CRISPR/Cas9 mediated deletion of genes encoding putative cell cycle regulators in *Plasmodium falciparum*

by

Hilde von Grüning

Submitted in fulfilment of the requirements for the degree:

Magister Scientiae Biotechnology
in the Faculty of Natural and Agricultural Sciences
University of Pretoria
Pretoria

September 2017

UNIVERSITY OF PRETORIA
DECLARATION OF ORIGINALITY

This document must be signed and submitted with every
essay, report, project, assignment, dissertation and/or thesis.

Full names of student:
..……
Student number:
……

Declaration

1. I understand what plagiarism is and am aware of the University's policy in this regard.
2. I declare that this …… (eg essay, report, project, assignment, dissertation, thesis, etc) is my own original work. Where other people's work has been used (either from a printed source, Internet or any other source), this has been properly acknowledged and referenced in accordance with departmental requirements.
3. I have not used work previously produced by another student or any other person to hand in as my own.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

SIGNATURE

STUDENT:.....……
SIGNATURE SUPERVISOR:…….

Contents

List of Figures

List of Tables

List of Abbreviations

ACT	Artemisinin-based combination therapies
AdoMetDC	*S*-Adenosylmethionine decarboxylase
ApiAP2	Apicocomplexan APETALA2 domain protein
Amp	Ampicillin
BSD	Blasticidin S deaminase protein
Cas	CRISPR associated genes
CDC	Cell division cycle
CDK	Cyclin-dependent kinase
CDT	Chromatin licensing and DNA replication factor
CRISPR	Clustered regularly spaced palindromic repeats
DFMO	D-L-α-difluoromethylornithine
DSB	Double-strand break
gDNA	Genomic DNA
GFP	Green fluorescent protein
GO	Gene ontology
GTS	Global technical strategy for malaria
hDHFR	Human dihydrofolate reductase
HDR	Homology directed repair
HPI	Hours post invasion
HR	Homology region
IDC	Intra-erythrocytic developmental cycle
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
IRS	Indoor residual spraying
ITN	Insecticide-treated mosquito nets
KAHRP	Knob associated histidine rich repeat protein
KO	Knockout
MAPK	Mitogen-activated protein kinases
MCM	Minichromosome maintenance complex
NIMA	Never in mitosis gene *a*
ODC	Ornithine decarboxylase
ORC	Origin of replication complex
PCR	Polymerase chain reaction
RNAi	RNA interference
gRNA	Guide RNA
TALEN	Transcription activator-like effector nucleases
WHO	World health organisation
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
ZNF	Zinc finger nucleases

Acknowledgements

I would like to express my profound gratitude to Prof. Lyn-Marie Birkholtz (University of Pretoria, SA), my supervisor. I am incredibly thankful for her mentorship, her time, heartfelt motivation, enormous dedication and the extraordinary contributions she made to help me become an independent scientist. I thank her for believing in me, for giving me opportunities that would otherwise have been outside my reach and for being a guiding inspiration to me.

I sincerely thank Dr. Jandeli Niemand, my co-supervisor, for her insightful advice, support and effort in proofreading of this dissertation.

I thank all the members of the Wellcome Trust Sanger Institute Malaria Programme (Cambridge, UK), with my sincere appreciation extended to Dr. Marcus Lee, for welcoming me to his fantastic laboratory, his kindness and involvement in this project. I thank the remarkable Dr. Sophie Adjalley, for mentoring me and for all the enlightening discussions we had. Without their passionate participation and input, this project could not have been successfully conducted. Finally, to Bettina Lengger, for all the fun adventures we had in Cambridge and beyond.

To all the members from the M^2PL laboratory, thank you for all the technical and moral support you have given me.

I am grateful for the financial support given to me by the National Research Foundation and University of Pretoria by funding this degree through the NRF Grantholder-linked bursary and the University of Pretoria Study Abroad bursary, respectively.

To every person who was involved in this journey, I am deeply grateful for your support. Specifically, I thank my partner in life, Pierre Hugo, for his continuous enthusiasm, understanding and encouragement given for me to pursue my dreams. My mother, Marida von Grüning, for her persistent motivation and for teaching me to always persevere and to strive to be better. To “Tannie” Louise Fest, for her consolation and empathy during each moment of this journey. It is my greatest pleasure to just make each of them proud.

Finally, I am thankful for science and scientists across the world pursuing to make the world a better place. We have the privilege to study the magnificent natural world with our inborn curiosity for which I am truly grateful.

Wait | Three days | Meridian

Summary

Functional genomic tools can be used to interrogate unique features of parasite biology that could be marked for future intervention strategies. One such unique biological feature is the atypical cell cycle of the *Plasmodium falciparum* parasite, with a particular interest in the key regulators that modulate the parasite's cell cycle progression. The orderly progression of the cell cycle of the parasite allows it to undergo a vast expansion of parasite numbers during a malaria infection and cause pathologies in infected individuals. However, the precise control mechanisms and functional cascades involved in the *P. falciparum* parasite's unusual cell cycle have not yet been fully elucidated. A recent study showed that gene expression profiles of the parasite switch between states of cell cycle arrest (quiescence) to cell cycle re-entry (proliferation). Global transcriptome expression data from this study was used to identify potential cell cycle regulators. The study identified a number of putative cell cycle regulators and their predicted functional interacting partners, particularly transcription factors and several members of the origin of replication complex. The study subsequently aimed to functionally validate these putative regulators in a reverse genetics approach by creating targeted disruption of the putative regulatory genes using the CRISPR-Cas9 system. Preliminary knockout studies indicated possible essential phenotypes of these putative cell cycle regulators in *P. falciparum* parasites. Given that cell cycle elements of the *P. falciparum* parasites are unique and divergent from those of the human host, the information provided in this study dissects the role of regulators in cell cycle modulation in the parasite. Not only is this important to understand parasite biology, but these cell cycle regulators are attractive potential sites for chemical interference of parasite proliferation and thereby provide as novel drug targets for antimalarial discoveries.

Chapter One: Literature review

1.1 Malaria as infectious disease

Malaria is a preventable and treatable parasitic disease, and significant progress has been made to combat malaria since the year 2000 with case indices and mortality rates reduced by 41% and 62%, respectively. However, malaria remains one of the most deadly infectious diseases threatening global health (1) with a staggering 212 million cases of malaria infections reported in 2015, with ~429 000 cases resulting in death (1). Vulnerable populations, especially pregnant woman, young children and immunocompromised individuals carry the heaviest burden of malaria as evidenced by ~70% of all malaria deaths occurring in children under the age of five (1). Adverse effects of the disease include the characteristic cyclic fever and chills, anaemia and low neonatal birth weight (2–4).

Sub-Saharan Africa remains the primary region affected by malaria with 114 million cases of malaria infection and where 92% of all malaria-caused deaths reported in 2015 for the World Health Organisation (WHO) African Region occurred (Fig. 1) (1). This translates to 90% of all malaria cases occurring in Africa, followed by the WHO South-East Asia (7%) and Eastern Mediterranean Regions (2%) (1).

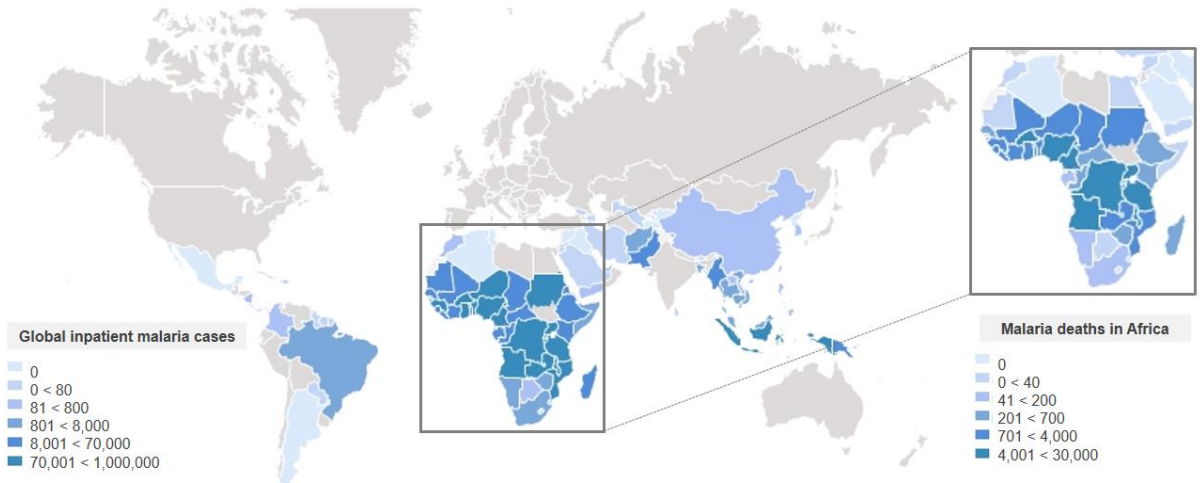

Figure 1. Number of global inpatient malaria cases. Inset: Number of deaths in Africa caused by malaria. Source: map generated *de novo* from the WHO malaria mapper (http://www.who.int/malaria/publications/world_malaria_report/global_malaria_mapper/en/).

The socio-economic impacts of malaria pose a substantial strain on developing countries where basic health care is typically substandard; there is a close relationship between malaria incidence, transmission of malaria parasites, susceptible populations, and poverty, thus classifying malaria as a poverty related disease (5). Malaria foci correlate with a large loss in gross domestic product *per capita* (6). Despite the rapid and dynamic population changes that occur in sub-Saharan Africa due

to, e.g. rapid urbanisation, depopulation of rural communities, economic development and increases in non-communicable diseases, progress in malaria control has been insufficient due to a multitude of factors. These include a lack of, or restricted access to, preventative measures such as mosquito repellent bed nets, poor public health systems and education measures and inadequate healthcare facilities (7,8). However, a number of countries are focusing their efforts towards moving from malaria control to active strategies aimed at malaria elimination, defined by the WHO as regions with zero local cases of malaria transmission for a period of three years. Currently, malaria is endemic to 91 countries in the world, down from 108 countries in 2000 (1). This global elimination goal has galvanised renewed efforts geared toward malaria control and generation of effective and multifaceted intervention strategies, deployed in an integrative fashion.

Key intervention strategies include the use of long lasting insecticide-treated mosquito nets (ITNs) and indoor residual spraying (IRS) for vector control and artemisinin-based combination therapies (ACT) for chemotherapeutic parasite control (1). Vector control is however threatened by the development of insecticide resistance in mosquitoes in 60 countries, most predominantly to the pyrethroids currently used in ITNs (1). Fortunately, a WHO-coordinated five-country evaluation of pyrethroid resistance reported a standing and continuous efficacy of ITNs (1).

Since 2010, the WHO has recommended fixed-dose ACTs as the frontline, gold standard in the treatment of malaria (9). ACTs consist of an artemisinin derivate with rapid action, which quickly clears the majority of circulating parasites, combined with a long-lasting partner drug to completely remove residual parasites after the patient has recovered (10). However, similar to all antimalarials previously used, parasite resistance has developed to artemisinins as well as its partner drugs. This is particularly evident in countries from the Greater Mekong subregion (Cambodia, Lao People's Democratic Republic, Myanmar, Thailand and Vietnam) (1,11) where a 26% treatment failure of dihydroartemisinin-piperaquine was seen in 2015 in Vietnam (12), and resistance to four different ACTs (1) has been reported in Cambodia. Of great concern is the recent emergence of mutant alleles associated with artemisinin resistance in Africa, confounded by reports suggesting delayed parasite clearance after ACT treatment (13) on the continent. Therefore, the efficacy of ACTs is being monitored closely in Africa since there are no therapeutic alternatives ready for clinical deployment for the next ~5 years.

The WHO has set targets in the Global Technical Strategy for Malaria 2016–2030 (GTS), which include 1) the reduction of malaria incidence and mortality globally by at least 90%, 2) to eliminate malaria in at least 35 countries and 3) to prevent re-emergence of malaria in malaria-free countries (1). For these targets to be achieved, total funding for malaria programmes and research requires an estimated US$ 6.4 billion by 2020 (1). Considering the potential threats to the significant progress

made in the past two decades in reducing incidences and mortality in malaria, new strategies are required to manage and optimise current antimalarial drugs. This entails not only maximising the life span of currently used ACTs but also to developing alternative chemotherapies (14). Transmission blocking drugs are top priority; drugs targeting the asymptomatic liver stages and gametocytes are essential (15).

Due to the development of resistance, there is an urgent need for new antimalarial drugs with novel mechanisms of action for effective control and eradication of malaria (16). The Medicines for Malaria Venture (a product development partnership coordinating global malaria drug discovery efforts) predicts that 48 new chemical entities with new modes of action are required for a >90% probability of obtaining a combination of 2 new drugs useful in the clinical context (17). Therefore, we need a massive effort to increase our knowledge base on the basic biology of the parasite. We need to identify essential but parasite-specific processes in the malaria parasite that could be exploited in drug discovery and development programmes. Thus, in the following sections, unique and atypical features of the parasite's biology will be introduced concerning the parasite life cycle, cell cycle and strategies to study the parasite biology and essential processes.

1.2 The malaria causing parasite, *Plasmodium*

The malaria disease is caused by intra-erythrocytic infection of parasites from the genus *Plasmodium* including the human parasites *P. malariae*, *P. ovale*, *P. knowlesi*, *P. vivax* and *P. falciparum. P. falciparum* parasites are the most lethal and cause 99% of all malaria mortalities (1). Plasmodia are eukaryotic, endoparasitic protists from the phylum *Apicomplexa*, of which all members grow and replicate in a specialised compartment within erythrocytes known as the parasitophorous vacuole (18) and are characterised by the presence of an apicoplast, a structurally unique organelle similar to chloroplasts (19).

1.2.1 *Plasmodium* life cycle

Plasmodium parasites have a complex life cycle involving two hosts. In the female *Anopheles* mosquito, sexual development of the parasite occurs, while in the human host, asexual proliferation takes place during the intra-erythrocytic developmental cycle (IDC, Fig. 2) (20). Infection is initiated when 15-123 sporozoites are injected by the feeding mosquito as it takes a blood meal, after which sporozoites glide through the dermis into peripheral blood circulation and migrate to the liver sinusoids and infect hepatocytes (18). In the liver, exo-erythrocytic schizogony, an asexual replication process characterised by nucleic acid expansion, occurs (18). Thousands of hepatic merozoites that developed from a single sporozoite are consequently released into the blood stream in merosomes, vesicles filled with parasites (21,22). Blood stage infection is established when

merozoites are released to invade erythrocytes, thus initiating the asexual IDC. The pathophysiology and clinical manifestation of infection are presented during the asexual blood stages of the parasite. During the IDC, a single merozoite develops into a ring stage parasite within ~6 hours after erythrocyte invasion, during which time the parasitophorous vacuole forms. This is followed by development into the highly metabolically active trophozoite stage, during which new permeability pathways are induced that allow for substrate uptake from the surrounding nutrient rich medium and the efflux of metabolic waste products (23). Importantly, during this stage, erythrocyte haemoglobin is proteolysed in the acidic digestive vacuole of the parasite to provide free amino acids (except isoleucine) which are utilised in protein synthesis (24) in addition to maintaining osmolarity in the intra-erythrocytic environment. Crystalised haemozoin produced from haematin (derived from haemoglobin) results in pigmentation of trophozoite stage parasites (25).

P. falciparum parasites are obligate intracellular auxotrophs, as these parasites are entirely dependent on the host for purine nucleotides required for DNA replication due to their inability to synthesise purines *de novo* (26). Approximately 33-36 hours post-invasion (hpi), the schizont stage of the parasite appears as a result of asexual fission (i.e. schizogony). During schizogony, multiple rounds of DNA replication and asynchronous nuclei fission occur, resulting in a multi-nucleated, polyploid schizont (27). Subsequent segmentation and cytokinesis of each nucleus occur, leading to the formation of 16-32 haploid daughter merozoites that are released by erythrocyte rupture 42-48 hpi, completing a round of the IDC (28). Each released merozoite proceeds to infect a new erythrocyte, allowing for persistent cycles of infection. This is coordinated with the rapid amplification of up to 6×10^{10} parasites per infected individual in untreated, immune-naïve individuals (29).

Sexual differentiation of the *P. falciparum* parasite is fundamentally responsible for the continued transmission of the parasite between the host and the vector; it reinitiates the life cycle and ensures the spread of the disease (Fig. 2) (30). Typically, less than 10% of asexual schizonts are committed to gametocytogenesis from a previous IDC cycle (20). Differentiation of the parasite characterises gametocytogenesis, a process by which the parasites develop through five distinct morphological stages (31), ultimately resulting in mature, transmissible female and male gametocytes after ~14 days (32). Within the mosquito vector, gametogenesis is triggered by the drop in temperature, a rise in pH and the mosquito derived xanthurenic acid (33). Male gametogenesis ensues with three rapid rounds of DNA replication and flagella assembly (34) resulting in eight motile flagellated male microgametes (35,36). Each haploid microgamete fertilises a female gamete, generating a diploid zygote which rapidly undergoes meiotic recombination (33). The zygote subsequently develops into an ookinete (37). The ookinete transforms into oocysts that contain multiple haploid sporozoite progeny within the cell. After release from the oocysts, sporozoites are ready to reinitiate the life cycle after migrating to the mosquito's salivary glands, ready for depositing into a human dermis.

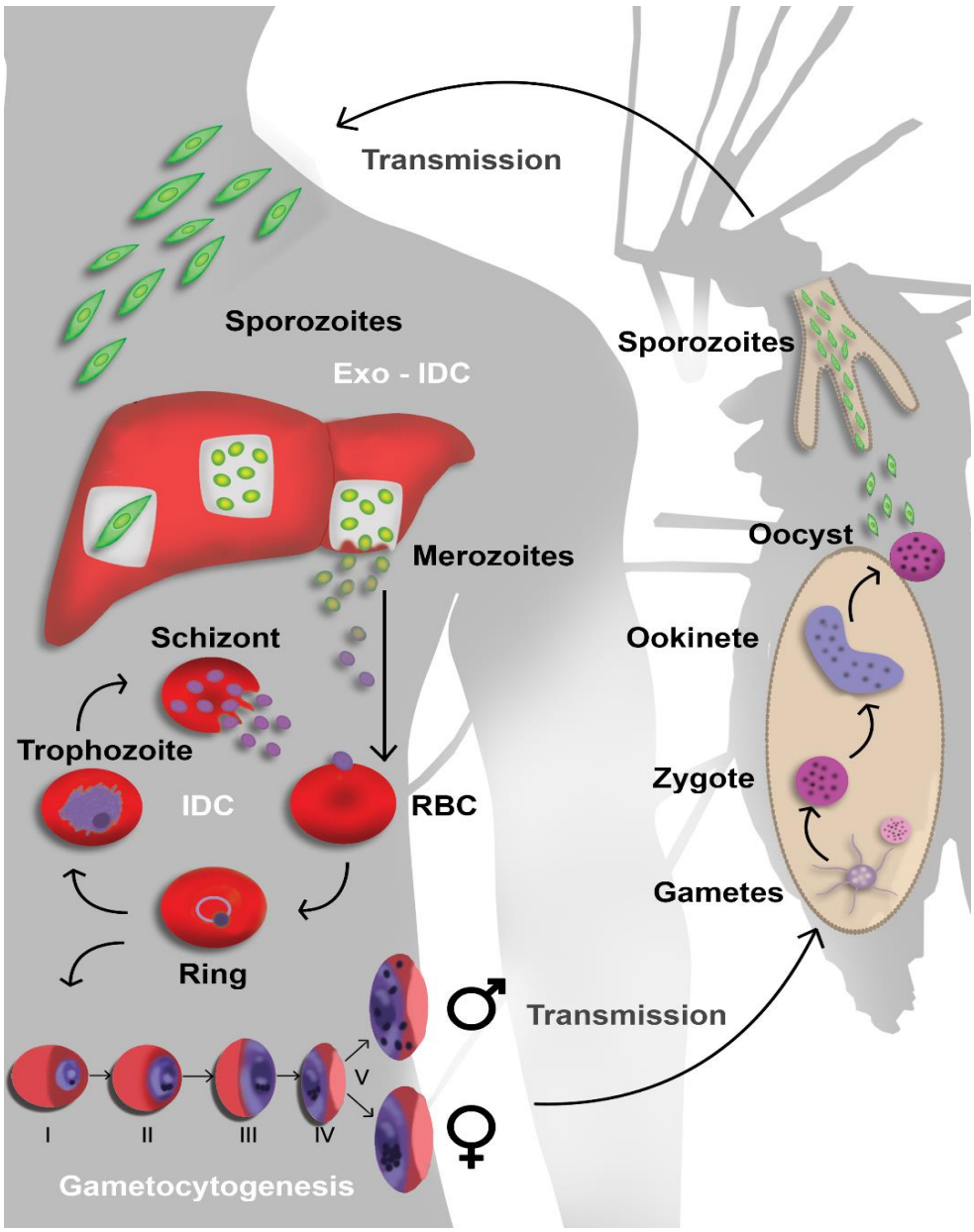

Figure 2. The *Plasmodium* parasite life cycle. During its life cycle, the parasite undergoes asexual and sexual development in a vertebrate host and an *Anopheles* mosquito vector respectively. Sporozoites delivered from the mosquito during a blood meal where it travels to the liver and undergoes asexual exoerythrocytic development. Merozoites released from the hepatocytes initiate the intra-erythrocytic developmental cycle (IDC). Ring stage parasites develop into trophozoite and schizont stages that release merozoites for re-infection in a cycle that last 48 hours. A subset of asexual parasites differentiates and commit to sexual development during gametocytogenesis. Male and female gametocytes are transmitted to a feeding mosquito, after which gametocytes differentiate into gametes. The microgamete (male) fertilises the macrogamete (female), producing a zygote that develops into an ookinete in the mosquito midgut lumen. After crossing the endothelial gut lining, oocysts release sporozoites that migrate to the mosquito salivary glands, for transmission to the human host.

A feature that contributes to the pathogenesis of a *P. falciparum* parasite's infection is its ability to proliferate and undergo a mass expansion of parasite numbers, ultimately causing adverse effects to the infected individual. Periods of intensive rounds of DNA replication and asynchronous nuclear divisions include stages of the sporogony within the vector and schizogony within the vertebrate host (38). Complex extrapolations can be made regarding the life cycle features and the associated cell cycle of the parasite. The eukaryotic and *P. falciparum* cell cycle and the regulation thereof will be examined in the next section.

1.3 The eukaryotic cell cycle

The eukaryotic cell cycle is extraordinarily controlled, resulting in a systematic replication of genetic material and subsequent cell division. In the classical cell cycle, four distinct phases occur (Fig. 3) (39) G_1 (initial gap phase), where cells physically increase in size and synthesise RNA and proteins required for the subsequent DNA synthesis steps in the S phase (40). This is followed by a second gap phase (G_2) where cells increase in size and produce organelles and proteins required for mitosis (M). Collectively, the G_1-, S- and G_2-phases are termed interphase. Mitosis is characterised by prophase, metaphase and anaphase where chromosomes condense, mitotic spindle bodies form and sister chromatids divide. This forms two new daughter nuclei, each inheriting a single copy of the genome from the original nucleus. Once segregation occurred, the mitotic spindle dissembles and chromosomes condense during telophase, after which the nuclear envelope restructures. Cytoplasmic division occurs (cytokinesis), resulting in two daughter cells (39).

Transition to the next cell cycle phase is governed by cell cycle checkpoints, critical boundaries that ensure completion of each particular cell cycle phase before allowing progression. The G_1/S checkpoint oversees cell growth and transcription of genes required for DNA synthesis in the proceeding step (41). The G_2/M checkpoint ensures completion of high fidelity DNA replication, while the mitotic M-phase (spindle) checkpoint ensures complete spindle formation and chromosome attachment at the metaphase plate. Non-functional checkpoints lead to uncontrolled cell proliferation, a typical trademark of cancer (42,43). Unlike terminally differentiated cells, cells confined to a quiescent (G_0) phase may re-enter the cell cycle upon external stimuli (44).

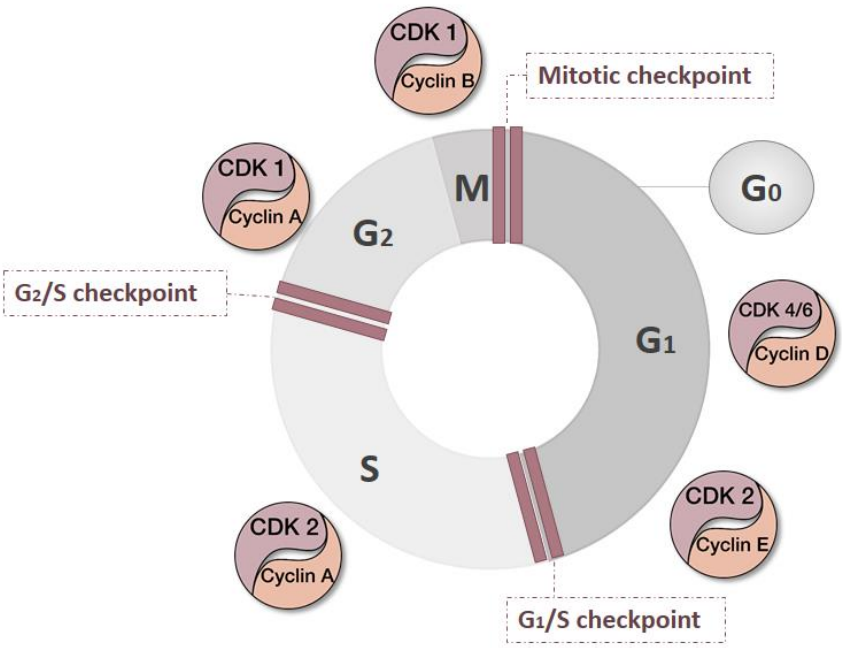

Figure 3. The eukaryotic cell cycle and associated regulatory cyclins and CDKs. Four major phases of the eukaryotic cell cycle include the first gap phase (G_1), the DNA synthesis phase (S), the second gap phase (G_2) and the mitotic phase (M). A quiescent, non-replicative G_0 phase is conserved after a cell exits the cell cycle. Cyclins fluctuate in specific cell cycle phases during cell cycle progression. As one cyclin is synthesised in a cell cycle phase, the previous phases cyclin is degraded. Congruently, each cyclin activates its corresponding cyclin-dependent kinase to modulate gene expression during a particular phase.

A family of protein kinases, termed the cyclin-dependent kinases (CKDs) modulate cell cycle progression (Fig. 3) by associating with their regulatory (cyclin) subunit and subsequently phosphorylating key proteins involved in the cell cycle. Each cell cycle phase associates with a specific CDK-cyclin complex that oscillates throughout the cell cycle where each complex stimulates the activity of the successor complex and represses the precursor complex (29). Progression through the cell cycle is a result of the concerted action of many other molecules, some of which act as cell cycle checkpoint inhibitors and others that indirectly affect cell cycle progression. The latter includes a group of ubiquitous organic cationic polyamines that have essential roles in cell differentiation and proliferation and subsequently cell cycle regulation (44). At physiological pH, the polycationic nature of polyamines enable interactions with polyanionic molecules in the cells, such as nucleic acids, thereby modulating their function (45). During the eukaryotic cell cycle, biosynthesis of polyamines occurs bicyclically with the activities of the two primary biosynthetic enzymes (ornithine decarboxylase [ODC] and *S*-adenosylmethionine decarboxylase [AdoMetDC]) peaking at the G_1/S and S/G_2 transitions (46). Cells have an absolute requirement for polyamines to transition through the G_1/S checkpoint into the S-phase (47). When polyamine biosynthesis is inhibited (e.g. by targeting either ODC or AdoMetDC activities), cytostasis and cell cycle arrest occur at this point in the cell cycle (48).

1.4. The *Plasmodium* cell cycle

1.4.1 Cell cycle progression in the *P. falciparum* parasite

The cell cycle during both the sexual and asexual development of the *P. falciparum* parasite is unusual and it is difficult to associate cell cycle compartments with the complex life cycle stages of the parasite. However, convincing evidence suggests that the parasite's cell cycle is atypical and includes unusual features such as asynchronous nuclear divisions within a given schizont, distinct mechanisms of organelle segregation and morphogenesis of daughter merozoites (reviewed in (49,50). During the IDC, single-nucleated rings and early stage trophozoites are considered to be in the G_1-phase as seen by the increase in RNA and protein levels (54,55), similar to the G_1a to G_1b transition in other eukaryotic cells (Fig. 4). DNA replication ensues in mature trophozoites and continues until late stage schizonts, concurrent with S-phase (53).

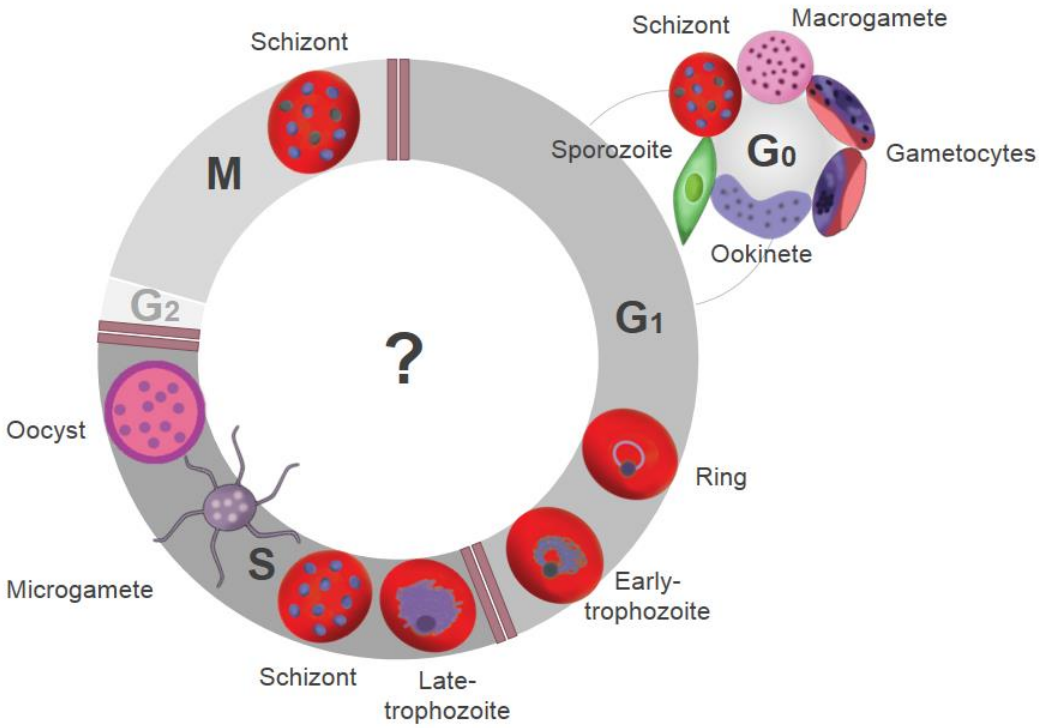

Figure 4. *P. falciparum* parasite associated life- and cell cycle stages. Sporozoites, ookinetes, male and female gametocytes, and macrogametes (female) are associated with the G_0-phase. Merozoites transition from ring stages to early trophozoite stages in the G_1-phase, while late trophozoites transition to S-phase. In the S-phase, late trophozoite stages, microgametes (male), schizonts and oocysts undergo DNA synthesis. Nuclei within a single schizont transition from S-phase to either a non-replicative G_0-phase or divide again to the M-phase.

DNA replication and nuclear division occur asynchronously in *P. falciparum* in a process known as endocyclic schizogony. The latter is characterised by 3-4 rounds of continuous DNA synthesis in the absence of cytodieresis and cytokinesis, alternating with a rapid M-phase without a much-extended G_2 phase. This results in a multinucleated schizont (>2N); nuclear division in a single schizont occur asynchronously (54) except for the last round of mitosis, which occurs synchronously for all nuclei, followed by a single segmentation step releasing haploid daughter merozoites after cytokinesis (55–57). Unlike other model organisms, *P. falciparum* undergoes endomitosis characterised by the lack of chromosome condensation and nuclear envelope disintegration as spindles remain intact within the nuclear envelope. For a nuclear division to occur in late schizogony, the kinetochores attached to the parasite's chromosomes connect to microtubules to form the mitotic spindles which originate from the centriolar plaques. This is also known as the *Plasmodium* microtubule organising centre embedded in the nuclear envelope (58). Centriolar plaque duplication occurs between 24-26 hpi, before the start of the S-phase. DNA is replicated up to three to four times during the S-phase. As the S-phase concludes, centriolar plaques disconnect as they migrate to opposite poles of the nucleus (characteristic of a G_2-phase), which allows spindle formation and nuclear division to occur. One nucleus inherits the 'mother' centriolar plaque and will re-enter S-phase before the daughter nuclei that inherited the 'daughter' centriolar plaque enters a lagging S-phase. This leads to the asynchronous nuclear division or odd numbers of nuclei in early schizonts

(58,59). However, an equal number of nuclei are produced at the final stages of schizogony, as 'lagging' nuclei are allowed to complete current nuclear divisions, suggesting the presence of a mitotic checkpoint. After several mitotic divisions, some nuclei become arrested in a G_0/G_1-like phase, further promoting the notion that schizogony may be controlled by DNA replication regulators (58). After a brief uncharacterised gap phase, nuclei within a single schizont undergo segmentation followed by mass cytokinesis whereby the cytoplasm of the schizont is divided. The number of merozoites produced from a single schizont may vary up to threefold, from 8-24 merozoites per schizont (54). The host erythrocyte ruptures and releases multiple individual merozoites for reinvasion, effectively permitting re-initiation of the blood stage *P. falciparum* infection.

A model proposed by Sinden and Smalley in 1979 draws parallels between gametocytogenesis and the cell cycle of *P. falciparum* parasites, however, the current view is that haploid gametocytes are arrested in the quiescent G_0-phase of the cell cycle as seen by the lack of DNA synthesis and incorporation of tritiated hypoxanthine, only undergoing genome replication once within the mosquito midgut (60). Beyond these proposed cell cycle compartments, the link between gametocytogenesis and cell cycle regulation is poorly understood.

The exact molecular mechanisms and signalling dynamics that regulate the cell cycle during the IDC remain underexplored. However, the parasite must tightly regulate its cell cycle in response to its environment to complete its life cycle.

1.4.2 Cell cycle regulators in the *P. falciparum* parasite

P. falciparum IDC development correlates with a tight control of phase-like transcriptional expression (61) and a dynamic epigenome (62). ApiAP2 DNA-binding transcription factors are thought to coordinate transcriptional control. These proteins form part of the APETALA2/ethylene response factor (AP2/ERF) family which consist of only 27 members, limited to the *Plantae* and *Apicocomplexa* phylums. This limited repertoire of transcription factors is thought to allow strict control of gene expression in the parasite (63). However, homologs of well-known eukaryotic cell cycle regulators have been identified. These include *P. falciparum* protein kinase 5 (PfPK5, homolog of human CDK1 and CDK5,) PfPK6 (human CDK1/MAPK), MO15-related kinases (PfMRK, CDK7 homologue) and several cdc2-related protein kinases (PfCRKs) (64). Some of these cell cycle regulatory genes have been investigated in a reverse genetic approach. Approximately half of the total *P. falciparum* kinome comprising of 65 eukaryotic protein kinases have been shown to be likely essential for parasite life cycle progression and proliferation (65). These included the DNA replication essential CDKs, Aurora and NIMA (never in mitosis gene *a*, Neks) kinase genes (66–68).

