University of Glasgow

Power, Bridget Joanne (2019) *Epigenomics and gametocytogenesis in the rodent malaria model, Plasmodium berghei*. PhD thesis.

https://theses.gla.ac.uk/71046/

Copyright and moral rights for this work are retained by the author

A copy can be downloaded for personal non-commercial research or study, without prior permission or charge

This work cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author

The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author

When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given

Enlighten: Theses
https://theses.gla.ac.uk/
research-enlighten@glasgow.ac.uk

Epigenomics and gametocytogenesis in the rodent malaria model, *Plasmodium berghei*

Bridget Joanne Power,
BSc (Hons), MRes

Thesis submitted in fulfilment of the requirements for the Degree of Doctor of Philosophy

The Institute of Infection, Immunity and Inflammation,
College of Medical, Veterinary and Life Sciences,
University of Glasgow

September 2018

Abstract

In 2017, an estimated 219 million people became infected with a *Plasmodium* parasite, resulting in malaria. Of the 435,000 deaths, 61% were in children aged 5 years and under. With increasing incidences of resistance to standard anti-malarial therapies, and the unprecedented failure of rapid diagnostic tests in Sub-Saharan Africa, there is an onus on the scientific community to identify new mechanisms by which parasites of the *Plasmodium* genus can be detected and targeted for eradication.

In the present study, the epigenetic mechanisms by which the rodent malaria parasite, *Plasmodium berghei*, regulates its transmission were examined. In malaria parasites, transmission from a mosquito vector to a mammalian host is mediated by a stage of sexual development known as gametocytogenesis. Using reverse genetics approaches, in combination with fluorescence-assisted cell sorting, a previously undetermined role for histone acetyltransferase 1 (HAT1) in *Plasmodium* intraerythrocytic growth and gametocytogenesis was determined. In a similar manner, histone deacetylase 1 (HDA1), a protein previously shown to be transcribed alongside proteins AP2-G and LSD2 prior to schizont egress, was determined to be essential to optimal asexual growth of *P. berghei*.

In addition to these functional studies, acid extraction of histones and tandem mass spectrometry were used to determine the array of histone post-translational modifications that define the mature asexual schizont from sexual-stage gametocytes. To examine some of these sex-specific histone modifications further, a novel chromatin-immunoprecipitation/high-throughput sequencing technique, called 'ChIPmentation', was used to identify the genomic targets of four epigenetic modifications in mature asexual parasites, with a view to future comparison of these histone marks to those found at gametocyte stage. This study also represents the first time that H3K122 acetylation has been examined at any stage in any *Plasmodium* species.

The insights gleaned from this study add to our current knowledge of the landscape of epigenetic modifications and their regulatory enzymes at asexual and sexual stages of *Plasmodium* parasite development.

Table of Contents

List of Tables

List of Figures

List of Supplemental Material

Supplemental Material 1 (S1): Supplemental Information

Supplemental Material 2 (S2): Supplemental Bioinformatic Text

Electronic Supplemental Material 1 (ES1): Co-immunoprecipitation

Electronic Supplemental Material 2 (ES2): ChIPmentation results- Epic

Electronic Supplemental Material 3 (ES3): ChIPmentation repeat results- Epic

Electronic Supplemental Material 4 (ES4): RNA-seq EdgeR_paired

Electronic Supplemental Material 5 (ES5): NGS result metadata

Glimma plots

Acknowledgments

I would first of all like to thank Prof Andy Waters at the Wellcome Centre for Integrative Parasitology for allowing me to join his lab for my thesis project. I am extremely grateful for the chance to work in a lab with such cutting-edge molecular technology and to be allowed to work alongside some of the best minds in both malaria research and molecular parasitology. And I would like to thank my second supervisor, Dr Richard Burchmore, for his wonderful expertise in both proteomic analysis and teaching. Some of the most enjoyable moments of my PhD experience were those in which I had the opportunity to teach the next generation of scientists, no matter how young or old.

I possibly should have thanked them first, but I would not have been here without the co-ordinators of the fantastic Wellcome Trust 4-Year PhD Programme at the University of Glasgow: Prof Darren Monckton and Dr Olwyn Byron. I will never be able to thank them enough for the opportunity that they gave me. I am eternally grateful to them for allowing me to be part of the Wellcome Trust PhD family. Not only did they teach me to write a scientific grant, budget for lab consumables, and organise research forums and retreats; they also provided me with stability for four years, and through them I met people who will be my friends for the rest of my life. I would like to thank many of those former Wellcome students also: Andrew Pountain, Síle Johnson, Andrei Kamenski, Alejandro Huerta Uribe, Hannah Wessel (a friend I made at the very first interview!), Kevin Rattigan, Nina Divorty, Marlène Corbet, Hussain Jaffery, Catrina Mullan, Pieter Steketee, Tamsin Zangerle-Murray, Daniel Nii Aryee Tagoe and Dominika Korzekwa (to name a few).

I would also like to thank my assessors, Dr Richard McCulloch and Dr David Vetrie for their advice, chat, and for the opportunity to travel to the Wellcome Sanger Institute for the amazing ‘Chromatin, Structure and Function’ course in 2015. Taking that advanced course helped me tremendously in my scientific endeavours, and again, I met another wonderful bunch of unique and talented scientists who have remained my friends ever since.

I would also like to thank members of the Waters, Barrett, and de Koning Labs for making my PhD experience more than just a degree. I am particularly grateful for

the help and friendship of Sebastian Kirchner, Heli Vaikkinen, Angela McBride, Julie Martin, Katie Hughes, Michael Rennie, Becca Lee, Robyn Kent, Harshal Patil, Anubhav Srivastava, Agnieszka Religa, Katarzyna Modrzynska, Scott Millar, Dafni Bechtsi, Titilola Kalejaiye Iliyemi, Gustavo Campagnaro, Clément Regnault, Matthew Gould, Federica Giordani, Anne Donachie, Gordon Campbell, Jean-Christophe Le Bayon, Isabel Vincent, and Stanley Jean-Vincent Le Bayon. And though you weren't in my building, a big thank you to Bhoomi Gor!

I would also like to thank a group of friends in Ireland who got me through my undergraduate degree in Dublin and are equally responsible for getting me this far: Graham Hughes, Thomas Schwarzl, Aisling McHugh, Niamh McCarry, and Finbarr Tarrant. And thank you to Prof Brendan Loftus at UCD for giving me my first ever chance to work in a lab.

And most importantly, thank you to my family for their help and support throughout my PhD in Scotland: my parents, Séamus and Kay Power, my older brother, James, and his wonderful fiancée Colette, and my not-so-little brother, Paul. And thank you to my Granny, aunts, uncles, and cousins, from the youngest to the oldest! And, of course, thank you so much to my wee 'Scottish' family, David and Milfred ('Milly') Meltzer, for making Glasgow feel like home. And thank you to Rita Walsh and Aimee O'Neill, my 'oldest' friends, for keeping me sane too!

And finally, I would like to dedicate this work to two of the 'Joanne Powers' who came before me: my aunt, Joan Power, for whom I was named, and my great-aunt, Johanna 'Josie' Grace (née Power), who would have loved this!

Author's Declaration

I declare that I am the sole author of this thesis and that all work carried out in the present study is my own. Experiments carried out in collaboration with others are clearly stated, and their contribution is declared. This work has not been submitted for a degree at any other institution.

Bridget Joanne Power

Abbreviations/Acronyms

°C	Degrees Celsius
μ	Micro
μg	Microgram
μl	Microlitre
μm	Micrometre
μM	Micromolar
μmol	Micromoles
3'cam	3'UTR of *P. berghei* calmodulin gene
3'p48/45	3'UTR of *p48/45* gene
3'UTR	Three prime untranslated region
5-FC	5-fluorocytosine
5-fluoro-UMP	5-fluoro-uridine monophosphate
5-FU	5-fluorouracil
5'UTR	Five prime untranslated region
6-Cys	Six cysteine
AdoMet	s-adenosyl methionine
AID	Auxin-inducible degron
Alba	Acetylation lowers binding affinity protein
AP2-G	Apetala 2-Gametocyte
AP2-G2	Apetala 2-Gametocyte 2
AP2-L	Apetala 2-Liver
AP2-O	Apetala 2-Ookinete
AP2-sp	Apetala 2-sporozoite
ASPA	Animals (Scientific Procedures) Act 1986
ATc	Anhydrotetracycline
BFP	Blue fluorescent protein
BSA	Bovine Serum Albumin
bp	Base pairs
CAF1	CCR4-associated factor 1
Cas	CRISPR-associated
CCR4	Carbon catabolite repressor 4
CDC	Centers for Disease Control and Prevention

CD	Cluster of Differentiation
CDPK	Calcium-Dependent Protein Kinase
CELF	CUG-BP, Elav-like family protein
CHD	Chromodomain Helicase DNA-binding protein
CI	Confidence Interval
circRNA	Circular RNA
CITH	Homolog of worm CAR-I and fly Trailer Hitch
CpG	5'-C-phosphate-G-3'
Cre	Creates recombination
CRISPR	Clustered, regularly interspersed, short palindromic repeat
CRISPR-Cas9	CRISPR-CRISPR-associated protein 9
crRNA	CRISPR RNA
DAPI	4',6-Diamidino-2-Phenylindole
DBD	DNA-binding domain
ddH_2O	Double Distilled Water
DHFR-TS	Dihydrofolate Reductase-Thymidylate Synthase
DHS	Deoxyhypusine synthase
DiCre	Dimerisable Cre recombinase
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DOZI	Development of zygote inhibited
DOT1	Disruptor of telomeric silencing 1
dTMP	Deoxythymidine monophosphate
DTT	Dithiothreitol
dTTP	Deoxythymidine triphosphate
DUF	Domain of Unknown Function
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediamintetraacetic acid
EGFP	Enhanced Green Fluorescent Protein
EGTA	Ethylene glycol-bis(ß-aminoethyl ether)-N,N,N',N'-tetraacetic acid
eIF	Eukaryotic translation initiation factor

ERF	Ethylene Response Factor
FACS	Fluorescence Activated Cell Sorting
FAD	Flavin Adenine Dinucleotide
FASP	Filter Aided Sample Preparation
FBS	Foetal Bovine Serum
FKBP	FK506-binding protein
Floxed	Flanked by *loxP* sites
FRB	FKBP12-rapamycin associated binding protein
FRG KO huHep	*Fah*$^{-/-}$, *Rag2*$^{-/-}$, *Il2rg*$^{-/-}$ knockout, human hepatocyte-engrafted
g	Grams
g	Relative centrifugal force
G	Gauge
G6PD	Glucose-6-Phosphate Dehydrogenase
GAF	Gametocyte Activating Factor
GBSSR	Genome Bootstrapping with Step-wise Search space Reduction
GCS1	Generative cell specific protein 1
gDNA	Genomic DNA
GEST	Gamete egress and sporozoite traversal protein
GFP	Green Fluorescent Protein
GFP-LUC	GFP-luciferase
GIMO	'Gene Insertion/Marker Out'
GNAT	Gcn5-related *N*-acetyltransferase
GOI	Gene Of Interest
GOMO	'Gene Out/Marker Out'
GPCR	G-protein coupled receptor
gRNA	'Guide' RNA
HAT1	Histone Acetyltransferase 1
HDA	Histone deacetylase
HDAC	Histone deacetylase
HDAC2/IPK1	Histone deacetylase 2/Inositol Phosphate Kinase 1
HDHFR	Human Dihydrofolate Reductase

