mosquito development (i.e. female gamete and ookinete) where AP2-O is naturally expressed. Therefore we conclude that the 3'UTR of AP2-O is not sufficient to trigger silencing in female gametocytes.

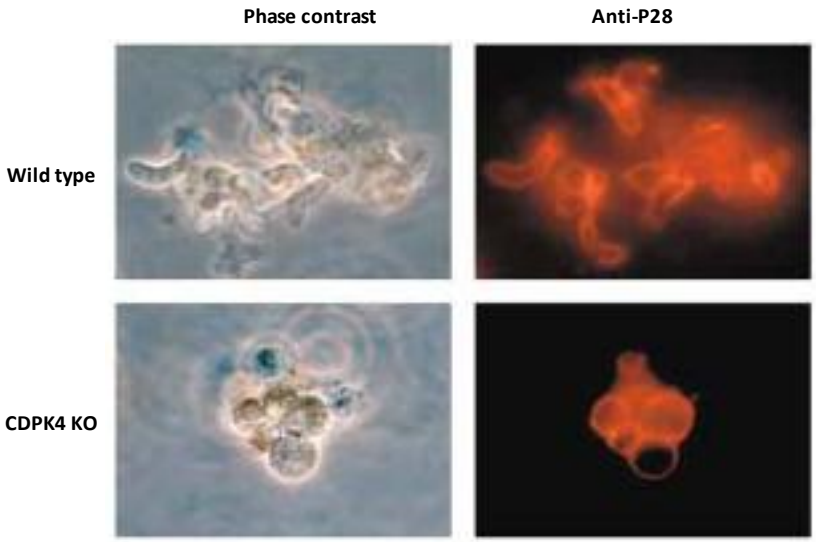

Fig. 16|Wild type and CDPK4 parasites express P28 but only wild type macrogametes differentiate into ookinetes. Gametocytes in infected blood were activated and cultured for 24 hr in vitro, when surface expression of P28 was assessed in live cells by direct immunofluorescence staining with the Cy3-labeled monoclonal antibody. (Adapted from Billker et al. 2004)

5- Appendix I

1- RNA extraction by TRIzol method:

RBC were harvested by centrifugation (600rcf), 4minutes at room temperature. 10v/v of the total pellet volume was added of TRIzol.(eg 1ml TRIzol to 0.1ml pellet). The pellet was resuspended and 0,2v of the original TRIzol volume in chloroform was added (eg 0.2 ml $CHCl_3$). Tubes were centrifuged for 30 minutes at 9000rcf. The aqueous layer was collected to a new tube and isopropanol was added (v/v) in order to precipitated total nucleic acids. Samples were incubated overnight at 4°C. On the following day, tubes were centrifuged to pellet nucleic acids, 30 minutes at 12000rcf. The supernatant was discarded. Pellet was resuspended and DNAseI treated (Invitrogen, 3U/sample) for 15 minutes at 37ºC. Reaction was stopped with 25mM RNAse-free EDTA and by heat inactivation, 10minutes at 65°C. RNA was stored in 100% ethanol/ 3M sodium acetate (pH=5.6).

2- Preparation of cDNA (First strand) with SuperScript II:

cDNA was prepared according to Invitrogen guidelines. Therefore the RNA was resuspended in 10μL of RNAse free water and 1μL (0,5μg) of OligodT primers (Invitrogen) were added followed by an incubation of 10 minutes at 70ºC. On ice the following reagents were added: 4μL of 5X first strand buffer (Invitrogen), 2μL 0,1M DTT (Invitrogen), 1μL (40U) RNase out (Invitrogen) and1μL 10mM dNTP (Invitrogen). The mix was preincubated for 2minutes at 42ºC and then 0,5μL (100U) of Superscript II (Invitrogen) were added. Samples were incubated 50minutes at 42ºC and then reaction was stopped after incubation at 70ºC for 15 minutes.