Several atypical CDKs, CDK-related proteins, cyclins and other CDK regulators have also been described (52) in addition to novel cell cycle related proteins in *P. falciparum* that do not share homology to known eukaryotic cell cycle regulators. These include the cyclin-dependent serine-threonine kinases (2), cGMP-dependent protein kinase (PKG) and calcium dependent protein kinase 1 (CDPK1) (69). Numerous proteins involved in phosphorylation signalling, DNA replication and repair, and transcription have also been proposed to function as cell cycle regulators in addition to their innate roles (70). It is rather postulated that the parasite employs evolutionary distinct cell cycle regulatory mechanisms. However, the functional involvement and underlying mechanisms of many of the molecular components needed for cell cycle regulation is not clear in *P. falciparum* parasites.

1.4.3 Investigating cell cycle regulation of *P. falciparum*: work preceding this study

One of the primary methods to study cell cycle regulators in other organisms is to synchronise cells to a particular cell cycle phase/compartment, then reverse this block and observe various features of both the arrested and the emerging cell. However, the inability to tightly synchronise the parasite to a particular cell cycle compartment *in vitro,* with a very narrow window and age-range of parasites (57,71), has complicated such mechanistic evaluation of cell cycle progression in *P. falciparum*.

Disruption of polyamine biosynthesis (regulators of normal cell cycle progression) allows defined cell cycle compartmentalisation in cancer cells, inducing a G_1 arrest. This arrest is reversible and has allowed cell cycle kinetic evaluations in these cells (46,72). The method to disrupt polyamine biosynthesis has been applied successfully to tightly (3 to 5 h window) synchronise *P. falciparum* parasites to early trophozoite stages (73), and this block is known to be cytostatic and reversible (48).

In a recent study by Niemand *et al*., this system was applied to *P. falciparum* parasites, allowing comprehensive profiling of a parasite under cell cycle arrest and re-entry (Niemand *et al.*, submitted Nature Communications). D-L-α-difluoromethylornithine (DFMO) is an anti-proliferative drug used as anti-cancer (45) and anti-parasitic agent in diseases such as colorectal cancer and African sleeping sickness (caused by parasites from the *Trypanosoma* genus) (74–76). DFMO, which selectively inhibits the *P. falciparum* bifunctional *S*-AdoMetDC/ODC, resulting in complete inhibition of polyamine biosynthesis (77). Parasites were morphologically arrested 26 hpi in an early trophozoite stage before DNA replication. This life cycle arrest corresponded to a cell cycle arrest as the block 1) was reversible, parasite proliferation resumed upon addition of putrescine to parasite cultures (even in the presence of DFMO); 2) did not cause overt stress responses with no statistical significant difference in the ability of parasites to induce gametocytogenesis and 3) cells remained viable in the presence of putrescine. The cell cycle compartment associated with the arrest was

identified as the G_1/S transition checkpoint, based on DNA and RNA content analyses. Global expression profile analyses were used to evaluate the processes affected as a result of the compartmentalised cell cycle arrest. This was followed by a temporal evaluation of the molecular processes involved in enabling parasites to resume their cell cycle progression (Fig. 5).

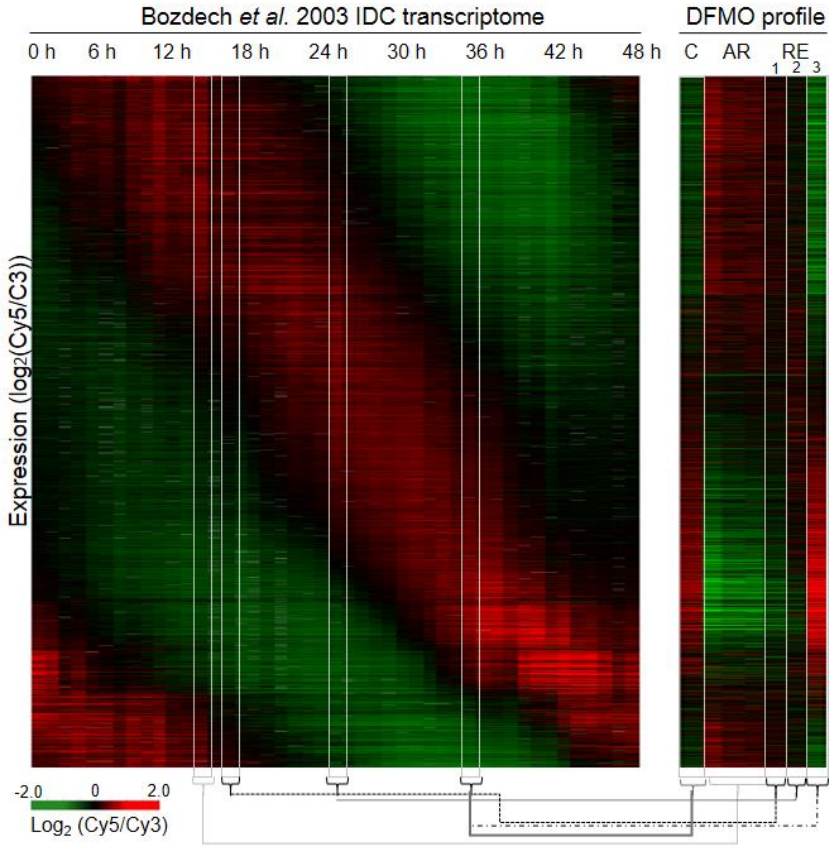

Figure 5. The transcriptomic fingerprint of cell cycle arrest and re-entry in *P. falciparum* 3D7 parasites. Global comparison of the first published IDC by Bozdech *et al*. (78) reveal a correlation between each of the treatment conditions and the IDC. The global comparison of the normal IDC gene expression profile as compared to DFMO arrested and putrescine treated gene expression profile is shown. The control (C) transcriptome correlates to parasites that have passed the G_1/S checkpoint at 36 h (solid dark grey line), while the arrested (AR, solid light grey line) correlates to a morphological early trophozoite stage that has not entered S-phase. After re-entry of the cell cycle, parasites resumed proliferation with gene expression profiles at 3 different time points shown (dashed black lines) as putrescine parasites progressed through the expected G_1/S checkpoint.

The data from this study provides the first view of the parasite under compartmentalised cell cycle arrested conditions, as well as the mechanisms induced to allow the cells to re-enter the cell cycle. The regulatory involvement of known eukaryotic cell cycle control genes was identified in addition to the first association of parasite-specific genes with potential cell cycle regulation. As expected, genes associated with DNA replication showed decreased expression during cell cycle arrest at the G_1 phase, and significantly increased expression after cell cycle re-entry, even from the earliest time point evaluated. Notably, genes involved in the pre-replicative complex that functions in licensing DNA for replication (including members of the origin of replication complex [ORC] and the heterohexameric minichromosome maintenance complex [MCM]) was profoundly affected, suggesting that these genes may modulate cell cycle progression apart from serving merely as DNA replication initiation factors.

Genes with potential cell cycle regulatory function include members of the kinase superfamily that showed marked transcriptional responses and were also present as some of the earliest genes responding to the cell cycle re-entry. Additionally, several other gene families identified include members of the ApiAP2 family of transcription factors. The functional validation of these putative cell cycle regulators in *P. falciparum* parasites is imperative. Modern reverse genetics tools allow association of single gene function to a specific phenotype. Therefore, genes implicated as putative cell cycle regulators can be functionally validated through the use of reverse genetic screening via gene knockout.

1.5 Functional validation of cell cycle regulators

1.5.1 Genome editing as a tool to study gene functions

Tools for efficient genome editing have become increasingly accessible and robust (79,80). However, genome editing in *P. falciparum* parasites has typically not been as successful as in other organisms with various techniques such as Bxb1 integrase system (81), piggyback transposon mutagenesis (82), FKBP destabilisation domain post-translation modulation (83,84), FRT/FLP-recombinase system (85), diCre conditional gene knockout system (86), the riboswitch post-transcriptional knockdown system (87), zinc-finger nucleases (88) and the TetR-aptamer system (89) employed to various degrees of success (see glossary). *P. berghei* parasites, in a murine malaria model, is often used as a tool in reverse genetic studies to study pathogenesis as it is considered to be substantially more permissive to genetic manipulation with higher success rates than in *P. falciparum* parasites (80). However, for the human parasite, gene manipulation has been inefficient and restricted due to the lack of selectable markers, ineffective delivery of foreign DNA into the parasite and inherent difficulties arising from the genome plasticity of the parasite's AT-rich genome (90).

The discovery of the genome editing tool, CRISPR-Cas9, has revolutionised multiple academic fields within medicine and biotechnology (91). Clustered regularly interspaced short palindromic repeats (CRISPR) and their associated proteins (Cas) collectively form an inheritable acquired immune system in prokaryotes and nanoarchea (92,93). These organisms evolved a nucleic acid based adaptable immunity that records a 'memory' of previous infections (i.e. from phages) and subsequently recognises and degrades exogenous genetic elements upon re-infection through RNA-guided endonuclease activity (94–102).

The natural CRISPR-Cas9 mediated defence system functions by the acquisition of 30-40 bp protospacers, i.e. nucleotide sequences from the foreign DNA, which integrates into the CRISPR

locus through Cas1 or 2 integrase activity. These protospacers are typically separated by ~26 bp repetitive protospacer adjacent motifs (PAM) nucleotide sequences. The complete CRISPR locus includes an AT-rich leader sequence, CRISPR RNA (crRNA) and associated Cas enzymes. crRNAs are integrated into effector Cas9 endonucleases via the PAM sequence and then guide this complex to new invading complimentary DNA or RNA sequences. Subsequent DNA cleavage occurs 3-8 bp upstream of the PAM sequence which effectively silence these genes (Fig. 6) (101).

After spacer integration, a pre-crRNA molecule is transcribed from the CRISPR locus in addition to the *cas9* gene. A *trans*-activating CRISPR RNA (tracr-RNA) hybridises with direct repeats of the crRNA, forming an RNA duplex that is processed and cleaved by endogenous RNase III in the presence of the Cas9 endonuclease, completing crRNA maturation (99,101,103,104). Next, the tracrRNA:crRNA:Cas9 effector complex is directed towards target DNA where PAM licensing initially occurs. If the PAM is confirmed, the non-template strand is displaced, forming an R-loop (105,106). Next, the single guide RNA (gRNA) interrogation occurs, where Watson-Crick base pairing occurs between the template and the guide sequence (107,108). When the correct targeting sequence is confirmed, the Cas9 endonuclease cleaves site-specific DNA strands through two nuclease domains, RuvC-like (RNase H fold from the retroviral integrase superfamily) and a HNH domain (cleaves DNA strand complementary to guide) (109).

There are two classes, six types, and nineteen subtypes of CRISPR systems (108), each defined by CRISPR sequence and length, *cas* genes, locus architecture, biosynthesis and composition of crRNA, and the effector molecule. Of all the CRISPR systems, class two, type II system have been characterised in detail (Fig. 6) (103).

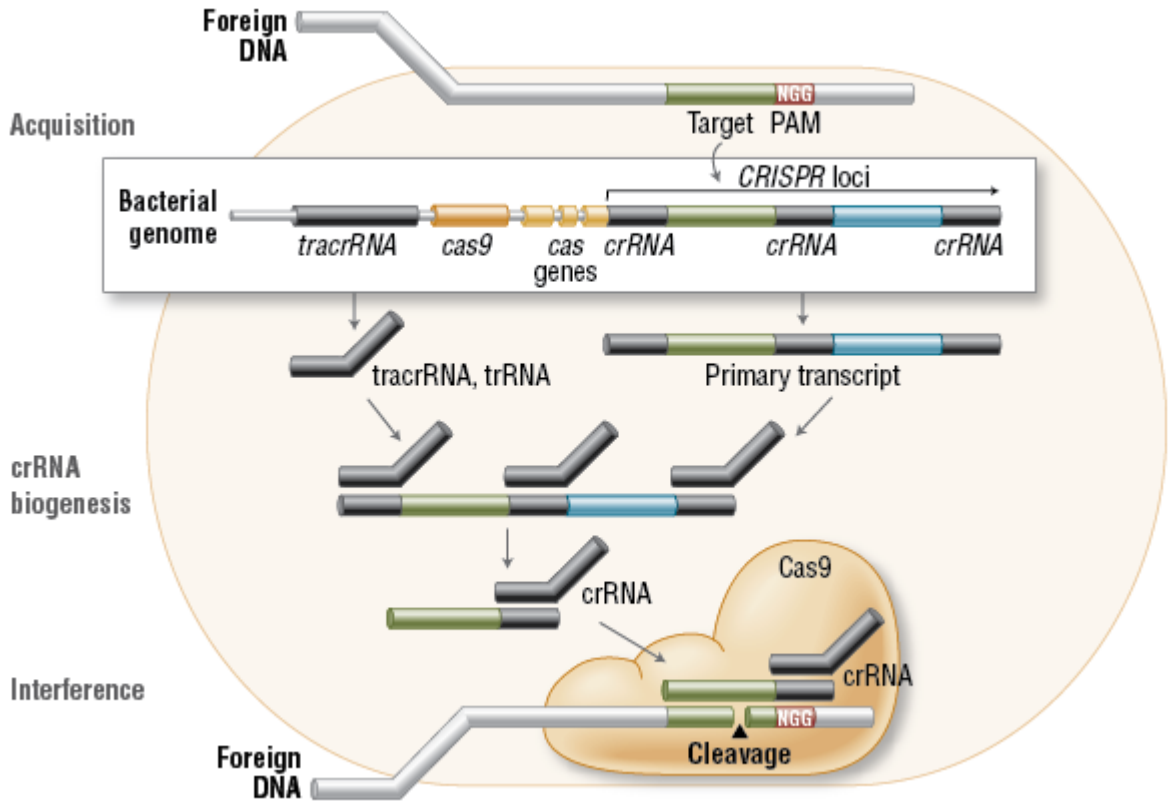

Figure 6. Overview of the type II CRISPR-Cas9 mediated bacterial adaptive immune system *in vivo*. During the acquisition phase, spacer sequences (green) derived from foreign DNA are excised and inserted into the CRISPR array on the bacterial genome. During crRNA biogenesis, *cas* genes are transcribed and translated. Trans-activating crRNA (tracrRNA) are transcribed to base pair with repeat sequences of the primary transcript, followed by host RNA III cleavage within the repeat region. The interference phase is characterised by the dual crRNA:tracrRNA duplex that guides the Cas9 endonuclease to induce double-stranded breaks at the target DNA site (green) approximately three nucleotides 5' of the PAM sequence (red). Copyright permission granted NEB (https://www.neb.com/tools-and-resources/feature-articles/crispr-cas9-and-targeted-genome-editing-a-new-era-in-molecular-biology, accessed 10 June 2017).

In 2012, this nucleotide-based recognition system was successfully used to provide a straightforward approach to genome editing by generating customisable nucleases for gene targeting. By synthesising an active CRISPR-Cas endonuclease complex, Jinek *et al.* demonstrated that the introduction of an exogenous construct into a heterologous organism effectively simulates the RNA-guided endonuclease system to target almost any gene in the organism, initiating the field of CRISPR-Cas genome editing in biotechnology (101). One of the breakthroughs in the generation of CRISPR-Cas9 as genome editing tool was the realisation that the crRNA and tracRNA could be engineered and combined into a single chimeric guide RNA (101), which allows for enhanced targeted gene editing in the genomes of numerous eukaryotic organisms (110,111). The application of CRISPR-Cas to genome engineering has become ever increasingly applied since the first description in 2012 due to the perceived efficacy, specificity, versatility, reversibility, and its inexpensive ease of use in various biotechnological settings and organisms (112,113). The CRISPR-Cas9 system allows any number of modifications, from gene knockouts and knockdowns, gene insertions, fusions and tagging for visualisation to epigenetic activation or repression.

1.5.2 CRISPR-Cas9 as genome editing tool

As a genome editing tool, the host cell under study is transfected with constructs including all the genetic elements necessary for CRISPR-mediated engineering of a particular gene in the host cell (Fig. 7). Transfection should include vectors that will enable the following functionality in the host cell once expressed: 1) the Cas9 endonuclease (most frequently used of the Cas members); 2) a CRISPR 'cassette' that contains the guide RNAs to target the Cas complex to the specific gene of interest close to PAM sites; and 3) the specific gene modification (e.g. knockout or site-specific mutation) that will be substituted for the normal gene into the host genome through normal recombination events (Fig. 7). From the two CRISPR classes, the *Streptococcus pyogenes* class II system is the most widely used in biotechnology applications (103).

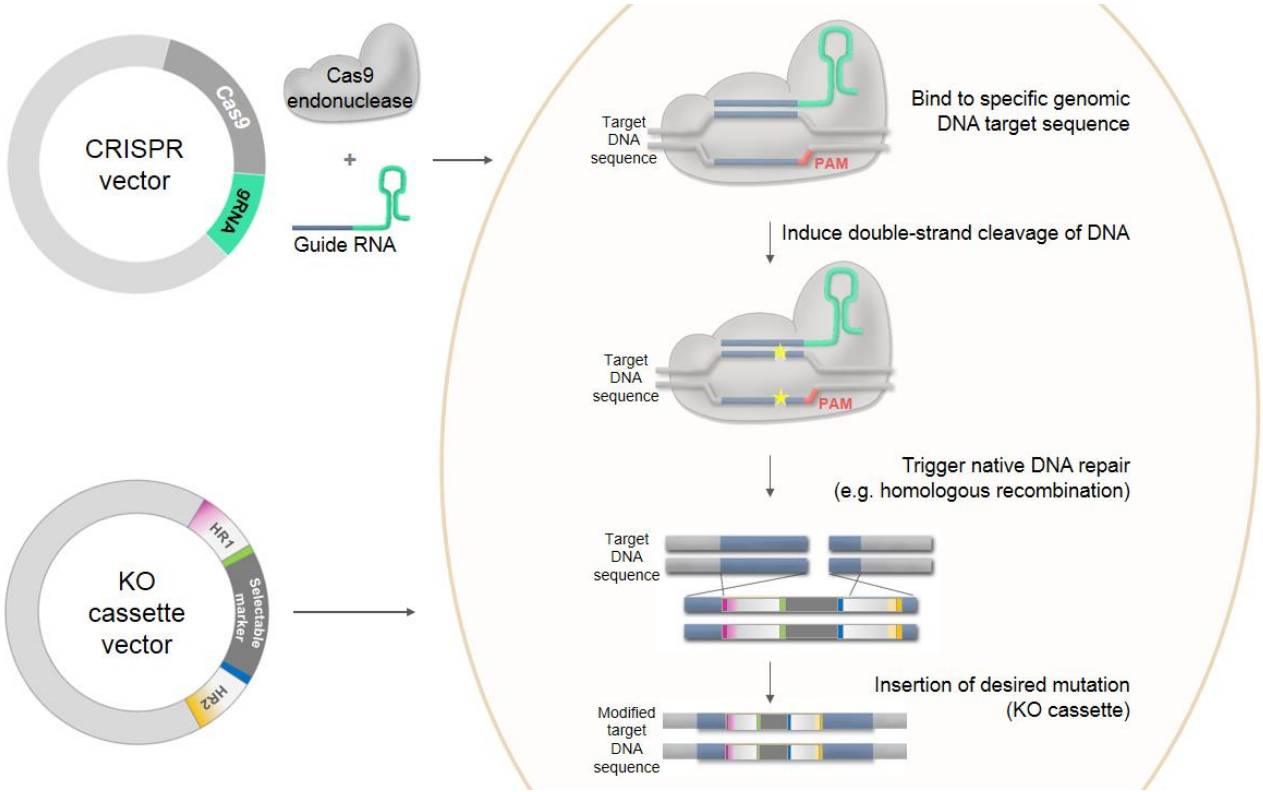

Figure 7. The CRISPR-Cas9 genome editing system strategy. The Cas9 endonuclease and gRNA is delivered through transfection of a 'CRISPR vector' after which the Cas9 protein-guide RNA duplex is formed. The synthetic guide RNA chimaera that mimics the native tracrRNA and (green) and crRNA (blue) elements are engineered as shortened ~20 bp DNA sequence that is transcribed to produce a RNA molecule. Guide interrogation of the target sequence occurs, inducing a double-strand break in the DNA (yellow stars), which is repaired according to the host native DNA break repair pathways such as homologous recombination (HR). A second vector, the 'KO cassette vector', delivers the knockout cassette (KO cassette) containing the desired mutations or gene, inserted at the double-strand break region through homologous recombination, the preferred DNA repair pathway in *P. falciparum* for example.

In June 2014, the first successful genome modification using the CRISPR-Cas9 system for *P. falciparum* was published (114). The authors not only demonstrated that single nucleotide substitutions and specific gene knockouts are possible but occurred with exceptionally high

efficiency in a marker free approach (114). Examples of CRISPR-Cas9 genome modification in *P. falciparum* include the disruption of the non-essential knob-associated histidine-rich protein-coding gene (*kahrp*) resulting in a lack of structural knobs protruding from infected erythrocytes (115). Introduction of the artemisinin resistance mutation (C580Y) into *P. falciparum* PF3D7_1343700 *kelch* propeller domain (K13-propeller) increased ring stage survival rates by ~13.5% (114). The highly conserved gene knockout of the *Pfvap1* (*P. falciparum* parasites virulence associated protein 1) gene indicated the importance of this protein in adhesion of parasites to the host cells (116). Furthermore, UDP-galactose and acetyl-CoA transporters were recently implicated as *Plasmodium* multidrug resistance genes using the CRISPR-Cas9 system (117). Together, these results show the possibilities that the CRISPR-Cas9 system holds for the study of various biological systems in *P. falciparum* parasites.

Here, we used CRISPR-Cas genome editing to elucidate the functional roles of proposed cell cycle regulators of *P. falciparum* parasites. This study may potentially identify an ideal target for drug intervention purposes and contributes to the knowledge of the dynamics of the *P. falciparum* parasite life cycle, cell cycle progression and phenotypes associated with the knockout of essential transcription factors.

Hypothesis

Putative cell cycle regulator genes can be functionally characterised using the genome-editing tool, CRISPR-Cas9, in *P. falciparum* parasites.

Aim and Objectives

Aim

To functionally characterise and evaluate the importance of putative cell cycle regulator genes to parasite cell cycle progression and survival in *P. falciparum* parasites by CRISPR-Cas9 mediated gene disruption.

Objectives

1. Identify genes as putative cell cycle regulators in *P. falciparum* parasites from the transcriptome data profile of parasites under cell cycle arrest and during cell cycle re-entry.
2. Design and clone guide RNAs for site-specific DNA cleavage by Cas9.
3. Construction and cloning of knockout cassettes for each gene investigated to allow homologous recombination based integration.
4. Transfection of *P. falciparum* NF54 and Dd2 strain parasites.
5. Validation of gene knockout in recombinant *P. falciparum* parasites.

Outputs

Conference Oral Presentation:

Hilde von Grüning, Sophie Adjalley, Riëtte van Biljon, Jandeli Niemand, Marcus Lee and Lyn-Marié Birkholtz. “Toward CRISPR-Cas9 evaluation of putative cell cycle regulators of the *Plasmodium falciparum* parasite”. Gordon Research Seminar, Les Diablerets, Switzerland, 1-2 July 2017

Conference Poster Presentation:

Hilde von Grüning, Sophie Adjalley, Riëtte van Biljon, Jandeli Niemand, Marcus Lee and Lyn-Marié Birkholtz. “Toward CRISPR-Cas9 evaluation of putative cell cycle regulators of the *Plasmodium falciparum* parasite”. Gordon Research Conference, Les Diablerets, Switzerland, 2-7 July 2017

Hilde von Grüning, Riëtte van Biljon, Jandeli Niemand and Lyn-Marié Birkholtz. “CRISPR-Cas9 evaluation of putative cell cycle regulators in *P. falciparum* parasites” South African Medical Research Council Office of Malaria Research Annual International Conference, Pretoria, South Africa, 31st July - 2 Aug 2016.

Submitted manuscript:

Jandeli Niemand, Riette van Biljon, Roelof van Wyk, Katherine Clark, Bianca Verlinden, Hilde von Grüning, Karen van der Merwe, Werner Smidt, Annél Smit, Janette Reader, Teresa Carvalho, Heather Painter, Manuel Llinás, Christian Doerig and Lyn-Marié Birkholtz “Cell cycle arrest of the human malaria parasite *Plasmodium falciparum* reveals regulated quiescence-proliferation decision-making”. Submitted, Nature Communications.

Chapter 2: Experimental Procedures

2.1 Identification of putative cell cycle regulators in *P. falciparum* parasites

The global transcriptome profile of *P. falciparum* parasites response under cell cycle arrest and cell cycle re-entry induction was obtained from L. Birkholtz (Niemand *et al*., submitted manuscript). *P. falciparum* parasite gene expression profiles showing transcriptional responses during cell cycle arrest (24 h DFMO treatment) as well as at 3, 6 and 12 h after cell cycle re-entry (induced with putrescine dihydrochoride treatment) was obtained after initial data processing and normalisation (*P*-value >0.05). Log_2 fold changes (log_2FC) in transcript abundance between the cell cycle arrest transcriptome (Cy5/Cy3, untreated reference pool/DFMO-treated samples) and cell cycle re-entry transcriptomes, (Cy5/Cy3, DFMO-treated sample/putrescine dihydrochloride samples) were obtained.

2.1.1 Transcriptome data processing and selection of cell cycle regulatory genes

The complete transcriptome dataset of 5 297 genes from Niemand *et al.* (submitted manuscript), was filtered for differential expression with log_2FC thresholds at 0.5 and -0.5, refining the differentially expressed dataset to 2 903 genes. These were sorted for directionality in expression profiles, where transcripts were either decreased in abundance during cell cycle arrest with concomitant increased abundance upon cell cycle re-entry or *vice versa*. The resulting 2 295 genes were subsequently clustered using unsupervised, uncentered hierarchical clustering with Cluster 3.0 (118) (average linkage) and evaluated for the enrichment of Gene Ontology (GO) terms (www.plasmoDB.org, an extensive biological database with genome, transcriptome, proteome and metabolome information on *Plasmodium* parasites). Prior phenotypic data using PhenoPlasmDB (http://phenoplasm.org/) for *P. falciparum* genes were probed, and genes with previous knockout data were removed from the dataset. Functional association with cell cycle regulation as a biological process was performed based on search term association in PlasmoDB ("regulation of transcription", “transcript*”, "DNA replication initiation", "DNA repair", "protein phosphorylation", “chromatin*”, “kinas*” and “histon*”) or through associations as described in Butler *et al*., Kozlov *et al*., Campbell *et al*. and Cai *et al*. (69,70,119,120). Genes that overlapped in these total datasets were analysed and refined to 86 previously uncharacterised genes that may have functional involvement in cell cycle regulation. Target genes were selected based on their transcript regulation profiles and strong involvement of genes from a specific biological cluster.

For clarity in subsequent sections, the identity of the four target genes chosen following the above strategy are provided here: PF3D7_0613800 (transcription factor with AP2 domains, ApiAP2),

PF37D_1239200 (transcription factor with AP2 domains, ApiAP2), PF3D7_1211700 (DNA replication licensing factor MCM5) and PF3D7_0705300 (origin recognition complex subunit 2).

2.2 Cloning strategies for CRISPR-Cas9 targeted gene disruption in *P. falciparum* parasites

The following components are necessary for a successful CRISPR experiment: 1) guide RNAs (gRNAs), i.e. specific 20-23 nucleotides (nt) RNA oligos per gene, designed as DNA probes adjacent to associated PAM sites on the target DNA, 2) Cas9 endonuclease per targeting gene and 3) gene KO cassette; each gene under investigation is disrupted and replaced by a blasticidin-S deaminase [*bsd*] cassette selection marker flanked by sequences homologous to the endogenous gene sequence. A dual vector approach was used in this study to cause a gene knockout in the *P. falciparum* genome as previously reported (110,114,121). This strategy relies on the co-transfection of two vectors (Fig. 8): 1) the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* expression vector, containing the specific guide sequences (gRNA) for each gene under investigation and allowing Cas9 expression; and 2) a separate pCR2.1®-TOPO® vector containing the KO cassette for each gene of interest. This strategy was investigated in collaboration with Drs. Marcus Lee and Sophie Adjalley at the Wellcome Trust Sanger Institute, Cambridge, UK (3-months research visit by H. von Grüning to the UK) to allow more dogmatic cloning and quality evaluation as this approach is technically complex, novel and requires expertise knowledge.

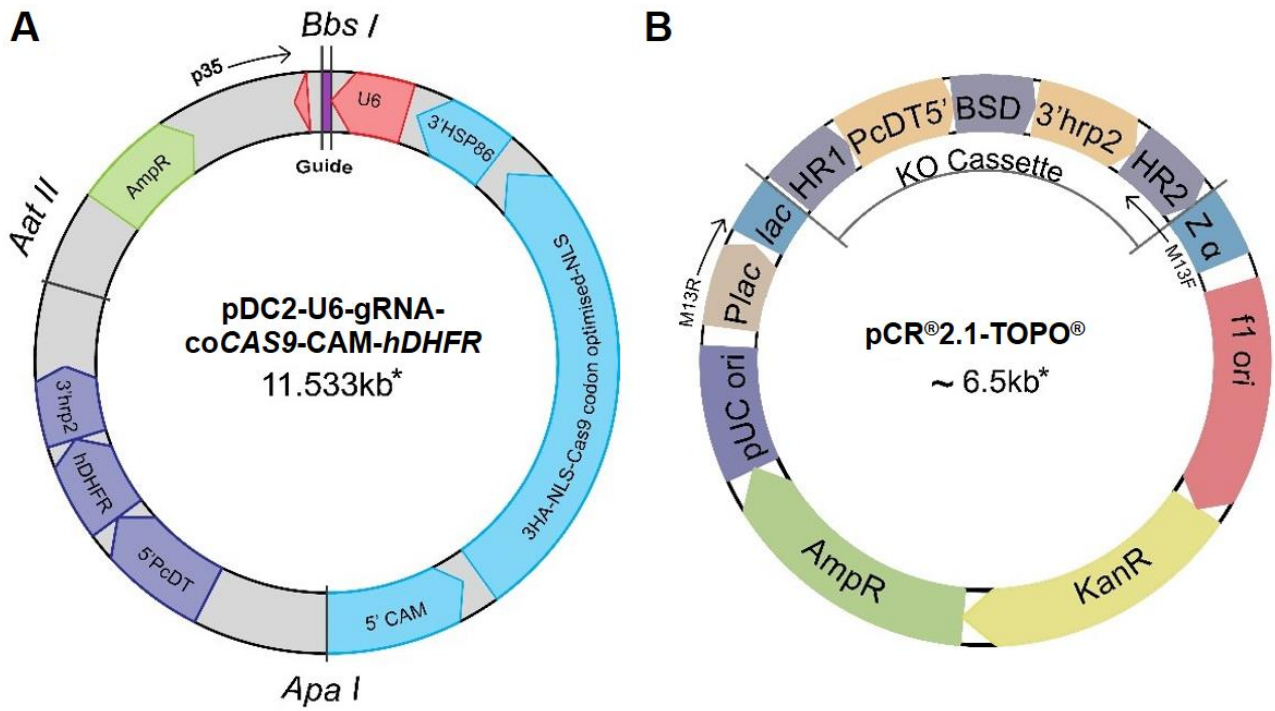

Figure 8. The dual vector approach for transfection of *P. falciparum* for CRISPR-Cas9. (A) The pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* expression vector. The guide RNA (dark purple) cloned between *BbsI* restriction sites and expression is driven by the U6 small nuclear RNA (snRNA) component of the spliceosome promoter (red). The *P. falciparum* parasite codon optimised Cas9 endonuclease (light blue) with nuclear localisation signals (NLS), and 3 hemagglutinin affinity epitopes (4.1 kb) is under control of a 5' calmodulin promoter (CAM, blue). A general KO cassette (dark blue) is present but was not used in the current cloning strategy. Prokaryotic propagation and selection occur through an ampicillin resistance gene (AmpR, green). The p35 sequencing primer is used for sequencing of the guide

in the reverse direction into the U6 promoter region. **(B)** TOPO-TA cloning vector, pCR®2.1-TOPO®. The vector contains the *lac* promoter and *lacZα* gene, two selection markers, AmpR and kanamycin (KanR) resistance genes, and two origins of replications (pUC ori and phage f1 ori). The KO cassette is cloned into the multiple cloning site located in the *lacZα* gene, allowing for selection by α-complementation. The KO cassette contains two homology regions (HR1 and HR2) that flank a central section containing a 5' *P. chabaudi* dihydrofolate reductase-thymidylate synthase (PcDT) UTR, a blasticidin S deaminase (*bsd*) selectable marker transgene and a *P. falciparum* hrp2 3′ UTR.

Within the 11.5 kb pDC2-U6-co*CAS9*-CAM-*hDHFR* vector (Fig. 8A, a kind gift from Dr. Marcus Lee, Wellcome Trust Sanger Institute, UK), the Cas9 expression cassette allows expression of Cas9 from a *P. falciparum* codon optimised *cas9* sequence under a parasite calmodulin promoter, with a *hsp86* terminator. Cas9 nuclear localisation is ensured through nuclear localisation signals tagged with a human influenza hemagglutinin (HA) epitope allowing antibody-based detection post-transfection. The RNA polymerase III dependent U6 small nuclear RNA (snRNA) component of the spliceosome promoter drives the expression of gRNAs. The *hDHFR* selectable marker gene, which confers resistance to the antifolate inhibitor WR99210, allows selection of parasites carrying the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector post-transfection. The general cloning strategy used in this study is outlined in Fig. 9.

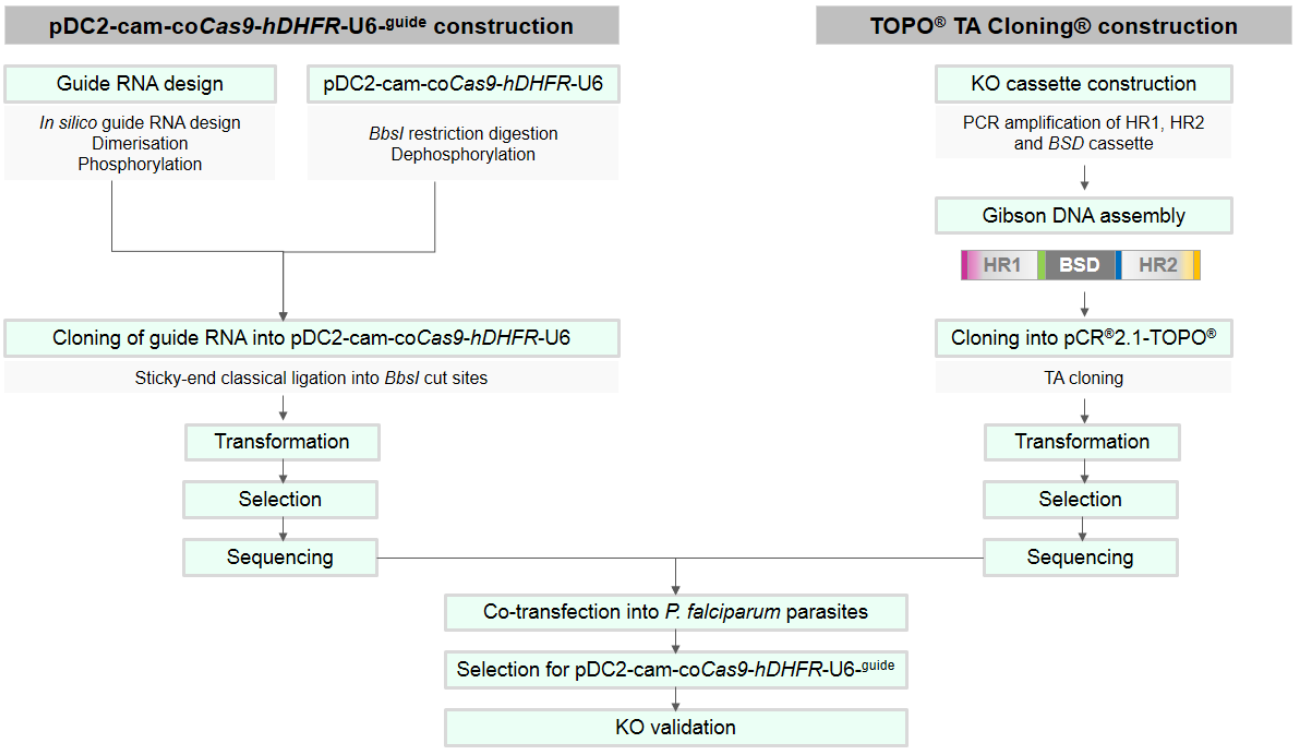

Figure 9. Cloning strategies for the dual vector CRISPR-Cas9 system applied in this study.