HEPES	4-(2-Hydroxyethyl)-1-Piperazineethanesulfonic acid
HKMT	Histone lysine methyltransferase
HMGB	High-mobility-group box
HMM	Hidden Markov Model
HP1	Heterochromatin Protein 1
HPLC	High Performance Liquid Chromatography
HRP	Horseradish Peroxidase
HT	Sodium Hypoxanthine and Thymidine
IAA	Indole-3-acetic acid
IDC	Intraerythrocytic Developmental Cycle
IFA	Immunofluorescence Assay
JGI	Jane Goodall Institute
JRF	Joint Research Facility
kb	Kilobases
KD	Knockdown
kDa	Kilodaltons
$KHCO_3$	Potassium Bicarbonate
KO	Knockout
KS	Knock-sideways
LCCL	Limulus coagulation factor C, Coch-5b2 and Lgl1 domain
lncRNA	Long non-coding RNA
loxP	Locus of crossover [x] in P1 bacteriophage
LysoPC	Lysophosphatidylcholine
m	Metre
m	Milli
M	Molar
MAOP	Membrane-attack ookinete protein
MDV1/PEG3	Male development protein 1/protein of early gametocyte 3
MEF	Mosquito Exflagellation Factor
mg	Milligram
$MgCl_2$	Magnesium Chloride
min	Minute

miRNA	Micro RNA
ml	Millilitre
ML	Maximum Likelihood
mM	Millimolar
mRNP	Messenger ribonucleoprotein
MS	Mass Spectrometry
MSP	Merozoite Surface Protein
MTOC	Microtubule Organising Centre
MYST	MOZ, Ybf1/Sas3, Sas2, and Tip60
n	Nano
NAA	1-naphthalenacetic acid
NaCl	Sodium Chloride
NAD	Nicotinamide Adenine Dinucleotide
NAT	*N*-terminal acetyltransferase
Nek	NIMA-related protein kinases
ng	Nanogram
NH_4Cl	Ammonium Chloride
NIMA	Never in mitosis/Aspergillus
NLS	Nuclear Localisation Sequence
nm	Nanometre
nM	Nanomolar
NOT	Negative on TATA
N-terminal	Amino terminal
ORF	Open Reading Frame
*Os*TIR1	*Oryza sativa* Transport Inhibitor Response 1
p230p	paralogue to the P230 6-cysteine male gametocyte surface protein
PABP	Poly(A)-binding protein
PAGE	Polyacrylamide Gel Electrophoresis
PAM	Photospacer Adjacent Motif
PBE	Puf-binding element
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
Pen-Strep	Penicillin Streptomycin
PIC	Pre-Initiation Complex

PIL	Procedure Individual Licence
pir	*Plasmodium* interspersed repeat
PKG	Protein Kinase G
PLA	Phospholipase A
PLC	Phospholipase C
pmol	Picomole
PPL	Procedure Project Licence
PREBP	Prx regulatory element binding protein
Puf	Pumilio and fem-3 binding factor
PVM	Parasitophorous Vacuolar Membrane
RACK1	Receptor for activated C kinase 1
RBC	Red Blood Cell
rDNA	Ribosomal DNA
RFP	Red Fluorescent Protein
rif	Repetitive interspersed family
RMV	Infected red blood cell-derived microvesicle
RNA	Ribonucleic Acid
RNAse	RNA-specific endonuclease
RPMI-1640	Roswell Park Memorial Park medium 1640
s	Seconds
SAP	Sporozoite asparagine-rich protein
SAR	Stramenopiles, Alveolates, and Rhizaria
SCF	Skp1, Cullin and F-box complex
SD	Standard Deviation
SDS	Sodium Dodecyl Sulphate
SET	Su(var)3-9, Enhancer of zeste, and Trithorax
SIP2	SPE2-interacting protein
SIR2A	Silent Information Regulator 2A
SIR2B	Silent Information Regulator 2B
Sf9	Spodoptera frugiperda clonal line 9
SNF2	Sucrose non-fermentation 2
TBB	Transmission bite back
TBE	Tris-Borate containing EDTA
TBP	'TATA-box'-binding protein
TE	Tris containing EDTA

TEMED	N,N,N′,N′-Tetramethylethylenediamine
tetO	Tetracycline operator
TetR	Tetracycline-repressible
TIR1	Transport Inhibitor Response 1
TM	Transmembrane
TmcA	tRNA(met) cytidine Acetyltransferase
TO	Theilers Original
tracrRNA	Trans-activating crRNA
TRAD	Tetracycline repressor-activating domain
TRAP	Thrombospondin-related anonymous protein
TSS	Transcription Start Site
UI	Uncertainty Interval
UTR	Untranslated Region
V	Volt
v/v	Volume per volume
w/v	Weight per volume
WHO	World Health Organisation
WT	Wild-type
XA	Xanthurenic Acid
yDHODH	yeast dihydroorotate dehydrogenase
yfcu	Uridyl Phosphoribosyl Transferase

1. General Introduction

1.1 Malaria

Malaria is a term used to describe the intermittent, recurring fever caused by infection with protozoan parasites of the *Plasmodium* genus. In 2017 alone, an estimated 219 million people worldwide were infected with a *Plasmodium* parasite, resulting in illness. Of these infected individuals, approximately 435,000 died, 61% of these being in children under 5 years of age. By the end of 2017, 90 countries and territories had ongoing malaria transmission, down from 108 countries in 2000. The majority of malaria cases (80%), were concentrated in 17 countries of the WHO African Region and India, with 53% of all deaths occurring in the 7 countries of Nigeria, Democratic Republic of the Congo, Burkina Faso, India, United Republic of Tanzania, Niger, and Sierra Leone. (World Malaria Report 2018, World Health Organisation, 2018) **(Figure 1.1).**

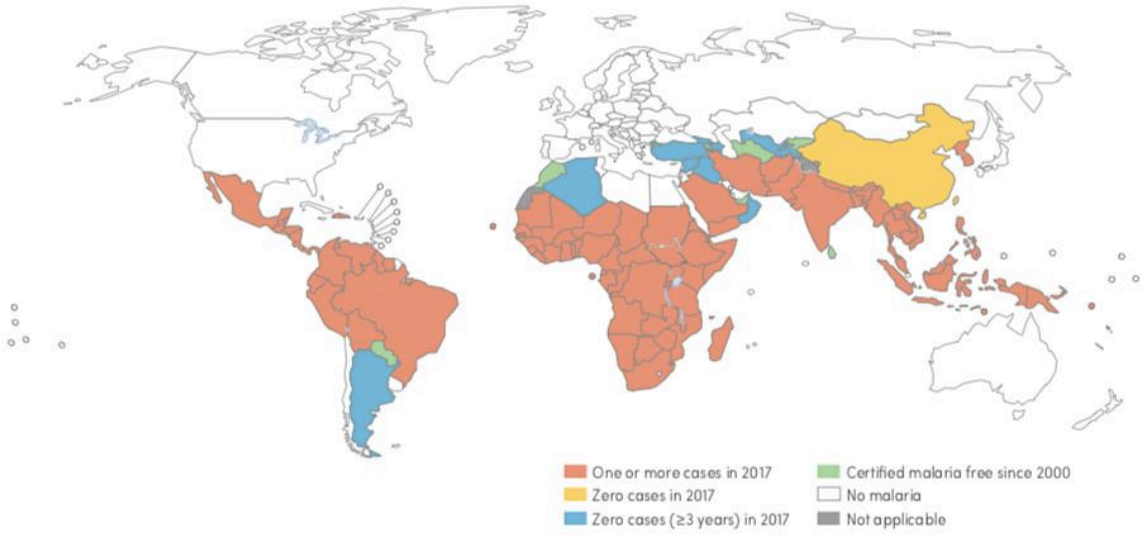


Figure 1.1: Countries and territories with indigenous cases of malaria in 2000 and 2017. According to the WHO World Malaria Report 2018, malaria was considered endemic to 90 countries worldwide (shown in orange and yellow), compared to 108 countries in the year 2000 (shown in orange, yellow, blue, and green). Countries with zero indigenous cases of malaria over a period of at least the past three consecutive years are eligible to request a status of "malaria-free' from the WHO. In this figure, countries shown in green have been awarded 'malaria-free' status since the year 2000. Countries shown in blue in this image are either awaiting certification from the WHO, or have not requested certification despite zero indigenous cases having been recorded in the past three consecutive years. Source: Image reproduced from the World Malaria Report 2018 (WHO, 2018).

1.2 *Plasmodium*

Plasmodium is a genus of parasitic protozoans (single-celled microscopic organisms of the *Protista* kingdom) in the phylum, Apicomplexa. Protists are an informal group of diverse eukaryotic organisms, i.e. those members of the Eukaryota taxon that are not animals, plants or fungi. Outside of animals, plants and fungi, are five 'protist' super-groups: Amoebozoa, Opisthokonta, Excavata, Archaeplastida, and 'SAR' (Stramenopiles (heterokonts), Alveolates, and Rhizaria). *Plasmodium* falls under the SAR super-group, within the Apicomplexa phylum (Adl *et al.*, 2012).

This large phylum of obligate intracellular parasites is distinguished from other Alveolata by the presence of a unique organelle, called an apicoplast; and an apical complex structure, consisting of the anterior structures of the parasite cell: the rhoptries, micronemes, conoid, subpellicular tubules, and polar ring (Levine, 1988). The apicoplast is a small organelle, a vestigial plastid, that is derived evolutionarily from a chloroplast; a remnant to past endosymbiosis between the ancestral Apicomplexan (debated as being of red algal origin) and a photosynthetic organism, the identity of which has yet to be discovered (Fast *et al.*, 2001; Funes *et al.*, 2002.; Janouskovec *et al.*, 2010). The most probable purpose of the apical complex in *Plasmodium* species and other Apicomplexa is to facilitate the invasion of host cells, most probably by actin-based gliding motility; rhoptries and micronemes are specialised secretory organelles that secrete proteins involved in adhesion during invasion and formation of the parasitophorous vacuole; the conoid is a cone-shaped spiral of filaments thought to play a role in mechanical invasion in some Apicomplexans (not present in *Plasmodium* or *Theileria*); the apical polar ring is a structural 'microtubule-organising centre' (MTOC) to which subpellicular tubules are laterally attached, giving motile Apicomplexan 'zoites' their elongated and polarised shape (Russell & Burns, 1984; Morrissette & Sibley, 2002). **(Figure 1.2).**

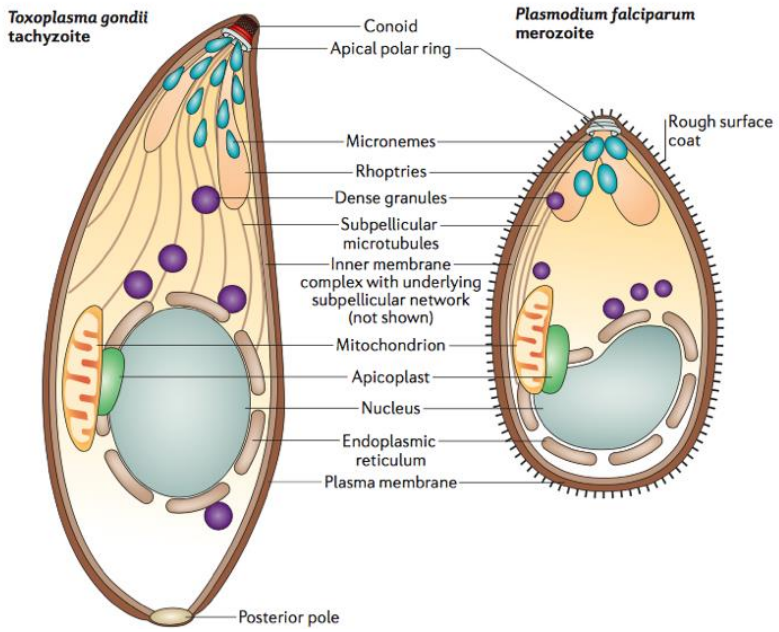


Figure 1.2: The apicoplast and apical complex structure of the Apicomplexa, using the *Toxoplasma gondii* tachyzoite and *Plasmodium falciparum* merozoite as examples. This comparison of two 'zoites' of the Apicomplexa phylum clearly shows placement of the apical complex structure elements (the micronemes, rhoptries, subpellicular tubules, and apical polar ring), in addition to the apicoplast in the motile *Toxoplasma* tachyzoite and *Plasmodium falciparum* merozoite. Note the absence of a 'conoid' in the *Plasmodium* merozoite. Source: (Baum *et al.*, 2006).

The phylum, Apicomplexa, consists of approximately 5,000 species, with 60,000 more yet to be named. Within this phylum are many genera of obligate endoparasites of medical importance, among others: *Babesia, Toxoplasma, Cryptosporidium, Eimeria,* and *Plasmodium*. With comparative analyses of small subunit rRNA genes, it has been concluded that the *Plasmodium* lineage of Apicomplexa diverged from its nearest relative parasites several hundred million years ago, perhaps even before the Cambrian Period (543 to 485.4 million years ago) during which it is hypothesised that vertebrates (chordates) emerged from their ancestral invertebrate lineage (Escalante & Ayala, 1995; Knoll *et al.*, 2004; Morrison & Ellis, 1997).