3- Genomic DNA isolation of malaria parasites (phenol-chloroform method):

After blood collection by cardiac puncture RBC are lysed and the pellet is resuspended in 700 µL of TNE buffer (10 mM Tris pH 8.0, 5 mM EDTA pH 8.0, 100 mM NaCl). The following supplements were added: 200 µg RNase, 1% (v/v) SDS and demineralised water up to 1 ml. Samples were incubated for 10 minutes at 37ºC and 200 µg proteinase K were added. After a 1-hour incubation buffered phenol:chloroform:isoamylalcohol (25:24:1) was added up to 1.5 ml. Samples were inverted several times and centrifuge for 5 minutes at 10000 rcf . The aqueous upper phase was transferred to a new tube and chloroform:isoamylalcohol (24:1) was added up to 1.5 ml. Again samples were inverted several times and centrifuged for 5 minutes at 10000 rcf. The new aqueous upper phase was transferred to a new tube and DNA was precipitated overnight at -20ºC with 3M sodium acetate/100% Ethanol, pH=5,6. On the following day DNA was pelleted, air-dried and resuspended in 500µl of DNAse-free water.

4- PCR Mix:

In each 50μL reaction the following solution was used: 1x Taq Fermentas buffer without $MgCl_2$ (5μL-10x), 2mM of $MgCl_2$ (4 μL-25mM), 20mM of each primer (2μL-10mM), DNA/cDNA sample, usually 10-50ng, 20mM of dNTPS (2μL-10mM), 1unit of Taq Polimerase (Fermentas). DNAse free water was added up to 50μL.

5- Chemically-competent DH5-α cells:

A streak of DH5-α bacteria was inoculated in 25mL of SOB medium (supplemented with 10mM $MgCl_2$ and 10mM $MgS0_4$) and incubated at 37ºC for 6h. 10mL of the primary culture were subcultured in 250mL of SOB medium under the same conditions, followed by overnight incubation at 18ºC at 200-250rpm. At an optical density of 0.6, cells were incubated on ice for 10 minutes and then centrifuged for 10 minutes at 4ºC, 2500G. Then the pellet was gently resuspended in 40mL of transformation buffer ([$MnCl_2.H_2O$]=55mM; [$CaCl_2.2H_2O$]= 15mM; [KCl]=250mM; [Pipes (pH=6,7)]= 10μM) and incubated 10 minutes on ice. After centrifugation for 10 minutes at 4ºC, 2500G, the pellet was once more resupended gently in 10mL of transformation buffer and, slowly, 750μL of sterile DMSO (7%) were added. Finally cells were aliquoted on ice, snap-frozen and stored at -80°C.

Aliquots were only thawed just before bacterial transformation.

6- Plasmid extraction by Boiling Minipreps:

After overnight growth in LB medium containing ampicillin (50-100ug/ml), mini-preparation cultures were centrifuged for 5 min and the pellet resuspended in 360μL of STET (Sucrose=8g; TritonX-100=5g; [0,5M]EDTA=10mL; [1M]Tris-HCL(pH 8)=5mL). After addition of 40μl lysozyme and boiling for 1 minute, they were centrifuged for 15min at 4°C and the pellet discarded. Subsequently 400μL of isopropanol were added to the supernatant, followed by vortexing and 10mins of plasmid DNA precipitation. Finally, the dried pellet was resuspended in 100μL water containing 20μl/mL RNAseA.

6- Appendix II – Tables

Table 3: Summary of AP2 genes in *P. falciparum* and *P.berghei* orthologues. Note the presence of one paralogue in *P. falciparum* (PFL1075W). The number of AP2 domains and the de-regulation profile in several KOs are also depicted. ID= Accession number