The KO cassette cloned into the pCR2.1®-TOPO® vector serves as the homologous template for DNA repair through subsequent double-crossover homologous recombination. Gene disruption or knockout is achieved by the insertion of a *bsd* sequence within the target gene, which can be chemically selected for using blasticidin S. Blasticidin S deaminase converts toxic blasticidin S to non-toxic deaminohydroxyblasticidin and ammonia (122). The KO cassette contains two

independent homology regions (HR1 and HR2), each 0.5-1 kb in size corresponding to 5´ and 3´ regions of the gene, respectively, to ensure replacement of the endogenous gene.

2.2.1 Design of guide RNAs

Specific guide sequences are required for each gene investigated to enable Cas9 to generate the site-specific double strand breaks. To analyse possible guide sequences within a gene, the CRISPR design tool within Benchling (www.benchling.com, open source) was used. *P. falciparum* was selected as organism along with the gene ID of each gene of interest. The genome coordinates are automatically annotated based on PlasmoDB information on chromosomal location. Selection parameters were set as a single guide, wild-type Cas9 option for a ~20 bp guide upstream of a 5´-NGG-3´ PAM sequence. Possible guide sequences within the open reading frame of the gene are listed as either sense or antisense but do not affect the efficiency of Cas9 mediated cleavage. A guanine in the first position of the ~20 bp is typically preferred for the initiation of transcription by RNA polymerase III. In this case, the G was already incorporated in the U6 promoter region of the pDC2-U6-co*CAS9*-CAM-*hDHFR* vector and was not included in the guide design.

Two guides were chosen per gene from the output list generated by Benchling with the following selection criteria: 1) high on- and off-target scores (>50), representing the cleavage efficiency of Cas9 (116) and the inverse probability of Cas9 off-target binding, respectively (117); 2) proximity to the 5´ region of gene as introducing a frame shift mutation in a constitutively expressed exon will likely create a non-functional protein provided that there are no critical functional domains upstream of the target region which could lead to a partially active protein; 3) G+C% content of ~40%, and 4) the frequency of repeating nucleotides <5. Once selected, the influence of guide position within the gene and homology region sequences were further interrogated for the presence of known alternative splice sites, skipped exons or known single nucleotide polymorphism (SNPs) by performing a gene query in PlasmoDB and Interpro (www.ebi.ac.uk/interpro/) to avoid a possibility of false-positive phenotypic outcomes. All guide sequences were evaluated for non-specific recognition and annealing by conducting a BLAST (https://blast.ncbi.nlm.nih.gov/) search and sequence alignment to the *P. falciparum* NF54 reference genome (optimised for megablast against highly similar sequences). Guide RNA candidates that have >85% similarity to any unintended sequence (with e-values above 1) were excluded.

The resultant prioritised guides were subsequently modified by addition of a forward 5´-ATTG overhang, with the reverse complementary strand to the guide carrying a 3´-CAAA overhang to allow cloning into the pDC2-U6-co*CAS9*-CAM-*hDHFR* vector. The final single stranded oligonucleotide pairs for each guide (sense and antisense compliments for the guides) were purchased from

Integrated DNA Technologies (IDT, USA) and subsequently suspended to 100 µM stock solutions in 1x TE-buffer (10 mM Tris, 1 mM EDTA, pH 8.0).

2.2.2 Cloning the guide RNAs into the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* expression vector

To enable downstream ligation of the guide oligonucleotides to the vector, the 5´ ends of the guide oligonucleotides were phosphorylated. Oligonucleotides (100 µM) were phosphorylated in a 10x T4 ligation buffer (66 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, 1 mM ATP, 7.5% (w/v) polyethylene glycol [PEG6000], pH 7.6), (New England Biolabs, UK) with 5 Richardson units of T4 polynucleotide kinase (New England Biolabs, UK), (see glossary) in a thermocycler at 37˚C for 30 min. Secondary structures were completely denatured at 94˚C for 5 min and oligonucleotides allowed to anneal into their complementary double stranded configuration by gradual cooling to 25˚C at a rate of 5˚C/min, and stored at -20˚C.

The pDC2-U6-co*CAS9*-CAM-*hDHFR* vector was prepared for cloning by digesting ~4 µg of pDC2-U6-co*CAS9*-CAM-*hDHFR* vector propagated and purified from *Escherichia coli* XL10-Gold bacteria with 4 U of *Bbs*I restriction enzyme (palindromic recognition site 5´-GAAGAC(N2)-3´, New England Biolabs, UK) at 37˚C for 1 h. The vector was treated with 2 U of Shrimp Alkaline Phosphatase (rSAP, New England Biolabs, UK) for 1 h at 37˚C to dephosphorylate 5´ DNA termini and prevent re-circularisation. After heat inactivation (65˚C for 5 min), the dephosphorylated, linearised vector was immediately column purified using a NucleoSpin® Gel and PCR Clean-up (Macherey-Nagel GmbH & Co.KG, Germany) and stored at -20˚C.

For ligation, the guide oligonucleotides were diluted in 1:200 in nuclease-free dddH_2O of which 1 µL was ligated into the *Bbs*I cut sites (with compatible sense 5´-ATTG and antisense 3´-CAAA overhangs) of 50 ng dephosphorylated pDC2-U6-co*CAS9*-CAM-*hDHFR* vector with 1 U of T4 DNA ligase in T4 DNA ligase buffer at room temperature for 1 h. The ligated construct was stored at -20˚C.

Competent XL10-Gold *E. coli* was prepared from a saturated culture (grown overnight in Luria-Bertani [LB] broth [1% w/v tryptone, 0.5% w/v yeast extract, 171.23 mM NaCl, pH 7.5], at 37˚C at 180 rpm). A secondary culture was inoculated from the overnight culture at an optical density (OD_{600}) of ~0.025 OD_{600}/mL, and grown at 18˚C (180 rpm) until an OD_{600} of 0.4 (log phase growth). Cells were chilled on ice for 20 min, centrifuged for 15 min at 3200x*g* at 4˚C in an Eppendorf centrifuge. Cells were suspended in ice-cold TB-buffer (10 mM HEPES, 15 mM $CaCl_2$, 250 mM KCl, 55 mM $MnCl_2$, pH 6.7) after which dimethyl sulfoxide (DMSO) was added drop wise to a final concentration of 7% (v/v) to enhance crystallisation of membrane regions and vector uptake (123). Cells were

incubated on ice for 10 min before aliquoted and flash frozen with liquid nitrogen before storing at -80˚C.

The 5 µL of the ligation mixture (i.e. 25 ng/µL of the pDC2-U6-co*CAS9*-CAM-*hDHFR* vector) was added to competent *E. coli* XL10-Gold cells (50 µL) and incubated on ice for 20 min. Cells were heat shocked for 30 s at 42˚C (Eppendorf™ ThermoStat Plus Interchangeable Block Heater) and transferred to ice for 2 min, after which 150 µL pre-warmed SOC medium (2% w/v tryptone, 0.5% w/v yeast extract, 0.05% w/v NaCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM glucose) was added for cells to recover (20 min at 37˚C with agitation, 180 rpm). The complete mixture was plated onto LB-agar-Amp plates (LB-broth, 1% w/v agar, 100 µg/mL Amp) and incubated overnight at 37˚C.

Single colonies were subsequently inoculated in LB-Amp broth (50 µg/mL Amp) grown overnight with agitation (180 rpm) at 37˚C. Plasmid DNA was purified using the NucleoSpin® Plasmid Isolation Kit (Macherey-Nagel GmbH & Co.KG, Germany) based on alkaline denaturation of DNA (see glossary) (124,125) and its subsequent purification through a silicon matrix as per manufacturer's instructions. Eluted DNA was evaluated for concentration and purity with a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, USA) at A_{260} (for the carbon-nitrogen π-bonds of nucleotides) and A_{260}/A_{280} and A_{260}/A_{230}, respectively. Values of 1.8-2.0 and 2.0-2.5 were indicative of pure dsDNA free of contaminants such proteins (A_{260}/A_{280}) and phenolic compounds (A_{260}/A_{230}), respectively. Purified plasmid DNA was stored at -20˚C. The resultant constructs (termed pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR*) were evaluated by automated Sanger di-deoxynucleotide sequencing at GATC Biotech (Germany) using an ABI 3730xl DNA Analyser. Sequences were analysed using Lasergene 14 SeqMan Pro (v.14) (MAFFT Multiple Sequence Alignment algorithm) and aligned to the U6 consensus sequence.

Positive pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* constructs for each of the four genes of interest were designated PF37D_1239200-^{guide}, PF3D7_0613800-^{guide}, PF3D7_1211700-^{guide} and PF3D7_0705300-^{guide}.

2.3 Construction of the KO cassettes for each gene

HR1 and HR2 regions of each gene, as well as the *bsd* insert (amplified from the pDC2-eGFP-CAM-BSD), were created for each KO cassette. Each of these regions was amplified separately for subsequent complete assembly of the KO cassettes using a Gibson hifi assembly reaction. This allows the simultaneous assembly of multiple fragments of DNA of various lengths with overlapping sequences in a single-tube isothermal reaction. This involves three simultaneous enzymatic reactions whereby the four DNA fragments are annealed: 1) single stranded 3´ overhangs are generated by an exonuclease to allow annealing of DNA fragments with complementary ends, 2)

the proprietary polymerase fills gaps within an annealed fragment, and finally, 3) the nicks are sealed by DNA ligase (Fig. 10).

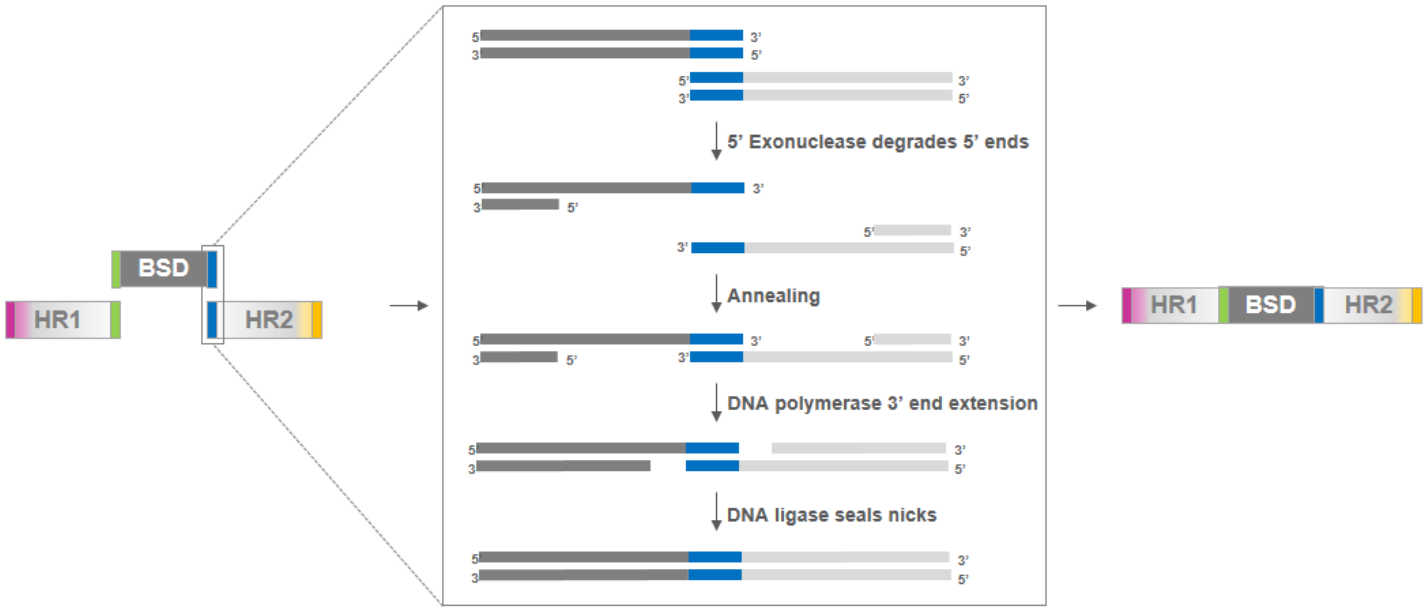

Figure 10. Gibson DNA assembly reactions. Three DNA fragments (PCR products), including the drug selection maker *bsd* flanked by two homology regions are assembled and joined through overlapping sequences (shown in purple, green, blue and orange). The assembly occurs through the catalysing reactions of a 5´ exonuclease, a ligase and a polymerase as shown.

Primers consisted of 18-24 nt complementary to the gene of interest as well as 15-40 nt sequence to facilitate overlap with adjacent regions in the Gibson assembly, including overlap with *Aat*II restriction enzyme digested vector, overlap between HR1 and BSD, overlap between BSD and HR, and finally overlap at the *Apa*I restriction site of the vector. The HR1 and HR2 regions of each gene were amplified from *P. falciparum* genomic DNA, necessitating the *in vitro* cultivation of the parasites and DNA extraction.

2.3.1 *In vitro* cultivation of *P. falciparum* parasites for genomic DNA isolation

Ethics approval for the cultivation of human malaria parasites was obtained (University of Pretoria 120821–077). Intra-erythrocytic *P. falciparum* parasites (NF54 [drug sensitive line] and Dd2 [drug resistant]) were cultivated in fresh human erythrocytes (either A^+ or O^+) in RPMI-1640 culture medium supplemented with 25 mM HEPES (pH 7.5, Sigma Aldrich, USA), 0.2 mM hypoxanthine (Sigma Aldrich, USA), 0.024 mg/mL gentamycin (Hyclone, USA), 5g/L Albumax II (Invitrogen, USA), 23.81 mM sodium bicarbonate and 0.2% w/v D-glucose. Cultures were maintained with daily media change at 5% haematocrit, 2% parasitaemia under hypoxic conditions (5% O_2, 5% CO_2, 90% N_2) with moderate shaking at 37˚C (126,127). Parasitaemia was assessed through light microscopy of Giemsa stained parasites.

Genomic DNA was isolated from a late trophozoite (>98% population, 34-38 hpi) *P. falciparum* NF54 parasite culture (13% parasitaemia) for optimal gDNA yield as parasites are actively dividing nuclei. Parasites were released with 0.06% (v/v) saponin (5 min incubation and centrifuged for 3 min at 1890*xg*, see glossary). Parasites were washed three times with phosphate buffered saline (1xPBS, 137 mM NaCl, 2.7 mM KCl, 12 mM Na_2HPO_4, 1.7 mM KH_2PO_4, pH 7.2) and frozen at -80˚C for 48 h. Parasites were thawed on ice before genomic DNA was isolated with the Quick-gDNA™ Blood MiniPrep kit (Zymo Research, USA). Genomic DNA was eluted in proprietary elution buffer and quantified using a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, USA) as before.

2.3.2 PCR amplification of gene homology regions and *bsd* cassette

Homology regions were amplified for each of the targeted genes (PF37D_1239200, PF3D7_0613800, PF3D7_1211700 and PF3D7_0705300) using KAPA *Taq* DNA polymerase master mix that contains KAPA *Taq* DNA polymerase (0.5 U), KAPA *Taq* buffer (proprietary), dNTPs (0.2 mM of each dNTP, $MgCl_2$ [1.5 mM] and proprietary stabilisers [Kapa Biosystems, South Africa]) added to 10 ng template gDNA and 10 pmol of each primer. An initial denaturation at 94˚C for 2 min was followed by 30 cycles of denaturation at 94˚C for 40 s, primer annealing at the particular annealing temperature (Ta) for 30 s, extension at 68˚C for 45 s and a final extension at 65˚C for 2 min (see glossary). Primer sequences and Ta's used for each primer pair are provided in Table 1. The complete *bsd* cassette was amplified from 100 ng pDC2-*eGFP*-CAM-*BSD* vector with 10 pmol primer each in a PCR containing 2x KAPA *Taq* DNA polymerase master mix, cycling 30x at 94˚C for 40 s, 56˚C for 30 s, 68˚C for 45 s followed by a final extension at 65˚C for 2 min.

Table 1. Primers and cycling conditions used for homology region amplification and sequencing.

Cassette	Primers					Ta (ºC)
bsd cassette (from pCC1-*hDHFR-kahrp*KO vector)	F	5´-GGGCCCGCATGCTTAGCTAATTCG-3´		R	5´-CTCGGTACCTCTAGATTTCTCTGC GG-3´	56
M13 sequencing primers (KO cassette in pCR®2.1-TOPO® vector)	F	5´-CAGGAAACAGCTATGAC-3´		R	5´-CCCTATAGTGAGTCGTATTA-3´	55
Gene	**Homology region 1**		**Ta (ºC)**	**Homology region 2**		**Ta (ºC)**
PF3D7_0613800	F	5´-CGAAAAGTGCCACCTGACGTC GTCCTTCATCATAGAGGACTCC-3´	60	F	5´-TTAGCTAAGCATGCGGGCCC GATGATGAAAGAGATTATGATAGG-3´	60
	R	5´-AGAAATCTAGAGGTACCGAG CAAATGTCCATGCTATATCTTC-3´		R	5´-GGGAATTTCCTTATAGGGCCC GTAACACGATTAATCACATGTTC-3´	
PF3D7_1239200	F	5´-CGAAAAGTGCCACCTGACGTC ATGAGTGATAAAGGTGAAGG-3´	65	F	5´-TTAGCTAAGCATGCGGGCCCGA AATAATATCTTCAAATTGTAGCAATC-3´	65
	R	5´-AGAAATCTAGAGGTACCGAG TTAGTCTTCTCAAGTTTAAGACAATTC-3´		R	5´-GGGAATTTCCTTATAGGGCCC CAGAACATTATTCAGAAGATTAGC-3´	
PF3D7_1211700	F	5´-CGAAAAGTGCCACCTGACGTCAT GATAGGAATACAAGAAGGTAG-3´	67.5	F	5´-TTAGCTAAGCATGCGGGCCCGG TACTAGAAAAAGAATTGGTGAGG-3´	65
	R	5´-AGAAATCTAGAGGTACCGAGCTAT CCATAGCATTACCTAATCCC-3´		R	5´-GGGAATTTCCTTATAGGGCC CTTACTGATCTTCTATCTCCTCT-3´	
PF3D7_0705300	F	5´-CGAAAAGTGCCACCTGACGTCATG CTGAAAAACTTTGAAGTAA-3´	65	F	5´-TTAGCTAAGCATGCGGGCCC TATTCAGAAGATATGAAAAG-3´	55
	R	5´-AGAAATCTAGAGGTACCGAGCATT GTTTGACACATTGGATTCG-3´		R	5´-GGGAATTTCCTTATAGGGCCCC AATTCAAAAGATACATCATC-3´	

In addition to the above, the *P. falciparum* knob-associated histidine-rich protein (*kahrp*) gene (PF3D7_0202000) KO cassette was amplified from the pCC1-*hDHFR-kahrp*KO construct (Fig. 11, a kind gift from Zenon Zenonous, Wellcome Trust Sanger Institute, Cambridge, UK). The *kahrp* gene was included as a positive control for Gibson assembly and transfection due to prior evidence of its effective disruption resulting in a traceable phenotype (121). In all PCR reactions, appropriate controls were included, including a no primer control and a no template control.

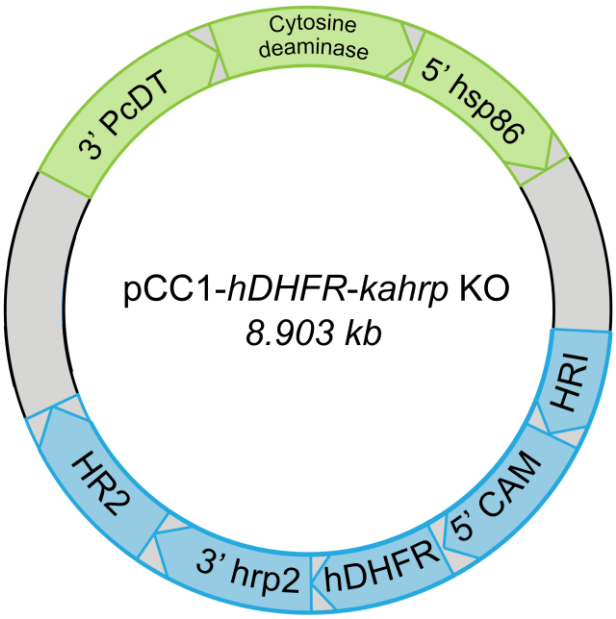

Figure 11. The pCC1-*hDHFR-kahrp*KO vector from Dr. Zenon Zenonos (Wellcome Trust Sanger Institute, Cambridge, UK). The complete 3 895 bp KO cassette was amplified from the ~9.3 kb pCC1-*hDHFR-kahrp*KO.

All PCR products were purified using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel GmbH & Co.KG, Germany) and stored at -20˚C. PCR products were analysed on 0.8% (w/v) agarose gels in 1x TAE (0.04 M Tris-acetate, 1 mM EDTA, pH 8)(Sigma-Aldrich, UK) at 6-10 V/cm in a BioRad Sub-Cell® GT Cell electrophoresis chamber. Samples were loaded with 5x DNA Loading Buffer Blue (Bioline, UK), with the HyperLadder™ 1kb (Bioline, UK) used as a molecular marker. Agarose/TAE gels were stained using the SYBR® Safe DNA Gel Stain (Thermoscientific, USA) and DNA visualised with a BioDoc-It Imaging System (Analytik Jena, USA) and photographed with a Monochrome CCD camera (Analytik Jena, USA).

2.3.3 Gibson assembly of KO cassette (HR1-*bsd* cassette-HR2)

Equal amounts (0.067 pmol per fragment, up to ~0.2 pmol for 2-3 fragments of total DNA) of the homology region PCR products and the *bsd* cassette PCR product were added to the 2x Gibson Assembly master mix containing proprietary exonuclease, polymerases and ligases (New England Biolabs, UK) to a final volume of 20 µL. The mixture reaction was incubated at 50˚C for 1 h and then transformed into *E. coli* XL10-Gold bacteria as previously described.

The efficacy of the Gibson assembly reaction was evaluated by PCR amplification and nucleotide sequencing of the complete KO cassette (HR1, *bsd* and HR2). For this, the HR1 forward and the HR2 reverse primers were used to amplify the ~3.2 kb KO cassettes for PF37D_1239200, PF3D7_0613800, PF3D7_1211700 and PF3D7_0705300 using Kapa *Taq* DNA polymerase (Kapa Biosystems, USA) with the same cycling conditions as previously described. The PCR product was column purified and subsequently analysed by 1% (w/v) agarose gel electrophoresis. The band size corresponding to the KO cassette size was gel extracted using the X-tracta Gel Extraction Tool (Sigma-Aldrich), quantified using a NanoDrop® ND-1000 spectrophotometer (Thermo Fisher Scientific, USA) and 20-80 ng/µL subsequently sequenced with Sanger sequencing at GATC Biotech (Germany).

2.3.4 TOPO® TA Cloning® of the KO cassette in the pCR®2.1-TOPO® vector

The pCR®2.1-TOPO® vector allows for TA cloning in addition to a covalently bound topoisomerase I to reversibly wind and unwind supercoiled DNA (see glossary). The pCR®2.1-TOPO® (50 µg), the KO cassette PCR product (HR1-*bsd*-HR2 for each gene), 2 mM of dATP and salt solution (200 mM NaCl, 10 mM $MgCl_2$) were incubated at room temperature for 30 min. The complete ligation mixture was transformed into competent *E. coli* XL10-Gold for blue-white selection (see glossary) and plated onto LB-agar-Amp plates containing 100 mM IPTG and 20 mg/mL X-gal. As a pre-screen for recombinant bacteria, positive white colonies from transformation (resuspended in 5 µL PBS and 2 µL subsequently used in PCR) were colony PCR amplified using 10 pmol each of the vector specific

M13 forward and reverse primers (Table 1) and Go*Taq*® Long PCR master mix containing a proprietary combination of optimised buffer components and concentrations, dNTPs, $MgCl_2$ and a blend of hot start-, *Taq*-, and a proprietary proof reading DNA polymerases (Promega, USA). Cycling conditions were 94˚C for 2 min, 94˚C for 40 s, 55˚C-67.5˚C for 30 s, 68˚C for 45 s followed by a final extension at 65˚C for 2 min, for 30 cycles. PCR products were analysed using DNA gel electrophoresis as described before. The remaining 3 µL of the positive colonies were subsequently re-inoculated in LB-broth and grown at 37˚C overnight with agitation (180 rpm). Plasmid DNA was purified as described before and Sanger sequenced at GATC Biotech (Germany) using an ABI 3730xl DNA Analyser, using the M13 forward and reverse primers. Lasergene software package SeqMan Pro (v.14) was used to analyse sequence and chromatogram results received as .ab1 files.

The resultant pCR®2.1-TOPO®-^{KO} positive constructs for each of the four genes of interest were termed PF37D_1239200-^{KO}, PF3D7_0613800-^{KO}, PF3D7_1211700-^{KO} and PF3D7_0705300-^{KO}.

2.4. Transfection of *P. falciparum* parasites with guide and KO constructs

2.4.1 Large scale vector DNA isolation for all guide and KO constructs

Vector DNA was isolated using a NucleoBond® Xtra Midi / Maxi purification kit (Macherey-Nagel GmbH & Co.KG, Germany) according to manufacturer's specification from saturated cultures of each of the four guide constructs and four KO constructs in their individual *E. coli* XL10-Gold cells. Eluted vector DNA was concentrated by adding 2.5 volumes of room temperature isopropanol to precipitate eluted vector DNA (see glossary), centrifuged at 15000x*g* for 30 min at 4˚C, followed by a 70% ethanol wash of pellet DNA (centrifuged at max speed for 5 min at room temperature) after which the pellet DNA was allowed to dry at room temperature. DNA was reconstituted to 1 µg/µL in Cytomix (120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 10 mM K_2PO_4, 25 mM HEPES, pH 7.6) for subsequent transfection and stored at -20˚C.

2.4.2 Parasite transfection

Both NF54 and Dd2 *P. falciparum* parasite lines were transfected to exclude possible ambiguity with regards to growth rates, as the Dd2 strain is known for an accelerated life cycle and enhanced fitness (128). Trophozoite cultures (2% parasitaemia, 3% haematocrit) of *P. falciparum* NF54 and Dd2 parasites were synchronised with 70% Percoll resulting in a ~5% ring parasite population. Ring stage parasites (4.3-5.2% parasitaemia, 3% haematocrit, 15 mL) were harvested by centrifugation (1000x*g,* 3 min at room temperature) after which the spent media was removed and the ~100 µL parasite infected erythrocyte pellet washed once in ice-cold cytomix and subsequently mixed with 50 µg each of pCR®2.1-TOPO®-^{KO} and pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* constructs (i.e. 100 µg total) for each individual gene of interest and reconstituted with Cytomix to a final volume of 400 µL. The mixture was transferred to pre-chilled electroporation cuvettes and electroporated using

a BioRad Gene Pulser Xcell™ Electroporation System at a high capacitance of 950 µF with a low voltage of 310 V, ensuring optimal time constants (see glossary) of 10-15 ms.

Electroporated, infected erythrocytes were mixed with 1 mL pre-warmed culture medium transferred to a tube containing 4 mL pre-warmed culture medium after which parasites were allowed to recover at 37˚C for 1 h. The recovered parasite containing mixture was centrifuged at 1000x*g* for 3 min after which lysed erythrocyte debris were aspirated. Transfected cells were resuspended in 5 mL pre-warmed culture medium supplemented with 100 µL fresh erythrocytes (50% haematocrit) to compensate for lysed cells during electroporation. The parasite suspension (~5.1 mL final volume, 3% haematocrit) was transferred to a single well of a tissue culture 6-well plate and incubated at 37˚C stationary with 5% O_2, 5% CO_2, 90% N_2 gas supplementation.

2.4.3 Drug selection of recombinant parasites

Transfected *P. falciparum* parasites (minimum 0.5% parasitaemia) were maintained under selective single drug pressure in selection media (RPMI-1640 media containing 5% human serum, 2 mM GlutaMAX [Life Technologies] and 2.5 nM WR99210 [Jacobus Pharmaceutical Company Inc.]). WR99210 is a potent inhibitor of *P. falciparum* DHFR, however, *hDHFR* confers resistance to WR99210 (129). Selection of *P. falciparum* parasites successfully transformed with the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector started one day after transfection. Parasites transformed with the pCR®2.1-TOPO®-KO was not selected for due to the enhanced strain on parasite growth in the presence of two or more drugs. Parasite cultures were cloned by diluting the culture 5-fold and transferred to 96-well plates at a 3% haematocrit, ~4-6 days post-transfection while under WR99210 drug selection. Parasite cultures were analysed microscopically to verify parasitaemia twice a day for the first 1-7 days post-transfection and drug selection. Media was changed every 24 h for the first 6-8 days after which drug pressure was removed, and complete media changed twice weekly until parasites recovered or ~40 days post-transfection.

2.5 Screening for recombinant knockout *P. falciparum* parasites

Once an increase in parasitaemia to 1% was observed, the ~100 µL infected erythrocyte cell cultures were expanded to a 20 mL culture each by supplementation of culture medium and fresh erythrocytes (3% haematocrit). Genomic parasite DNA was extracted from parasite cultures with parasitaemias of ~3-5% using the Qiagen DNeasy® Blood & Tissue kit (Qiagen, USA). To determine whether the KO cassette was integrated into the genome of *P. falciparum* parasites, amplification of the whole KO cassette was performed with 100 ng template DNA using HR1 and HR2 primers for each gene of interest: initial denaturation at 94˚C for 2 min, denaturation at 94˚C for 40 s, 50˚C for 30 s for annealing of primers and 62˚C for 3 min extension for 30 cycles. PCR products were

analysed using gel electrophoresis as previously described. The screening of recombinant parasites was performed by Dr. Sophie Adjalley (Wellcome Trust Sanger Institute, Cambridge, UK).

2.6. Constructing a single-vector aimed at CRISPR-Cas9 targeted gene disruption in *P. falciparum* parasites

An alternative cloning strategy was investigated using a single plasmid vector to allow the envisaged advantage of a simplified and more efficient transfection and clone selection. The pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR*-KO vector consists all the components to allow such a single vector strategy. The KO cassettes for each gene investigated and constructed in section 2.6, could, therefore, be cloned into their corresponding pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* (Fig. 12).

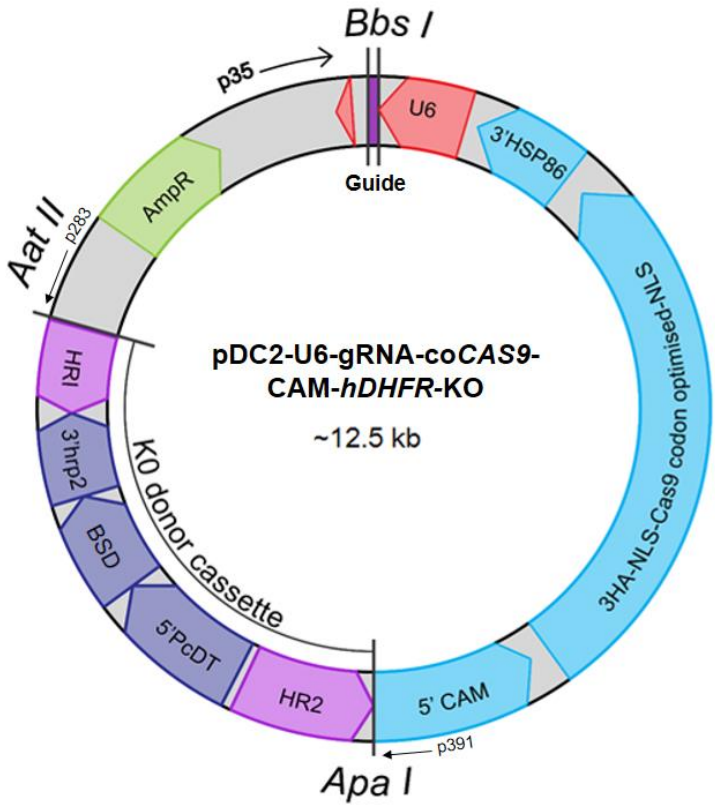

Figure 12. Single vector approach to gene KO of putative cell cycle regulators in *P. falciparum* parasites. Here, the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector carries all three components of a CRISPR KO experiment, 1) the guide sequences expressed under control of the U6 promoter, 2) the codon optimised Cas9 endonuclease expressed under control of the calmodulin promoter of *P. falciparum*, and 3) the KO cassette as cloned at the *Aat*II/*Apa*I restriction sites.

The required cloning strategy for construction of such single vectors is outlined below. Approximately 12 µg of the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector for each gene were sticky-end digested with 50 U of *ApaI* (New England Biolabs, UK) for 3 h at 25˚C in the presence of CutSmart buffer (50 mM K acetate, 20 mM Tris-acetate, 10 mM Mg^{2+} acetate, 100 µg/mL BSA, pH 7.9), followed by a 40 U of *Aat*II digestion overnight at 37˚C. The enzymes were heat inactivated at 65˚C for 5 min and the linearised vector treated with 4 U of rSAP for 1 h at 37˚C, followed by heat inactivation at 65˚C for 5 min. The digested DNA was analysed on 0.8% (w/v) agarose gels as

before. Bands corresponding in size to the KO cassettes was gel extracted using the X-tracta Gel Extraction Tool (Sigma-Aldrich) and a NucleoSpin® Gel and PCR Clean-up Kit (Macherey-Nagel GmbH & Co.KG, Germany) and quantified using the NanoDrop® ND-1000 spectrophotometer (Thermo Fischer Scientific, USA).

For use as a positive control in the transfections, the PF3D7_0202000 *kahrp* gene homology regions were amplified as a whole cassette including the *hDHFR* gene from the pCC1-*hDHFR-kahrp*KO vector DNA (Fig. 11) using the forward primer 5´-GGGAATTTCCTTATAGGGCCCTCCTGCTTCTTGTAATTG-3´ and reverse primer 5´-CGAAAAGTGCCACCTGACGTCGTGCAATAATGGAAACGG-3´, bearing the same 5' overlap sequences aimed at hybridising the KO cassette to the vector backbone. Cycling conditions were 95˚C for 3 min, 98˚C for 20 s, 52˚C for 30 s, 62˚C for 4 min, 65˚C for 5 min for 30 cycles and held at 4˚C.

Gibson assembly proved to be challenging. Alternatively, overlapping sequences aimed at hybridisation to the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector was investigated by adding 5 nucleotides to each overlap at their respective 5´ regions, thus increasing the overlap length from 21 to 26 nucleotides. All HR regions for all genes, therefore, had to be re-amplified (PCR reaction and cycling conditions as previously described with primers for HR amplification and assembly are shown in Table 2).