Plasmodium is a member of the order Haemosporida (also known as Haemosporidia), a diverse group of parasites with heteroxenous life cycles, alternating between a vertebrate host and a Dipteran insect host that acts as a vector for parasite transmission (Martinsen *et al.*, 2008). *Plasmodium* share many

traits with other haemosporidians: they are obligate vertebrate endoparasites with a complex life cycle that involves a blood-feeding Dipteran host. All haemosporidians use host erythrocyte haemoglobin as a major nutrient source, though members of this group often store haemoglobin degradation products in a different manner; a morphological characteristic often used to separate and define species. Haemosporidians can also be differentiated by variations in their complex life cycles, and by their host and insect vector specificities (Goldberg *et al.*, 1990; Blanquart & Gascuel, 2011). The *Plasmodium* genus contains approximately 200 species, infecting a range of vertebrate hosts, including bats, rodents, lizards, birds, non-human primates, and humans (Hayakawa, *et al.*, 2008; Duval *et al.*, 2010; Schaer *et al.*, 2013). The *Plasmodium* genus itself is split into a number of subgenera, with groups based on similarities in both morphology and host preference **(Figure 1.3).**

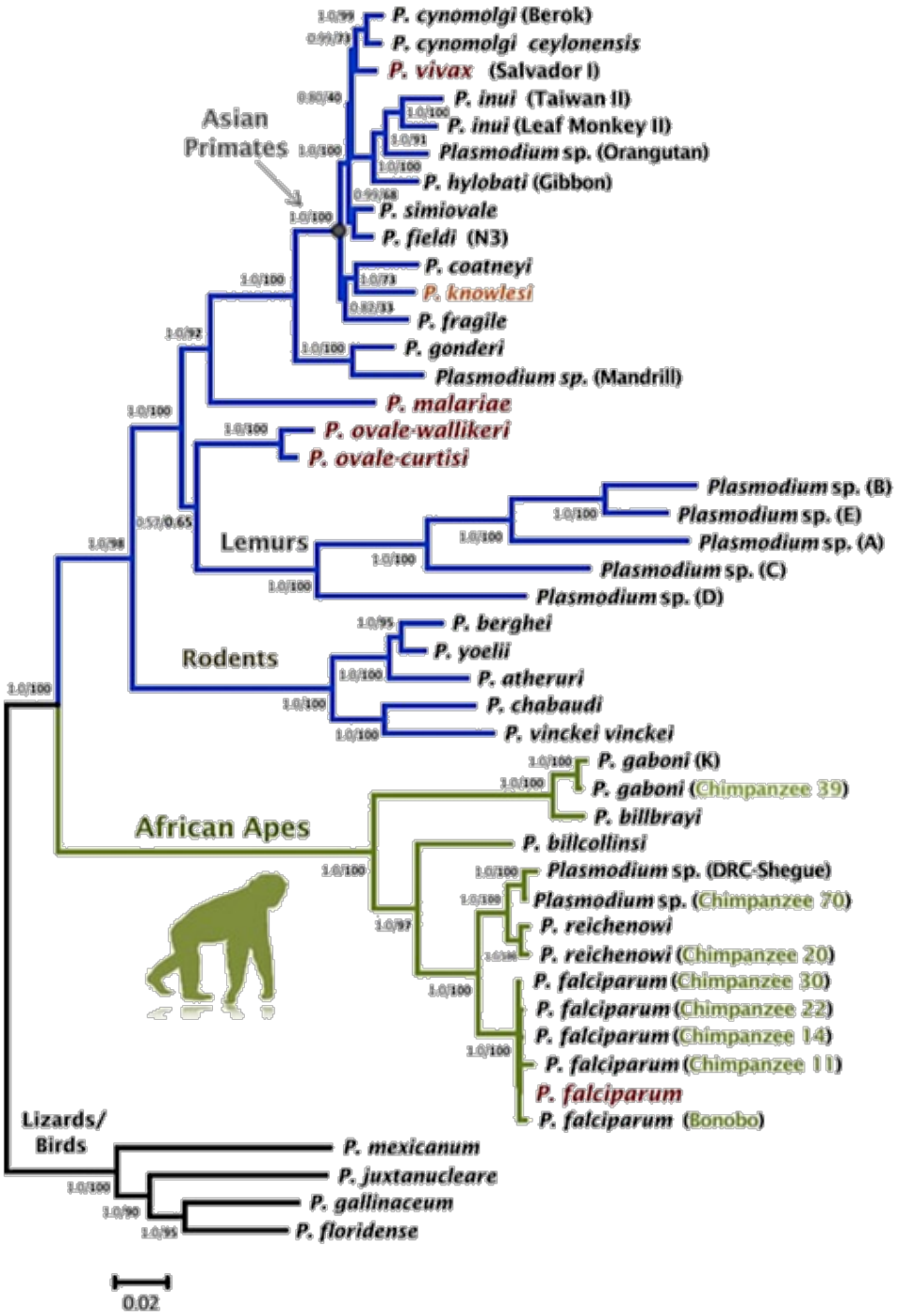


Figure 1.3: Phylogenetic tree of *Plasmodium* species. This image shows a Bayesian phylogenetic tree of *Plasmodium* spp. based on complete mitochondrial genomes taken as part of routine chimpanzee (*Pan troglodytes*) health examinations at the Jane Goodall Institute (JGI) Tchimpounga Chimpanzee Rehabilitation Center in the Republic of Congo (RC), in combination with other *Plasmodium* species' sequences available from GenBank. Values above branches are posterior probabilities together with bootstrap values (in bold text) as a percentage obtained for a phylogenetic tree using Maximum Likelihood (ML) methods. Human *Plasmodium* spp. are shown in red text (orange indicates both human- and primate- infective) and chimpanzee data is indicated in green text. Green branches indicate African ape *Plasmodium* species; blue branches correspond to other mammal *Plasmodium* species; and black branches indicate non-mammalian *Plasmodium* species. Source: (Pacheco *et al.*, 2013).

Of these *Plasmodium* subgenera, it is only one member of the *Laverania* subgenus, *Plasmodium falciparum*, and 4 members of the *Plasmodium* subgenus: *Plasmodium vivax, Plasmodium ovale, Plasmodium malariae*, and *Plasmodium knowlesi* that are responsible for malarial illness in humans. In recent times however, further research has determined that in fact *P. ovale* consists of two distinct but sympatric subspecies, *P. ovale curtisi* and *P. ovale wallikeri*, both of which occur throughout Africa and Asia-Pacific (Hay *et al.*, 2004; Mueller, Zimmerman, & Reeder, 2007; Sutherland *et al.*, 2010). In addition, it has been postulated that six further species of non-human primate malaria are transmissible to humans via insect bite: *Plasmodium cynomolgi, Plasmodium brasilianum, Plasmodium eylesi, Plasmodium inui, Plasmodium simium,* and *Plasmodium schwetzi*. In 2014, the first case of naturally acquired human infection with *P. cynomolgi* was confirmed in Malaysia, highlighting the importance of PCR diagnosis in determining malaria parasite species in humans, as opposed to diagnosis based purely on blood film microscopy, as morphology between *Plasmodium* species can be ambiguous (Ta *et al.*, 2014)

1.3 The *Plasmodium* life-cycle

1.3.1 Sporozoite infection and pre-erythrocytic development

Plasmodium parasites are obligate endoparasites that alternate between a vertebrate host and a Dipteran insect vector (Cox, 2010). The duration of time spent at each developmental stage of the *Plasmodium* life-cycle differs between species, though a *Plasmodium* mammalian infection always begins at the same point: infection into the vertebrate skin with *Plasmodium* sporozoites from the insect vector salivary glands, following by migration of sporozoites through the bloodstream, to the liver sinusoids (Kappe *et al.*, 2004; Medica & Sinnis, 2005). When the *Plasmodium* sporozoite reaches the liver, it traverses a number of hepatocytes, often permanently wounding or killing Kupffer cells within the liver parenchyma, before selecting a single hepatocyte within which a parasitophorous vacuole is formed to establish sporozoite residence (Mota *et al.*, 2001; Tavares *et al.*, 2013). Selection of a single host hepatocyte within which to begin schizogony is a choice that is still not understood. One hypothesis is that invasion begins with sporozoite interaction with the hepatocyte-of-choice via the surface circumsporozoite protein, which is cleaved upon contact with highly sulphated proteoglycans at the hepatocyte surface (Coppi *et al.*, 2011). However, the mechanisms by which this cleavage initiates formation of a parasitophorous vacuole remains to be seen (Kaushansky and Kappe, 2015).

Once hepatocyte invasion by the Plasmodium sporozoite has occurred, this parasite develops into an exo-erythrocytic form (EEF), also known as the liver stage (LS), within an intrahepatic parasitophorous vacuole. This EEF, or LS, of the malaria parasite expresses a unique set of virulence proteins that it shares with the sporozoite; a set of proteins necessary for hepatocyte invasion that are the subject of vaccine development research (Jaijyan, Singh and Singh, 2015). In *P. falciparum,* a single intrahepatic sporozoite grows into a liver-stage trophozoite (1 nucleus (1n) to 2 nuclei (2n)) before undergoing 13 to 14 rounds of DNA synthesis, mitosis, and nuclear division, to produce a schizont containing approximately 10,000 to 30,000 merozoites (Singh *et al.*, 2007; Gerald, Mahajan and Kumar, 2011). At 2-16 days post-invasion of the hepatocyte (depending on *Plasmodium* species), the parasitophorous vacuolar membrane (PVM) is breached,

with the release of thousands of free merozoites, and merozoites that are encased in membrane-bound vesicles devoid of cell nuclei, a parasite-filled vesicle known as a merosome (Sturm *et al.*, 2006). Little is known about how the PVM is ruptured at the liver stage, though it has been demonstrated that a *Plasmodium* phospholipase (labelled PbPL and encoded by gene PBANKA_1128100 in *P. berghei*) plays a role in PVM disruption (Burda *et al.*, 2015) **(Figure 1.4).**

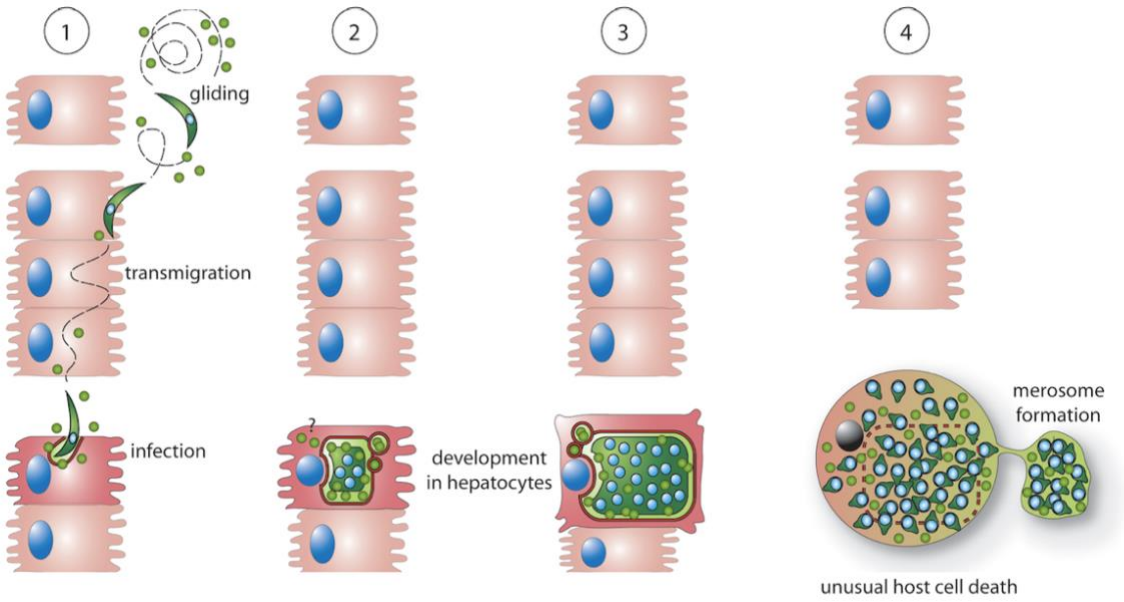


Figure 1.4: Cartoon showing *Plasmodium* sporozoite infection, invasion, and development to a merozoite or merosome, based on current knowledge of these life-cycle stages. (1) depicts invasion of a host hepatocyte by a *Plasmodium* sporozoite (in this case, *P. berghei*). As the sporozoite moves, it excretes a potent *P. berghei* inhibitor of cysteine proteases (PbICP) that is thought to play a role in hepatocyte invasion, as well as sporozoite egress from mosquito salivary glands, and liberation of the exo-erythrocytic merozoites from the liver. In (2) and (3), schizogony occurs. The question mark in (2) indicates the possibility of PbICP in the host cell cytoplasm. At present, it is postulated that the majority of PbICP is confined to the parasite cytosol and parasitophorous vacuole. In (4), schizogony is completed and a merosome emerges from the infected host cell, with free merozoites (presumably released upon PVM rupture) still inside the schizont. The host cell nucleus is depicted as black at this point to indicate the possible activity of PbICP in inhibition of host cell cysteine proteases, to allow for a well-timed host cell death without disruption to merozoite and merosome formation. Source: (Rennenberg *et al.*, 2010).