P. falciparum ID	exons	protein length	DOZIko/WT (Ref 27)	CITHko/WT (Ref 36)	PUF2ko/WT (Ref 35)	domains	P. berghei ID
MAL8P1.153	2	2577	nd	nd	nd	1x AP2	PBANKA_122810
PF07_0126	1	1331	nd	nd	nd	2x AP2	PBANKA_021440
PF10_0075	2	1597	nd	nd	nd	3x AP2	PBANKA_120590
PF11_0091	2	1828	1,5	nd	nd	1xAP2, 1xp53-like	PBANKA_093910
PF11_0163	4	231	nd	nd	nd	2x AP2	PBANKA_093230
PF11_0404	3	2653	nd	nd	nd	2x AP2	PBANKA_090960
PF11_0442	1	1604	2,5	1,5	nd	1x AP2	PBANKA_090590
PF13_0026	1	328	1,3	nd	nd	1x AP2	PBANKA_140370
PF13_0097	1	2672	nd	nd	nd	1x AP2	PBANKA_141570
PF13_0235	1	3858	nd	nd	nd	4x AP2	PBANKA_135600
PF13_0267	1	521	nd	nd	nd	1x AP2	PBANKA_136370
PF14_0079	1	1702	nd	nd	nd	1x AP2	PBANKA_103430
PF14_0271	1	1186	nd	nd	nd	1x AP2	PBANKA_101550
PF14_0471	2	715	-23	nd	nd	1x AP2	PBANKA_131320
PF14_0533	2	1374	-2,3	nd	nd	1x AP2	PBANKA_131970
PF14_0633	1	813	nd	nd	nd	1x AP2	PBANKA_132980
PFD0200c	2	2249	nd	nd	nd	3x AP2	PBANKA_100180
PFD0985w	1	3473	nd	nd	-2,4	2x AP2	PBANKA_052170
PFE0840c	1	2378	nd	nd	nd	2x AP2	PBANKA_123160
PFF0200c	1	1979	1,6	nd	nd	2x AP2	PBANKA_010290
PFF0550w	3	278	nd	nd	1,6	1x AP2	PBANKA_010950
PFF0670w	2	4109	nd	nd	nd	3x AP2	PBANKA_011210
PFF1100c	1	1989	nd	nd	nd	1xAP2	PBANKA_112180
PFI1665w	1	200	nd	nd	1,4	1x AP2	PBANKA_083520
PFL1075w	1	2558	nd	nd	nd	1x AP2, 1x Ac-coA	na
PFL1085w	1	2432	nd	nd	nd	1x AP2	PBANKA_143750
PFL1900w	1	2577	nd	nd	nd	2x AP2	PBANKA_145370

Table 4: Primers used to confirm AP2 gene models (section 2.5). Fwd = forward primer; Rev= reverse primer

Primer	sequence (5'-3')	target	Fwd/Rev
78	AAACAATTGAGATAATGGATTAATTTG	PBANKA_090590	Fwd
79	AAGCGGCCGCAAATGTTATCAACAGTGATAC	PBANKA_090590	Rev
80	AATAATAGAAACTATTCG	PBANKA_090590	Rev
87	AAGATGTGATATTGCTGG	PBANKA_090590	Fwd
88	TGCGTTTAATTCGGTTCC	PBANKA_090590	Rev
100	ACAAGGAAATAAAGCTGC	PBANKA_090590	Rev
107	AAAGATATCATGGTTAATACACATAAC	PBANKA_140370	Fwd
108	AAAGCGGCCGCTTTTTAGGTGTATTCGTCC	PBANKA_140370	Rev
114	ATTGGAATATATATGTGC	PBANKA_140370	Rev
117	TTCGATATCATAATAACTCGAACAATC	PBANKA_093910	Fwd
118	AAAGCGGCCGCATTCATCATATCCCC	PBANKA_093910	Rev
119	TTCATCATATCCCTTATTTG	PBANKA_093910	Rev
120	TTGCAATATCCCAGATAC	PBANKA_093910	Rev
144	TTTATTTTTGAAGTTTCC	PBANKA_010290	Rev
145	AAAGCGGCCGCTTTCATCATTCAATATATG	PBANKA_010290	Rev
146	AGAGAATTCCCCCAAAAAAC	PBANKA_010290	Fwd
142	GTATCGATATCTAACACAAATATTATAGG	PBANKA_131320	Fwd
143	AAAGCGGCCGCATAATTTGTCTAAATA	PBANKA_131320	Rev

Table 5: Primers used to amplify AP2 gene in RT-PCR (sections2.6 and 2.7). Fwd = forward primer; Rev= reverse primer