Table 2. Revised PCR primers for HR amplification and ligation to the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vectors

Gene	Homology region 1		Ta (ºC)	Homology region 2		Ta (ºC)
PF3D7_0613800	Forward	5'- TTCCCCGAAAAGTGCCACCTGAC GTC GTCCTTCATCATAGAGGACTCC-3'	60	Forward	5'-TTAGCTAAGCATGCGGGCCC GATGATGAAAGAGATTATGATAGG-3'	60
	Reverse	5'-AGAAATCTAGAGGTACCGAG CAAATGTCCATGCTATATCTTC-3'		Reverse	5'- TATATGGGAATTTCCTTATAGGGC GTAACACGATTAATCACATGTTC-3'	
PF3D7_1239200	Forward	5'-TTCCCCGAAAAGTGCCACCTGACG TCATGAGTGATAAAGGTGAAGG-3'	65	Forward	5'-TTAGCTAAGCATGCGGGCCCGAA ATAATATCTTCAAATTGTAGCAATC-3'	65
	Reverse	5'-AGAAATCTAGAGGTACCGAGTTAG TCTTCTCAAGTTTAAGACAATTC-3'		Reverse	5'- TATATGGGAATTTCCTTATAGGGC CAGAACATTATTCAGAAGATTAGC-3'	
PF3D7_1211700	Forward	5'-TTCCCCGAAAAGTGCCACCTGACG TCATGATAGGAATACAAGAAGGTAG-3'	67.5	Forward	5'-TTAGCTAAGCATGCGGGCCCGG TACTAGAAAAAGAATTGGTGAGG-3'	65
	Reverse	5'-AGAAATCTAGAGGTACCGAGCTAT CCATAGCATTACCTAATCCC-3'		Reverse	5'- TATATGGGAATTTCCTTATAGGGC CCTTACTGATCTTCTATCTCCTCT-3'	
PF3D7_0705300	Forward	5'-TTCCCCGAAAAGTGCCACCTGACG TCATGCTGAAAAACTTTGAAGTAA-3'	65	Forward	5'-TTAGCTAAGCATGCGGGCCC TATTCAGAAGATATGAAAAG-3'	55
	Reverse	5'-AGAAATCTAGAGGTACCGAGCATT GTTTGACACATTGGATTCG-3'		Reverse	5'- TATATGGGAATTTCCTTATAGGG CCCAATTCAAAAGATACATCATC-3'	

Homology regions were re-amplified by PCR followed by repeating the Gibson assembly reaction using the newly synthesised HR regions. The KO cassette was successfully cloned into the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector and sequences confirmed by Sanger sequencing using purified 80 ng/µL vector DNA and sequencing primers p283 (forward) 5´-AACATATGTTAAATATTTATTTCTC-3´ and p391 (reverse) 5´-AATAAACCAATAGATAAAATTTGTAGAG-3´. This provides the first single vector KO strategy in *P. falciparum* that will be transfected to assess the improved KO efficiency in a future study.

Chapter Three: Results

3.1 Identification of putative cell cycle regulators in *P. falciparum* parasites

Cell cycle control in *P. falciparum* parasites, imposed at the point of G_1/S transition, was shown to result in the differential expression of 2 295 genes (log_2FC threshold of ≤0.5 and ≥0.5, at any time point) (Fig. 13) (Niemand *et al*., submitted manuscript). Strong differential responses were observed for the majority of genes, specifically at the point of cell cycle arrest as well as at the last time point of cell cycle re-entry. The earliest re-entry time point was associated with the least number of differentially expressed genes, but this steadily increased as the parasite re-entered its cell cycle. Unsupervised, uncentered hierarchical clustering yielded two clear clusters: 1 161 transcripts increase abundance in cell cycle arrest and a subsequent decrease in abundance upon cell cycle re-entry, and 1 134 transcripts had the opposite expression profile (Fig. 13).

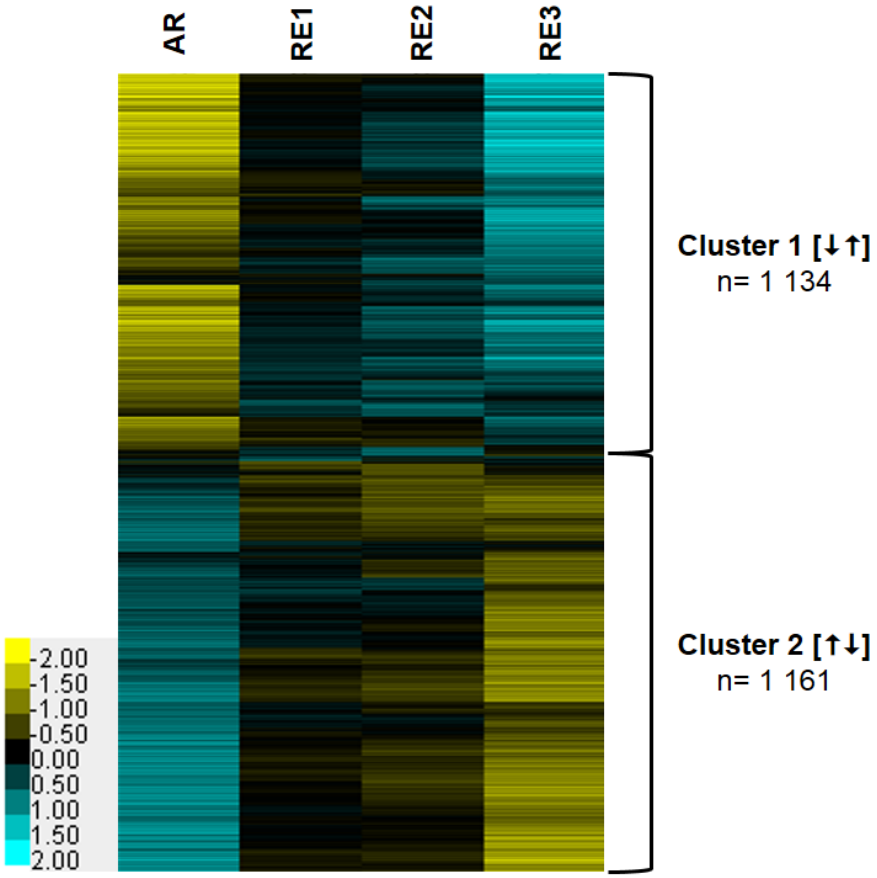

Figure 13. Hierarchical clustering of genes with directional differential expression greater than a log_2FC of 0.5. Two main clusters are shown: Genes (1 134) that showed significantly decreased expression under arrest and increased expression upon re-entry of the cell cycle, and another set of genes (1161) that increased expression under arrest and decreased expression in re-entry. Hierarchical clustering was performed using Cluster 3.0 (average linkage) and heatmap generated in TreeView 3.0. (AR, arrested cell cycle; RE1, re-entry time point 1; RE2, re-entry time point 2; RE3, re-entry time point 3). Yellow: decreased abundance; Blue: increased abundance.

To determine the biological signatures associated with these two clusters, GO enrichments were obtained. GO enrichment showed a general, systematic response to cell cycle perturbation involving numerous essential biological processes. Genes that showed decreased abundance during cell cycle arrest were involved in signal transduction, mitosis and nuclear division, chromatin

modifications, cellular respiration and immunoglobulin production (Fig. 14 A). Comparatively, genes showing increased abundance upon re-entry into the cell cycle were related to transcription, ribosome biosynthesis and assembly and protein folding (Fig. 14 B).

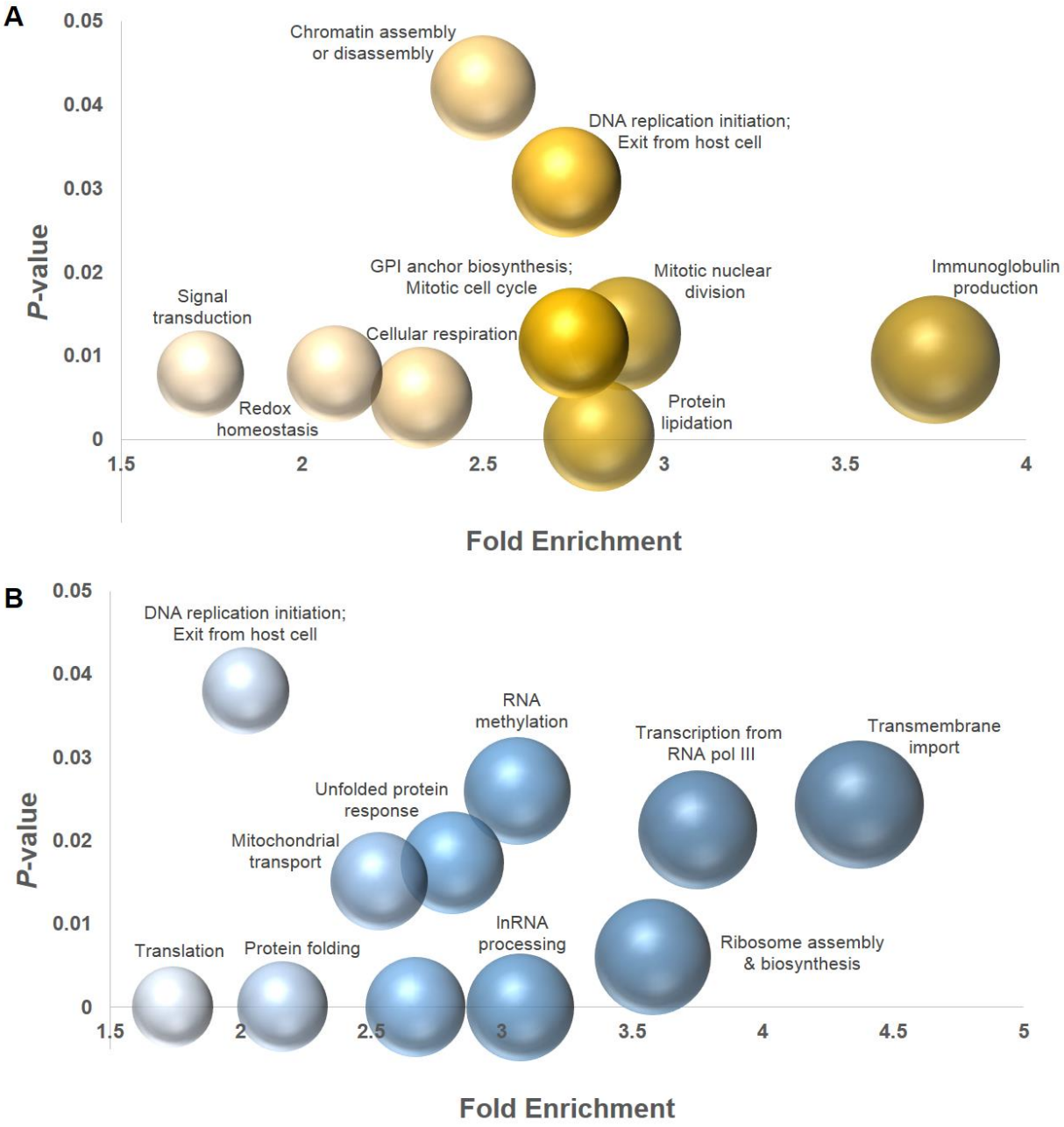

Figure 14. Gene ontology enrichments for differentially expressed genes during cell cycle control. Bubble plots showing gene ontology enrichment of gene cluster with **(A)** decreased transcript abundance upon arrest and transcript increase upon re-entry or **(B)** increase in transcript abundance upon arrest and decrease in abundance upon re-entry. Biological processes that were significantly enriched for are shown as bubbles plotted against the fold enrichment over the associated *P*-value<0.05. Sizes of bubble correspond to the percentage of genes related to a gene ontology term found in the data set.

The differentially expressed genes were subsequently evaluated for prior genetic attenuation of their activities *in vitro*, revealing functional information of the biological consequences of gene knockout. Only ~10% (203 of 2 295) of the genes have been previously phenotyped in *P. falciparum,* with phenotypic outcomes ranging from refractory to genetic attenuation (implying essential functions in

the *P. falciparum* parasites) to no discernable difference in mutant lines compared to wild-type parasites (Fig. 15).

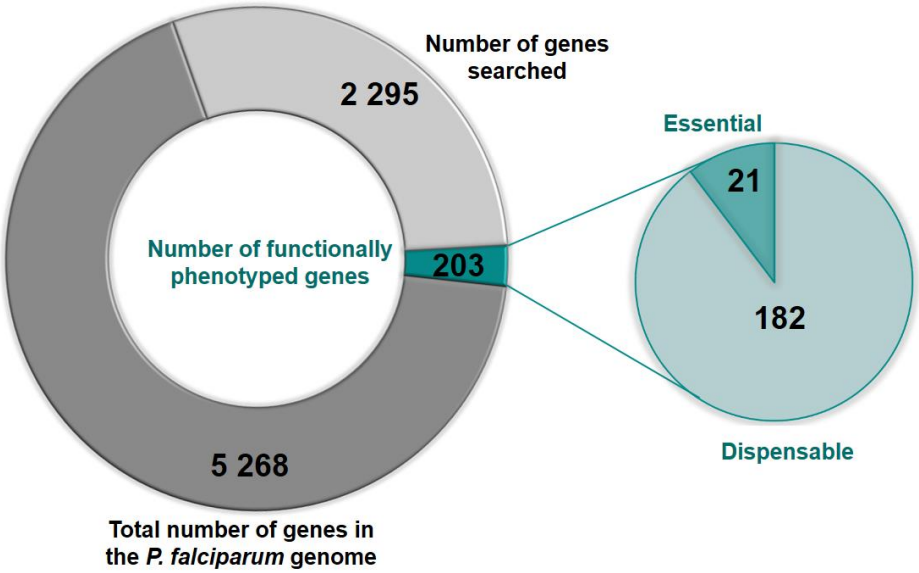

Figure 15. Assessment of functional phenotypes in *P. falciparum*. The total number of genes in the *P. falciparum* genome (dark grey) that have been searched (light grey) for functional phenotypes from the curated PhenoPlasmDB database are shown in proportion to the number of genes with phenotype data (dark teal). Viable phenotypes are sorted according to phenotypes that showed a difference in the normal wild-type (WT) phenotype, i.e. likely essential (teal) or showed no difference in the parasite stages under investigation, i.e. dispensable genes (light teal)

Such genes for which prior evidence of functional validation in *P. falciparum* existed were removed from the dataset, yielding 2 092 genes that have yet to be functionally phenotyped. To associate these genes with putative functions in cell cycle regulation, their presence in annotation clusters related to the cell cycle from PlasmoDB and literature (69,119,120,130) was determined. Subsequently, a total of 86 genes were linked to previous descriptors of their involvement in the cell cycle (Table 3). This included 66 genes for which no prior evidence of functional validation through genetic manipulation is available. Additionally, 20 genes, which have been previously functionally validated in literature, were included in the table for comparison.

Table 3. Genes from cell cycle arrest and re-entry study that shows overlap with cell cycle association.

Gene ID	Old ID	Description	Biological Process	Expression				Phase	Ref	GM[a]	Phenotype	Ref
				AR	RE1	RE2	RE3					
PF3D7_0312400	PFC0525c	Glycogen synthase kinase 3	Protein phosphorylation	-0.7976	0.0745	0.2075	0.6353	G_0-phase	(130,131)	Yes	Refractory	(65,132)
PF3D7_1323900	PF13_0136	Protein BCP1, putative	Protein transport	0.5829	-0.2564	-0.3805	-0.5649	G_1-phase	(69)	No		
PF3D7_0527000	PFE1345c	DNA replication licensing factor MCM3, putative	DNA replication initiation	-0.9672	0.1143	0.2373	0.8442	G_1-phase	(130)	No		
PF3D7_0705300	**MAL7P1.21**	**Origin recognition complex subunit 2, putative**	**DNA replication initiation**	**-0.632**	**0.011**	**0.012**	**0.736**	**G_1-phase**	(130)	**No**		
PF3D7_0705400	PF07_0023	DNA replication licensing factor MCM7	DNA replication initiation	-0.7794	0.0291	0.0159	0.4925	G_1-phase	(130,133)	No		
PF3D7_1203000	PFL0150w	Origin recognition complex subunit 1	DNA replication initiation	-0.9622	0.2634	0.2899	0.685	G_1-phase	(120,130, 133)	No		
PF3D7_1211300	PFL0560c	DNA helicase MCM8, putative	DNA replication initiation	-1.1308	-0.1727	0.1087	0.6414	G_1-phase	(130)	No		
PF3D7_1211700	**PFL0580w**	**DNA replication licensing factor MCM5, putative**	**DNA replication initiation**	**-1.04**	**-0.166**	**0.213**	**0.924**	**G_1-phase**	(130)	**No**		
PF3D7_1317100	PF13_0095	DNA replication licensing factor MCM4	DNA replication initiation	-1.2115	0.0442	0.2567	0.9728	G_1-phase	(130,134)	Yes	Refractory	(65)
PF3D7_1355100	PF13_0291	DNA replication licensing factor MCM6	DNA replication initiation	-0.707	-0.1599	0.3226	0.556	G_1-phase	(130,133)	No		
PF3D7_1361800	MAL13P1.308	Glideosome-associated connector	Entry into host cell	-3.3566	-0.1847	0.347	2.6763	G_1-phase	(135)	No		
PF3D7_0503200	PFE0155w	Conserved *Plasmodium* protein, unknown function	DNA replication	-0.6834	0.184	0.2588	0.8906	G_1-phase	(120,130, 133)	No		
PF3D7_0615500	PFF0750w	Cdc2-related protein kinase 5	Protein phosphorylation	-1.2317	0.0868	0.2432	1.028	G_1-phase	(65)	Yes	Dispensable	(65)
PF3D7_1334100	PF13_0189	Conserved *Plasmodium* protein, unknown function	DNA replication	-0.956	0.1142	0.4242	0.6733	G_1-phase	(130)	No		
PF3D7_0213700	PFB0620w	Conserved protein, unknown function	Unknown	0.9931	0.3161	0.1544	-1.2271	G_1-phase	(69)	Yes	Refractory	(136)
PF3D7_1014600	PF10_0143	Transcriptional coactivator ADA2	Regulation of transcription, DNA-templated	0.0802	-0.1493	-0.5416	-0.0748	G_1-phase	(137)	No		
PF3D7_0415300	PFD0740w	Cdc2-related protein kinase 3	Protein phosphorylation	0.8621	-0.1246	0.075	-0.6761	G_1/S-phase	(49,130)	Yes	Refractory	(65)
PF3D7_0415300	PFD0740w	Cdc2-related protein kinase 3	Protein phosphorylation	0.8621	-0.1246	0.075	-0.6761	G_1/S-phase	(49,130)	Yes	Refractory	(65)
PF3D7_0218000	PFB0840w	Replication factor C subunit 2, putative	DNA replication	-0.7024	0.0643	0.5343	0.3808	S-phase	(130)	No		
PF3D7_0317200	PFC0755c	Cdc2-related protein kinase 4	Protein phosphorylation	-0.771	-0.1403	-0.0169	0.5367	S-phase	(120,130)	Yes	Cell cycle arrest	(65,68)
PF3D7_0509100	PFE0450w	Structural maintenance of chromosomes protein 4, putative	Chromosome organization	-1.7252	0.0583	0.9059	1.2867	S-phase	(130,138)	No		
PF3D7_1107400	PF11_0087	DNA repair protein RAD51	DNA repair	-1.2913	0.0533	0.7354	0.5134	S-phase	(130,139)	No		
PF3D7_1227500	PFL1330c	Conserved *Plasmodium* protein, unknown function	Unknown	-2.4989	-0.0875	0.2969	1.4657	S-phase	(130,140)	No		
PF3D7_1361900	PF13_0328	Proliferating cell nuclear antigen 1	Regulation of DNA replication	-0.587	0.1808	0.2978	0.3172	S-phase	(130,141)	No		
PF3D7_1241700	PFL2005w	Replication factor C subunit 4, putative	DNA replication	-0.4436	0.1624	0.5872	0.1412	S-phase	(130)	No		
PF3D7_1337100	PF13_0206	Protein kinase 6	Protein phosphorylation	1.1248	0.091	0.5273	-1.3489	S-phase	(130,142)	No		
PF3D7_1463200	PF14_0601	Replication factor C subunit 3, putative	DNA replication	-0.2473	0.2555	0.5407	0.2021	S-phase	(130)	No		

PF3D7_1107400	PF11_0087	DNA repair protein RAD51	DNA repair	-1.2913	0.0533	0.7354	0.5134	S-phase	(130,139)	No		
PF3D7_1107400	PF11_0087	DNA repair protein RAD51	DNA repair	-1.2913	0.0533	0.7354	0.5134	S-phase	(130,139)	No		
PF3D7_0302100	PFC0105w	Serine/threonine protein kinase	Protein phosphorylation	0.8584	-0.064	-0.2921	-0.9778	S-phase	(137)	Yes	Refractory	(65)
PF3D7_1145200	PF11_0464	Serine/threonine protein kinase, putative	Protein phosphorylation	-1.2738	0.1858	-0.0118	1.3509	S-phase	(137)	Yes	Refractory	(65)
PF3D7_0719200	MAL7P1.100	NIMA related kinase 4	Protein phosphorylation	0.4454	0	1.9899	0.6662	S-phase, G_2-phase	(143)	Yes	Cell cycle arrest	(143,144)
PF3D7_1356900	MAL13P1.279	Protein kinase 5	Protein phosphorylation	-0.7563	0.2239	0.6749	0.5166	S-phase, G_2-phase, spindle checkpoint	(29,130)	Yes	Refractory	
PF3D7_0719200	MAL7P1.100	NIMA related kinase 4	Protein phosphorylation	0.4454	0	1.9899	0.6662	S-phase, G_2-phase	(143)	Yes	Cell cycle arrest	(143,144)
PF3D7_1356900	MAL13P1.279	Protein kinase 5	Protein phosphorylation	-0.7563	0.2239	0.6749	0.5166	S-phase, G_2-phase, spindle checkpoint	(29,130)	Yes	Refractory	
PF3D7_1228300	PFL1370w	NIMA related kinase 1	Protein phosphorylation	-0.775	0.1485	0.2336	0.5935	G_2-phase	(130,145)	Yes	Refractory	(65)
PF3D7_0309200	PFC0385c	Serine/threonine protein kinase, putative	Protein phosphorylation	-0.7974	0.2703	0.1644	0.7899	G_2-phase	(130,146)	Yes	Refractory	(65)
PF3D7_1305500	PF13_0027	Mitogen-activated protein kinase phosphatase 1, putative	Cell cycle	0.6857	-0.2651	-0.4121	-0.7088	G_2-phase	(130,147)	No		
PF3D7_1246900	PFL2250c	RAC-beta serine/threonine protein kinase	Protein autophosphorylation	-0.0439	0.1908	0.5318	0.6386	G_2-phase, spindle checkpoint	(120,130)	Yes	Refractory	(65)
PF3D7_0723300	MAL7P1.125	Conserved *Plasmodium* protein, unknown function	Regulation of transcription, DNA-templated	-1.1242	-0.1036	0.1843	1.0414	S/M-phase	(69)	No		
PF3D7_0903600	PFI0175w	Conserved *Plasmodium* protein, unknown function, unspecified product	ATP activity	0.0294	0.1785	0.6061	0.1221	S/M-phase	(69)	No		
PF3D7_1409400	PF14_0092	Conserved *Plasmodium* membrane protein, unknown function	Unknown	-1.4866	-0.1104	0.3976	1.4317	S/M-phase	(69)	No		
PF3D7_0204100	PFB0190c	Conserved *Plasmodium* protein, unknown function	Unknown	-1.9161	0.0983	0.2118	2.2945	S/M-phase	(69)	No		
PF3D7_0210600	PFB0475c	Conserved *Plasmodium* protein, unknown function	Unknown	-0.716	-0.1324	0.2939	0.6641	S/M-phase	(69)	No		
PF3D7_0419700	PFD0955w	Apical merozoite protein	Host cell entry	-1.4817	0.0253	0.1935	1.6249	S/M-phase	(69)	No		
PF3D7_1136200	PF11_0373	Conserved *Plasmodium* protein, unknown function	Unknown	-1.8017	0.2117	-0.273	1.9726	S/M-phase	(69)	No		
PF3D7_1145600	PF11_0467	Conserved protein, unknown function	Unknown	-1.6602	0.0996	0.2057	0.7227	S/M-phase	(69)	No		
PF3D7_1221300	PFL1025c	Conserved *Plasmodium* protein, unknown function	Unknown	-0.1742	0.1309	0.3268	0.8129	S/M-phase	(69)	No		
PF3D7_1313600	PF13_0079	Conserved *Plasmodium* protein, unknown function	Unknown	-0.3893	-0.1213	-0.9055	0.9038	S/M-phase	(69)	No		
PF3D7_1327500	MAL13P1.152	Conserved *Plasmodium* protein, unknown function	Unknown	-2.393	0.3315	0.6423	2.0958	S/M-phase	(69)	No		
PF3D7_1435200	PF14_0333	Conserved *Plasmodium* protein, unknown function	Unknown	-0.8255	0.2664	0.3026	0.7769	S/M-phase	(69)	No		
PF3D7_1463900	PF14_0607	EF-hand calcium-binding domain-containing protein, putative	Calcium ion binding	-1.0336	0.0269	0.1629	1.8413	S/M-phase	(69)	No		
PF3D7_0803200	MAL8P1.146	Filament assembling protein, putative	Structural constituent of cytoskeleton	-2.6774	0.0822	0.5373	2.2411	M-phase	(137)	No		
PF3D7_0515300	PFE0765w	Phosphatidylinositol 3-kinase	Intracellular transport	-0.5252	0.1789	-0.0536	0.4479	Checkpoint	(70,120)	No		
PF3D7_0803400	PF08_0126	DNA repair and recombination protein RAD54, putative	DNA recombination	-0.8423	0.0418	0.0863	0.6055	Checkpoint	(49,130)	No		
PF3D7_0206000	PFB0265c	DNA repair protein RAD2, putative	DNA repair	0.7225	-0.1609	-0.3295	-0.4722	Checkpoint	(70)	No		
PF3D7_0803400	PF08_0126	DNA repair and recombination protein RAD54, putative	DNA repair	-0.8423	0.0418	0.0863	0.6055	Checkpoint	(49,130)	No		

PF3D7_0803400	PF08_0126	DNA repair and recombination protein RAD54, putative	DNA repair	-0.8423	0.0418	0.0863	0.6055	Checkpoint	(49,130)	No		
PF3D7_1238900	PFL1885c	Protein kinase 2	Protein phosphorylation	1.0545	-0.8474	-1.3629	-1.2866	Spindle checkpoint	(120,130)	Yes	Refractory	(65)
PF3D7_0420300	PFD0985w	AP2 domain transcription factor, putative	Regulation of transcription, DNA-templated	1.3254	-0.307	-0.6483	-1.2942	Cell cycle progression	(137)	No		
PF3D7_0516800	PFE0840c	AP2 domain transcription factor AP2-O2, putative	Regulation of transcription, DNA-templated	-2.2543	-0.0133	0.0258	1.2879	Cell cycle progression	(137)	No		
PF3D7_0730300	PF07_0126	AP2 domain transcription factor AP2-L, putative	Regulation of transcription, DNA-templated	1.6463	0.5317	-0.164	-2.1034	Cell cycle progression	(137)	No		
PF3D7_1007700	PF10_0075	AP2 domain transcription factor, putative	Regulation of transcription, DNA-templated	0.5305	0.0419	-0.6487	-0.7487	Cell cycle progression	(137)	Yes	Refractory	(148)
PF3D7_1143100	PF11_0442	AP2 domain transcription factor AP2-O, putative	Regulation of transcription, DNA-templated	-0.6213	0.2281	0.3742	0.7521	Cell cycle progression	(137)	No		
PF3D7_1239200	**PFL1900w**	**AP2 domain transcription factor, putative**	**Regulation of transcription, DNA-templated**	**-0.895**	**0.076**	**-0.358**	**-0.14**	**Cell cycle progression**	(137)	**No**		
PF3D7_1454400	PF14_0517	Aminopeptidase P	Proteolysis	0.5182	0.5231	0.3386	-0.6925	Cell cycle progression	(137)	No		
PF3D7_1456000	PF14_0533	AP2 domain transcription factor, putative	Regulation of transcription, DNA-templated	-0.8013	-0.6125	-0.0275	0.4649	Cell cycle progression	(119)	No		
PF3D7_0613800	**PFF0670w**	**AP2 domain transcription factor, putative**	**Regulation of transcription, DNA-templated**	**-2.749**	**-0.197**	**-0.061**	**1.757**	**Cell cycle progression**	(119)	**No**		
PF3D7_1107800	PF11_0091	AP2 domain transcription factor, putative	Regulation of transcription, DNA-templated	1.517	-0.3791	-0.8155	-1.3641	Cell cycle progression	(119)	No		
PF3D7_1317200	PF13_0097	AP2 domain transcription factor, putative	Regulation of transcription, DNA-templated	0.8289	0.1445	-0.3694	-1.0783	Cell cycle progression	(119)	No		
PF3D7_1139300	PF11_0404	AP2 domain transcription factor, putative	Regulation of transcription, DNA-templated	0.9094	0.177	0.2401	-1.0453	Cell cycle progression	(119)	No		
PF3D7_1342900	PF13_0235	AP2 domain transcription factor, putative	Regulation of transcription, DNA-templated	0.5719	-0.0084	-0.466	-1.8088	Cell cycle progression	(119)	No		
PF3D7_1408200	PF14_0079	AP2 domain transcription factor AP2-G2, putative	Regulation of transcription, DNA-templated	1.1163	-0.259	-0.2871	-1.2515	Cell cycle progression	(119)	No		
PF3D7_1209300	PFL0465c	Zinc finger transcription factor, putative	Regulation of transcription, dna-templated	0.2285	-0.3491	-0.7662	-0.1639	Cell cycle progression	(137)	No		
PF3D7_0802100	MAL8P1.153	AP2 domain transcription factor, putative	Regulation of transcription, DNA-templated	-2.3139	-0.0952	0.2314	1.7239	Cell cycle progression	(119)	Yes	Refractory	(132)
PF3D7_0604100	PFF0200c	AP2 domain transcription factor	Regulation of transcription, DNA-templated	-2.9004	-0.2325	-0.4352	2.7892	Cell cycle progression	(119)	Yes	Refractory	(149)
PF3D7_1466400	PF14_0633	AP2 domain transcription factor AP2-SP	Regulation of transcription, DNA-templated	0.1262	0.1231	-0.2437	-0.6199	Cell cycle progression	(119)	No		
PF3D7_0622900	PFF1100c	AP2 domain transcription factor AP2-SP3, putative	Regulation of transcription, DNA-templated	0.0915	-0.1664	-0.009	-0.5063	Cell cycle progression	(119)	No		
PF3D7_1350900	PF13_0267	AP2 domain transcription factor AP2-O4, putative	Regulation of transcription, DNA-templated	-0.5898	0.147	0.5031	0.7344	Cell cycle progression	(119)	No		
PF3D7_0708400	PF07_0029	Heat shock protein 90	Protein folding	0.8097	0.6068	0.3532	-0.6634	Other	(137)	No		
PF3D7_1118200	PF11_0188	Heat shock protein 90, putative	Protein folding	1.4827	-0.023	-0.0901	-1.7432	Other	(137)	No		
PF3D7_0213100	PFB0595w	Protein SIS1	Protein folding	1.3459	-0.3062	-0.5767	-1.4846	Other	(137)	No		
PF3D7_0711000	PF07_0047	AAA family atpase, CDC48 subfamily	ER-associated ubiquitin-dependent protein catabolic process	0.395	-0.0042	0.0243	-0.625	Other	(137)	No		
PF3D7_0818900	PF08_0054	Heat shock protein 70	Protein folding	0.4605	0.7242	-0.302	-0.8676	Other	(137)	No		
PF3D7_0917900	PFI0875w	Heat shock protein 70	Protein folding	-0.6362	-0.0147	-0.1237	0.5964	Other	(137)	No		
PF3D7_1015300	PF10_0150	Methionine aminopeptidase 1b, putative	Proteolysis	1.2964	0.0111	-0.1709	-1.4755	Other	(137)	No		
PF3D7_1134000	PF11_0351	Heat shock protein 70	Protein folding	1.4924	0.1685	0.1369	-1.7208	Other	(137)	No		
PF3D7_1337000	MAL13P1.184	Mitochondrial intermediate peptidase, putative	Proteolysis	1.5357	-0.2917	-0.2441	-1.3297	Other	(137)	Yes	Refractory	(65)

[a]GM= genetically modified.
Genes selected for further investigation are shown in bold.

From the above, the four target genes prioritised for functional validation included two members of the ApiAP2 transcription factor family, PF37D_1239200 and PF3D7_0613800, and two members of the pre-replicative complex, PF3D7_1211700 (DNA replication licensing factor MCM5, putative) and PF3D7_0705300 (origin of replication complex subunit 2, putative). These choices were made for three main reasons: 1) transcription factors have obvious functions as potential cell cycle regulators and the role of these transcription factors has yet to be explored 2) components of the pre-replicative complex may either directly act as regulators (causal) or be downstream indicators of important regulation events (effect) and 3) strong differential expression was observed for all these genes in question.

Both PF3D7_0613800 and PF37D_1239200 are DNA-binding transcription factors from the AP2 family of transcription factors that have previously been implicated as major role players for gene regulation during parasite development. The ApiAP2 genes express at distinct times during the 48 h IDC, notably in early to late schizont stages; PF37D_1239200 mRNA abundance peaks at 32 hpi, while PF3D7_0613800 mRNA abundance peaks at 42-48 hpi when the parasite cell cycle is alternating between S- and M-phases. In response to cell cycle arrest, PF3D7_0613800 had a significant decrease (log_2 FC of -2.75) in mRNA abundance but after cell cycle re-entry, significantly increased (log_2 FC of 1.75). Interestingly, PF37D_1239200 showed a significant decrease in mRNA abundance upon cell cycle arrest, while upon re-entry were present at steady transcript levels. It is currently hypothesised that members of the AP2 family of transcription factors function synergistically in a hierarchical network, allowing a complex program of gene regulation that enables the parasite to progress and develop during the asexual and sexual stages (150). Therefore, these genes are potential candidate cell cycle regulators and were thus selected for further investigation.

Genes involved in the origin of replication complex, i.e. MCM5 subunit (PF3D7_1211700) and ORI complex subunit 2 (PF3D7_0705300) function in licencing DNA for DNA replication typically in the G_1-phase of the cell cycle. Both of these genes show peak gene expression between 42-48 hpi. Upon cell cycle arrest, both genes showed significant decreases in expression, but after cell cycle re-entry they significantly increase to normal steady state levels. The successful assembly of the pre-replicative complex is critical in G_1/S cell cycle progression, and the essential role of each component of the re-replicative complex must be investigated for their possible alternative capacity in cell cycle regulation.

3.2. Design of guide RNAs for targeted gene disruption by Cas9

3.2.1 Guide design

Two gRNAs were designed and selected per target gene to compare efficiency upon successful knockout, with the two gRNAs differing only in location within the gene (on- and off-target scores were negligible). BLAST nucleotide homology search of each guide within the *P. falciparum* genome revealed significant sequence homology to the target genome region (*e*-value <0.0004), and *e*-values >23 for off-target regions. Guides meeting optimal design scores and criteria (as per section 2.2) were found in the intragenic regions, and thus, homology regions flanking the gRNA sites with functional domains annotated. Relative positions of guide sequences, homology regions for DNA repair, exons, active domains and gene loci in the chromosomes of the *P. falciparum* parasites genome are schematically represented in Fig. 16.