1.3.2 Host erythrocyte invasion and tropism

In the absence of hypnozoite formation, merosomes are released and traverse through the liver sinusoids, protecting merozoites from phagocytosis before they reach the host bloodstream. Each merozoite is then free to invade a host

erythrocyte. This process marks the beginning of the recurrent malaria parasite infection (Prudêncio *et al.*, 2006). The process of red blood cell recognition, attachment, and invasion of host erythrocytes is postulated to take mere seconds, as the merozoite surface antigens are particular susceptible to immune attack (Cowman & Crabb, 2006). *Plasmodium* species also differ in age preference of erythrocytes to which they attach and invade. A merozoite of *P. falciparum* is capable of invading erythrocytes at all stages of maturation, while *P. vivax* and *P. ovale* species show a marked preference for the invasion of young erythrocytes (reticulocytes), and *P. malariae* for mature erythrocytes (Kerlin *et al.*, 2013).

Primary *Plasmodium* merozoite attachment is thought to be mediated by glycosylphosphatidyl inositol (GPI)-anchored merozoite surface proteins (MSPs) located on the 'fuzzy' fibrillar merozoite surface coat. Numerous GPI-anchored MSPs have so far been characterised in *P. falciparum*: MSP1, MSP2, MSP4, MSP5, MSP10, and 6-cysteine domain family proteins, Pf92, Pf38, and Pf12 (Sanders *et al.*, 2005). These proteins, in combination with other GPI-anchored proteins such as rhoptry-associated membrane antigen (RAMA) and GPI-anchored micronemal antigen (GAMA), anchor the merozoite to the erythrocyte surface and initiate tight junction formation (Beeson *et al.*, 2016; Arumugam *et al.*, 2011; Topolska *et al.*, 2004; Lamarque *et al.*, 2011). From here, the attached merozoite traverses the erythrocyte membrane via a conserved actin-myosin motor complex (Baum *et al.*, 2006).

Once merozoite invasion has occurred, the infected erythrocyte undergoes brief echinocytosis (red blood cell deformation characterised by 'thorny' membrane protrusions), before normal shape is resumed and the parasite, now referred to as a 'ring' stage, has been completely internalised (Gilson & Crabb, 2009). Within the erythrocyte, this ring stage parasite is encased within a parasitophorous vacuolar membrane and modifies host cell permeability to provide adequate nutrition for growth and development to a trophozoite (Maier *et al.*, 2009). The trophozoite stage of the *Plasmodium* life-cycle is characterised by rapid growth and the appearance of non-toxic haemozoin, sometimes referred to as 'malaria pigment'. Haemozoin is an inorganic and insoluble crystal containing haem (or 'heme') that is released from the parasite digestive vacuole upon haemoglobin degradation (Bakar *et al.*, 2010). The highly permeabilised erythrocyte membrane

is prevented from haemolysis during the growth of the trophozoite by the decline in intracellular haemoglobin, reducing colloid osmotic pressure, and therefore reducing swelling of the host red blood cell (Lew *et al.*, 2003).

The final intraerythrocytic stage of the *Plasmodium* parasite is the schizont stage, during which the intracellular parasite generates up to 32 daughter merozoites. It is at this final stage that the infected red blood cell will rupture to release free merozoites into the host blood stream. These merozoites in turn will reinvade more host erythrocytes, continuing the asexual cycle of the malaria parasite (Grüring *et al.*, 2011). The pathogenicity of *Plasmodium* parasites is often a result of the sequestration of trophozoites and schizonts to the endothelia of small blood vessels, causing obstruction of host tissue perfusion. The cycles of host red blood cell rupture may also manifest clinically in a *Plasmodium*-infected human as anaemia (Miller *et al.*, 2002; Dondorp *et al.*, 2000). Severe anaemia and the blockage of microvasculature by infected red blood cells in malaria patients can hinder tissue perfusion and oxygen delivery, leading to metabolic acidosis and respiratory distress (Sasi *et al.*, 2007; Olliaro, 2008). Frontline antimalarial drugs primarily target the asexual cycle of *Plasmodium* parasites to alleviate clinical symptoms of infection. However, despite this, cerebral malaria and metabolic acidosis carry a mortality rate of 15-20%, with survivors often suffering long-term neurological sequelae (Miller *et al.*, 2013).

1.3.3 *Plasmodium* sexual development: Gametocytogenesis

From invasion of the sporozoite, through *Plasmodium* development in the host liver, to intraerythrocytic parasite growth; all of these stages of the malaria parasite life-cycle undergo asexual replication. Though these stages are responsible for morbidity and mortality in the *Plasmodium* host, it is only with sexual development of the parasite, a process known as gametocytogenesis, that transmission from a vertebrate host to a mosquito vector can occur, therefore ensuring the spread of the malaria parasite from one host to another (Baker, 2010). Gametocytogenesis is the process by which a male or female sexual-stage *Plasmodium* parasite (a gametocyte) is produced. This process takes approximately 10-12 days in *P. falciparum* and is a process marked by molecular

mechanisms that are distinct from parasites of the asexually replicating stages (Josling, Williamson and Llinás, 2018) **(Figure 1.5).**

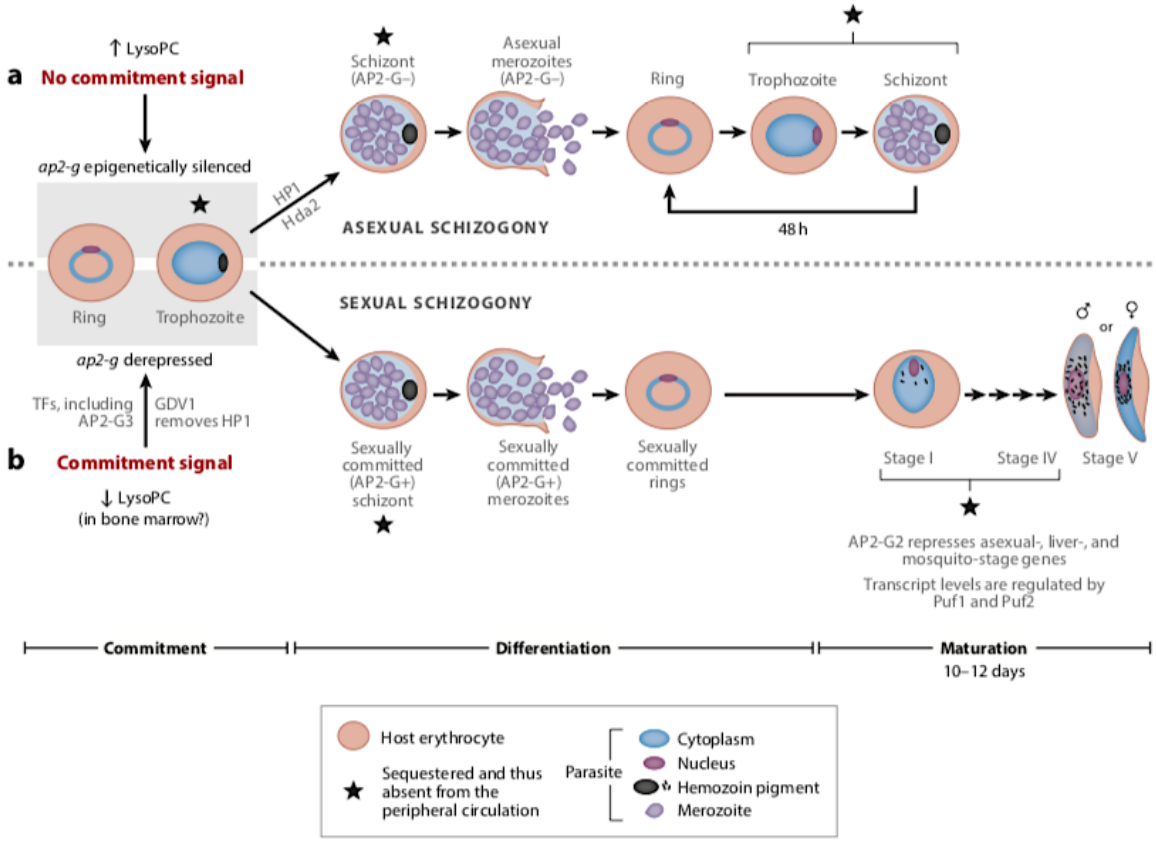


Figure 1.5: Commitment and maturation of the gametocytes of *P. falciparum*. The images above depict the life-cycle of a *P. falciparum* parasite about to undergo asexual replication to produce exclusively asexually-committed parasites **(a),** and the separate life-cycle progression of a parasite that has become committed to producing only sexually-committed merozoites **(b).** At present, it is unknown whether commitment to gametocytogenesis is pre-determined at the parasite 'ring' stage, or whether environmental factors during ring-to-trophozoite development are responsible for gametocyte conversion in the second reproductive cycle. In **(a),** an abundance of lysophosphatidylcholine (LysoPC) and the repression of *ap2-g* result in downstream commitment to asexual development. In parasites without *ap2-g* expression (AP2-G-), alterations in the activities of both heterochromatin protein 1 (HP1) and histone deacetylase 2 (Hda2) ensure that the trophozoite develops into a schizont in which all resulting merozoites will invade new host erythrocytes to continue the asexual proliferation of *Plasmodium* parasites. In **(b),** in the absence of LysoPC (possibly in parasitised cells sequestered in the host bone marrow), the ring or early trophozoite becomes committed to gametocytogenesis. This process is also marked by alterations in the expression of transcription factors (TFs) such as AP2-G3 and, in *P. falciparum* and a number of other human malaria parasite spp., by the removal of HP1 from genomic loci by gametocyte development protein 1 (GDV1). This parasite then undergoes what has been termed above as 'sexual schizogony', with all merozoites of this particular schizont producing either male or female gametocytes exclusively. In gametocytes, AP2-G is expressed (AP2-G+ cells) and the maturation of the gametocyte occurs (taking 10-12 days in *P. falciparum*). Source: (Josling, Williamson and Llinás, 2018).

1.3.3.1 Species-specific gametocyte morphologies

Different species of the *Plasmodium* genus differ in their gametocyte morphologies. In *P. falciparum*, gametocytogenesis is divided into five stages (stages I-V) that can be distinguished by their morphology (Carter & Miller, 1979) **(Figure 1.5)**. At stage I, 24 h to 30 h post-invasion, the *P. falciparum* gametocyte is indistinguishable from a mononucleated asexual trophozoite. As stage II of gametocyte development begins (Day 2), the parasite cell can be seen to elongate and form subpellicular microtubules, with an accumulation of F-actin at opposite ends of the parasite periphery, elongating the cell into a half-moon, or D-shaped cell (Hliscs *et al.*, 2015; Tibú Rcio *et al.*, 2015).

From approximately Day 4 to Day 8 of gametocyte development in *P. falciparum*, the parasite passes through gametocyte stages III-IV. Stage III gametocytes can be distinguished by rounded ends and further elongation of the cell, with more noticeable subpellicular tubule development. At stage IV, the *P. falciparum* gametocyte has pointed ends and is slightly curved, with the host erythrocyte reduced to a thin layer around the edge of the parasitophorous vacuole. This crescent or sickle-shaped form of the *P. falciparum* parasite is the form from which the species was named (the Latin word '*falx*' meaning 'sickle', and noun '*parere*' meaning 'to give birth') (Josling & Llinas, 2015; Dixon *et al.*, 2012).

From stages III-IV in *P. falciparum*, discrimination between male and female gametocytes is not possible by microscopy. However, from late stage IV onwards, it has been shown that discerning gametocyte sex is possible prior to induction of male gametocyte exflagellation and female gamete emergence by light microscopy and the use of antibodies recognising protein *Pf*g377 (primary marker) and α-tubulin II (secondary marker). Using the well-characterised female-specific *Pf*g377 protein (encoded by gene PFL2405c) and male-specific α-tubulin II (gene PFD1050w), male/female sex ratios could be derived, with *Pf*g377 recognising female gametocytes and α-tubulin II showing decreased abundance in female gametocytes relative to males (Schwank *et al.*, 2010; Severini *et al.*, 1999; Rawlings *et al.*, 1992).