Primer	sequence (5'-3')	Target	Fwd/Rev
649	GGTTAATGCAAGTGACCAG	PBANKA_010290	Fwd
650	CTAGATAAATTTGAACCCACATC	PBANKA_010290	Rev
651	GAAGAGCTATTTATTCAAATAGGG	PBANKA_010950	Fwd
652	GTCTTCTTGTTCTTCATCAGAG	PBANKA_010950	Rev
653	CGACAATTATGATGGAAATTG	PBANKA_011210	Fwd
654	CTCTCCATGCTTGTACTACTCC	PBANKA_011210	Rev
573	CAAAATTGGAAAGAAGCAGC	PBANKA_021440	Fwd
574	GCACAACAAATGTACTTACCC	PBANKA_021440	Rev
611	CATTGGATTACGGAAATGC	PBANKA_052170	Fwd
612	CCTCGCACCTAGTATTTCC	PBANKA_052170	Rev
657	GATGGAGGTAATGTAAATCCC	PBANKA_090590	Fwd
658	CCTTAATAGCCCAAAATTTTG	PBANKA_090590	Rev
659	CATGATAAAAGATGGGGGAAG	PBANKA_090960	Fwd
660	CTTTTGGGAATAGCATTTTCC	PBANKA_090960	Rev
661	CGGTGTGTGAGCAAAATG	PBANKA_093230	Fwd
662	CGTAGTAAACTCCTTCTACCCC	PBANKA_093230	Rev
663	GGTGGTGTTAGAAACAGTTGTG	PBANKA_093910	Fwd
664	CGCGCCAGTTGAATAACTAC	PBANKA_093910	Rev
665	GGAGGAAGATGGAACTTCATG	PBANKA_100180	Fwd
666	CCCCCTGTTTGATATTGTTG	PBANKA_100180	Rev
667	GTTTCAGGGGTTTGGTATG	PBANKA_101550	Fwd
668	GTTCCCAATTCCATCTTTC	PBANKA_101550	Rev
693	CAACAAAGATTGTGGGAATTTAG	PBANKA_112180	Fwd
694	CCATCCATTGTTTGGATG	PBANKA_112180	Rev
671	CTCCTAATGATGGCATTCG	PBANKA_120590	Fwd
672	CCATCTTCATTTCCATATCC	PBANKA_120590	Rev
673	CAAATTGAATCAAAGGGGG	PBANKA_122810	Fwd
674	GGGATTCATAATTTTCATTATCCC	PBANKA_122810	Rev
675	CAGGAATTTCGAACAAATGG	PBANKA_123160	Fwd
676	CGCCATTTTCTTCATTTGTG	PBANKA_123160	Rev
677	GATTTTCAAATTGATTCCCC	PBANKA_131320	Fwd
678	CTTTTTTCAATTCCCACAGG	PBANKA_131320	Rev
679	GTTGGGGGACTTAAAATACC	PBANKA_131970	Fwd
680	CCTTCCAATTAGCTCTCCATAAC	PBANKA_131970	Rev
681	GACGGAATTAGACACTGAAGC	PBANKA_135600	Fwd
682	CCCCAACAATTCTTGGTAG	PBANKA_135600	Rev
683	GCATTCAGGTTTTCCGAAG	PBANKA_140370	Fwd
684	GTCCACTACGCACTCTATTCC	PBANKA_140370	Rev
685	GCCCTCAAAGAAGTTACCC	PBANKA_141570	Fwd
686	CCCTTTTTCACTTGTGGG	PBANKA_141570	Rev
687	GAGGAAGAAGGCAAAAAATTATG	PBANKA_143750	Fwd
688	CTTCCATTTTCCCACCATC	PBANKA_143750	Rev
689	GATAATGTTAATTGCAATGGGG	PBANKA_145370	Fwd
690	CTTCGGCAACTGAATATTCG	PBANKA_145370	Rev
575	CATTATCAACAAATTCAACCAAC	PBANKA_132980	Fwd
576	ATACTGTGGAGATATATTACTCATTCC	PBANKA_132980	Rev
613	ATGTTGCGAAATATTATTAACC	PBANKA_083520	Fwd
614	AAACTGACCCTTTGATATTTTC	PBANKA_083521	Rev

Table 6: Primers used to amplify DNA FISH probes (section 2.8). Fwd = forward primer; Rev= reverse primer

Primer	sequence (5'-3')	target	Fwd/Rev
78	AAACAATTGAGATAATGGATTAATTTG	PBANKA_090590	Fwd
79	AAGCGGCCGCAAATGTTATCAACAGTGATAC	PBANKA_090590	Rev
107	AAAGATATCATGGTTAATACACATAAC	PBANKA_140370	Fwd
108	AAAGCGGCCGCTTTTTAGGTGTATTCGTCC	PBANKA_140370	Rev
117	TTCGATATCATAATAACTCGAACAATC	PBANKA_093910	Fwd
118	AAAGCGGCCGCATTCATCATATCCCC	PBANKA_093910	Rev
145	AAAGCGGCCGCTTTCATCATTCAATATATG	PBANKA_010290	Rev
146	AGAGAATTCCCCCAAAAAAC	PBANKA_010290	Fwd
140	AAAATTTAAATATTAATGATGG	PBANKA_131320	Fwd
143	AAAGCGGCCGCATAATTTGTCTAAATA	PBANKA_131320	Rev