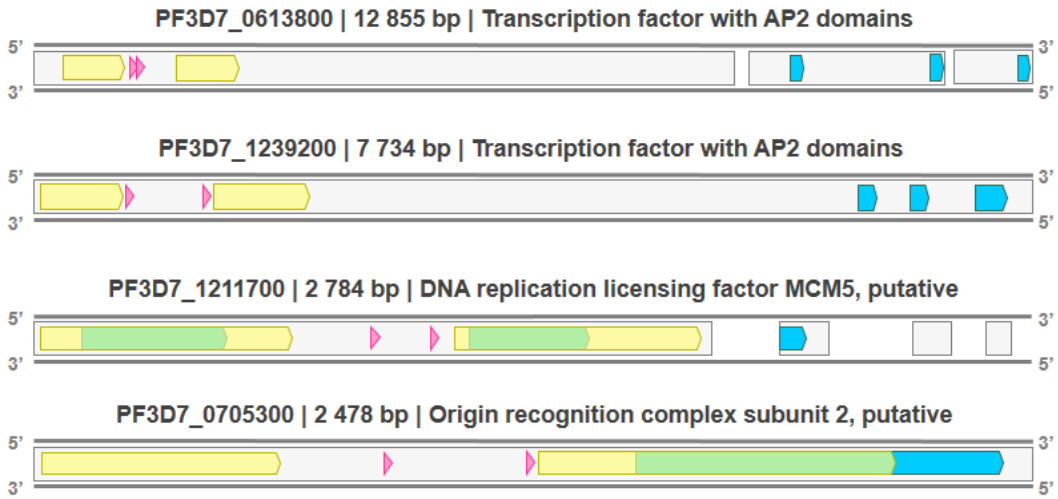

Figure 16. Schematic of genes PF3D7_1239200, PF3D7_0613800, PF3D7_1211700, and PF3D7_0705300, showing the relative proximity of guide, homology regions, exons and predicted functional domains. The complete gene is shown, including exon(s) (grey blocks), homology regions (yellow), guide sequences Guide 1 and Guide 2 (pink) and predicted functional domains (blue).

gRNAs were designed specifically between flanking homology regions associated as close to the 5´ region of each gene as possible. The final guide sequences and their characteristics are provided in Table 4.

Table 4. Single guide RNAs designed for a targeted knockout of cell cycle related genes.

Gene	Possible no. of gRNAs	Guide	Sequence	PAM	GC%	Strand	Nucleotide position	On-T[a]	Off-T[a]	BLAST (e-value)[b]
PF3D7_1239200	215	1	TCTGGAATTCAACAAAACGT	TGG	35	+1	1635343-1635362	62.1	49.9	$3x10^{-4}$
		2	GGTGTATTAAATAATTCGTC	TGG	30	+1	1636043-1636062	50.3	49.9	$3x10^{-4}$
PF3D7_0613800	403	1	GAGCTCTCCCGATACATCGT	TGG	55	-1	567353-567372	64.9	50.0	$3x10^{-4}$
		2	AGAGCTCCCTGATGTTGCTA	TGG	50	+1	567366-567385	57.4	100	$3x10^{-4}$
PF3D7_1211700	121	1	CCTAGACCTACCAGAGCAGA	TGG	55	+1	522629-522648	57.9	50	$3x10^{-4}$
		2	TGCCCCCGAATTATATGGAA	TGG	45	+1	522799-522818	53.4	50	$3x10^{-4}$
PF3D7_0705300	67	1	TTTAGCCTTATCGTCGACTT	TGG	40	-1	318217-318236	44.7	100	$3x10^{-4}$
		2	CCAACATTATATGGTATAGA	AGG	30	+1	318570-318589	51.4	99.2	$3x10^{-4}$

[a] On-T / Off-T = On-target / Off-target. Scores closer to 100 is preferred.
[b] e-value: Expectant value, closer to zero means more significant homology.

Sense and antisense guide oligonucleotide pairs were annealed to obtain double stranded guide oligonucleotides per gene for downstream cloning.

3.2.2 Cloning and sequencing of guide oligonucleotides into the pDC2-U6-co*CAS9*-CAM-*hDHFR* vector

The pDC2-U6-co*CAS9*-CAM-*hDHFR* vector was digested with *Bbs*I restriction enzyme digested and dephosphorylated, after which the vector was column purified, yielding DNA concentrations of 179-345.3 ng/µL, and purity indicators of A_{260}/A_{280} and A_{260}/A_{230} ranging between 1.76-1.92. Guide oligonucleotides were phosphorylated and the sense and antisense guide oligonucleotide pairs annealed and subsequently ligated and transformed into homemade competent *E. coli* XL10-Gold cells. Recombinant colonies were inoculated, and the purified vector DNA purified was determined to be ~220-400 ng/µl and purity indicators of A_{260}/A_{280} and A_{260}/A_{230} ranging from 1.82-2.01. Sanger sequencing confirmed successful cloning of the guide sequences under the U6 promoter region for every gene investigated (Fig. 17).

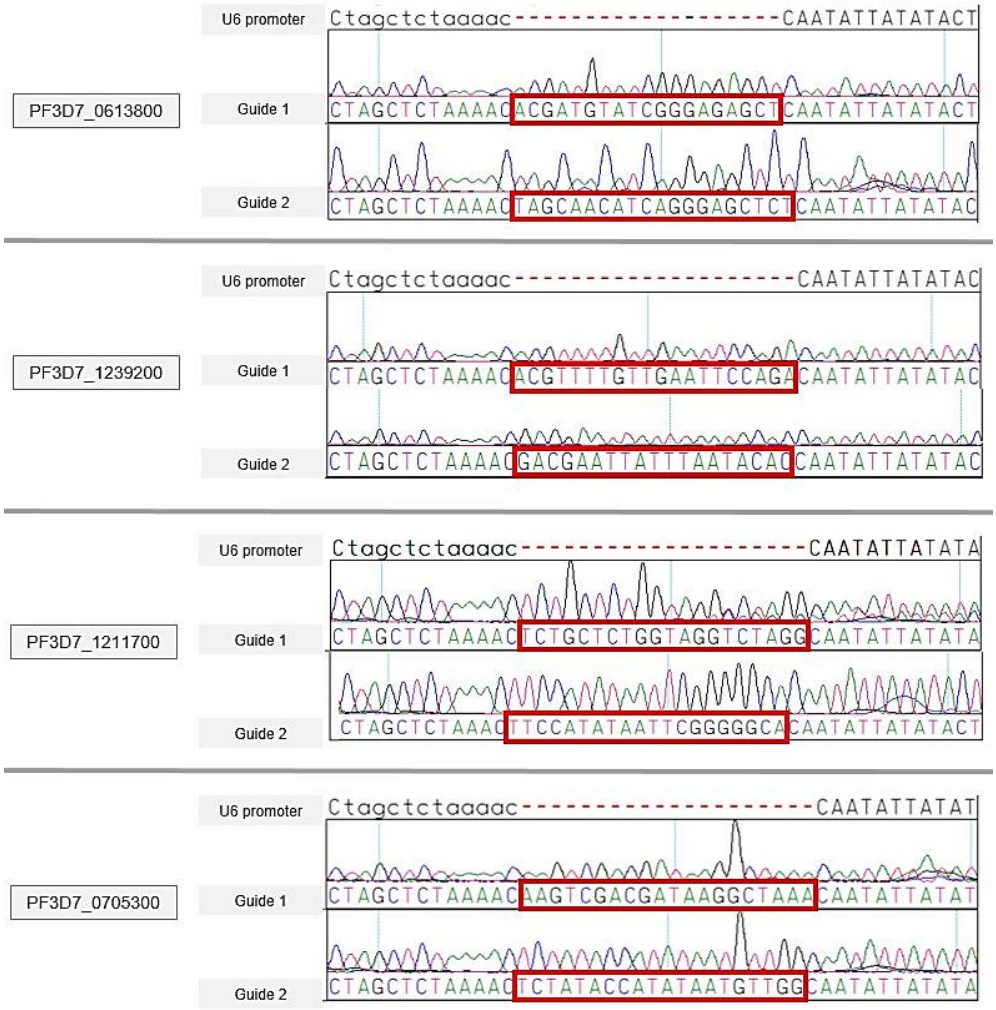

Figure 17. Sequence chromatograms showing multiple sequence alignments of the cloned guide oligonucleotides into the U6 promoter cassette of the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector. Oligonucleotide guide pairs were annealed, phosphorylated and ligated into the *BbsI* restriction enzyme generated cut sites. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm). Successfully cloned guide oligonucleotides (Table 4, reverse guide of each) are shown in a red square.

Successful cloning of the gRNAs and complete construction of the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector signified that two of the three components required for CRISPR-Cas9 mediated KO was achieved and ready for transfection.

3.3 Design and construction of the KO cassettes

KO cassettes were constructed from homology regions for each of the four genes under investigation. These were amplified from genomic DNA from *in vitro* cultures of *P. falciparum* parasites (Fig. 18). Parasites matured completely to the metabolically active trophozoite stage, as evident from the presence of haemozoin crystals in the parasites. Moreover, these parasites have not yet undergone nuclear division as evidenced by large nuclear stains on Giemsa microscopy and the absence of visible punctate daughter cells.

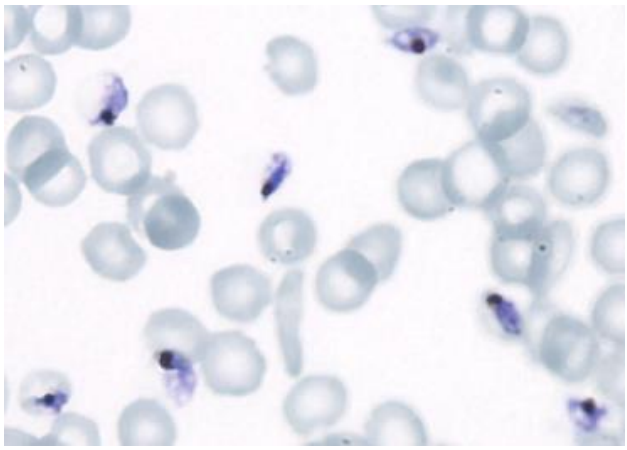

Figure 18. Genomic DNA extraction from *P. falciparum* NF54 parasites culture. Genomic DNA was extracted from a 13.4% parasitemia 34-38 h trophozoite culture (98% trophozoite, 2 % rings) using a silica column based method.

High yield, pure genomic DNA (71.2 ± 11.4 ng/µL) was extracted from these late trophozoite stages (~34-38 hpi). The purity of the genomic DNA extracted from *P. falciparum* parasites was spectrophotometrically determined, with A_{260}/A_{280} and A_{260}/A_{230} values ranging from 1.81-1.99.

Gene-specific regions were successfully amplified from the genomic DNA. PCR conditions were optimised for DNA template quantity, denaturation and annealing temperatures and duration, and extension times (Fig. 19).

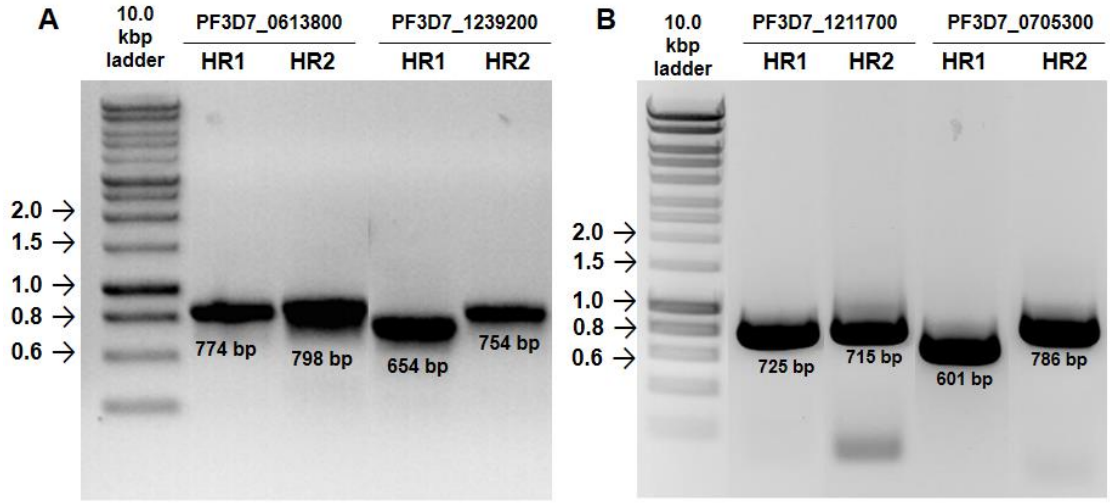

Figure 19. PCR amplification of gene specific homology regions. (A) Homology region amplification of PF3D7_0613800 and PF3D7_1239200. Expected sizes are shown in black below the DNA bands. **(B)** Homology region amplification of PF3D7_1211700 and PF3D7_0705300. PCR products were visualised with a 0.8% (w/v) agarose/TAE gel with SYBR® Safe DNA Gel Stain, under UV light. Expected sizes based on designed fragments are indicated, actual sizes based on Rf values were not determined.

All PCR products obtained were of the expected relative size with a faint non-specific band only observed in HR2 of PF3D7_1211700. All DNA bands of the correct size were gel extracted and purified, yielding 40-70 ng/µL DNA with A_{260}/A_{280} and A_{260}/A_{230} ratios ranging from 1.8-2.1.

To complete the construction of the KO cassette, the *bsd* selectable marker transgene, flanked by the *hrp2* 3´ UTR and *PcDT* 5´ UTR (see section 2.3), was amplified from the pDC2-*eGFP*-CAM-*BSD* vector (Fig. 20).

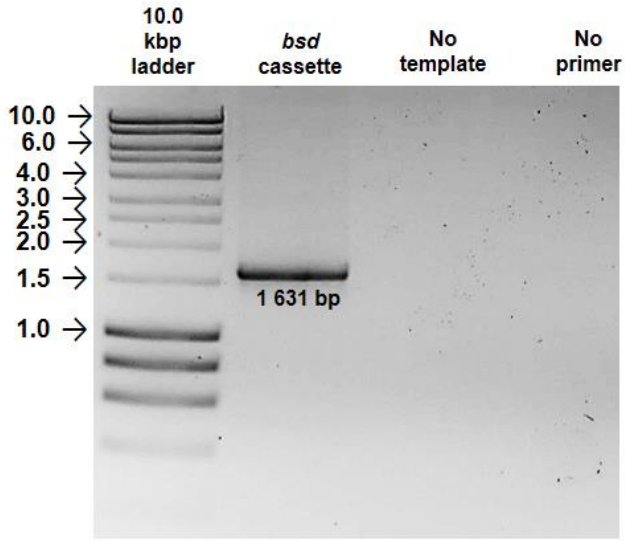

Figure 20. PCR amplification of the *bsd* cassette from pDC2-*eGFP*-CAM-*BSD* vector. Approximately 280 ng DNA was separated on a 0.8% (w/v) agarose/TAE gel and visualised with SYBR® Safe DNA Gel Stain. Expected sizes based on designed fragments are indicated, actual sizes based on Rf values were not determined.

The amplified *bsd* cassette separated according to its expected size of 1 631 bp, made up of the 585 bp *hrp* 3´ UTR, 412 bp *bsd* and 599 bp *PcDT* 5´ UTR. No non-specific amplification was observed, therefore, the primers and PCR conditions were considered optimal. This PCR product was also gel extracted and purified to yield 138 ng/µL DNA with A_{260}/A_{280} and A_{260}/A_{230} ratios of 1.94 and 2.1, respectively.

All PCR products (HR1 and HR2 per gene and the *bsd* cassette insert) were subsequently combined for Gibson DNA assembly. Immediately after the assembly reaction, the mixture was subjected to long-range PCR amplification using the HR 1 forward and HR2 reverse primers, for every gene (Fig. 21).

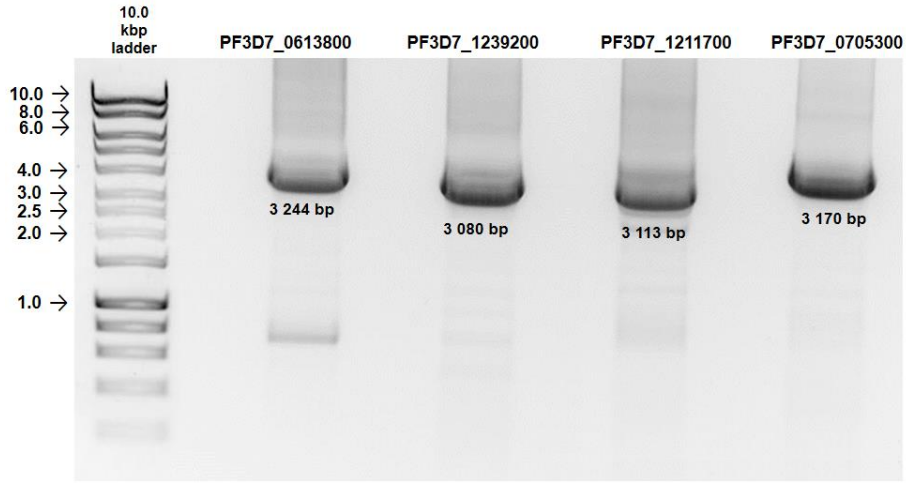

Figure 21. PCR amplification of KO cassettes for each of the 4 genes investigated. PCR amplification of the KO cassette immediately after Gibson assembly show the assembly of the three DNA fragments, yielding expected 3 244 bp (PF3D7_0613800), 3 080 bp (PF3D7_1239200), 3 113 bp (PF3D7_1211700) and 3 170 bp (PF3D7_0705300) DNA fragments. DNA was separated on a 0.8% (w/v) agarose/TAE gel and visualised with SYBR® Safe DNA Gel Stain. Expected sizes based on designed fragments are indicated, actual sizes based on Rf values were not determined.

PCR amplification of the assembled KO cassettes indicated the correct expected size ranges for each KO for each gene investigated. Expected bands were extracted and purified to yield 42.8-47.5 ng/µL DNA with A_{260}/A_{280} and A_{260}/A_{230} ratios ranging between 1.6-1.84. The assembled KO cassettes were subsequently Sanger sequenced from these purified PCR products to confirm the correct orientation following Gibson DNA assembly. In Fig. 22, a diagrammatic example is provided for gene PF3D7_0613800 whereas the sequences for the other three KO cassettes are supplied as Supplementary data.

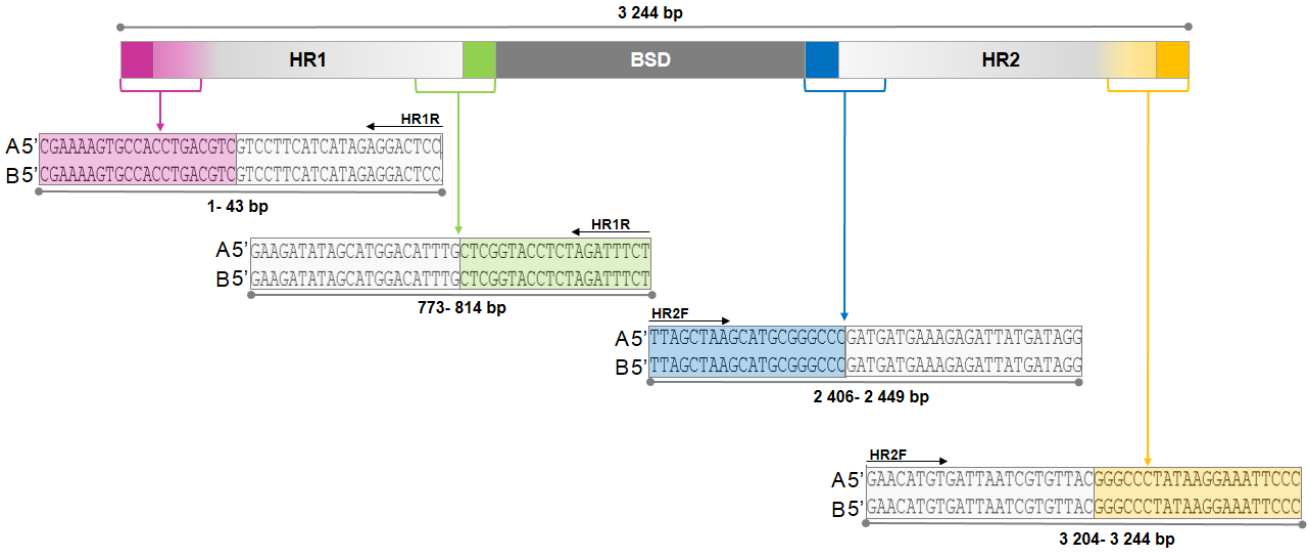

Figure 22. Example of PF3D7_0613800 PCR amplified KO cassette Sanger sequencing results. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Relative positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers HR1R (homology region 1 reverse primer) and HR2F (homology region 2 forward primer) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

Each KO cassette was cloned into a pCR®2.1 TOPO® expression vector under the *lacZ'* promoter, thus enabling blue white selection of recombinant bacteria episomally maintained in the vector with the integrated KO cassette. Plasmid DNA was extracted after white colonies were inoculated and grown overnight, followed by PCR screening to confirm the presence of the cloned KO cassette (Fig. 23).

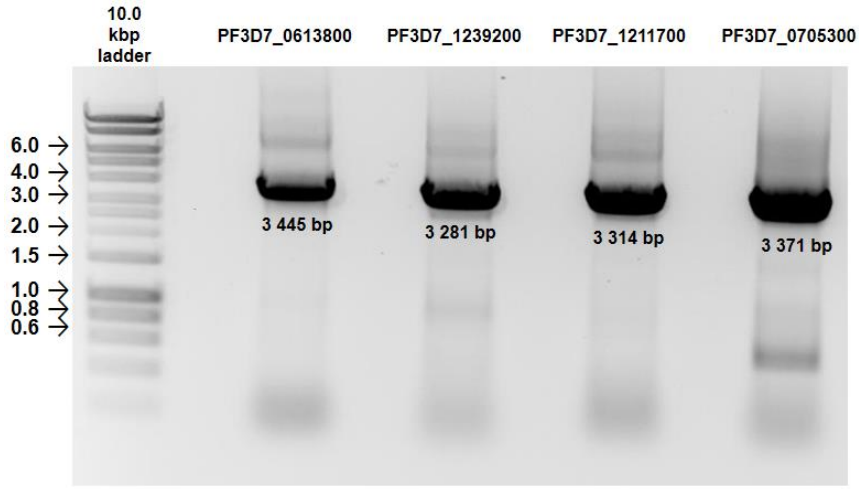

Figure 23. PCR amplification of assembled KO cassette for each gene investigated cloned into pCR®2.1 TOPO® expression vector. The pCR®2.1 TOPO® DNA was extracted from *E. coli* XL10 Gold cells with the KO cassette subsequently PCR amplified using vector sequencing primers M13F and M13R. The expected KO cassette band size was obtained, 3 445 bp (PF3D7_0613800), 3 281 bp (PF3D7_1239200), 3 314 bp (PF3D7_1211700) and 3 371 bp (PF3D7_0705300) DNA fragments. DNA was separated on a 0.8% (w/v) agarose/TAE gel and visualised with SYBR® Safe DNA Gel Stain.

The expected DNA sizes were observed, confirming the positive integration of each KO cassette into pCR®2.1 TOPO® vectors. Bacteria containing the vector were propagated from which the vector DNA was purified to yield 93-121 ng/µL DNA (A_{260}/A_{280} and A_{260}/A_{230} ratios ranging between 1.78-1.89). The integration of the KO cassettes was confirmed by Sanger sequencing, and as before, a diagrammatic representation of the data for one gene is provided as an example (Fig. 24, PF3D7_1239200). The sequences for the other three KO cassettes are supplied as supplementary data.

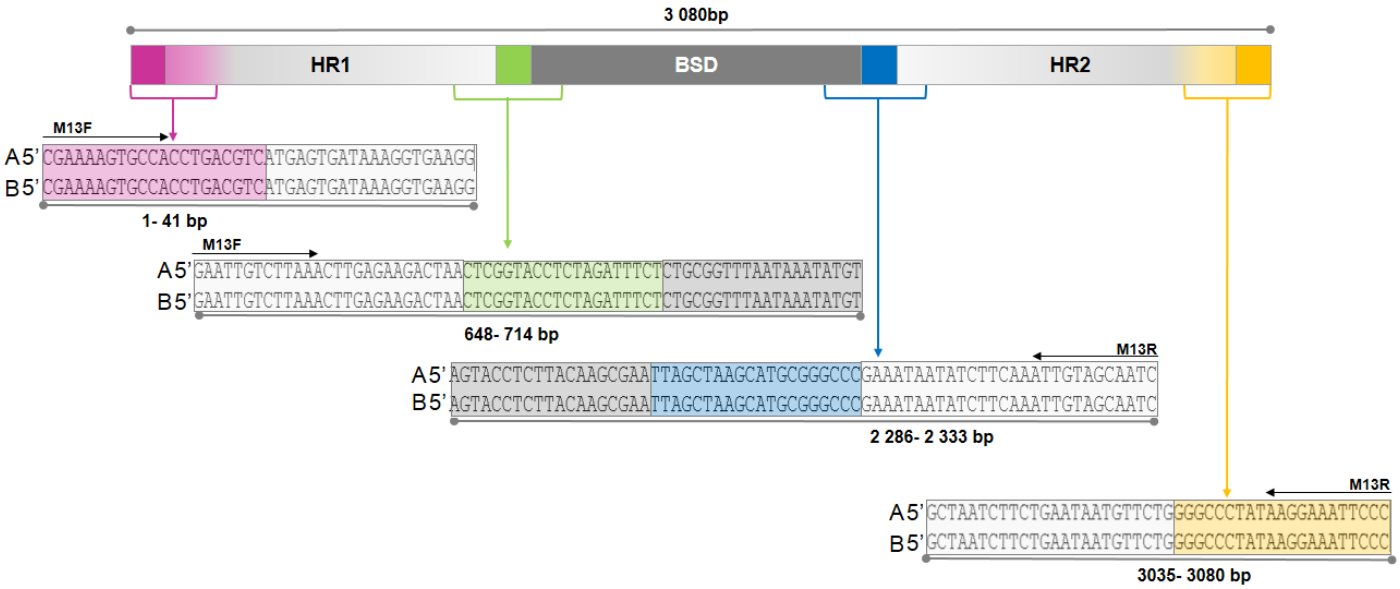

Figure 24. Sanger sequence results of PF3D7_1239200 KO cassette cloned into the pCR®2.1 TOPO® expression vector. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Relative positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers M13F (forward) and M13R (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

Sanger sequencing of the KO cassette from the pCR®2.1 TOPO® expression vector confirmed that the homology regions were assembled correctly and inserted into the vector as expected. The overlapping regions were sequenced, with 100% identity observed to the specific homology region for each gene and identity to the full *bsd* gene.

3.4 Transfection of *P. falciparum* NF54 and Dd2 parasites

The double transfection strategy required transfection of *P. falciparum* parasites (NF54 and Dd2 strains) with both the pCR®2.1TOPO®-^{KO} constructs as well as the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* constructs for each gene. Positive transfectants were selected only with WR99210, which selected for parasites carrying the *hDFHR* marker present on the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vectors. Double drug selection for both vectors negatively compromised parasite viability and could not be performed (personal communication, Marcus Lee, Wellcome Trust Sanger Institute). The selection was continued for 9-10 days post-transfection after which drug pressure was removed and parasites monitored for 6 weeks post transfection. Parasite recovery, i.e. expansion of potential recombinant parasites, was noted between 16-27 days post-transfection for both the NF54 and Dd2 strains of *P. falciparum* (Fig. 25). Transfection into the NF54 strain of *P. falciparum* yielded no recombinant parasites that recovered after WR99210 selection, compared to the more virulent Dd2 strain.

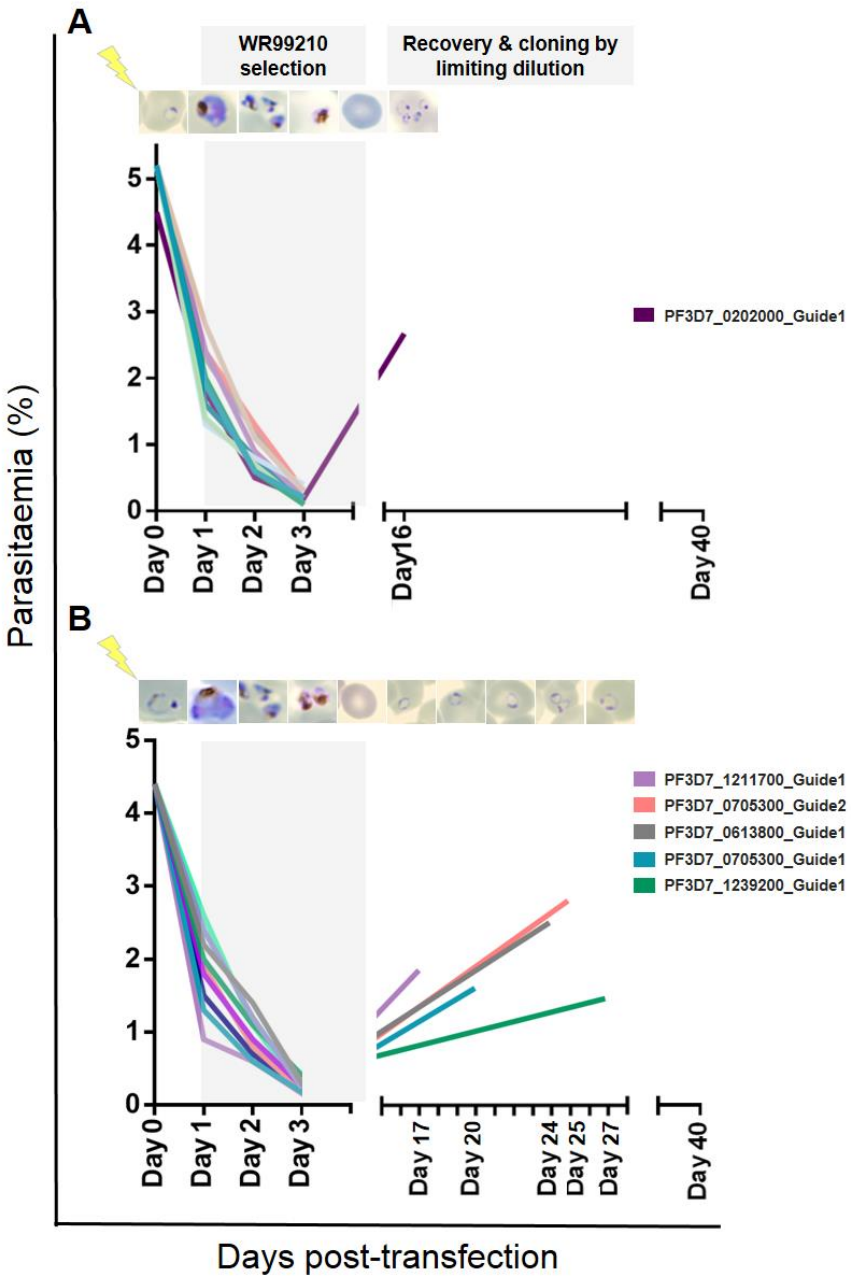

Figure 25. Recovery and selection of recombinant *P. falciparum* parasites after transfection with the dual vector CRISPR-Cas9 strategy. (A) *P. falciparum* NF54 parasites and **(B)** *P. falciparum* Dd2 strain post-transfection. Parasitaemia of *P. falciparum* NF54 and Dd2 parasites were determined before transfection (day 0) and daily post-transfection up to 40 days. Initial ring stage parasite populations at 4.3-5.2% parasitaemia were rapidly decreased to ~1-2% due to consequent lysis following electroporation. As of day 1 post-transfection, parasitaemia decreased due to WR99210 drug selection to <1% by day 3. Parasites morphologically became more stressed and died by day 2-3 post-transfection. WR99210 selection pressure was applied for 9-10 days post-transfection. Parasite stage and viability was evaluated daily by Giemsa stained thin blood smears. All recovered parasite populations were cloned by limiting dilution.

As seen in Fig. 25, electroporation of both the NF54 and Dd2 strains resulted in a rapid drop in parasitaemia (from ~4-5% to 1-2%) due to erythrocyte lysis. WR99210 selection further caused a decrease in parasitaemia (from 1-2% to ~0.5%), causing parasites to stress and die after 3 days of WR99210 selection.

In the *P. falciparum* NF54 strain (Fig. 25 A), parasites transfected with the pCC1-*hDHFR-kahrp*KO vector targeting the non-essential PF3D7_0202000 control gene successfully recovered. Parasites were visible on day 16 post-transfection and showed a steady increase in parasitaemia up to 2.7%, as previously described (121). After 40 days of evaluation of parasites in the NF54 strain, no other parasites recovered that had been transfected with the pCR®2.1TOPO®-^{KO} and pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* constructs targeting genes PF3D7_0613800, PF3D7_1239200, PF3D7_1211700 and PF3D7_0705300 for knockout.

By contrast, *P. falciparum* Dd2 strain parasites (Fig. 25 B) showed recovery between days 17 and 27 for parasites transfected with pCR®2.1TOPO®-^{KO} and pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* constructs for knockout. This included parasites transfected with constructs targeting PF3D7_0613800^{guide1}, PF3D7_1239200^{guide1}, PF3D7_1211700^{guide1}, PF3D7_0705300^{guide1} and PF3D7_0705300 ^{guide2} recovered on days 17, 20, 24, 25 and 27 post-transfection, respectively. No parasite recovery was observed for parasites transfected with constructs targeting PF3D7_0613800^{guide2}, PF3D7_1239200^{guide2}, and PF3D7_1211700^{guide2}.

If disruption was successful and parasites did not recover, this may indicate an essential nature of these genes to parasite survival. By contrast, if parasites did recover after successful gene disruption, the gene is most likely not essential. To assess whether successful gene disruption occurred, parasite lines that recovered after transfection were expanded and the complete KO cassette (HR1-*bsd* cassette-HR2) amplified from genomic DNA from these parasites, as an indicator of genomic integration of the KO cassettes into the recovered parasite lines (Fig. 26). The transfection control targeting the *kahrp* (PF3D7_0202000) was shown to have successfully integrated into the genome of NF54 transfected *P. falciparum* parasites, as was observed before (151).

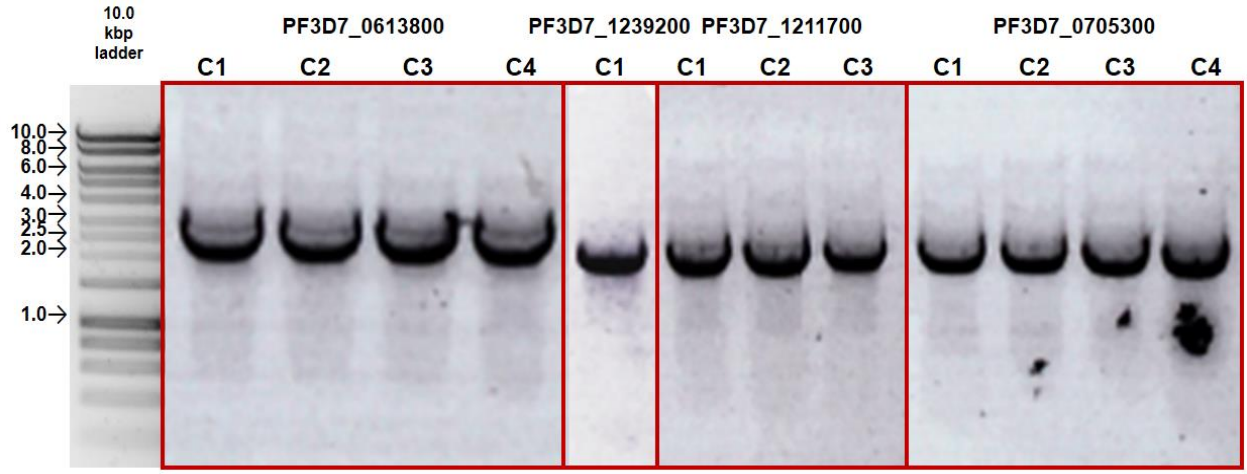

Figure 26. PCR amplification of the KO cassette from gDNA isolated from cloned out transfected *P. falciparum* Dd2 strain parasites. Approximately 100 ng of template was used to PCR amplify selected regions of the KO cassette The PCR and gel electrophoresis was done by Dr. Sophie Adjalley at the Wellcome Trust Sanger Institute, Cambridge, UK. (C1= clonal population 1).