At Days 10-14 of gametocyte development, mature stage V male and female *P. falciparum* gametocytes are permissible to uptake from host peripheral blood to the mosquito vector, though the optimum day at which infectivity is at its peak may be specific to *P. falciparum* strains (Lensen *et al.*, 1999). Stage V *P. falciparum* gametocytes are readily identifiable as male or female. Stage V female gametocytes are more elongated and curved, whereas mature male gametocytes tend to be straight, shorter and thicker than female gametocytes (Josling & Llinas, 2015) **(Figure 1.6).**

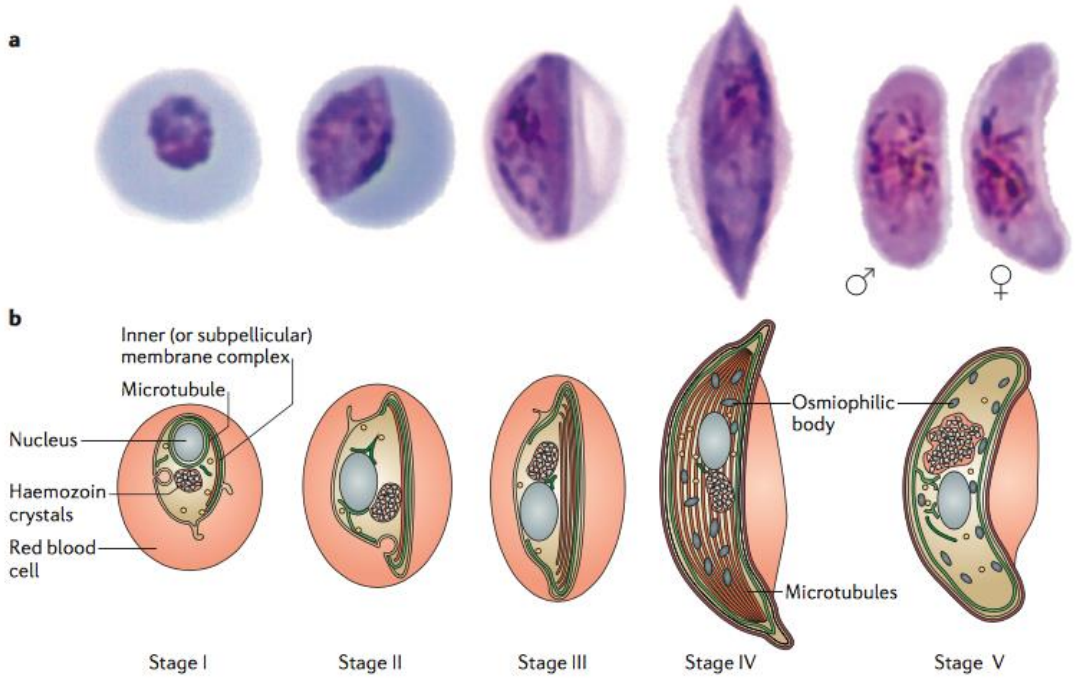


Figure 1.6: Stages I-V of gametocyte development in *P. falciparum*. Row (**a**) shows *P. falciparum* (3D7 strain) gametocytes at stages I-V of development, corresponding to the cartoon depictions in row (**b**). Photographs also enable observation of the different morphologies between female and male gametocytes at stage V. Row (**b**) depicts stages I-V of *P. falciparum* gametocytogenesis. Stage I is indistinguishable from trophozoites of the same age. However, stage II gametocytes become distinguishable from asexual parasites as a result of their elongation, becoming D-shaped and forming subpellicular tubules. Stage III gametocytes are further elongated, with visible subpellicular tubules. At stage III, the gametocyte body becomes visibly larger than the remaining host cell cytosol. At stage IV, gametocytes have lost their rounded ends; showing pointed ends and elongation to the point that only a sliver of host cell cytoplasm can be discerned. At stage IV, an abundance of osmiophilic bodies and extensive rough endoplasmic reticulum become visible in female gametocytes (not shown; only a few osmiophilic bodies are shown in this diagram). At stage V, *P. falciparum* male and female gametocytes have their characteristic crescent shape. Male and female gametocytes can be differentiated morphologically, with males being straighter and thicker. Not shown is the fact that at stage V, female and male gametocytes of *P. falciparum* can

be differentiated by Giemsa stain, with females staining blue, and males staining pinker in colour (Josling & Llinas, 2015). Source: (Josling & Llinas, 2015).

Outside of *P. falciparum*, much less is understood about gametocyte development in human-infective *Plasmodium* species that are not part of the *Laverania* subgenus. In *Plasmodium vivax*, considered to be the most widely distributed malaria parasite throughout temperate and tropical climates, difficulties in maintaining continuous cultures have severely hampered the ability of researchers to document changes in parasite biology throughout gametocytogenesis (Gething *et al.*, 2012; Roobsoong *et al.*, 2015). *P. vivax* parasites preferentially invade young red blood cells; reticulocytes expressing high levels of Cluster of Differentiation protein 71 (CD71) that are most abundant in host bone marrow (Martín-Jaular *et al.*, 2013). So far, it has been observed that *P. vivax* gametocytes do not exhibit distinct gametocyte morphological stages such as *P. falciparum* stages I-V. Both sexes of gametocytes are large and round (sometimes oval) and fill the entirety of the stippled host red blood cell. At a molecular level, little can be determined until advances in *P. vivax* in vitro culture has been established, possibly by usage of genetically modified erythrocytes that express *P. vivax* reticulocyte binding proteins on their cell surface (Bousema & Drakeley, 2011; Udomsangpetch *et al.*, 2008).

Even less is known about malaria parasite species, *P. knowlesi, P. malariae* and *P. ovale* (either subspecies). *P. knowlesi* has the shortest intraerythrocytic life-cycle of all five human-infective Plasmodium species, at 24 h. For *P. falciparum, P. vivax*, and *P. ovale*, this cycle lasts approximately 48 h; while in *P. malariae*, the cycle takes approximately 72 h (Garnham, 1966). In blood samples from human infections, *P. knowlesi* and *P. malariae* gametocytes appear similar in morphology to those of *P. vivax*: round, filling the host erythrocyte, and with both sexes indistinguishable from one another (Singh & Daneshvar, 2013). For *P. ovale curtisi* and *P. ovale wallikeri*; gametocytes from a human *P. ovale* infection (subspecies unknown) showed round or oval morphologies, with immature gametocytes showing more pointed ends. All *P. ovale* gametocytes filled the host erythrocyte cell, with the host cell again stippled with darkened dots (most likely haemozoin crystals) (Garnham, 1966; Ribaut *et al.*, 2008). Again, sex determination at a

molecular level between *P. ovale* gametocyte males and females is not yet possible. **(Figure 1.7).**

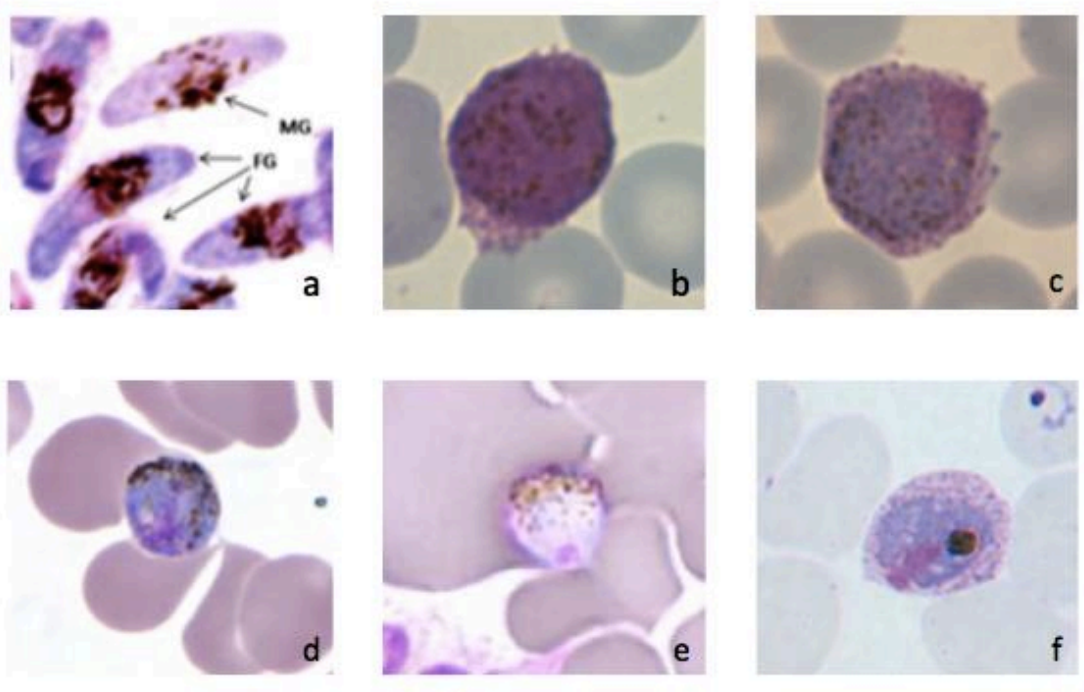


Figure 1.7: Comparison of gametocyte stage morphology between human-infective *Plasmodium* species. (a) shows a Giemsa stain of a *P. falciparum* male gametocyte (MG) next to *P. falciparum* female gametocytes (FGs) (adapted from (Schwank *et al.*, 2010)). **(b)** and **(c)** show Giemsa stains of male and female *P. vivax* gametocytes respectively (both adapted from (Bousema & Drakeley, 2011)). **(d)** depicts a Giemsa stain of a *P. malariae* gametocyte (sex undetermined) (adapted from the Center for Disease Control (CDC) DPDx Laboratory Guidelines (CDC, 2017)). **(e)** shows a Giemsa stain of a *P. ovale* (subspecies unknown) gametocyte of undetermined sex (adapted from (Ribaut *et al.*, 2008)). **(f)** shows a Giemsa stain of a *P. knowlesi* gametocyte of undetermined sex (adapted from (Lee, Cox-Singh & Singh, 2009)).

1.3.3.2 Molecular mechanisms of sexual development in *Plasmodium*

The exact moment at which a ring-stage or early trophozoite-stage *Plasmodium* parasite commits to sexual development in any species of *Plasmodium*, is still unknown and the subject of much research (Josling and Llinas, 2015; Josling, Williamson and Llinás, 2018). However, in 2014, it was demonstrated that the production of gametocytes in *Plasmodium* (in these studies, *P. berghei* and *P. falciparum* were used), is dependent upon the expression of a conserved member of the Apetala 2 (AP2) family of apicomplexan DNA-binding proteins, AP2-gametocyte (AP2-G). In *P. berghei*, mutations in the *ap2-g* gene

(PBANKA_1437500) generated parasite lines unable to produce male or female gametocytes, with repair of these mutations reversing this gametocyte non-producer phenotype (Rovira-Graells *et al.*, 2012; Kafsack *et al.*, 2014; Sinha *et al.*, 2014).

In a recent study, the role of gametocyte development 1 (GDV1) protein in *P. falciparum* as an antagonist of heterochromatin protein 1 (HP1)-mediated silencing of *pfap2-g* has been demonstrated (Filarsky *et al.*, 2018). Though undoubtedly an upstream regulator of sexual commitment in *P. falciparum* in this manner, the rodent malaria parasite clade (encompassing *P. berghei, P. chaboudi, P. yoelii* and *P. vinckei*) does not possess a GDV1 gene, therefore suggesting a divergence in the regulation of sexual development in these *Plasmodium* species.

Despite these milestones in our understanding of *Plasmodium* sexual development, there is still much that we don't understand. In a natural *P. falciparum* infection, the gametocyte to asexual stage parasite ratio is less than 1:10, and has been documented as low as 1:156 (Eichner *et al.*, 2001; Talman *et al.*, 2004; Sinden, 1983). *In P. falciparum*, it has been demonstrated that sexual development choice for merozoites of a single schizont are mutually exclusive, with one pre-committed schizont generating either sexual or asexual stage parasites, and not the other (Bruce *et al.*, 1990).