Table 7: Primers used to amplify the cloning inserts (section 2.9). Restriction sites are highlighted. Fwd = forward primer; Rev= reverse primer

Primer	sequence (5'-3')	target	Fwd/Rev
78	AAA**CAATTG**AGATAATGGATTAATTTG	PBANKA_090590	Fwd
79	AA**GCGGCCGC**AAATGTTATCAACAGTGATAC	PBANKA_090590	Rev
107	AAA**GATATC**ATGGTTAATACACATAAC	PBANKA_140370	Fwd
108	AAA**GCGGCCGC**TTTTTAGGTGTATTCGTCC	PBANKA_140370	Rev
117	TTC**GATATC**ATAATAACTCGAACAATC	PBANKA_093910	Fwd
118	AAAGCGGCCGCATTCATCATATCCCC	PBANKA_093910	Rev
145	AAA**GCGGCCGC**TTTCATCATTCAATATATG	PBANKA_010290	Rev
146	AGA**GAATTC**CCCCAAAAAAC	PBANKA_010290	Fwd
140	AAA**ATTTAAAT**ATTAATGATGG	PBANKA_131320	Fwd
141	TTAGATATCGATACATGATTATC	PBANKA_131320	Rev
142	GTATCGATATCTAACACAAATATTATAGG	PBANKA_131320	Fwd
143	AAA**GCGGCCGC**ATAATTTGTCTAAATA	PBANKA_131320	Rev

Table 8: Primers used to genotype amplified parasites (section 2.11). Fwd = forward primer; Rev= reverse primer

Primer	sequence (5'-3')	target	Fwd/Rev
84	AAAGAATTCTGATGGTTTACAATCACC	RNApol II (ctrl)	Fwd
85	AAAGCGGCCGCTTTCTTCCTGCATCTCCTC	RNApol II (ctrl)	Rev
100	ACAAGGAAATAAAGCTGC	PBANKA_090590 (WT)	Rev
182	AAAGGATCCAATAACAATTTAGCGGG	PBANKA_090590 (WT)	Fwd
408	GTATGTTGCATCACCTTC	GFP (mutant)	Rev
58	AATGGCTAGTATGAATAGCC	Plasmid (mutant)	Fwd
250	CCGGTGTGTCTGGTCGTCG	Toxoplasma dhfr (mutant)	Fwd
251	GCCCAGAGACAGCGCTGCC	Toxoplasma dhfr (mutant)	Rev
87	AAGATGTGATATTGCTGG	PBANKA_090590 (WT)	Fwd
423	CAATTTGTTCTCTTTCTGAATTAC	PBANKA_050120 (uis4 - ctrl)	Fwd
444	CCAAACCAAGCGATCATACATACAG	PBANKA_050120 (uis4 - ctrl)	Rev

Table 9: Primers used in RT-PCR to amplify ap2-o::gfp transcript (section 2.13). Fwd = forward primer; Rev= reverse primer

Primer	sequence (5'-3')	target	Fwd/Rev
87	AAGATGTGATATTGCTGG	PBANKA_090590 (WT)	Fwd
416	CAAGGAAATAAAGCTGCCTC	PBANKA_090590 (WT)	Rev
100	ACAAGGAAATAAAGCTGC	PBANKA_090590 (WT)	Rev
408	GTATGTTGCATCACCTTC	GFP (mutant)	Rev
330	TAATACGACTCACTATAGGGAAGGTGATGCAACATACGG	GFP (mutant)	Fwd
331	TAATACGACTCACTATAGGGAATGGTCTGCTAGTTGAACG	GFP (mutant)	Rev
375	CGGAATTCCAATAAACTATCCTACTAATTCG	PBANKA_093910 (RT control)	Fwd
419	AATATTGCAATATCCCAGATAC	PBANKA_093910 (RT control)	Rev

7- References

1 WHO, The World Health Report. (2008).

2 WHO, The World Health Report. (2003).

3 Matuschewski, K., Getting infectious: formation and maturation of Plasmodium sporozoites in the Anopheles vector. *Cell Microbiol* 8 (10), 1547-1556 (2006).