Expected sizes of the successfully integrated KO cassettes into the Dd2 parasite line are 3 244 bp for PF3D7_0613800, 3 080 bp for PF3D7_1239200, 3 113 bp for PF3D7_1211700 and 3 170 bp for PF3D7_0705300. Wild-type genotypes would yield expected sizes of 2 217 bp for PF3D7_0613800, 2 136 bp for PF3D7_1239200, 1 894 bp for PF3D7_1211700 and 2 031 bp for PF3D7_0705300. Since the band sizes obtained corresponded to the latter size of each set (~2000 bp), none of the theoretical recombinant parasites showed gene disruption (Fig. 26), which indicates the presence of only wild-type genomic regions. Therefore, it was concluded that no successful KO of the target genes using PF3D7_0613800^{guide1}, PF3D7_1239200^{guide1}, PF3D7_1211700^{guide1}, PF3D7_0705300^{guide1} and PF3D7_0705300 ^{guide2} was achieved.

However, parasites transfected with vectors targeting PF3D7_0613800^{guide2}, PF3D7_1239200^{guide2}, PF3D7_1211700^{guide2} did not recover in the Dd2 strain of the parasite, and no parasites were recovered from the NF54 strain parasites. This could thus imply essentiality to parasite survival. However, this requires technically stringent future validation.

3.5 Construction of a single vector for targeted gene knockout

One of the reasons for the poor integration of the KO cassette into the *P. falciparum* genome could be because selection pressure for the pCR®2.1 TOPO®-^{KO} constructs could not be performed since double drug selection negatively impacts parasite survival. As an initial step to attempt to solve this problem, the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector was used as a single vector that would contain both the guides and KO cassettes for each gene investigated. This necessitated re-cloning of the KO cassettes for each gene into the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* construct for each gene.

The pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* constructs were prepared for Gibson assembly by *Apa*I and *Aat*II restriction digestion to provide clear overhangs (Fig. 27). This yielded the expected two fragments of 9 255 bp ~2 200 bp, suggesting efficient digestion by the restriction enzymes.

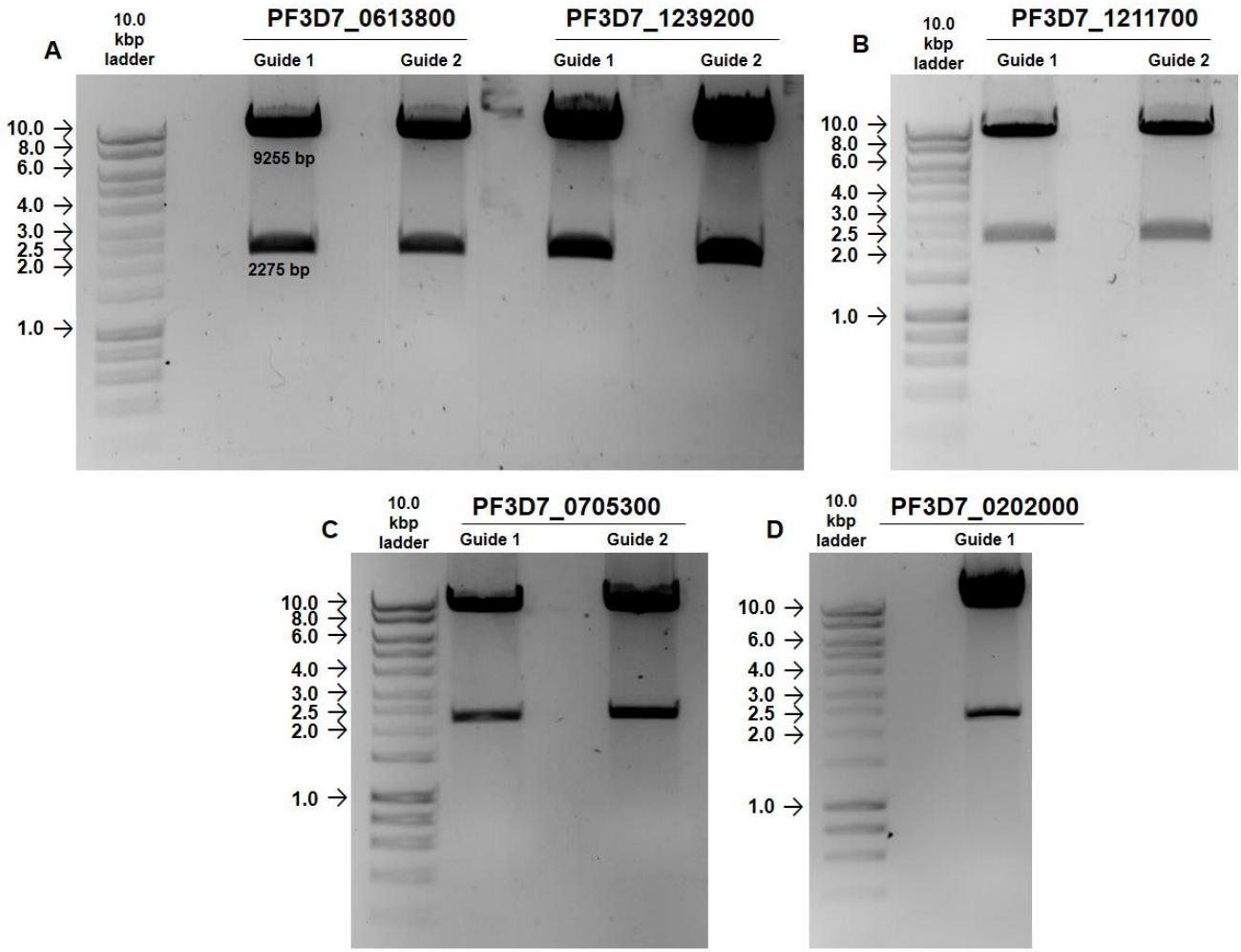

Figure 27. Vector pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* restriction enzyme digested DNA between *AatII* and *ApaI* sites for KO cassette assembly and integration. Approximately 4 µg of plasmid DNA was digested and alkaline phosphatase treated. Cut plasmid DNA was separated on a 0.8% (w/v) agarose/TAE gel and visualised with SYBR® Safe DNA Gel Stain. Expected sizes of the DNA fragments 2 275 bp and 9 255 bp the backbone size. **(A)** Cut plasmid DNA of PF3D7_0613800 and PF3D7_1239200 with guides cloned in, **(B)** PF3D7_1211700, **(C)** PF3D7_0705300 and **(D)** PF3D7_0202000 with the guide cloned in.

The KO cassettes had to be re-amplified from the pCR®2.1 TOPO®-^{KO} constructs and 5 additional nucleotides were included on the 5' ends of the primers to ensure a longer overlap between the restriction sites in the KO cassettes and the digested pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vectors. This was also performed for the control *kahrp* KO cassette from the pCC1-*hDHFR-kahrpKO* vector (Fig. 28).

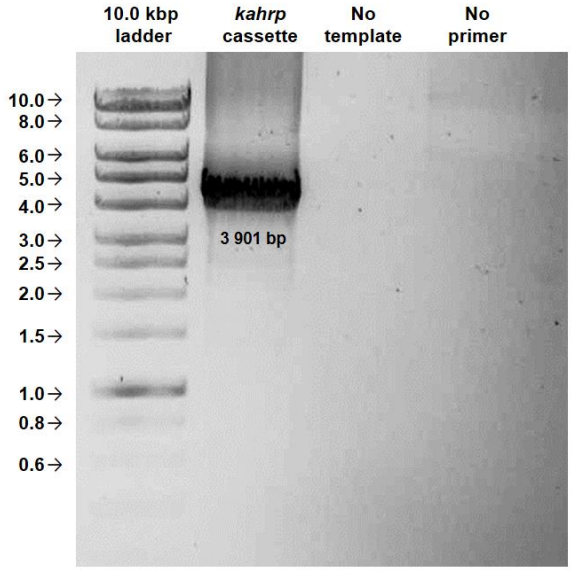

Figure 28. PCR amplification of the PF3D7_0202000 *kahrp* knockout cassette. The amplification of the knockout cassette including homology region one, the *hDHFR* and homology region two was analysed via loading of 240 ng of PCR product DNA and was separated on a 0.8% (w/v) agarose/TAE gel and visualised with SYBR® Safe DNA Gel Stain. The band corresponding to 3 901 bp KO cassette was gel extracted from the agarose gel.

Gel electrophoresis of the amplified PCR product indicated to its expected size of 3.9 kbp, including the 869 bp HR1, 2.1 kbp and 889 bp HR2. The PCR product was subsequently gel extracted and purified to yield 120 ng/µL DNA with A_{260}/A_{280} and A_{260}/A_{230} ratios of 1.89 and 2.27, respectively.

The successfully amplified KO cassettes were used directly in Gibson assembly reactions, Gibson assembly of the three fragments was performed, and subsequently PCR amplified. The KO cassette PCR products were purified and cloned into the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector and then transformed into *E. coli* XL10-Gold bacteria. Transformation of vectors targeting PF3D7_0613800^{guide1}, PF3D7_0613800^{guide2}, PF3D7_1239200^{guide1}, PF3D7_1211700^{guide1} and PF3D7_1211700^{guide1} into the bacteria was successful but not so for vectors targeting PF3D7_1239200^{guide2}, PF3D7_0705300^{guide1} and PF3D7_0705300^{guide2} which could not be successfully cloned into pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR*. Sanger sequencing confirmed the correct sequencing primers and overlaps (Fig. 29) while restriction mapping of the vectors confirmed correct KO cassettes sizes.

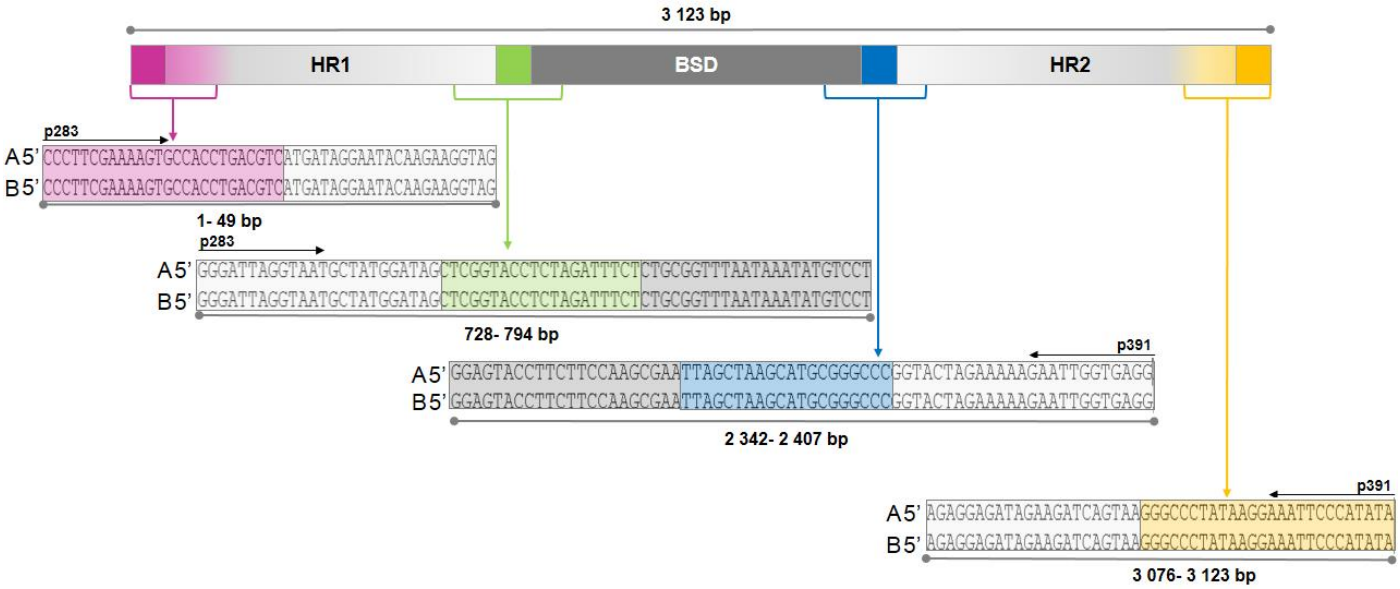

Figure 29. Sanger sequence results of the KO cassette targeting PF3D7_1211700^{guide1} cloned into pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Relative positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers p283 (forward) and p391 (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

Cloning of the KO cassettes into the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector was successful. Therefore, transfection and subsequent selection of recombinant parasites are possible with the expected advantage of single drug selection and stable integration of the KO cassette. These constructs can, therefore, be used in a single transfection strategy, where these vectors, as well as successful genomic integration, can be selected for in the parasites with blasticidin S drug pressure.

Chapter Four: Discussion

The *Plasmodium* parasite has the unique ability to regulate its genome in such a way that allows it to progress through a complex life cycle. This extraordinary feature of parasite biology implies that the cell cycle of the parasite is unique and it is therefore of particular scientific interest. Cell cycle progression in *Plasmodium* diverges from the classical eukaryotic cell cycle, and this suggests regulatory dexterity that can be potentially targeted within a chemotherapeutic strategy. Despite the presence of known classical cell cycle regulators, their exact role in the cell cycle of *Plasmodium* has not been clarified. Moreover, the possibility of additional regulatory elements guiding cell cycle control in the parasite in a unique manner remains unclear and unexplored. Therefore, elucidation of the master regulators of the cell cycle in *Plasmodium* could provide essential information that can be used in alternative approaches to control parasite development. This study aimed to identify and functionally characterise putative cell cycle regulators in *P. falciparum* parasites.

The study was successful in identifying genes that we associate a novel function to as potential cell cycle regulators due to their unique gene expression responses during induced cell cycle arrest and re-entry events. These potential cell cycle regulators can, therefore, be functionally characterised in the context of cell cycle regulation in the parasite. This signifies that precise control elements in the cell cycle (and therefore life cycle) control can be delineated. This is a biological feature of the parasite that has yet to be extensively described yet holds great potential in knowledge on how the parasite progresses through its life cycle that essentially allows transmission.

Genes from the pre-replicative complex were differentially expressed upon cell cycle arrest and re-entry. This relates to the biological function that the pre-replicative complex (i.e. the complex formed at the origins of DNA replication sites) fulfil during the cell cycle. During late M- and G_1-phases of the cell cycle, the pre-replicative complex functions primarily to license the DNA for replication during the next S-phase. Oscillations of cell cycle factors are present: during the G_1 phase, where CDK activity is low, and the pre-replication complex factors are high, while during the S-, G_2- and M-phases, CDK activity is high while pre-replication complex proteins are ubiquitinated and degraded (142). In this study, genes from the pre-replicative complex were shown to be likely essential as parasites that have undergone possible gene knockout did not survive. Critical importance of each protein element of the pre-replicative complex is, however, yet to be confirmed with additional experimentation.

Formation of the pre-replicative complex involves the sequential assembly of the origin of replication complex (ORC1-6 subunits), Cdc6, Cdt1 and the MCM proteins 2-7 to DNA (152). In most studied

eukaryotes organisms, the G_1/S checkpoint marks the signal for the pre-replication complex disassembly and inactivation of the involved factors. These factors are only replenished when the cell undergoes mitosis – as demonstrated by the transcriptional responses to cell cycle arrest and re-entry (153). It is, however, questionable whether these pre-replicative complex factors have any regulatory capacity over cell cycle progression or whether these genes merely modulate DNA replication as an effect caused by another cell cycle regulator's signal.

ORCs bind to high AT-rich sequences in a length (rather than sequence) dependent manner and exhibit ATPase activity (AAA^+, ATPases associated with diverse cellular activities), implying that ORCs hydrolyse ATP to replicative and transcribe DNA apart from their involvement in DNA replication by binding other DNA binding proteins (139). *Plasmodium* appears to have a minimal ORC, involving three putative ORC subunits, namely ORC1, ORC2 and ORC6 (154). ORC1, ORC2 and ORC6 bear consensus sequences for phosphorylation by CDKs for cell cycle dependent regulation where only ORC2 can be cross-linked to the DNA during S-phase, consistent with the disassembly of the pre-replicative complex (147). Homologues of Cdc6 and Cdt1 are absent in the *Plasmodium* genome. Other unknown CDKs may be responsible for the phosphorylation of ORC and MCM subunits. ORC components are thus potentially are regulated by cell cycle associated kinases rather than by autoregulation. The functional consequences ORC2 depletion is chromosome duplication, as well as abnormalities in chromosome structure and centrosome copy number (146). The phenotypic effects of ORC2 depletion include cell cycle arrest (particularly in S-phase where DNA replication occurs) and an aberrant mitotic-like state with abnormal chromosome condensation (148). If the subunits of the pre-ORC do regulate the cell cycle, it would support the hypothesis that by the lack of ORC2-mediated recognition of origins of replication prevented parasite growth as observed in this study. In *P. berghei* parasites, the ORC2 knockout reduced the growth rate significantly (155), demonstrating its function and importance in parasite survival. Given the importance and role of the ORC and its subunits, the ORC2 gene disruption in *P. falciparum* parasites may cause a slow growth phenotype or absolute essentiality for parasite survival.

The eukaryotic MCM complex, again part of the pre-replicative complex, is a hetero-hexameric complex with subunits MCM2 to MCM7 loaded onto DNA by the ATP hydrolysis dependent Cdc6-like kinase. It remains inactive up to the G_1/S-phase checkpoint and is then activated by CDKs (153,154). Upon DNA replication licencing, the MCM complex functions as a helicases, unwinding double stranded DNA ahead of polymerases to create the replication fork (155). It unclear whether DNA replication will continue without subunit MCM5. The loss of MCM5 could potentially stall the replication fork and cause DNA damage or chromosomal rearrangements. An incomplete MCM complex may inactivate DNA replication origins to become dormant and unable to licence DNA for

replication, effectively regulating cell cycle events. If individual components of the pre-replication complex are confirmed to regulate DNA replication, this would provide a range of drug targets to interfere with parasite proliferation and transmission.

ApiAP2 transcription factors, by contrast, are known to regulate the cell cycle and modulate the transcription of many thousands of genes targets of the IDC (63,119). Both ApiAP2 transcription factor genes investigated here are expressed at distinct time points during the IDC, and once expressed, can cause a cascade of gene expression of 617 (PF37D_1239200) and 722 (PF3D7_0613800) target genes, respectively, based on predicted DNA binding motifs. ApiAP2 transcription factors interact antagonistically or synergistically (119), creating a complex network of autoregulatory positive feedback loops to regulate parasite gene expression. PF3D7_0613800 and PF37D_1239200 would thus be essential for parasite survival, confirming that a knockout would be lethal as observed in this study. A corresponding knockout screen of ten ApiAP2 genes in *P. berghei* parasites confirm that they are essential for parasite development, mosquito transmission and the formation of ookinetes and sporogony (156).

Although ApiAP2 genes regulate *cis*-genes transcriptionally, ApiAP2 gene regulation is currently unknown. This raises an intriguing possibility that ApiAP2 genes are controlled directly or indirectly on an epigenetic level, like AP2-G, the master regulator of sexual differentiation, which is reversibly silenced by a repressive chromatin structures brought about by silencing factor heterochromatin protein 1 binfing to a histone 3 lysine 9 trimethylation mark (H3K9me3) associated with transcriptional repression (20). Modrzynska *et al*. proposed that there is a transcriptional network whereby ApiAP2 interact stage-specifically and may co-regulate the transcription of other ApiAP2 genes (156). This means that disruption of a single transcription factor of an already small repertoire of transcription factor family may cause a cascade of negative gene regulation and ultimately be detrimental to parasite survival.

Assuming that the ApiAP2 transcription factors investigated here are essential, these may be targeted with drug intervention. ApiAP2 proteins consist of functionally distinct domains that could be targeted for drug intervention including a ligand-binding, a trans-effector, a DNA binding and a nuclear localisation domain. Alternatively, the post-translational phosphorylation of ApiAP2 could be targeted. Transcription factor activity may be modulated by affecting the expression of each of the transcription factors, targeting their ability to bind or recognise DNA sequences by small molecules by binding to the DNA binding domain of the protein or its allosteric domain, or by disruption of the protein-protein interactions by small peptides of peptidomimetics (157). Each approach has advantages, such as high selectivity, bioavailability and slight to no side-effects. However,

disadvantages include pro-inflammatory, drug stability or off-target effects. Transcription factor inhibitors are effective against, e.g. Stat-3, a human transcription factor (158,159). Targeting transcription factors that modulate multiple signalling pathways may be a risky approach, but has proven effective in cancer and may be an ideal target given its strong possible outcome. This thus supports the notion of ApiAP2 proteins as potential antimalarial targets.

In this study, components of the CRISPR system, namely the guide sequences and KO cassette, were designed and assembled in a dual vector approach for *in vitro* characterisation by transfecting two *P. falciparum* parasite strains. No recovery of recombinant was observed for the sensitive NF54 strain. The few parasite clones that did recover in the Dd2 strain were non-recombinant. Thus, either the integration of the KO cassette into the parasite genome did not occur or the genes may be essential. Molecular events leading to integration and their control remain enigmatic in *P. falciparum* parasites. This may include single- or double crossover recombination through the orderly transcription of the Cas9 endonuclease and Cas9-guide, or duplex formation and effective site-specific DNA cleavage. Drug selection of a sensitive recombinant parasite challenging as two or more selection drugs may limit parasite growth or cause to false-positive phenotypic observations due to drug responses rather than loss-of-function mutations. Uneven chromosome segregation between daughter merozoites may cause variable copy numbers and defective replication, leading to loss of the episome four weeks after removal of drug pressure and reverts to the drug sensitive genotype (160–162). Extending of the drug selection period may not prevent loss of episomal plasmid DNA as, in *P. falciparum* parasites, extended periods of drug selection of non-integrated DNA may result in the formation of concatamers, tandem copies of DNA that are associated with the unstable replicating forms of the plasmid and non-specific integration in the genome. Thus, the few Dd2 strain transfected parasites may have reverted to the wild-type genotype, which does not necessarily indicate non-essentiality of the genes. This postulation is further supported by discrepancies in the recovery between the NF54 and Dd2 parasite strains. All transfected parasites from the NF54 strain and a few transfected parasite clones from the Dd2 strain did not recover, supporting an essential phenotype for these genes.

Although more studies are required to validate these findings, the preliminary results reveal two important features of malaria parasite biology, 1) that we are only at the beginning of characterising essential putative cell cycle regulatory genes crucial to parasite survival, life and cell cycle progression, and 2) that *P. falciparum* may regulate its cell cycle in unique ways such that genes normally involved in cell cycle regulation adopt additional pleiotropic functions to what is currently annotated. The fact that the physical and functional removal of a single gene could lead to lethal effects in the parasite poses an interesting scenario. It is questionable whether the presence of the

genes might act as enhancers for other genes, or hypothetically, could provide an additional level of information regarding the functional or behavioural aspects of the gene, which only becomes evident if the sequence is disrupted.

Most *P. falciparum* parasite genes remain functionally uncharacterised, partly due to low sequence homology to other eukaryotic genomes (163), and the lack of tools to successfully interrogate essential gene functions. Here, the biological consequence of disrupting putative cell cycle regulator genes was probed, and therefore, the novel CRISPR-Cas9 system was chosen as the preferred platform for gene knockout given its robust application in other model organisms and intended use for downstream studies of putative cell cycle regulatory genes. However, due to the complexity of the cloning strategy followed and the inherent difficulty working with a mostly uncharacterised plastid and extreme (>80%) AT-bias genome of *P. falciparum* parasites, many technical challenges were faced trying to establish the technique, which required multiple optimisation steps and adjustment in a limited time period. However, this study offers a more simplified, streamlined approach to the technicalities of gene knockouts in *P. falciparum* and opened a platform for more genetic modification studies to identify the key biological elements in the parasite.

In future, it would be preferable to investigate gene essentiality through a conditional knockdown approach. Instead of disrupting the genomic sequence of a gene, conditional knockdown approaches focus on inhibiting the functionality of a gene product through, e.g. inhibition of transcription or directing the protein product to another organelle where it cannot fulfil its function. This will give insights as to the protein interaction and possible inclination points in the cell cycle regulation network *in vitro*, in addition to showing the biological consequence clearer as opposed to mere lethal phenotype. Furthermore, it may be beneficial to include more transfection controls, including a negative control where parasites are transfected with no plasmid DNA, but instead with a plasmid that does not target any gene but carries a marker for transfection (like a GFP gene). This would thus enable transfection efficiency calculations. An additional negative control could include a vector that does not contain a guide sequence but only the Cas9 enzyme alone. Positive controls would not only include the transfection non-essential *kahrp* gene again, but also a plasmid targeting an essential gene in the IDC, such as basigin, a receptor of the PfRh5 ligand that is essential for merozoite invasion of the erythrocyte (164). Additionally, determination of possible off-target activity of each guide targeting a gene would be assessed through whole genome sequencing.

Given the renewed focus on malaria elimination and eradication, specific properties of next generation antimalarial drugs have been outlined. The emphasis is on, among other, targeting multiple *Plasmodium* life cycle stages to prevent transmission and to reduce the parasite load

burden during the asexual (symptomatic) stages (17). The proteins investigated in this study offer molecular targets that mirror the desired properties. Firstly, proteins ORC2 and MCM5 are involved in DNA replication which occurs in the liver, blood and mosquito stages of the parasites life cycle and could therefore act as possible transmission blocking sites (165). Secondly, targeting DNA replication during the intra-erythrocytic stage would prevent parasite proliferation and ultimately reduce parasite load during infection. Thirdly, targeting the parasite specific ApiAP2 proteins would repress many processes directly linked to parasite development and differentiation, ultimately, providing the ideal drug targets for chemo-intervention.

Towards elucidation of the unique cell cycle regulators coordinating the unusual cell cycle of the *P. falciparum* parasites, putative cell cycle regulators were studied with respect to their essentiality or role in the progression and ultimate multiplication of parasite numbers. This aims to contribute to the understanding of a non-conformant cell cycle in eukaryotes for which regulatory elements remain to be defined. The uniqueness of the *Plasmodium* cell cycle implies exceptional control and regulation of cell cycle events such that it functionally validates core biological principles of the malaria parasite development and pathogenesis.

Chapter Five: Conclusion

Malaria incidence and mortality rates have decreased profoundly since the year 2000, falling by an estimated 37% and 60% respectively (1). This is due the concerted global effort and investment toward malaria control and the development of current existing antimalarial drugs (17). However, with the recent development of malaria parasite drug resistance to first-line drugs, the significant progress that has been made to control malaria is threatened (166). Thus, there is an urgent need to identify novel antimalarial drug targets that target distinct processes in the parasite (167). Therefore, interrogation of the malaria parasite biology is crucial to identifying molecular targets that can be exploited in future rational drug design and elimination strategies.

Parasitic protozoa, such as the *P. falciparum* parasite, have evolved unusual biological features that allows them to survive in contrasting host environments. The life cycle is exquisitely controlled, oscillating between quiescent stages (e.g. sporozoites or gametocytes) and stages of rapid proliferation during the IDC, as reflected by the orderly progression of the cell cycle. However, the precise control mechanisms and functional cascades involved in the atypical *P. falciparum* parasite cell cycle have not yet been fully elucidated. Therefore, in-depth characterisation of parasite cell cycle regulation and identification of essential regulators is necessary.

In this study, gene expression profiles of a parasite under cell cycle arrest and re-entry were interrogated to reveal distinct gene clusters that modulate parasite biological responses on the transcriptional, post-transcriptional and epigenetic levels. Four genes were subsequently assessed for their potential regulatory roles in the cell cycle of the parasite. A dual vector CRISPR-Cas9 system was employed that required vector construction, including the design and cloning of highly specific guide RNAs and the knockout cassette for DNA repair. Both drug sensitive and drug resistant strains of the *P. falciparum* parasite were transfected, without significant recovery of parasites being observed. This finding may imply an essential nature to the gene studies, therefore, provisionally confirming the hypothesis of this study. However, this finding needs to be thoroughly interrogated in future research. If essential, these gene products are potential antimalarial targets as they are specific to the malaria parasite, target processes that occur in multiple parasite life stages, and would prevent parasite proliferation, differentiation and survival.

Additionally, an alternative and more robust strategy was developed, which combines all necessary components of a CRISPR-system in a single vector. In future, this will potentially enhance transfection efficiencies and simplify the complexity involved in gene function interrogation of *P. falciparum* parasites.

Functional validation of essential cell cycle regulators in *Plasmodium* is critical to prove that cell cycle regulators are attractive antimalarial drug targets. These cell cycle regulators are unique to *Plasmodium* species and could potentially be considered in future as a novel targetable processes in *Plasmodium*.

References

1. World Health Organization. World Malaria Report. 2016.
2. Perez ML, Pacheco MA, Buriticá L, Escalante AA, Herrera S, Herrera MA. Malaria in pregnancy: a passive surveillance study of pregnant women in low transmission areas of Colombia , Latin America. Malar J. BioMed Central; 2016;15(66):1–10.
3. Rogerson SJ, Hviid L, Duff PE, Leke RFG, Taylor DW. Malaria in pregnancy: pathogenesis and immunity. Lancet. 2007;7(2):105–17.
4. Amexo M, Tolhurst R, Barnish G, Bates I. Malaria misdiagnosis: Effects on the poor and vulnerable. Lancet. 2004;364(9448):1896–8.
5. Dupas P. Health Behavior in Developing Countries. Annu Rev Econom. 2011;3(1):425–49.
6. Gallup JL, Sachs JD. The economic burden of malaria. Am J Trop Med Hyg. 2001;64(1):85–96.
7. Hay SI, Guerra CA, Tatem AJ, Atkinson PM, Snow RW. Urbanization, malaria transmission and disease burden in Africa. Nat Microbiol. 2005;3(January):1–10.
8. Manego RZ, Witte M, Held J, Gmeiner M, Gebru T, Tazemda B, *et al*. Demography, maternal health and the epidemiology of malaria and other major infectious diseases in the rural department Tsamba-Magotsi, Ngounie Province, in central African Gabon. BMC Public Health. BMC Public Health; 2017;17(130):1–7.
9. Wells TN, Poll EM. When is enough enough? The need for a robust pipeline of high-quality antimalarials. Discov Med. 2010;9(48):389–98.
10. Hott A, Casandra D, Sparks KN, Morton LC, Castanares G-G, Rutter A, *et al*. Artemisinin-resistant *Plasmodium falciparum* exhibit altered patterns of development in infected erythrocytes. Antimicrob Agents Chemother. 2015;(March):AAC.00197-15.
11. The threat of artemisinin resistant malaria in Southeast Asia. Vol. 14, Travel Medicine and Infectious Infectious. 2016.
12. Thanh NV, Nhien NT, Thi N, Tuyen K, Tong NT, Thuy N, *et al*. Rapid decline in the susceptibility of *Plasmodium falciparum* to dihydroartemisinin – piperaquine in the south of Vietnam. Malar J. BioMed Central; 2017;16(27):1–10.
13. World Health Organization. Artemisinin and artemisinin-based combination therapy resistance. 2017.
14. malERA Consultative Group on Drugs. A research agenda for malaria eradication: drugs. PLoS Med. 2011;8(1).
15. Alonso PL, Brown G, Arevalo-Herrera M, Binka F, Chitnis C, Collins F, *et al*. A research Agenda to underpin Malaria Eradication. PLoS Med. 2011;8(1).
16. Biagini G, Fisher N, Shone E, Mubaraki M, Srivastava A, Hill A, *et al*. Generation of quinolone antimalarials targeting the *Plasmodium falciparum* mitochondrial respiratory chain for the treatment and prophylaxis of malaria. Proc Natl Acad Sci. 2012;109:8298–303.
17. Burrows JN, Duparc S, Gutteridge WE, Huijsduijnen RH Van, Kaszubska W, Macintyre F, *et al*. New developments in anti - malarial target candidate and product profiles. Malar J. BioMed Central; 2017;1–29.
18. Garcia JE, Puentes A, Patarroyo ME. Developmental biology of sporozoite-host interactions in *Plasmodium falciparum* malaria: Implications for vaccine design. Clin Microbiol Rev. 2006;19(4):686–707.
19. Lim L, McFadden GI. The evolution, metabolism and functions of the apicoplast. Philos Trans R Soc Lond B Biol Sci. 2010;365(1541):749–63.
20. Josling G, Llinás M. Sexual development in *Plasmodium* parasites: knowing when it's time to commit. Nat Rev Microbiol. 2015;13(9):573–87.
21. Prudêncio M, Rodriguez A, Mota MM. The silent path to thousands of merozoites: the *Plasmodium* liver stage. Nat Rev Microbiol. 2006;4(November):849–56.
22. Tarun AS, Baer K, Dumpit RF, Gray S, Lejarcegui N, Frevert U, *et al*. Quantitative isolation and in vivo imaging of malaria parasite liver stages. Int J Parasitol. 2006;36(12):1283–93.
23. Kirk K. Membrane transport in the malaria-infected erythrocyte. Physiol Rev. 2001;81(2):495–537.
24. Francis SE, Sullivan DJ, Goldberg DE. Hemoglobin metabolism in the malaria parasite *Plasmodium falciparum*. Annu Rev Microbiol. 1997;51:97–123.
25. Garcia CRS, de Azevedo MF, Wunderlich G, Budu A, Young JA, Bannister L. *Plasmodium* in the postgenomic era: new insights into the molecular cell biology of malaria parasites. Int Rev Cell Mol Biol. 2008;266(7):85–156.
26. Downie MJ, Kirk K, Mamoun C Ben. Purine salvage pathways in the intraerythrocytic malaria parasite *Plasmodium falciparum*. Eukaryot Cell. 2008;7(8):1231–7.
27. Arnot DE. Encyclopedia of Malaria. New York: Springer Science+Business Media; 2013. 1-9 p.
28. Bannister L, Mitchell G. The ins, outs and roundabouts of malaria. 2003;19(5):209–13.
29. Doerig C, Endicott J, Chakrabarti D. Cyclin-dependent kinase homologues of *Plasmodium falciparum*. Int J Parasitol. 2002;32(13):1575–85.
30. Bousema T, Drakeley C. Epidemiology and infectivity of *Plasmodium falciparum* and *Plasmodium vivax* gametocytes in relation to malaria control and elimination. Clin Microbiol Rev. 2011;24(2):377–410.
31. Kariuki MM, Kiaira JK, Mulaa FK, Mwangi JK, Wasunna MK, Martin SK. *Plasmodium falciparum*: Purification of the various gametocyte developmental stages from *in vitro*-cultivated parasites. Am J Trop Med Hyg. 1998;59(4):505–8.