Individual pre-committed *P. falciparum* schizonts are also mutually exclusive in their generation of male or female gametocyte populations. Each schizont that is pre-conditioned to commit to gametocytogenesis will produce progeny of the same sex (Silvestrini, Alano, & Williams, 2000). Of all pre-committed schizonts, 67-71% will produce female progeny, with this female-biased sex ratio remaining even under field conditions (Smith *et al.*, 2000; Robert *et al.*, 2003; Gbotosho *et al.*, 2011). At present, it is still unknown at which stage a *Plasmodium* parasite commits to sexual development, the complete array of environmental conditions underlying this choice, or how these conditions dictate male/female sex determination.

Despite the many still-unanswered questions regarding commitment to, and initiation of, *Plasmodium* gametocytogenesis, implementation of a novel flow-

cytometry-based assay for the examination of early gametocyte development in *P. falciparum* (Brancucci *et al.*, 2015) led to the identification of one environmental factor responsible for regulating sexual development upstream of *ap2-g* transcription. The exogenous, host-derived lipid, lysophosphatidylcholine ('LysoPC'), was identified as a serum factor responsible for the reversible repression of gametocyte commitment in *P. falciparum* (Brancucci *et al.*, 2017). Exposing *P. falciparum* parasites to 'parasite-conditioned' medium, as opposed to freshly-prepared culture medium, had previously been shown to increase gametocyte numbers by approximately 5.9-fold to 7.6-fold (Williams, 1999), with further experiments demonstrating that exposing *P. falciparum* parasites to medium that was free of human serum (mimicking restricted host factor availability *in vivo*) also resulted in the induction of sexual commitment *in vitro*. Ultimately, LysoPC was determined an environmental repressor of gametocytogenesis, inhibiting sexual development at a 50% inhibitory concentration of 1.73 μM *in vitro* in both Pf2004 (derived from NF54) and HB3 strains of *P. falciparum* (Brancucci *et al.*, 2017).

1.3.4 *Plasmodium* development within the mosquito

At the end of the *Plasmodium* gametocyte maturation process, male gametocytes (microgametocytes) and female gametocytes (macrogametocytes) are amenable to uptake by female mosquitoes during a blood meal (Paul *et al.*, 2002). A female-biased sex ratio, with the generation of one microgametocyte to every five macrogametocytes results in an approximate 1:1 ratio of male to female gametes in the mosquito midgut. This is because each microgametocyte produces 8 microgametes upon exflagellation within the mosquito midgut, with the female macrogametocyte differentiating to a single macrogamete (Reece *et al.*, 2008). Within seconds of entering the mosquito midgut, gametogenesis is induced by a drop in temperature to 5°C lower than that of the host vertebrate blood, and by a change in pH from 7.4 in the vertebrate to 8-8.2 in the Dipteran vector (Sinden *et al.*, 1996; Nijhout & Carter, 1978). A secondary requirement for gametogenesis has also been documented. At a lowered temperature, but in the absence of an increased pH, gametogenesis (of both sexes) can be stimulated by the addition of xanthurenic acid (XA) derived from crude insect homogenates. Previous studies in

gametocyte-to-gamete activation using insect secreted factors and homogenates have referred to this factor as mosquito exflagellation factor (MEF), and later, gametocyte activating factor (GAF). The activating factor involved has now been identified definitively as XA. (Nijhout, 1979; Billker *et al.*, 1998; Billker *et al.*, 1997).

Upon induction of gametogenesis, exflagellation of the male microgametocyte results in 8 motile flagellar microgametes that detach from the residual activated male gametocyte. Microgamete exflagellation and macrogamete emergence are both dependent upon XA, though so far, no ligand for XA has been identified on the *Plasmodium* gametocyte surface. In mammals, it has been suggested that XA acts through a G-protein coupled receptor (GPCR), coupled with a GPCR-induced cationic channel modulatory mechanism (Taleb *et al.*, 2012). So far, what we do know of microgamete exflagellation is that XA triggers an increase in intracellular calcium ion ($Ca2^{+}$) concentration that is in part mediated by calcium-dependent protein kinase 4 (CDPK4) (Billker *et al.*, 2004). Evidence for the role of second messengers phospholipase A (PLA), phospholipase C (PLC), and cyclic GMP has also been revealed; with inhibitors of PLA and PLC disrupting microgamete exflagellation, and agents that increase cGMP levels enhancing exflagellation (Ogwan'g *et al.*, 1993; Martin *et al.*, 1994; Kawamoto *et al.*, 1990). In fact, cGMP-dependent protein kinase, Protein Kinase G (PKG), has been shown as essential to the activation of microgamete exflagellation (McRobert *et al.*, 2008). In summary, *Plasmodium* microgamete exflagellation is in all probability, subject to both calcium-dependent and calcium-independent effector pathways, though the timing of activation of these pathways in relation to one other is so far not understood (Kuehn & Pradel, 2010).

Female gamete formation appears much less arduous when compared to microgamete exflagellation, though it is equally as important to malaria transmission. Macrogametes egress from the host erythrocyte upon uptake into the mosquito midgut, first by $Ca2^{+-}$independent rupture of the parasitophorous vacuolar membrane (PVM) at multiple sites, and followed by rupture of the erythrocyte membrane at a single point. This form of egress is shared between male and female gametocytes (Sologub *et al.*, 2011). During both male and female gametocyte egress, electron-dense secretory vesicles known as osmiophilic bodies

are observed to accumulate at rupture sites, often disintegrating simultaneously with the PVM. However, osmiophilic bodies are consistently observed in much greater number in female gametocytes and are associated with gametocyte-specific proteins *Pf*g377, PfMDV1/PfPEG3, and gamete egress and sporozoite traversal (GEST) protein (De Koning-Ward *et al.*, 2007; Talman *et al.*, 2011; Olivieri *et al.*, 2015; Lal *et al.*, 2009; Ponzi *et al.*, 2009). Upon rounding and emergence of the female macrogamete, translation of transcripts occurs which, until now, been translationally repressed by a messenger ribonucleoprotein (mRNP) silencing complex that includes both the DDX6 class RNA helicase, DOZI (development of zygote inhibited) and the Sm-like factor, CITH (homolog of worm CAR-I and fly Trailer Hitch) (Mair *et al.*, 2006; Mair *et al.*, 2010). This undoing of translational repression in *Plasmodium* macrogametes has been shown to be initiated by calcium-dependent protein kinase 1 (CDPK1) (Sebastian *et al.*, 2012).

Once fertile male and female gametocytes have emerged from the host erythrocyte into the mosquito midgut lumen, the motile microgamete must attach to the surface of a macrogamete to initiate zygote production. A number of proteins have been implicated in this attachment process, including gamete-associated Limulus coagulation factor C, Coch-5b2 and Lgl1 (LCCL)- domain-containing proteins (termed PfCCp proteins in *P. falciparum*), and proteins containing a six cysteine (6-Cys) structural domain, such as microgamete-associated protein P48/45 and P230, and macrogamete-associated P47 (Pradel *et al.*, 2004; Delrieu *et al.*, 2002; van Dijk *et al.*, 2001; van Dijk *et al.*, 2010; van Schaijk *et al.*, 2006). The Armadillo repeat domain-containing protein, PF16, has also been shown as essential for correct functioning of microgamete flagellar morphology via its role in axoneme assembly. In *Plasmodium* parasites, the axoneme is made up of central C1 and C2 microtubules, surrounded by nine doublet microtubules and is fundamental to motility of the male gamete (Straschil *et al.*, 2010).

Zygote formation in *Plasmodium* occurs upon fusion of the male and female gamete plasma membranes and uptake of the male nucleus and axoneme into the cytoplasm of the macrogamete. Gamete fusion is mediated in part by male-specific generative cell specific 1 (GCS1) protein (also referred to as plant sterility

protein HAP2), which nuclear fusion ensuing thereafter by meiosis (Janse *et al.*, 1986; Liu *et al.*, 2008; Hirai *et al.*, 2008). Two NIMA (Never in mitosis/Aspergillus)-related protein kinases (Neks), Nek-2 and Nek-4 are vital for *Plasmodium* gamete nuclear fusion and the parasite tetraploidy that precedes ookinete conversion (Guttery *et al.*, 2012; Reininger *et al.*, 2005; Reininger *et al.*, 2009).

1.3.5 The final stretch: From ookinete to sporozoite

Soon after nuclear fusion, the zygote in the mosquito midgut lumen matures into a motile ookinete, the development of which is directed by a plethora of ookinete-specific proteins necessary for midgut invasion, such as membrane-attack ookinete protein (MAOP) and chitinase (Vinetz *et al.*, 2000; Vlachou *et al.*, 2006; Kadota *et al.*, 2004). Expression of many genes responsible for ookinete development are, as in gametocyte formation, regulated by an AP2 DNA-binding factor; in this case named AP2-ookinete (AP2-O) (Yuda *et al.*, 2009). Once matured, motile ookinetes exit the mosquito midgut lumen, only to traverse it once more to facilitate oocyst formation. Mechanisms for ookinete midgut epithelium and peritrophic membrane traversal are controversial and require further elucidation (Baton & Ranford-Cartwright, 2005; Bennink, Kiesow, & Pradel, 2016).

The *Plasmodium* oocyst develops beneath the basal lamina of the midgut epithelium. Beneath the basal lamina of the mosquito midgut, a motile ookinete transforms into an immobile (or 'sessile') spherical oocyst via a proposed intermediate stage known as a 'took' stage, with oocyst development requiring only a minimal environmental stimulus, namely lowering of pH to 7.0 (Carter *et al.*, 2007). The production of the *Plasmodium* oocyst is the longest stage of the parasite's life-cycle, taking 10-15 days, depending on species. During this time, the oocyst grows from approximately 5 μm to 50-60 μm in size and undergoes multiple rounds of nuclear division to produce up to 8,000 haploid nuclei (Canning & Sinden, 1973). This period of nuclear division and sporozoite formation is called *sporogony*, with the oocyst cytoplasm subdivided by cleft formation into separate *sporoblasts*, though without loss of outer oocyst membrane integrity (Terzakis *et al.*, 1967; Nacer *et al.*, 2008).

Oocyst-derived sporozoites bud from sporoblast bodies once development has been completed. $1x10^3$ to $1x10^4$ sporozoites are released from each oocyst, which travel through the midgut to the salivary glands. This process of sporozoite invasion into the mosquito salivary glands is inefficient, with only 25% of sporozoites making it to the salivary glands; the rest remaining, and subsequently degrading, within the mosquito haemocoel (Baton & Ranford-Cartwright, 2005; Beier, 1998). Only a relatively small number of sporozoites will eventually be inoculated upon a mosquito blood meal of a host vertebrate, typically fewer than 100, which rapidly lose infectivity after leaving the mosquito (Pumpuni & Beier, 1995; Frischknecht *et al.*, 2004). Once inoculated into a vertebrate host, the *Plasmodium* cycle begins anew **(Figure 1.8).**