4 WHO, World Malaria Report 2009. (2010).

5 Engwerda, C.R. & Good, M.F., Disarming the malaria parasite. *Nat Med* 14 (9), 912-913 (2008).

6 Saxena, A.K., Wu, Y., & Garboczi, D.N., Plasmodium p25 and p28 surface proteins: potential transmission-blocking vaccines. *Eukaryot Cell* 6 (8), 1260-1265 (2007).

7 Carter, R., Transmission blocking malaria vaccines. *Vaccine* 19 (17-19), 2309-2314 (2001).

8 Thera, M.A. *et al.*, Safety and immunogenicity of an AMA1 malaria vaccine in Malian children: results of a phase 1 randomized controlled trial. *PLoS One* 5 (2), e9041.

9 Hisaeda, H. *et al.*, Antibodies to malaria vaccine candidates Pvs25 and Pvs28 completely block the ability of Plasmodium vivax to infect mosquitoes. *Infect Immun* 68 (12), 6618-6623 (2000).

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

11 Medica, D.L. & Sinnis, P., Quantitative dynamics of Plasmodium yoelii sporozoite transmission by infected anopheline mosquitoes. *Infection and Immunity* 73 (7), 4363-4369 (2005).

12 Prudencio, M., Rodriguez, A., & Mota, M.M., The silent path to thousands of merozoites: the Plasmodium liver stage. *Nature Reviews Microbiology* 4 (11), 849-856 (2006).

13 Khan, S.M. & Waters, A.P., Malaria parasite transmission stages: an update. *Trends Parasitol* 20 (12), 575-580 (2004).

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

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

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

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

18 Khan, S.M. *et al.*, Proteome analysis of separated male and female gametocytes reveals novel sex-specific Plasmodium biology. *Cell* 121 (5), 675-687 (2005).

19 Lasonder, E. *et al.*, Proteomic profiling of Plasmodium sporozoite maturation identifies new proteins essential for parasite development and infectivity. *PLoS Pathog* 4 (10), e1000195 (2008).

20 Balaji, S., Babu, M.M., Iyer, L.M., & Aravind, L., Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res* 33 (13), 3994-4006 (2005).

21 Iyer, L.M., Anantharaman, V., Wolf, M.Y., & Aravind, L., Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. *International Journal for Parasitology* 38 (1), 1-31 (2008).

22 Chervitz, S.A. *et al.*, Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* 282 (5396), 2022-2028 (1998).

23 Templeton, T.J. *et al.*, Comparative analysis of apicomplexa and genomic diversity in eukaryotes. *Genome Res* 14 (9), 1686-1695 (2004).

24 Aravind, L., Iyer, L.M., Wellems, T.E., & Miller, L.H., Plasmodium biology: genomic gleanings. *Cell* 115 (7), 771-785 (2003).

25 Lespinet, O., Wolf, Y.I., Koonin, E.V., & Aravind, L., The role of lineage-specific gene family expansion in the evolution of eukaryotes. *Genome Res* 12 (7), 1048-1059 (2002).

26 Hakimi, M.A. & Deitsch, K.W., Epigenetics in Apicomplexa: control of gene expression during cell cycle progression, differentiation and antigenic variation. *Curr Opin Microbiol* 10 (4), 357-362 (2007).

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

28 Coleman, B.I. & Duraisingh, M.T., Transcriptional control and gene silencing in Plasmodium falciparum. *Cell Microbiol* 10 (10), 1935-1946 (2008).

29 Braks, J.A., Mair, G.R., Franke-Fayard, B., Janse, C.J., & Waters, A.P., A conserved U-rich RNA region implicated in regulation of translation in Plasmodium female gametocytes. *Nucleic Acids Res* 36 (4), 1176-1186 (2008).

30 de Moor, C.H., Meijer, H., & Lissenden, S., Mechanisms of translational control by the 3' UTR in development and differentiation. *Semin Cell Dev Biol* 16 (1), 49-58 (2005).

31 Mazumder, B., Seshadri, V., & Fox, P.L., Translational control by the 3'-UTR: the ends specify the means. *Trends Biochem Sci* 28 (2), 91-98 (2003).

32 Wickens, M., Bernstein, D.S., Kimble, J., & Parker, R., A PUF family portrait: 3'UTR regulation as a way of life. *Trends Genet* 18 (3), 150-157 (2002).