32. Taylor LH, Read AF. Why so few transmission stages? Reproductive restraint by malaria parasites. Parasitol Today. Elsevier Ltd; 1997;13(4):135–40.
33. Billker O, Shaw MK, Margos G, Sinden RE. The roles of temperature, pH and mosquito factors as triggers of male and female gametogenesis of *Plasmodium berghei in vitro*. Parasitology. 1997;115(February):1–7.
34. Guttery DS, Holder AA, Tewari R. Sexual development in *Plasmodium*: Lessons from functional analyses. PLoS Pathog. 2012;8(1):8–10.
35. Gerber JE, Ukena TE, Cote L, Wyllie JM, Winn Jr. WC. Exflagellation of malarial parasites in human peripheral blood. J Clin Microbiol. 1981;13(1):236–7.
36. McRobert L, Taylor CJ, Deng W, Fivelman QL, Cummings RM, Polley SD, *et al*. Gametogenesis in malaria parasites is mediated by the cGMP-dependent protein kinase. PLoS Biol. 2008;6(6):1243–52.
37. Delves MJ, Ruecker A, Straschil U, Lelièvre J, Marques S, López-Barragán MJ, *et al*. Male and female *Plasmodium falciparum* mature gametocytes show different responses to antimalarial drugs. Antimicrob Agents Chemother. 2013;57(7):3268–74.
38. Hammarton TC, Mottram JC, Doerig C. The cell cycle of parasitic protozoa: potential for chemotherapeutic exploitation. Prog Cell Cycle Res. 2003;5:91–101.
39. Voet, Donald., Voet, Judith G., Pratt CW. Fundamentals of Biochemistry. New Jersey: John Wiley & Sons, Inc.; 2013. 71-73 p.
40. Bertoli C, Skotheim JM, Bruin RAM De, Street G. Control of cell cycle transcription during G_1 and S phases. Nat Rev Mol Cell Biol. 2015;14(8):518–28.
41. Foster DA, Yellen P, Xu L, Saqcena M. Regulation of G_1 Cell Cycle Progression: Distinguishing the Restriction Point from a Nutrient-Sensing Cell Growth Checkpoint(s). Genes Cancer. 2010;1(11):1124–31.
42. Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. Nat Rev Cancer. 2009;9(3):153–66.
43. Renthal W, Lee E. Cancer and the role of cell cycle checkpoints. Undergrad Res. 2002;1:1–7.
44. Alm K, Oredsson S. Cells and polyamines do it cyclically. Essays Biochem. 2009;46:63–76.
45. Wallace HM, Fraser A V, Hughes A. A perspective of polyamine metabolism. Biochem J. 2003;376(Pt 1):1–14.
46. Oredsson SM. Polyamine dependence of normal cell-cycle progression. Biochem Soc Trans. 2003;31(2):366–70.
47. Bettuzzi S, Davalli P, Astancolle S, Pinna C, Roncaglia R, Boraldi F, *et al*. Coordinate changes of polyamine metabolism regulatory proteins during the cell cycle of normal human dermal fibroblasts. FEBS Lett. 1999;446(1):18–22.
48. Assaraf Y., Golenser J, Sira D., Messer G, Bachrach U. Cytostatic effect of DL-α -difluoromethylornithine against *Plasmodium falciparum* and its reversal by diamines and spermidine. Parasitol Res. 1987;73:313–8.
49. Doerig C, Chakrabarti D, Kappes B, Matthews K. The cell cycle in protozoan parasites. In: Progress in Cell Cycle Research. 2000. p. 163–83.
50. Doerig C, Chakrabarti D. Cell Cycle Control in *Plasmodium falciparum*: A Genomics Perspective. In: Waters AP, Janse CJ, editors. Malaria Parasites: Genomes and Molecular Biology. Leiden: Caister Academic Press; 2004. p. 29208.
51. Francia ME, Striepen B. Cell division in apicomplexan parasites. Nat Rev Microbiol. Nature Publishing Group; 2014;12(2):125–36.
52. Doerig C. Protein kinases as targets for anti-parasitic chemotherapy. Biochim Biophys Acta - Proteins Proteomics. 2004;1697(1–2):155–68.
53. Inselburg J, Banyal HS. Synthesis of DNA during the asexual cycle of *Plasmodium falciparum* in culture. Mol Biochem Parasitol. 1984;10(1):79–87.
54. Gerald N, Mahajan B, Kumar S. Mitosis in the human malaria parasite *Plasmodium falciparum*. Eukaryot Cell. 2011;10(4):474–82.
55. Francia ME, Jordan CN, Patel JD, Sheiner L, Demerly JL, Fellows JD, *et al*. Cell Division in Apicomplexan Parasites Is Organized by a Homolog of the Striated Rootlet Fiber of Algal Flagella. PLoS Biol. 2012;10(12):1–14.
56. Janse CJ, Haghparast A, Sperança MA, Ramesar J, Kroeze H, Portillo HA, *et al*. Malaria parasites lacking eef1a have a normal S/M phase yet grow more slowly due to a longer G 1 phase. Mol Microbiol. 2003;50(5):1539–51.
57. Naughton J, Bell A. Studies on cell-cycle synchronization in the asexual erythrocytic stages of *Plasmodium falciparum*. Parasitology. 2007;134(Pt 3):331–7.
58. Arnot DE, Ronander E, Bengtsson DC. The progression of the intra-erythrocytic cell cycle of *Plasmodium falciparum* and the role of the centriolar plaques in asynchronous mitotic division during schizogony. Int J Parasitol. 2011;41(1):71–80.
59. Gerald N, Mahajan B, Kumar S, Gerald N, Mahajan B, Kumar S. Mitosis in the Human Malaria Parasite *Plasmodium falciparum*. Eukaryot Cell. 2011;10(4):474–82.
60. Raabe AC, Billker O, Vial HJ, Wengelnik K. Quantitative assessment of DNA replication to monitor microgametogenesis in *Plasmodium berghei*. Mol Biochem Parasitol. 2009;168(2):172–6.
61. Bozdech Z, Llinás M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. PLoS Biol. 2003;1(1):85–100.
62. Coetzee N, Sidoli S, Van Biljon R, Painter H, Llinás M, Garcia BA, *et al*. Quantitative chromatin proteomics reveals a dynamic histone post- translational modification landscape that defines asexual and sexual

Plasmodium falciparum parasites. Nat Sci Reports. 2017;607(7):1–12.
63. Painter HJ, Campbell TL, Llinas M. The Apicomplexan AP2 family: Integral factors regulating *Plasmodium* development. Mol Biochem Parasitol. Elsevier B.V.; 2011;176(1):1–7.
64. Doerig C, Billker O, Haystead T, Sharma P, Tobin AB, Waters NC. Protein kinases of malaria parasites: an update. Trends Parasitol. 2008;24(12):570–7.
65. Solyakov L, Halbert J, Alam MM, Semblat J-P, Dorin-Semblat D, Reininger L, *et al*. Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*. Nat Commun. Nature Publishing Group; 2011;2(565):1–12.
66. Carvalho TG, Morahan B, John von Freyend S, Boeuf P, Grau G, Garcia-Bustos J, *et al*. The ins and outs of phosphosignalling in *Plasmodium*: Parasite regulation and host cell manipulation. Mol Biochem Parasitol. Elsevier B.V.; 2016;208(1):2–15.
67. Carvalho TG, Doerig C, Reininger L. Nima- and Aurora-related kinases of malaria parasites. Biochim Biophys Acta. Elsevier B.V.; 2013;2–11.
68. Ganter M, Goldberg JM, Dvorin JD, Paulo JA, King JG, Tripathi AK, *et al*. *Plasmodium falciparum* CRK4 directs continuous rounds of DNA replication during schizogony. Nat Microbiol. Nature Publishing Group; 2017;17017(February):1–48.
69. Butler CL, Lucas O, Wuchty S, Xue B, Uversky VN, White M. Identifying novel cell cycle proteins in apicomplexa parasites through co-expression decision analysis. PLoS One. 2014;9(5).
70. Kozlov S, Waters NC, Chavchich M. Leveraging cell cycle analysis in anticancer drug discovery to identify novel plasmodial drug targets. Infect Disord Drug Targets. 2010;10(3):165–90.
71. Arnot DE, Ronander E, Bengtsson DC. The progression of the intra-erythrocytic cell cycle of *Plasmodium falciparum* and the role of the centriolar plaques in asynchronous mitotic division during schizogony. Int J Parasitol. Australian Society for Parasitology Inc.; 2011;41(1):71–80.
72. Landau G, Ran A, Bercovich Z, Feldmesser E, Horn-saban S, Korkotian E, *et al*. Expression Profiling and Biochemical Analysis Suggest Stress Response as a Potential Mechanism Inhibiting Proliferation of Polyamine-depleted Cells. J Biol Chem. 2012;287(43):35825–37.
73. Hoppe H, Verschoor J, Louw AI. *Plasmodium falciparum*: A Comparison of Synchronisation for *in vitro* Cultures. Exp Parasitol. 1991;72(1991):464–7.
74. Marı R, Tekwani BL, Balan R. Polyamine transport in parasites: A potential target for new antiparasitic drug development. Comp Biochem Physiol. 2005;140(2):151–64.
75. Müller S, Coombs GH, Walter RD. Targeting polyamines of parasitic protozoa in chemotherapy. Trends Parasitol. 2001;17(5):242–9.
76. Sharma RA, Manson MM, Gescher A, Steward WP. Colorectal cancer chemoprevention: Biochemical targets and clinical development of promising agents. Eur J Cancer. 2001;37(1):12–22.
77. Heby O, Gray J., Lindl PA, Marton LJ, Wilson C. Changes in L-Orntihine decarboxylase activity during the cell cycle. Biochem Biophys Res Commun. 1976;71(1):99–105.
78. Bozdech Z, Zhu J, Joachimiak M, Cohen F, Pulliam B, DeRisi J. Expression profiling of the schizont and trophozoite stages of *Plasmodium falciparum* with a long-oligonucleotide microarray. Genome Biol. 2003;4(2):1–15.
79. Crabb BS, Cooke BM, Reeder JC, Waller RF, Caruana SR, Davern KM, *et al*. Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress. Cell. 1997;89(2):287–96.
80. de Koning-Ward TF, Gilson PR, Crabb BS. Advances in molecular genetic systems in malaria. Nat Rev Microbiol. Nature Publishing Group; 2015;13(6):373–87.
81. Adjalley SH, Lee MCS, Fidock DA. A Method for Rapid Genetic Integration into *Plasmodium falciparum* Utilizing Mycobacteriophage Bxb1 Integrase. Methods Mol Biol. 2011;634(3):1–12.
82. Balu B, Shoue D a, Fraser MJ, Adams JH. High-efficiency transformation of *Plasmodium falciparum* by the lepidopteran transposable element piggyBac. Proc Natl Acad Sci U S A. 2005;102(45):16391–6.
83. Armstrong CM, Goldberg DE. An FKBP destabilization domain modulates protein levels in *Plasmodium falciparum*. Nat Methods. 2007;4(12):1007–9.
84. Ganesan SM, Falla A, Goldfless SJ, Nasamu AS, Niles JC. Synthetic RNA–protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. Nat Commun. Nature Publishing Group; 2016;7:1–10.
85. van Schaijk BCL, Vos MW, Janse CJ, Sauerwein RW, Khan SM. Removal of heterologous sequences from *Plasmodium falciparum* mutants using FLPe-recombinase. PLoS One. 2010;5(11).
86. Collins CR, Das S, Wong EH, Andenmatten N, Stallmach R, Hackett F, *et al*. Robust inducible Cre recombinase activity in the human malaria parasite *Plasmodium falciparum* enables efficient gene deletion within a single asexual erythrocytic growth cycle. Mol Microbiol. 2013;88(4):687–701.
87. Prommana P, Uthaipibull C, Wongsombat C, Kamchonwongpaisan S, Yuthavong Y, Knuepfer E, *et al*. Inducible Knockdown of *Plasmodium* Gene Expression Using the glmS Ribozyme. PLoS One. 2013;8(8):1–10.
88. Straimer J, Lee MCS, Lee AH, Zeitler B, Williams AE, Pearl JR, *et al*. Site-specific genome editing in *Plasmodium falciparum* using engineered zinc-finger nucleases. Nat Methods. 2012;9(10):993–8.
89. Goldfless SJ, Wagner JC, Niles JC. Versatile control of *Plasmodium falciparum* gene expression with an inducible protein-RNA interaction. Nat Commun. 2015;18(5):1067–73.

90. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, *et al*. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature. 2002;419:498–511.
91. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. Cell. Elsevier; 2014;157(6):1262–78.
92. Burstein D, Harrington LB, Strutt SC, Probst AJ. New CRISPR-Cas systems from uncultivated microbes. Nature. Nature Publishing Group; 2016;0:1–20.
93. Shmakov S, Smargon A, Scott D, Cox D, Pyzocha N, Yan W, *et al*. Diversity and evolution of class 2 CRISPR–Cas systems. Nat Rev Microbiol. Nature Publishing Group; 2017;15(3):1–14.
94. Lange SJ, Alkhnbashi OS, Rose D, Will S, Backofen R. CRISPRmap: An automated classification of repeat conservation in prokaryotic adaptive immune systems. Nucleic Acids Res. 2013;41(17):8034–44.
95. Marraffini L a, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in *Staphylococci* by targeting DNA. Science. 2008;322(5909):1843–5.
96. Brouns SJJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJH, Snijders APL, *et al*. Small CRISPR RNAs guide antiviral defense in prokaryotes. Science. 2008;321(5891):960–4.
97. Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, *et al*. RNA-Guided RNA Cleavage by a CRISPR RNA-Cas Protein. Cell. 2010;139(5):945–56.
98. Garneau JE, Dupuis M-È, Villion M, Romero D a, Barrangou R, Boyaval P, *et al*. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature. 2010;468(7320):67–71.
99. Deltcheva E, Chylinski K, Sharma CM, Gonzales K. Europe PMC Funders Group Europe PMC Funders Author Manuscripts CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. 2011;471(7340):602–7.
100. Sapranauskas R, Gasiunas G, Fremaux C, Barrangou R, Horvath P, Siksnys V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. Nucleic Acids Res. 2011;39(21):9275–82.
101. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science. 2012;337(August):816–22.
102. Gasiunas G, Barrangou R, Horvath P, Siksnys V. PNAS Plus: Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci. 2012;109(39):E2579–86.
103. Chylinski K, Makarova KS, Charpentier E, Koonin E V. Classification and evolution of type II CRISPR-Cas systems. Nucleic Acids Res. 2014;42(10):6091–105.
104. Bhaya D, Davison M, Barrangou R. CRISPR-Cas Systems in Bacteria and Archaea: Versatile Small RNAs for Adaptive Defense and Regulation. Annu Rev Genet. 2011;45(1):273–97.
105. Mekler V, Minakhin L, Semenova E, Kuznedelov K, Severinov K. Kinetics of the CRISPR-Cas9 effector complex assembly and the role of 3-terminal segment of guide. Nucleic Acids Res. 2016;44(6):2837–45.
106. Chylinski K, Rhun A Le, Charpentier E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. RNA Biol. 2013;10(5):726–37.
107. Rath D, Amlinger L, Rath A, Lundgren M. Biochimie The CRISPR-Cas immune system: Biology, mechanisms and applications. Biochimie. Elsevier B.V; 2015;117:119–28.
108. Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, *et al*. An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol. Nature Publishing Group; 2015;13(11):722–36.
109. Makarova KS, Grishin N V, Shabalina S a, Wolf YI, Koonin E V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. Biol Direct. 2006;1:7.
110. Zhang C, Xiao B, Jiang Y, Zhao Y, Li Z, Gao H, *et al*. Efficient Editing of Malaria Parasite Genome Using the CRISPR/Cas9 System. MBio. 2014;5(4):1–9.
111. Wright A V, Nunez JK, Doudna JA. Review Biology and Applications of CRISPR Systems: Harnessing Nature's Toolbox for Genome Engineering. Cell. 2016;164(1–2):29–44.
112. Sternberg SH, Doudna JA. Expanding the Biologist's Toolkit with CRISPR-Cas9. Mol Cell. Elsevier Inc.; 2015;58(4):568–74.
113. Barrangou R, Doudna JA. Applications of CRISPR technologies in research and beyond. Nat Biotechnol. 2016;34(9):17–20.
114. Ghorbal M, Gorman M, Macpherson CR, Martins RM, Scherf A, Lopez-Rubio J-J. Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. Nat Biotechnol. 2014;32(8):819–21.
115. Crabb BS, Cooke BM, Reeder JC, Waller RF, Caruana SR, Davern KM, *et al*. Targeted Gene Disruption Shows That Knobs Enable Malaria-Infected Red Cells to Cytoadhere under Physiological Shear Stress. Cell. 1997;89(2):287–96.
116. Nacer A, Claes A, Roberts A, Scheidig-Benatar C, Sakamoto H, Ghorbal M, *et al*. Discovery of a novel and conserved *Plasmodium falciparum* exported protein that is important for adhesion of PfEMP1 at the surface of infected erythrocytes. Cell Microbiol. 2015;(January).
117. Lim MY, Lamonte G, Lee MCS, Reimer C, Tan BH, Corey V, *et al*. UDP-galactose and acetyl-CoA transporters as *Plasmodium* multidrug resistance genes. Nat Microbiol. 2016;(September):1–12.
118. Hoon MJL De, Imoto S, Nolan J, Miyano S. Open source clustering software. Bioinforma Appl Note. 2004;20(9):1453–4.
119. Campbell TL, de Silva EK, Olszewski KL, Elemento O, Llinás M. Identification and Genome-Wide Prediction of DNA Binding Specificities for the ApiAP2 family of regulators from the malaria parasite. PLoS Pathog.

2010;6(10).
120. Cai H, Sanchez M, Wang Y, Gu J. Putative cell cycle related genes in *Plasmodium falciparum*. In: International Joint Conference on Bioinformatics, Systems Biology and Intelligent Computing, 2009. 2009. p. 230–6.
121. Wagner JC, Platt RJ, Goldfless SJ, Zhang F, Niles JC. Efficient CRISPR-Cas9 mediated genome editing in *Plasmodium falciparum*. Nat Methods. Nature Publishing Group; 2014;11(9):915–8.
122. Koning-ward TF De, Fidock DA, Thathy V, Menard R, Spaendonk RML Van, Waters AP, *et al*. The selectable marker human dihydrofolate reductase enables sequential genetic manipulation of the *Plasmodium berghei* genome. Mol Biochem Parasitol. 2000;106:199–212.
123. Chung CT, Niemela SL, Miller RH. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. Proc Natl Acad Sci. 1989;86(7):2172–5.
124. Birnboim HC, Doly J. A rapid alkaline extraction procedure for screeening recombinant plasmid DNA. Nucleic Acids Res. 1979;7(6):1513–23.
125. Brown TA. Gene Cloning & DNA Analysis. 6th ed. Sussex: John Wiley & Sons, Inc.; 2010. 85 p.
126. Trager W, Jensen JB. Human malaria parasites in continuous culture. Science. 1976;193(4254):673–5.
127. Verlinden BK, Niemand J, Snyman J, Sharma SK, Beattie RJ, Woster PM, *et al*. Discovery of novel alkylated (bis)urea and (bis)thiourea polyamine analogues with potent antimalarial activities. J Med Chem. 2011;54(19):6624–33.
128. Reilly HB, Wang H, Steuter JA, Marx AM, Michael T. Quantitative dissection of clone-specific growth rates in cultured malaria parasites. Int J Parasitol. 2008;37(14):1599–607.
129. Fidock DA, Ellems THEW. Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil. Proc Natl Acad Sci. 1997;94(September):10931–6.
130. Kozlov S, Waters NC, Chavchich M. Leveraging cell cycle analysis in anticancer drug discovery to identify novel plasmodial drug targets. Infect Disord Drug Targets. 2010;10(3):165–90.
131. Droucheau E, Primot A, Thomas V, Mattei D, Knockaert M, Richardson C, *et al*. *Plasmodium falciparum* glycogen synthase kinase-3: molecular model, expression, intracellular localisation and selective inhibitors. Biochem Biophys Acta. 2004;1697:181–96.
132. Maier AG, Rug M, Neill MTO, Brown M, Chakravorty S, Szestak T, *et al*. Exported Proteins Required for Virulence and Rigidity of *Plasmodium falciparum*- Infected Human Erythrocytes. Cell. 2008;134(1):48–61.
133. Patterson S, Robert C, Whittle C, Chakrabarti R, Doerig C, Chakrabarti D. Pre-replication complex organization in the atypical DNA replication cycle of *Plasmodium falciparum*: Characterization of the mini-chromosome maintenance (MCM) complex formation. Mol Biochem Parasitol. 2006;145:50–9.
134. Li J, Cox LS. Identification of an MCM4 homologue expressed specifically in the sexual stage of *Plasmodium falciparum*. Int J Parasitol. 2001;31:1246–52.
135. Lee AH, Symington LS, Fidock DA. DNA repair mechanisms and their biological roles in the malaria parasite *Plasmodium falciparum*. Microbiol Mol Biol Rev. 2014;78(3):469–86.
136. Birnbaum J, Flemming S, Reichard N, Soares AB, Mesén-ramírez P, Jonscher E, *et al*. A genetic system to study *Plasmodium falciparum* protein function. Nat Methods. 2017;(February).
137. Cai H, Hong C, Lilburn TG, Rodriguez AL, Chen S, Gu J, *et al*. A novel subnetwork alignment approach predicts new components of the cell cycle regulatory apparatus in *Plasmodium falciparum*. BMC Bioinformatics. BioMed Central Ltd; 2013;14(Suppl 12):S2.
138. Hammarton TC, Mottram JC, Doerig C. The cell cycle of parasitic protozoa: potential for chemotherapeutic exploitation. Prog Cell Cycle Res. 2003;5(April):91–101.
139. Bhattacharyya MK, Kumar N. Identification and molecular characterisation of DNA damaging agent induced expression of *Plasmodium falciparum* recombination protein PfRad51. Int J Parasitol. 2003;33:1385–92.
140. Merckx A, Le Roch K, Nivez M-P, Dorin D, Alano P, Gutierrez GJ, *et al*. Identification and initial characterization of three novel cyclin-related proteins of the human malaria parasite *Plasmodium falciparum*. J Biol Chem. 2003;278(41):39839–50.
141. Li X, Chen H, Rosa-Bahamontes N, Kun JFJ, Traore B, Crompton PD, *et al*. *Plasmodium falciparum* signal peptide peptidase is a promising drug target against blood stage malaria. Biochem Biophys Res Commun. 2014;380(3):454–9.
142. Bracchi-Ricard V, Barik S, Delvecchio C, Doerig C, Chakrabarti R, Chakrabarti D. PfPK6, a novel cyclin-dependent kinase/mitogen-activated protein kinase-related protein kinase from *Plasmodium falciparum*. Biochem J. 2000;347 Pt 1:255–63.
143. Reininger L, Billker O, Tewari R, Mukhopadhyay A, Fennell C, Dorin-Semblat D, *et al*. A NIMA-related protein kinase is essential for completion of the sexual cycle of malaria parasites. J Biol Chem. 2005;280(36):31957–64.
144. Solyakov L, Halbert J, Alam MM, Semblat J-P, Dorin-Semblat D, Reininger L, *et al*. Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*. Nat Commun. Nature Publishing Group; 2011;2(565):1–12.
145. Dorin D, Roch K Le, Sallicandro P, Alano P, Parzy D, Poullet P, *et al*. Pfnek-1 , a NIMA-related kinase from the human malaria parasite *Plasmodium falciparum* Biochemical properties and possible involvement in MAPK regulation. Eur J Biochem. 2001;2608:2600–8.
146. Miranda-Saavedra D, Gabaldón T, Barton GJ, Langsley G, Doerig C. The kinomes of apicomplexan parasites.

Microbes Infect. 2012;14(10):796–810.
147. Wilkes JM, Doerig C. The protein-phosphatome of the human malaria parasite *Plasmodium falciparum*. BMC Genomics. 2008;9(1):412.
148. Santos JM, Josling G, Ross P, Schieler A, Cristea IM, Llina M, *et al*. Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I Transcription Factor Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I Transcription Factor. 2017;731–41.
149. Flueck C, Bartfai R, Niederwieser I, Witmer K, Alako BTF, Bozdech Z, *et al*. A Major Role for the *Plasmodium falciparum* ApiAP2 Protein PfSIP2 in Chromosome End Biology. PLoS Pathog. 2010;6(2):1–15.
150. Sinha A, Hughes KR, Modrzynska KK, Otto TD, Pfander C, Dickens NJ, *et al*. A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. Nature. 2014;507(7491):253–7.
151. Wagner JC, Platt RJ, Goldfless SJ, Zhang F, Niles JC. Efficient CRISPR-Cas9–mediated genome editing in *Plasmodium falciparum*. Nat Methods. 2014;11(9):915–8.
152. Bell SP, Dutta A. DNA Replication in eukaryotic cells. Annu Rev Biochem. 2002;71:333–74.
153. Arias EE, Arias EE, Walter JC, Walter JC. Strength in numbers: preventing rereplication via multiple mechanisms in eukaryotic cells. Genes Dev. 2007;21(617):497–518.
154. Gupta A, Mehra P, Dhar SK. *Plasmodium falciparum* origin recognition complex subunit 5: functional characterization and role in DNA replication foci formation. Mol Microbiol. 2008;69(3):646–65.
155. Bushell E, Gomes AR, Sanderson T, Wengelnik K, Rayner JC, Billker O, *et al*. Functional Profiling of a *Plasmodium* Genome Reveals an Abundance of Essential Genes. Cell. Elsevier Inc.; 2017;170(2):260–72.
156. Modrzynska K, Pfander C, Chappell L, Rayner JC, Choudhary J, Modrzynska K, *et al*. A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the *Plasmodium* Life Cycle. Cell Host Microbe. 2017;21:11–22.
157. Fontaine F, Overman J, François M. Pharmacological manipulation of transcription factor protein-protein interactions: opportunities and obstacles. Cell Regen. 2015;4(2):1–12.
158. Nagel-wolfrum K, Buerger C, Wittig I, Butz K, Hoppe-seyler F. The Interaction of Specific Peptide Aptamers With the DNA Binding Domain and the Dimerization Domain of the Transcription Factor Stat3 Inhibits Transactivation and Induces Apoptosis in Tumor Cells. Mol Cancer Res. 2004;2(March):170–82.
159. Turkson J, Ryan D, Kim JS, Zhang Y, Chen Z, Haura E, *et al*. Phosphotyrosyl Peptides Block Stat3-mediated DNA Binding Activity , Gene Regulation, and Cell Transformation. J Biol Chem. 2001;276(48):45443–55.
160. Wu Y, Kirkman L a, Wellems TE. Transformation of *Plasmodium falciparum* malaria parasites by homologous integration of plasmids that confer resistance to pyrimethamine. Proc Natl Acad Sci U S A. 1996;93(3):1130–4.
161. van Dijk MR, Waters AP, Janse CJ. Stable Transfection of Malaria Parasite Blood Stages. Science (80-). 1995;268(5215):1358–62.
162. Dijk MR Van, Vinkenoog R, Ramesar J, Vervenne RAW. Replication , expression and segregation of plasmid-borne DNA in genetically transformed malaria parasites. Mol Biochem Parasitol. 1997;86:155–62.
163. Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, *et al*. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature. 2002;419:1–14.
164. Crosnier C, Bustamante LY, Bartholdson SJ, Bei AK, Theron M, Uchikawa M, *et al*. Basigin is a receptor essential for erythrocyte invasion by *Plasmodium falciparum*. Nature. 2011;480(7378):534–7.
165. White JH, Kilbey BJ. DNA replication in the malaria parasite. Parasitol Today. 1996;12(4):151–5.
166. Phillips MA, Burrows JN, Manyando C, Huijsduijnen RH Van, Voorhis WC Van, Wells TNC. Malaria. 2017;
167. Wells TNC, van Huijsduijnen RH, Van Voorhis WC. Malaria medicines: a glass half full? Nat Rev Drug Discov. Nature Publishing Group; 2015;14(6):424–42.
168. Nkrumah LJ, Muhle R a, Moura P a, Ghosh P, Graham F. Efficient site-specific integration in *Plasmodium falciparum* chromosomes mediated by mycobacteriophage Bxb1 integrase. Nat Methods. 2010;3(8):615–21.
169. Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, *et al*. Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. Science. 2014;343(6176):1247997.
170. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, *et al*. Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell. Elsevier Inc.; 2014;156(5):935–49.
171. Lacroix C, Giovannini D, Combe A, Bargieri DY, Späth S, Panchal D, *et al*. FLP/FRT-mediated conditional mutagenesis in pre-erythrocytic stages of *Plasmodium berghei*. Nat Protoc. 2011;6(9):1412–28.
172. Balu B, Chauhan C, Maher SP, Shoue D a, Kissinger JC, Fraser MJ, *et al*. piggyBac is an effective tool for functional analysis of the *Plasmodium falciparum* genome. BMC Microbiol. 2009;9:83.
173. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating. Proc Natl Acad Sci U S A. 1977;74(12):5463–7.
174. Ansorge I, Bhakdi S, Lingelbach K. Protein sorting in *Plasmodium falciparum*-infected red blood cells permeabilized with the pore-forming protein streptolysin O. Biochem J. 1996;315:307–14.
175. Allen RJW, Kirk K. Cell volume control in the *Plasmodium*-infected erythrocyte. Trends Parasitol. 2004;20(1):7–10.
176. Meissner M, Krejany E, Gilson PR, de Koning-Ward TF, Soldati D, Crabb BS. Tetracycline analogue-regulated transgene expression in *Plasmodium falciparum* blood stages using *Toxoplasma gondii* transactivators. Proc Natl Acad Sci U S A. 2005;102(8):2980–5.

Glossary

Alkaline denaturation	Briefly, bacterial cells are harvested by centrifugation at room temperature for 1 min at 11 000xg using an Eppendorf MiniSpin®. The supernatant is aspirated and the cell pellet is suspended in buffer A1 (with RNAse A to degrade cellular RNA, Tris and ethylenediaminetetraacetic acid [EDTA]). The EDTA in the buffer A1 functions to chelate divalent cations (Mg^{2+}, Ca^{2+}) essential for DNase activity, thereby preventing plasmid DNA damage and destabilisation of the bacterial cell wall. The bacterial cells were subsequently lysed with buffer A2 for 4 min at room temperature. This buffer likely contains sodium dodecyl sulphate to degrade the cell wall and membrane and sodium hydroxide to disrupt hydrogen bonds between DNA bases, producing single stranded DNA at a pH of 12.0-12.5. Buffer A3 (with potassium acetate) was added thereby neutralising the cell suspension and causing the re-naturing of the small circular plasmid DNA instead of the large genomic DNA fragments. Cells debris were pelleted by centrifugation and the supernatant transferred to the NucleoSpin® column where binding of DNA followed by 1 min centrifugation at 11 000x*g*. The column silica membrane was washed with buffer A4 (supplemented with ethanol) to remove contaminants such as salts, metabolites and soluble macromolecular molecules. The silica membrane was dried by centrifugation for 1 min at 11 000x*g*. Silica bound DNA was eluted with the slightly alkaline Buffer AE (5 mM Tris/HCl, pH 8.5).
Blue-white screening	Blue-white screening is a useful tool to identify recombinant bacteria and relies on the functionality of β-galactosidase, an enzyme which metabolises lactose to glucose and galactose. This gene is normally encoded by the *lac* operon of bacteria. *E. coli* strains, such as the XL10-Gold strain that is modified to only contain a *lacZΔM15* deletion mutation that only encodes for the Ω-peptide. The vector containing the ampicillin resistance gene and a gene *lacZ'* encodes for the α-peptide or first 146 amino acids of the β-galactosidase. Thus, β-galactosidase is only functional in the presence of recombinant bacteria. To select for recombinant bacteria containing vectors with the cloned insert, PCR products are cloned into the multiple cloning site in the *lacZ'* gene of the vector. Therefore, vectors with PCR product integration will not produce the α-peptide of β-galactosidase. *In vitro*, *Lac* operon transcription is induced by isopropyl β-D-1-thiogalactopyranoside (IPTG) spread onto agar bacterial growth plates. White colony formation indicates that a non-functional β-galactosidase was synthesised and likely contained the vector and cloned PCR product. If blue colonies are observed, the complete β-galactosidase molecule was synthesised, thus converted the lactose analog X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) to a dark blue product. Here, 100 µg/mL ampicillin supplemented agar plates was spread and allowed to dry with 100 mM IPTG first, and then with 20 mg/mL X-gal. Controls included a no insert (KO cassette) negative control and the pUC19 positive control.
BSD cassette	The selectable marker that is cloned into the gene target to disrupt the normal functioning of the target gene. The 5' *P. chabaudi* dihydrofolate reductase-thymidylate synthase (PcDT) UTR, a blasticidin-S deaminase (*bsd*) selectable marker transgene and a *P. falciparum* hrp2 3′ UTR make up the *bsd* cassette which can be selected for *in vitro* using blasticidin S hydrochloride. 
Bxb1 integrase system	The Bxb1 integrase isolated from mycobacteriophage catalyses the site-specific recombination between a chromosomally introduced *attB* site and the incoming *attP* site on the vector containing the desired transgene. This system is particularly useful for the efficient, stable introduction of a transgene into the parasite genome but is, however, limit to parasites strains that have been modified to contain the *attB* site (168).
Cas9	The wild-type *Streptococcus pyogenes* endonuclease used in Type II CRISPR systems induce double-strand breaks in DNA. The crescent-shaped molecule is approximately 100Å x 100Å x 50 Å in dimention and exhibits a bilobed architecture. The gRNA associated with the Cas9 endonuclease hybridise to homologous sequences on the target DNA strand. Once the sequence is confirmed and licensed for cleavage, the HNH and RuvC-like nuclease domains (approximately separated by 25Å apart) undergo conformational changes to cleave the annealed and displaced strands respectively, inducing a double strand break in the DNA about three

	nucleotides 5' of the PAM. The highly arginine rich region of the Cas9 endonuclease has been suggested to facilitate nucleic acid binding (169,170).
CRISPR	Clustered regularly interspaced short palindromic repeats is a locus in bacterial genomes that functions as a pathogen defence system where short nucleotide sequences are transcribed to direct degradation of invading foreign DNA sequences (pathogens such as phages).
crRNA	CRISPR RNA upstream sequences of protospacers; forms part of the CRISPR array in a CRISPR locus.
diCre conditional recombination	A codon-optimised version of the gene of interest and a partner selectable marker gene is flanked by *loxP* sites are delivered to parasites in addition to another vector that encodes for the *bsd* selectable marker and two inactive Cre1 and Cre2 polypeptides fused to rapamycin-binding proteins such as FK506-binding proteins (FKBP). Upon addition of rapamycin, the Cre polypeptides fuse to produce a diCre heterodimer wich allows excision of *loxP*-flanked sequences trough recombinase activity (86,87). This system offers advantages including that it is rapid and have negligible leakage; however, the system is not reversible, very costly and is not suitable for many applications.
Ethanol precipitation	Ethanol precipitation is a technique that purifies and concentrates nucleic acids. Nucleic acids are soluble in water with a high dielectric constant and can electrostatically interact based on the fact that water is polar and has partial negative charges surrounding the oxygen atom due to unpaired electrons and the polarity of nucleic acids solubilised in water Lowering of the dielectric constant of a solution (like ethanol) that contains the negatively charged nucleic acids (PO^{3-} backbone) and positively charged monovalent cations such as Na^{+} (salts like sodium acetate), increases the Coulomb force of attraction and allows efficient electrostatic interaction of the molecules. Nucleic acids are neutralised and subsequently precipitate out of solution.
FKBP destabilisation domain	The destabilisation domain, a mutated version of the rapamycin-binding FK506-binding protein (FKBP) C-terminus, is fused to a protein of interest and directs a protein for ubiquitylation and degradation. In the presence of ligand Shield1, a permeable small-molecule bind to the DD, subsequently reversing degradation and allowing the protein to be stabilised (83).
FRT/FLP recombinase system	Parasites strain is modified to carry the FLP recombinase where transgenes of interest flanked by *frt* sites are integrated into the genome. The outcome of the FLP-mediated recombination is determined by the relative position and orientations of the *frt* sites. This system has not been extensively used in *P. falciparum* due to spontaneous excision of essential genes that lead to lethal phenotypes. Furthermore, this system requires stage specific recombinase expression or additional inducible system integration, such as the use of FKBP destabilisation (171).
Gibson DNA assembly	Gibson DNA assembly allows for the assembly of multiple fragments of DNA of various lengths with overlapping sequences in a single-tube isothermal reaction. This involves three enzymatic reactions whereby the four DNA fragments are annealed: 1) Single stranded 3' overhangs are generated by an exonuclease to allow for annealing of DNA fragments with complementary ends, 2) the proprietary adds nucleotides in the holes within an annealed fragment, and finally, 3) the nicks are sealed by DNA ligase.
Guide RNA	TracrRNA:crRNA chimera guide RNA is a short nucleotide sequence that directs Cas9 to a particular site in the target genome to induce double-strand breaks. The gRNA functions to provide target specificity and binding ability to Cas9 endonuclease to a DNA region.
KO cassette	The KO cassette is the sequence template by which coordinated DNA repair of double-strand breaks will be repaired with carrying desired SNP mutations or gene replacements. 
Ligation	In the ligation reaction, T4 DNA ligase enzymatically catalyses the joining of the 3' hydroxyl terminus of the vector to the 5' phosphate terminus of the insert (annealed guide oligonucleotides). AMP is transferred from the lysine residue of the enzyme active site to the 5' phosphate; the AMP-phosphate bond is subsequently attacked by the 3' hydroxyl group, releasing the AMP and forming a new covalent phosphodiester linkage. The ATP in the T4 DNA ligase buffer serves to replenish the AMP in the enzymes active site (125).
Off-target activity	DNA double strand breaks by Cas9 at undesired target regions.