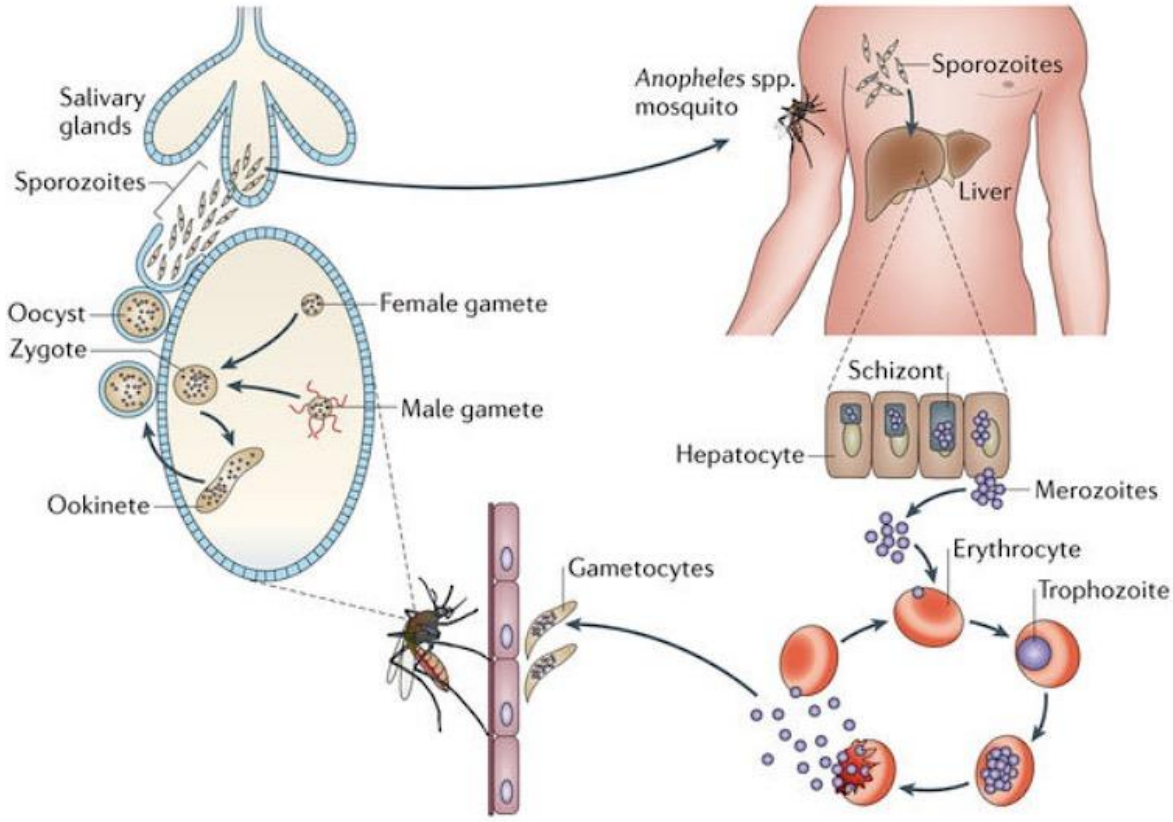


Figure 1.8: Malaria parasite development within the mosquito vector; from uptake of gametocytes to a new host infection. Transmission of malaria is facilitated by the uptake of mature *Plasmodium* gametocytes from the vertebrate host to a mosquito vector and activation of gametogenesis by environmental factors within the insect midgut. Zygote formation occurs upon nuclear fusion of a male microgamete with a female macrogamete, followed by development of the motile ookinete in the midgut lumen. Mature, motile ookinetes are capable of invading, and then traversing, the mosquito midgut epithelium whereupon an immobile oocyst can develop beneath the midgut basal lamina. The elongated ookinete becomes a spherical oocyst via a proposed intermediate 'took' stage. At oocyst stage, multiple rounds of nuclear division occur, producing thousands of haploid sporozoites. One oocyst subdivides into multiple sporoblasts,

followed by release of the sporozoites to the mosquito salivary glands. These sporozoites can then be passed to a vertebrate host when the female mosquito takes a blood meal, initiating a new host infection. Source: (Flannery *et al.*, 2013).

1.4 Regulation of gene expression in *Plasmodium*

1.4.1 Regulation of transcription initiation

To account for stage-specific transcriptomes and proteomes throughout the *Plasmodium* life-cycle (Bozdech *et al.*, 2003, 2008; Hall *et al.*, 2005; Young *et al.*, 2005; Foth *et al.*, 2011), a number of studies have been carried out to identify the underlying molecular mechanisms responsible for controlling gene expression and the translation of messenger RNA (mRNA) transcripts. Throughout the life-cycle of *Plasmodium* spp., as in other eukaryotes, mRNA transcription begins when RNA polymerase II is recruited to a gene promoter to synthesise pre-mRNA (also called 'nascent' RNA) that corresponds to the 5'-untranslated region (5'UTR), coding sequence (or 'exon'), introns (if present), and 3'-untranslated region (3'UTR) of the gene in question (Lenhard, Sandelin and Carninci, 2012; Vembar, Droll and Scherf, 2016). Even at this very early point in the transcription process, regulation of gene expression can be very complex. In protozoan parasites such as *Plasmodium*, as in metazoa, promoters of genes can consist of both 'core' and 'proximal' promoters which can be located 20 to 3000 base pairs upstream of the transcription start site (TSS) (Rodríguez *et al.*, 2010). In metazoa, the 'core' promoter is described as the promoter in the immediate vicinity, or overlapping, the TSS, and presumed to dock the pre-initiation complex (PIC).

Metazoan core promoters display a number of mechanisms for tight control of gene expression. Core promoters can be defined as 'TATA-box'-dependent promoters, with 'TATA-box' referring to a sequence of DNA containing TATAWAWR that is located about 30 base pairs upstream of the TSS. 'TATA-box'-binding proteins (TBPs) attach to this DNA sequence within a promoter, beginning the process of PIC assembly (Sainsbury, Bernecky and Cramer, 2015). TATA-dependent promoters are often tissue-specific in metazoa. However, core promoters may also be dispersed, with high CG content, or alternatively, placed within a region of DNA sequence with large amounts of cytosine nucleotides directly followed by

guanine nucleotides, areas known as CpG islands (where CpG stands for 5'-C-phosphate-G-3') (Lenhard, Sandelin and Carninci, 2012). A comprehensive review of regulatory elements surrounding promoters in eukaryotes is presented in figure **1.9.**

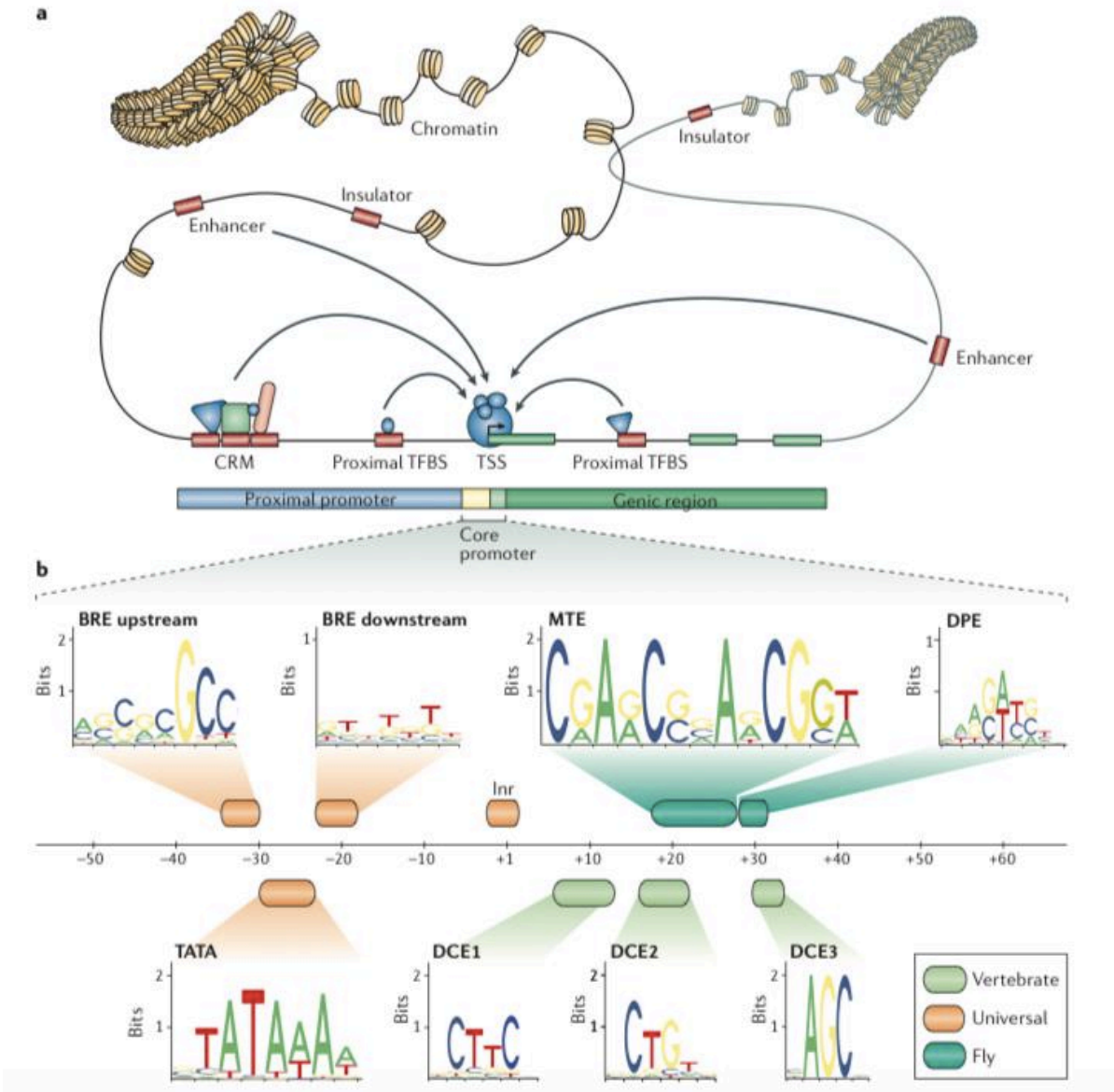


Figure 1.9: A summary of eukaryotic promoter elements and regulatory signals. (a) shows a cartoon depiction of a gene with the its upstream core promoter and proximal promoter regions outlined. Above the DNA sequence, molecular mechanisms for control of the gene's expression are illustrated. At the very top left is shown a length of chromatin that is then split up to reveal the separate nucleosomes. The nucleosomes in turn are comprised of eight core histones surrounded by a length of DNA (~147 base pairs of DNA). Chromatin can be found as either tightly-packed heterochromatin or loosely packed euchromatin, with its conformation heavily effecting whether the encased DNA (including promoter regions) are accessible to the transcriptional machinery of the cell. Long-range regulatory elements such as insulators or enhancers (both DNA sequences of

varied length) can also effect downstream gene expression and pre-initiation complex (PIC)-promoter binding. At the promoter site itself, the region around the transcription start site (TSS) is divided into a larger 'proximal' promoter upstream of the TSS, and the smaller 'core' promoter that surrounds and overlaps the TSS. To recruit RNA polymerase II, sequence-specific regulatory proteins (known as transcription factors) bind to transcription factor binding sites (TFBSs) that are located proximal to the TSS. TFBSs can occur in clusters to form *cis*-regulatory modules (CRMs). In **(b)** sequence patterns are shown that have been identified in eukaryotic 'core' promoters. Around the TSS, TATA-box and initiator (Inr) sequences are often found and over-represented. The locations of sequence patterns are shown relative to the TSS, with all sequence logos depicted other than the Inr sequence which varies considerably between eukaryotes. Most promoters contain only one or a few of these sequence patterns, and many appear to be species-specific. BRE: B recognition element; DCE: downstream core element; DRE: DNA recognition element; MTE: motif ten element. Source: (Lenhard, Sandelin and Carninci, 2012).

In *Plasmodium* spp., the transcription of protein-coding genes is thought to be generally monocistronic, though a bicistronic mRNA for the *maebl* gene and an adjacent putative mitochondrial ATP synthase has been identified in *P. falciparum* (Balu, Blair and Adams, 2009; Rodríguez *et al.*, 2010). Transcription of *Plasmodium* genes is carried out by RNA polymerase II, with homologues to all 12 subunit proteins identified in the genome, though only one third of the expected transcription-associated proteins (TAPs) were identified (Coulson, Hall and Ouzounis, 2004). The *P. falciparum* TATA-box binding protein (TBP) homologue also shows low sequence identity (42%) when compared at the primary sequence level to the archetypal yeast TBP, though it retains most residues known to be responsible for DNA-binding (McAndrew *et al.*, 1993).

With regard to the transcription initiation sites and TSSs within *Plasmodium* spp., the vast majority of *P. falciparum* genes have been shown to contain multiple transcription initiation sites, with initiation occurring primarily at adenine nucleotides (Watanabe *et al.*, 2002). An in-depth analysis of the dynamics of transcription initiation during the *P. falciparum* IDC further demonstrated that 81% of TSSs were positioned less than 1000 bp upstream of the start codon, and that distribution of TSSs were similar in genes displaying both tandem or bidirectional configuration (Adjalley *et al.*, 2016). Moreover, this study identified transcription initiation during the *P. falciparum* IDC at non-conventional sites, such as coding regions of single exon genes. In fact, 49% of all TSSs were located downstream of a start codon. In addition, evidence for stage-specific selection of

TSSs within a single gene, producing transcript isoforms, suggests that gene expression in *Plasmodium* is indeed subject to regulation at a transcriptional level (Adjalley *et al.*, 2016).