33 Cui, L.W., Fan, Q., & Li, J.F., The malaria parasite Plasmodium falciparum encodes members of the Puf RNA-binding protein family with conserved RNA binding activity. *Nucleic Acids Research* 30 (21), 4607-4617 (2002).

[34] Miao, J. *et al.*, The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite Plasmodium falciparum. *Journal of Cell Science* 123 (7), 1039-1049 (2010).

[35] Gomes-Santos, C.B., J; Prudêncio, M; Carret, C; Gomes, ARB; Pain, A; Feltwell, T; Khan, S; Waters, A; Janse, C; Mair, G and Mota, M, Developmental transition of *Plasmodium* sporozoites into liver-stage forms is regulated by the RNA binding protein pumilio. *MPMXXI, 2010, 12-16 September*.

[36] Mair, G.R. *et al.*, Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development. *PLoS Pathog* 6 (2), e1000767.

[37] Yuda, M. *et al.*, Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. *Mol Microbiol* 71 (6), 1402-1414 (2009).

[38] Lindner, S.E., De Silva, E.K., Keck, J.L., & Llinas, M., Structural Determinants of DNA Binding by a P. falciparum ApiAP2 Transcriptional Regulator. *Journal of Molecular Biology* 395 (3), 558-567 (2010).

[39] Magnani, E., Sjolander, K., & Hake, S., From endonucleases to transcription factors: evolution of the AP2 DNA binding domain in plants. *Plant Cell* 16 (9), 2265-2277 (2004).

[40] Riechmann, J.L. *et al.*, Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science* 290 (5499), 2105-2110 (2000).

[41] Allen, M.D., Yamasaki, K., Ohme-Takagi, M., Tateno, M., & Suzuki, M., A novel mode of DNA recognition by a beta-sheet revealed by the solution structure of the GCC-box binding domain in complex with DNA. *EMBO J* 17 (18), 5484-5496 (1998).

[42] Dutnall, R.N., Neuhaus, D., & Rhodes, D., The solution structure of the first zinc finger domain of SWI5: a novel structural extension to a common fold. *Structure* 4 (5), 599-611 (1996).

[43] Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T., & Kaneko, I., Transcription factor AP2-Sp and its target genes in malarial sporozoites. *Mol Microbiol* 75 (4), 854-863.

[44] Riechmann, J.L. & Meyerowitz, E.M., The AP2/EREBP family of plant transcription factors. *Biol Chem* 379 (6), 633-646 (1998).

[45] Jones, D.T., Taylor, W.R., & Thornton, J.M., The rapid generation of mutation data matrices from protein sequences. *Comput Appl Biosci* 8 (3), 275-282 (1992).

[46] LaCount, D.J. *et al.*, A protein interaction network of the malaria parasite Plasmodium falciparum. *Nature* 438 (7064), 103-107 (2005).

[47] De Silva, E.K. *et al.*, Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc Natl Acad Sci U S A* 105 (24), 8393-8398 (2008).

[48] Sakuma, Y. *et al.*, DNA-binding specificity of the ERF/AP2 domain of Arabidopsis DREBs, transcription factors involved in dehydration- and cold-inducible gene expression. *Biochem Biophys Res Commun* 290 (3), 998-1009 (2002).

[49] Cokol, M., Nair, R., & Rost, B., Finding nuclear localization signals. *EMBO Rep* 1 (5), 411-415 (2000).

[50] Franke-Fayard, B. *et al.*, A Plasmodium berghei reference line that constitutively expresses GFP at a high level throughout the complete life cycle. *Mol Biochem Parasitol* 137 (1), 23-33 (2004).

[51] Janse, C.J., Ramesar, J., & Waters, A.P., High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite Plasmodium berghei. *Nature Protocols* 1 (1), 346-356 (2006).

[52] Beier, J.C., Malaria parasite development in mosquitoes. *Annu Rev Entomol* 43, 519-543 (1998).

[53] Healer, J., Graszynski, A., & Riley, E., Phagocytosis does not play a major role in naturally acquired transmission-blocking immunity to Plasmodium falciparum malaria. *Infect Immun* 67 (5), 2334-2339 (1999).

[54] Kaushal, D.C. *et al.*, Monoclonal antibodies against surface determinants on gametes of Plasmodium gallinaceum block transmission of malaria parasites to mosquitoes. *J Immunol* 131 (5), 2557-2562 (1983).