On-target activity	Efficiency of DNA double-strand breaks induced by Cas9 at desired locations
PAM	The protospacer adjacent motif is the three nucleotide sequence upstream of the target DNA sequence that is essential for gRNA to direct Cas9 to induce double-strand breaks.
PCR	A typical polymerase chain reaction involves denaturing of double strand DNA by disruption of hydrogen bonds in high temperatures to allow two oligonucleotide primers to specifically hybridise to opposite strands of the double strand target DNA sequence intended for amplification. During the extension reaction, thermo stable high fidelity DNA polymerases such as *Taq* DNA polymerase (isolated from *Thermus aquaticus* bacteria) extends single strand DNA synthesised to double strand DNA (from a 3' hydroxyl group provided by the primer and deoxynucleosides triphosphates in the reaction) that can be denatured in the following cycle. This reaction occurs in a buffer that contains Mg^{2+} that serves as a co-factor for DNA polymerase and double-strand DNA stabiliser.
PiggyBac transposon mutagenesis	On one vector, the *piggyBac* class II integrase transposase is encoded to specifically excise a region where inverted terminal repeat (ITR1 and ITR2) flanking a transgene of interest on another vector. The transgene is inserted randomly at TTAA sites in the genome or can flank an open reading frame (82,172).
Pre-crRNA	Precursor CRISPR RNA are long, immature mRNA transcribed from CRISPR genes.
Riboswitch system	As a conditional knockdown system, the *glmS* ribozyme is placed between the 3' UTR and the stop codon in a target gene through homologous recombination at the C-terminus of the gene of interest. The ribozyme is activated by treatment of glucosamine-6-phosphate which allows removal of the 3' UTR, cleavage of the mRNA sequences and ultimately leading to its degradation.
Sanger di-deoxynucleotide sequencing	This is based on the chain termination method developed by Fredrick Sanger, where di-deoxynucleotide phosphates (ddNTPs) lacking 3' hydroxyl (OH) groups on the sugar carbon on are fluorescently labelled and incorporated at the terminal ends of newly synthesised DNA by DNA polymerases along with standard deoxynucleotides (dNTPs) and a single sequencing primer. Since the 3' OH group is required for the formation of a phosphodiester bond between two nucleotides, the DNA synthesis is terminated where the ddNTP is incorporated(173). The DNA sample is ran four times to detect each ddNTP, i.e. ddATP, ddGTP, ddCTP and ddTTP, separately. The samples are then detected by capillary gel electrophoresis and detection of the respective fluorescent ddNTP by exciting the dye to emit fluorescence which is detected.
Saponin lysis	Sapogenin, the active compound in used a process of saponin lysis, is a plant based glycoside structurally characterised by a tripernoid aglycone backbone with associated sugar chain. Saponin lysis allows for the isolation of intact parasites from the erythrocyte cytosol and parasitophorous vacuolar membrane trough permeabilising the membrane to macromolecultes, yet keeping the parasite membrane intact and able to maintain electrochemical gradients (174,175).
T4 DNA ligase	T4 DNA ligase enzymatically catalyses the joining of the 3' hydroxyl terminus of the vector to the 5' phosphate terminus of the insert (annealed guide oligonucleotides) in a 1X T4 DNA ligase buffer (50mM Tris-HCl, 10mM $MgCl_2$, 1mM ATP, 10mM DTT, pH 7.5 at 25˚C. AMP is transferred from the lysine residue of the enzyme active site to the 5' phosphate; the AMP-phosphate bond is subsequently attacked by the 3' hydroxyl group, releasing the AMP and forming a new covalent phosphodiester linkage. The ATP in the T4 DNA ligase buffer serves to replenish the AMP in the enzymes active site.
T4 polynucleotide kinase	T4 polynucleotide kinase (PNK) catalyse the transfer of an inorganic phosphate from the γ position of ATP to the 5' hydroxyl terminus of polynucleotides. Richardson units are defined as defined as the amount of enzyme catalysing the incorporation of 1 nmol of acid insoluble [^{32}P] in a total reaction volume of 50 µl in 30 minutes at 37˚C in 1X T4 Polynucleotide Kinase Reaction Buffer (70mM Tris-HCl, 10mM $MgCl_2$, 5mM DTT, pH 7.6 at 25˚C) with 66 µM [γ-32P] ATP (5 x 106 cpm/µmol) and 0.26 mM 5' hydroxyl-terminated salmon sperm DNA. Note that in the study, T4 polynucleotide was used in T4 DNA ligase buffer as PNK requires ATP and previous work proved that differences in Tris-HCl and DTT were negligible.
TA Cloning	The pCR®2.1-TOPO® is a vector is supplied with 3' thymidine overhangs suitable for the directional TA-cloning and topoisomerase I covalently bound the vector. This system works on the basis that Taq DNA polymerase adds a single deoxyadenosine nucleotide to the 3' ends of PCR products. The pCR®2.1-TOPO® is linearised by *EcoRI* restriction enzyme digestion with 3' deoxythymidine overhangs, therefore allowing for

	the ligation of the insert to the vector and prevention of the linearised plasmid vector to recircularise. The role of the topoisomerase I covalently bound to the vector facilitates the cleavage of the of the phosphodiester backbone at 5'-CCCTT in a single strand. This allows the formation of a 3' phosphodiester bond as the energy from the cleavage is conserved in the covalent bond formed between the 3' phosphate of the cleaved strand and a tyrosyl residue (Tyr-274) of topoisomerase I. The 5' hydroxyl attacks the covalent bond between the DNA and the protein, therefore reversing the reaction and releasing the enzyme.
TetR-aptamer system	The tetracycline-repressible conditional transcriptional knockdown system modulates transcription levels through binding of Tet repressor (TetR) that is bound to an aptamer to the operator sequences (TetO) of a gene of interest around the transcriptional start site. When TetR binds to the operator sequence, gene expression is repressed. However, in the presence of ligand binding, i.e. anhydrotetracycline (ATc), the repressor undergoes a conformational change, preventing operator binding, leading to gene expression (176). This system has recently been adapted to include a native small regulatory module (such as the evelopment of zygote inhibited or DOZI protein in *P. falciparum*) to enhance translational control and reduced leaky expression of the gene (84).
Time constant (electroporation)	The time constant is an exponential decay function where the release a preset voltage from the capacitor is measured as it is decayed over time (milliseconds).
Topoisomerase I	Topoisomerase I isolated from the *Vaccinia* virus, specifically recognises and cleaves at 5'-(C/T)CCTT-3' sites, unwinding the supercoiled DNA, and subsequently re-ligates the ends at the 3' phosphate group of a thymidine nucleotide. Given that gene amplification using a non-proofreading *Taq* DNA polymerase introduces a complementary adenine nucleotide to the 3' terminus of the PCR product, effective cloning of a PCR product into the vector.
tracrRNA	A trans-activating CRISPR RNA (tracr-RNA) hybridises with direct repeats of the crRNA, forming a RNA duplex is formed that is processed and cleaved by endogenous RNase III in the presence of the Cas9 endonuclease.
Zinc-finger nucleases	Protein pairs, (ZFNL and ZFNR) that are co-expressed under a single promoter (achieved through the inclusion of a viral 2A ribosomal skipping peptide on the peptide) explicitly recognises user-defined DNA sequences are linked to a FokI endonuclease that induces a double-strand break and thus triggers homologous recombination in the parasite. As with the CRISPR system, another vector that contains the transgene or desired mutation as a cDNA version (88).

Supplementary

1. Sanger sequence results of the KO cassette from the amplified PCR products following Gibson DNA assembly.

Figure 22 in the main text shows the Sanger sequencing results of PF3D7_0613800 PCR amplified KO cassette as an example. The figures below indicate the sequences for the remaining gene's KO cassettes' PCR products.

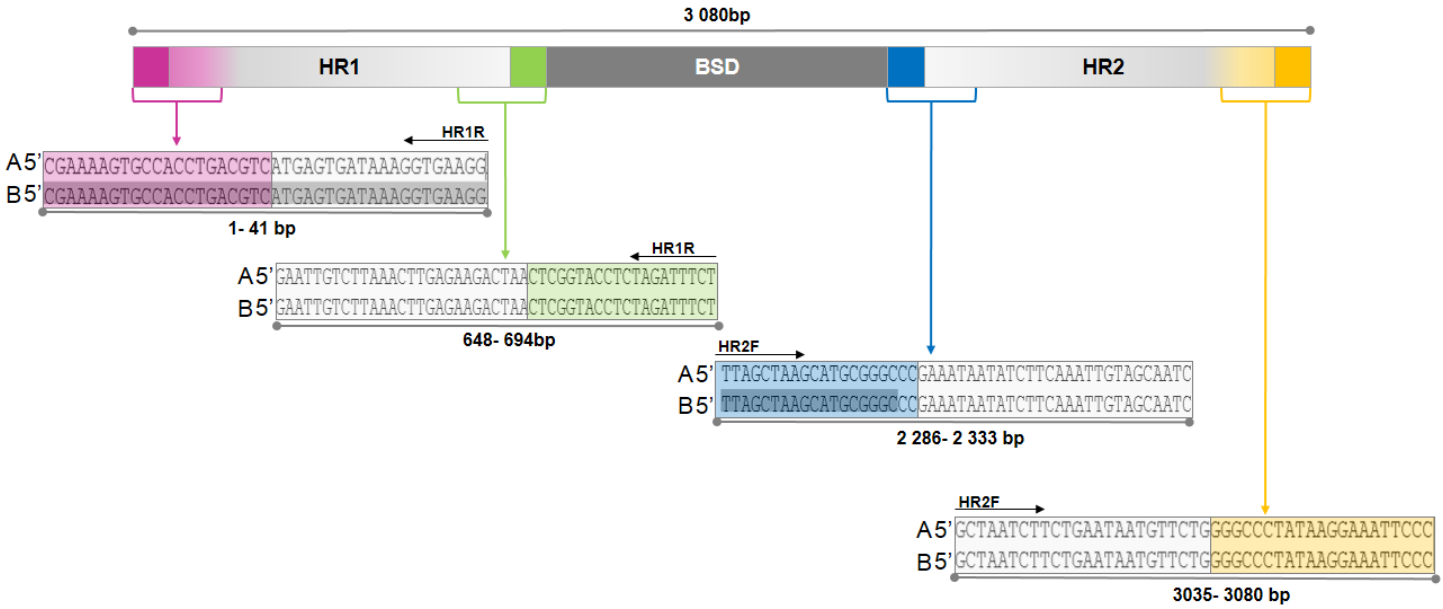

Figure S1. Sanger sequencing results of PF3D7_1239200 PCR amplified KO cassette. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers used to sequence are shown in the relative orientation as HR1R (homology region 1 reverese primer) and HR2F (homology region 2 forward primer). Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

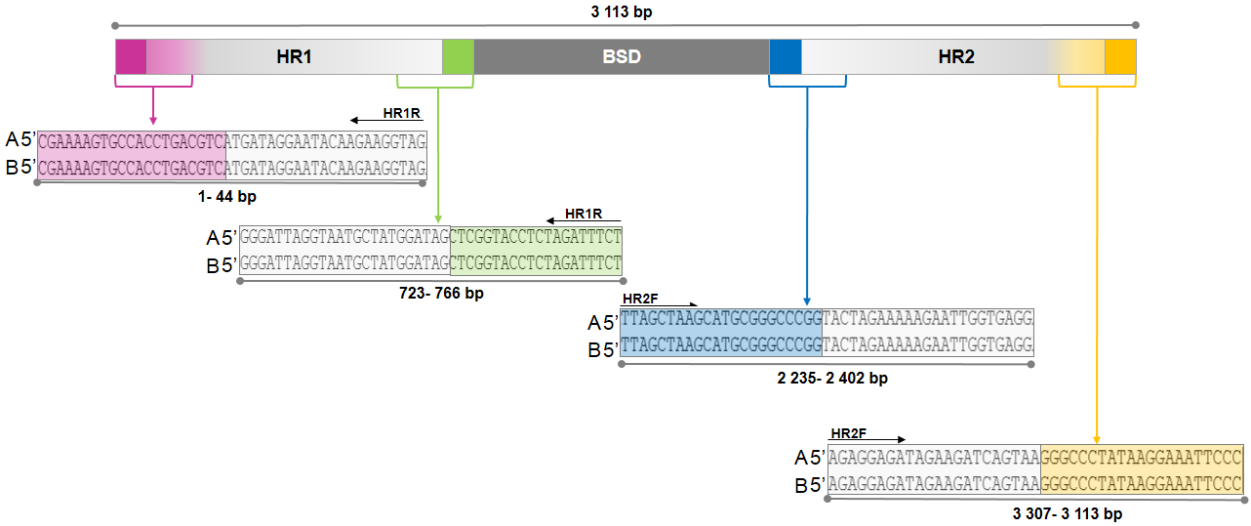

Figure S2. Sanger sequencing results of PF3D7_1211700 PCR amplified KO cassette. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers used to sequence are shown in the relative orientation as HR1R (homology region 1 reverese primer) and HR2F (homology region 2 forward primer). Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

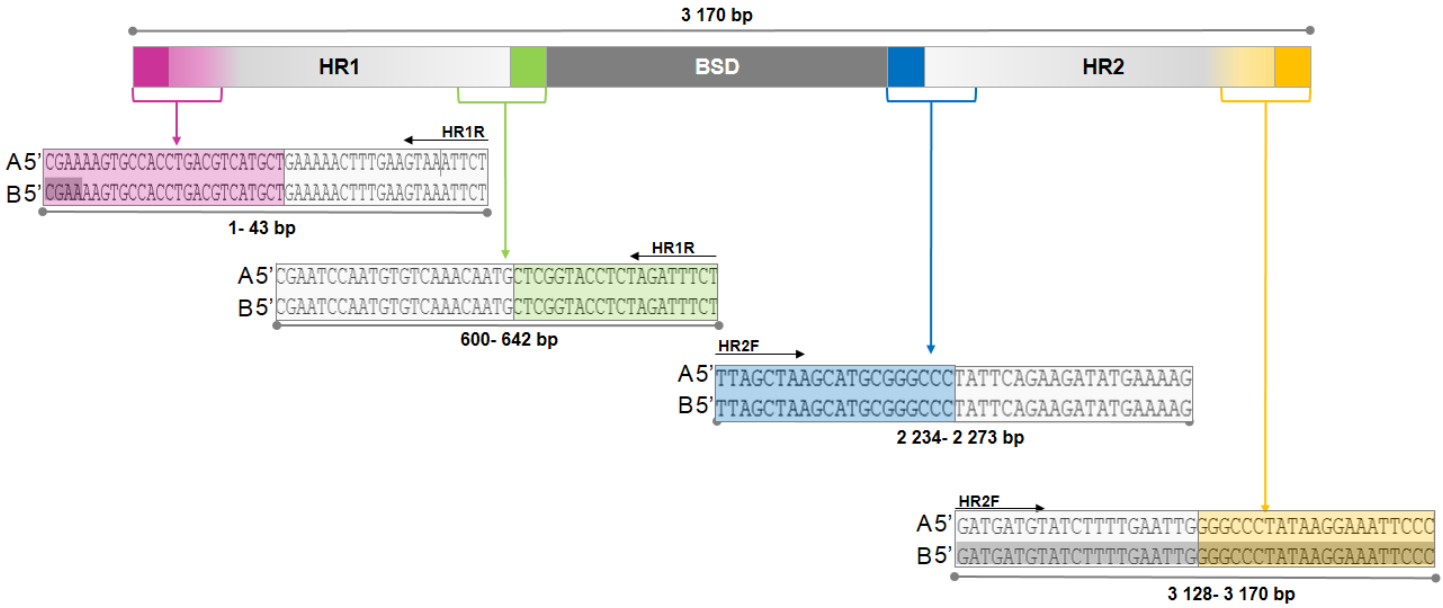

Figure S3. Sanger sequencing results of PF3D7_0705300 PCR amplified KO cassette. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers used to sequence are shown in the relative orientation as HR1R (homology region 1 reverese primer) and HR2F (homology region 2 forward primer). Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

2. Sanger sequence results of KO cassette of each gene as cloned into the pCR®2.1-TOPO® vector

Figure 24 in the text shows the Sanger sequence results of PF3D7_1239200 KO cassette cloned in to the pCR®2.1 TOPO® expression vector showing overlap and primer sequences as aligned to the (A) expected sequence and (B) sequence results. The figures below indicate the sequences for the remaining gene's KO casettes.

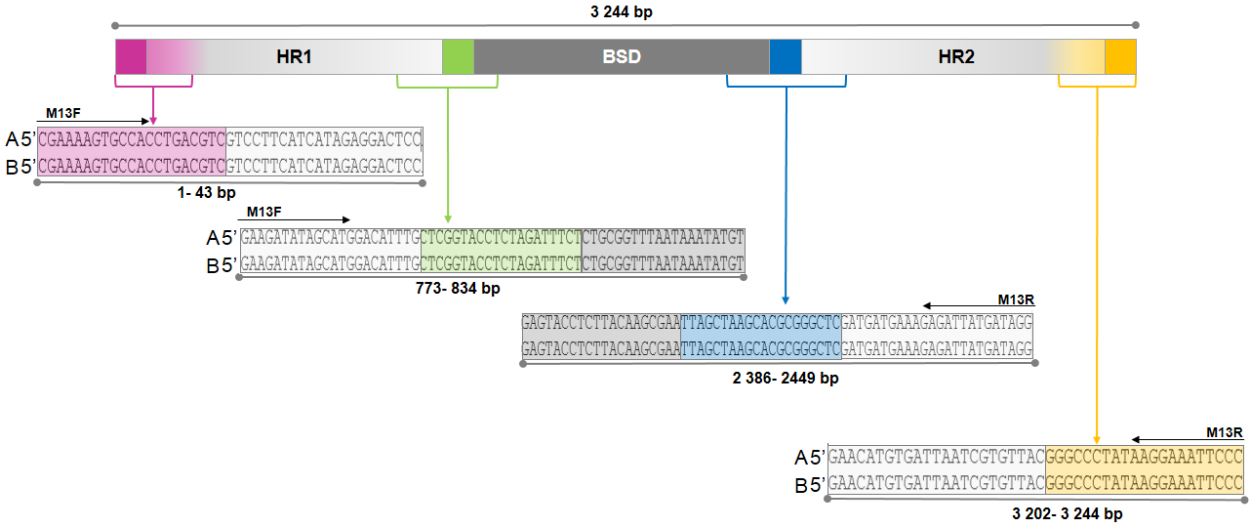

Figure S4. Sanger sequence results of PF3D7_0613800 KO cassette cloned in to the pCR®2.1 TOPO® expression vector. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers M13F (forward) and M13R (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

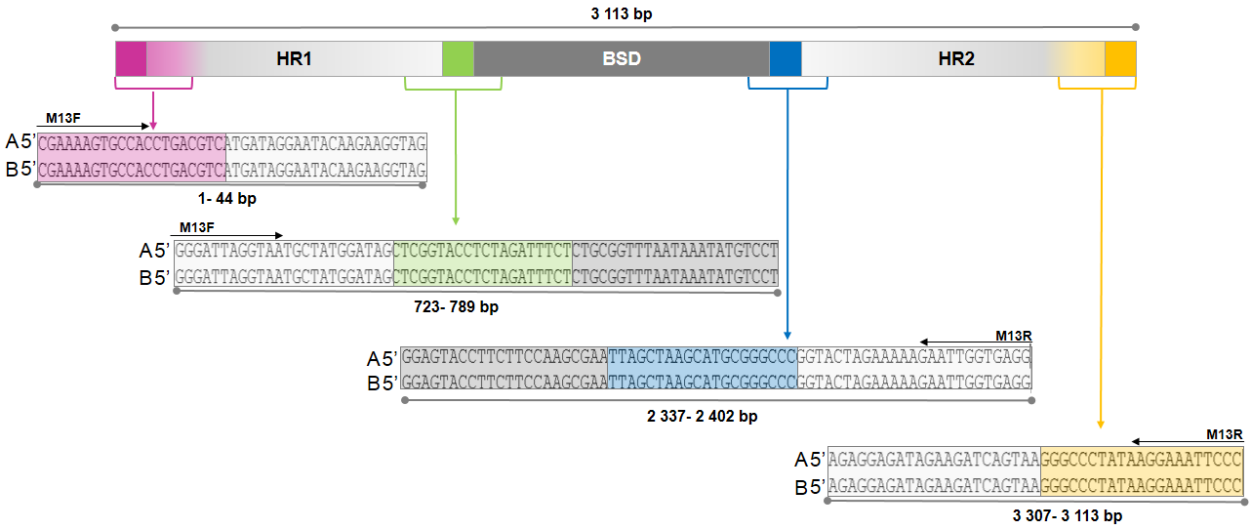

Figure S5. Sanger sequence results of PF3D7_1211700 KO cassette cloned in to the pCR®2.1 TOPO® expression vector. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the (A) expected sequence and (B) the sequencing product. Primers M13F (forward) and M13R (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

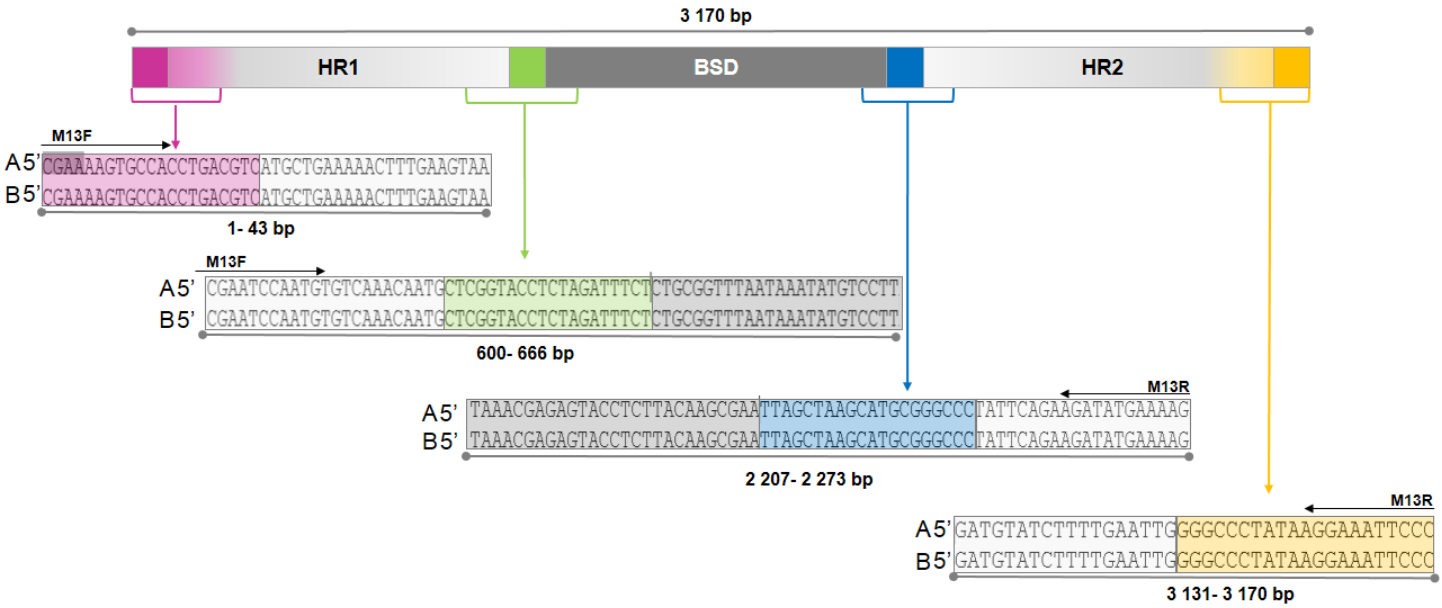

Figure S6. Sanger sequence results of PF3D7_0705300 KO cassette cloned in to the pCR®2.1 TOPO® expression vector. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers M13F (forward) and M13R (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm). Sanger sequence results of KO cassette as cloned into the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector.

3. Sanger sequence results of KO cassette of each gene as cloned into the pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector.

Figure 29 in the text shows the Sanger sequence results of an example of pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector targeting PF3D7_1211700[guide1].

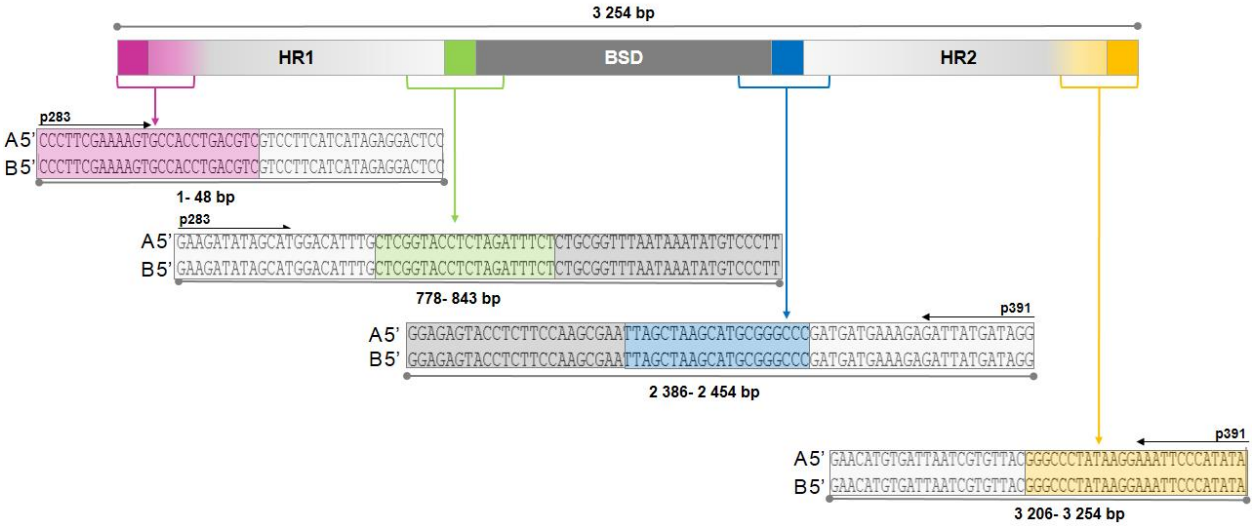

Figure S7. Sanger sequence results of the KO cassette cloned pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector targeting PF3D7_0613800[guide1]. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers p283 (forward) and p391 (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

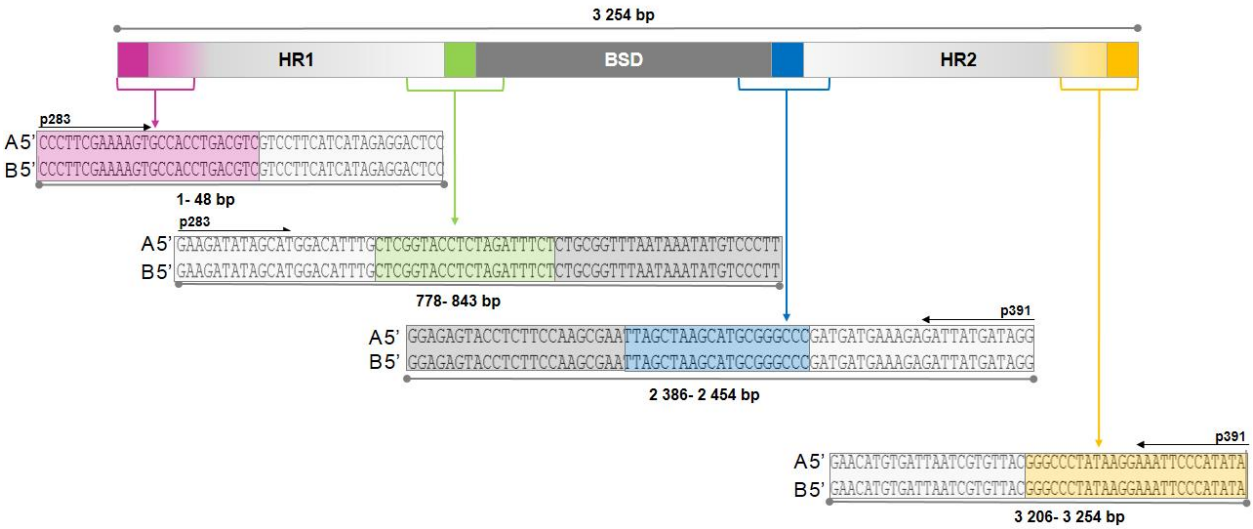

Figure S8. Sanger sequence results of the KO cassette cloned pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector targeting PF3D7_0613800[guide2]. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers p283 (forward) and p391 (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC

Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

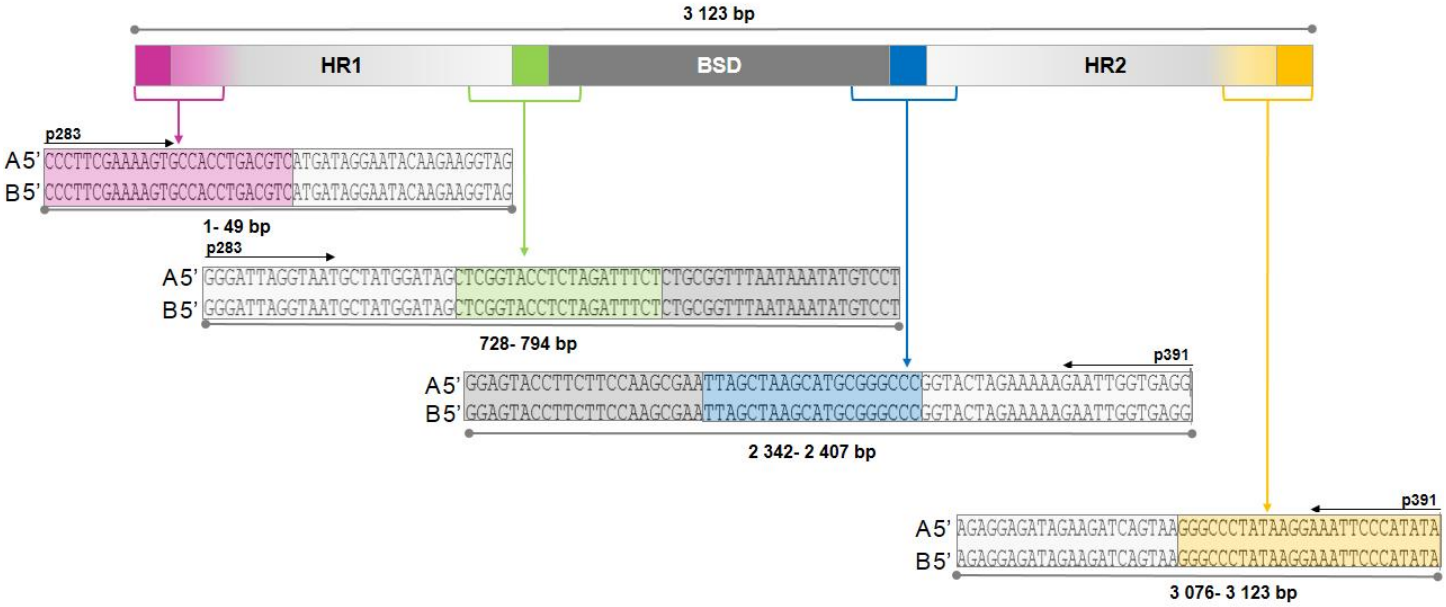

Figure S9. Sanger sequence results of the KO cassette cloned pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector targeting PF3D7_1211700^{guide2}. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers p283 (forward) and p391 (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).

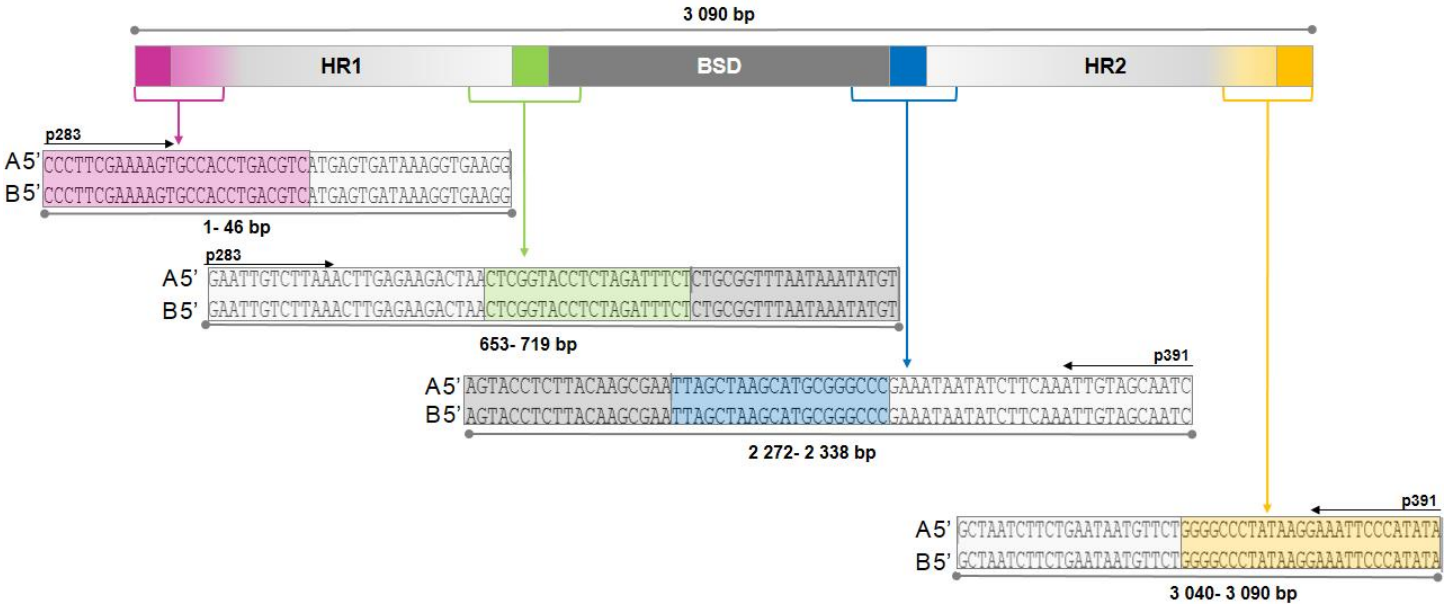

Figure S10. Sanger sequence results of the KO cassette cloned pDC2-U6-gRNA-co*CAS9*-CAM-*hDHFR* vector targeting PF3D7_1239200^{guide1}. The sequenced nucleotides are shown as fragments indicating the overlap sequences in colour while the gene specific primers are shown in light grey. Realtive positions of the sequences fragments are shown as aligned to the **(A)** expected sequence and **(B)** the sequencing product. Primers p283 (forward) and p391 (reverse) used to sequence are shown in the relative orientation of the template. Sequences received from GATC Biotech (.ab1 format) was analysed using Lasergene (v.14) SeqManPro multiple sequence alignment (MAFFT algorithm).