1.4.2 Regulation of gene expression by post-transcriptional mechanisms

In addition to the control of gene expression at the point of transcription initiation, nascent *Plasmodium* mRNA, once transcribed, is subject to a range of regulatory mechanisms. In *P. falciparum*, it has been shown that exoribonuclease-mediated degradation of nascent mRNA is involved in the silencing of *upsA*-type *var* genes that have been implicated in severe malaria (Zhang *et al.*, 2014). The chromatin-associated exoribonuclease, termed *Pf*RNaseII, controls the silencing of *upsA*-type *var* genes by marking the gene TSS and intronic promoter sites to produce short-lived cryptic mRNAs (both upstream and intronic promoter-derived antisense long non-coding RNAs (lncRNAs) being the main type in *upsA*-type *var* gene regulation) (Zhang *et al.*, 2014).

Furthermore, both RNA ligation-based and non-polyA-selected strand-specific RNA-sequencing studies revealed the presence of stage-specific independently-transcribed antisense transcripts, lncRNAs, and circular RNAs (circRNAs) during the IDC of *P. falciparum*, supporting the theory that non-coding regulatory elements play a role in the regulation of gene expression in *Plasmodium* spp. (Siegel *et al.*, 2014; Broadbent *et al.*, 2015). An examination of the transcriptional dynamics associated with early gametocyte formation in *P. falciparum,* using a novel nascent mRNA biosynthetic labelling approach, also identified gametocyte-specific transcription and mRNA stabilisation at a point earlier than, and independent of, the transcriptional regulator, Apetala 2-gametocyte (AP2-G) (Painter, Carrasquilla and Llinás, 2017). This study demonstrated enrichment of a *cis*-regulatory motif, AGACA, at the 5'UTR (within 1 kilobase (kb) upstream of the start codon) of 371 of 808 gametocyte-associated genes, and enrichment of a UGUR motif at the 3'UTR of the same gametocyte-associated genes, the latter of which resembles the RNA sequence motif that is recognised by the *P. falciparum* Puf2 RNA-binding protein (Miao *et al.*, 2013). The identification of sequence motifs with the potential for interaction with DNA/RNA-binding proteins in genes associated with early gametocyte development hints at a role for post-

transcriptional regulatory mechanisms in commitment to gametocytogenesis in *Plasmodium* (Painter, Carrasquilla and Llinás, 2017).

Upon completion of nascent mRNA (or pre-mRNA) synthesis, these mRNA molecules are then subject to RNA-processing mechanisms such as alternative splicing (to remove introns or to produce truncated mRNA transcripts) (Sorber, Dimon and Derisi, 2011), addition of the 7-methyl guanosine (m^7GpppN) cap structure to the 5' mRNA end (Calero *et al.*, 2002; Vembar, Droll and Scherf, 2016), and addition of a polyA-tail to the 3' mRNA end (Vembar, Droll and Scherf, 2016). This mRNA molecule is then suitable for export to the cytoplasm of the parasite cell where it can undergo an array of fates: i) translation by the 80S ribosome and tRNAs charged with amino acids; ii) degradation by exo- or endo-ribonucleases; or iii) sequestration by specific ribonucleoprotein (RNP) complexes to inhibit translation and/or decay, i.e. translational repression. The full array of post-transcriptional modifications available to *Plasmodium* spp. parasites is shown in figure **1.10.**

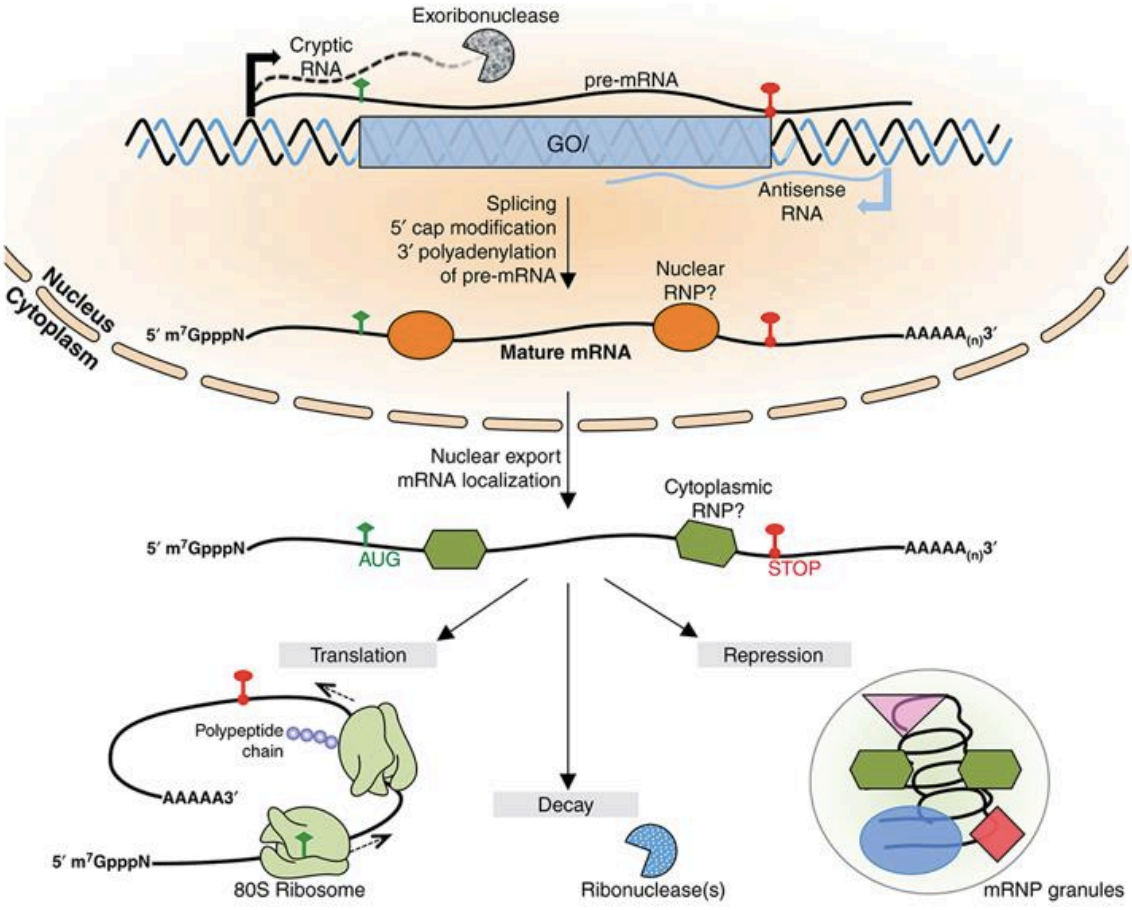


Figure 1.10: Cartoon summary of post-transcriptional modifications to which *Plasmodium* spp. mRNA is exposed. Starting from the top of the cartoon, the gene of interest (GOI) is depicted as

being transcribed into a pre-mRNA transcript, with the possibility of cryptic RNA generation (such as long non-coding RNAs) in an exoribonuclease-mediated fashion. In this image, green arrows/bars are used to depict the start codon (AUG) in the GOI or mRNA transcript, with red bars being used to indicate the stop codon of the same GOI or mRNA transcript. In addition to the generation of *cis* cryptic RNAs, the possible transcription of regulatory antisense RNA transcripts is also shown relative to the GOI. Below the GOI, the mechanisms that are responsible for pre-mRNA processing are listed (splicing, 5' m^7GpppN capping, and 3' polyA-tailing). The mature mRNA molecule is then shown to be subject to nuclear export and/or binding by a nuclear ribonucleoprotein (RNP) complex. Beyond this, exported mature mRNA can then be forwarded to the ribosome for translation, bound to cytoplasmic RNPs (for translational repression), or degraded by exo- or endo-ribonucleases. mRNP: mRNA-RNP granule. Source: (Vembar, Droll and Scherf, 2016).

1.4.3 Regulation of translation initiation, elongation, and termination

As mentioned in section **1.4.2**, mature mRNA molecules are not simply free to immediately begin translation. In addition to regulatory mechanisms at the point of transcription initiation and pre-mRNA processing, translation of mature mRNA transcripts is tightly controlled and stage-specific in nature. *Plasmodium* protein translation occurs in the parasite nucleus, mitochondrion, and apicoplast with the core translational machinery in *Plasmodium* spp. being highly conserved (Jackson *et al.*, 2011). The conserved *Plasmodium* spp. core translational machinery consists of a repertoire of ribosome components, translation factors, tRNA molecules, and aminoacyl-tRNA synthetases with certain features of translation being specific to parasite stage transitions.

In contrast to the ribosomal DNA (rDNA) clusters of other eukaryotes, *Plasmodium* spp. genomes contain 4-8 single-copy rDNA subunits on different chromosomes that are expressed in a stage-specific manner. These rDNAs are split into A-type (liver and intraerythrocytic asexual-stage), O-type (ookinete-specific), and S-type (sporozoite specific) rDNAs and are functionally divergent from the archetypal yeast rDNAs (McCutchan *et al.*, 1995; Vembar, Droll and Scherf, 2016). In addition, the *P. falciparum* 18S ribosomal RNA (rRNA) differs from other eukaryotes in that it contains large peripheral rRNA expansion elements that may be targeted by translational regulatory mechanisms (Wong *et al.*, 2014).

Strikingly, the *P. falciparum* intraerythrocytic parasite stages were shown to lack a RACK1 (receptor for activated C kinase 1) protein from a conserved binding site near the mature mRNA exit channel (Sengupta *et al.*, 2004; Wong *et al.*, 2014). In other eukaryotes, RACK1 is a 40S ribosome subunit-associated protein that is postulated as acting as a link between cell-signalling and translational machinery, with its absence suggesting a marked difference in *Plasmodium* translational regulation (Gallo and Manfrini, 2015; Vembar, Droll and Scherf, 2016). RACK1 was initially described as an anchoring protein for Protein Kinase C (PKC), a homologue of which is not present in the *Plasmodium* genome (Madeira *et al.*, 2003). Despite this, RACK1 expression is essential for *P. falciparum* intraerythrocytic proliferation, with the *Pf*RACK1 protein localised diffusely throughout the parasite cell cytoplasm (Blomqvist *et al.*, 2017). These findings may suggest that *Plasmodium* RACK1 is required for protein translation by interaction with the 40S ribosome subunit at stages outside of the intraerythrocytic cycle, stages that have yet to be examined. *Pf*RACK1 may also function as a key mediator of processes outside of translation, with an as-yet-undiscovered protein taking over the RACK1 role in *Plasmodium*. Of course, this mechanism of RACK1-40S ribosome interaction at the mature mRNA exit tunnel may be absent entirely from *Plasmodium* spp., with the parasite regulating translation by alternative means.

In eukaryotes, translation of a mature mRNA molecule to a polypeptide is divided into three stages: initiation, elongation, and termination. In *P. falciparum*, the conserved functions of eukaryotic translation initiation factor 4E (eIF4E), eukaryotic translation initiation factor 4F (eIF4F), and polyA-binding protein (PABP) have all been demonstrated *in vitro* (Shaw *et al.*, 2007; Tuteja and Pradhan, 2009). At initiation, a conserved method of translational repression of global mRNA levels has been demonstrated in *Plasmodium*, i.e. phosphorylation of the α-subunit of eukaryotic initiation factor 2 (eIF2α). Though repressing global protein synthesis, phosphorylation of eIF2α also acts to enhance gene-specific translation of a subset of mRNAs. In *Plasmodium*, as in *Toxoplasma*, elevated levels of phosphorylated eIF2α have been observed in latent parasite life stages, suggesting a role for this mechanism of translational repression in Apicomplexa dormancy (Zhang *et al.*, 2013). In addition, the three *P. falciparum* eIF2α serine/threonine kinases; eIF2α kinase 1 (*Pf*IK1), eIF2α kinase 2 (*Pf*IK2), and

eIF2α-kinase related protein 4 (*Pf*PK4), have been shown to be differentially expressed throughout the life cycle (Mohrle *et al.*, 1997; Fennell *et al.*, 2009; Zhang *et al.*, 2010; Solyakov *et al.*, 2011). As such, regulation of these enzymes in a stage-specific manner is indicative of regulatory mechanisms at the level of translation initiation.

In *P. falciparum* blood stages, structural subunits of the translation elongation complex have also been identified (elongation factor (EF) subunits 1α, β, γ, and δ), though they have yet to be functionally characterised (Takebe *et al.*, 2007; Vembar, Droll and Scherf, 2016). The *P. falciparum* translation elongation factor 2 (*Pf*EF2) has also been identified as the target of a novel antimalarial compound (DDD107498), but again, this protein has not been functionally characterised either *in vitro* or *in vivo* (Baragaña *et al.*, 2015). With regards to termination of translation by eukaryotic peptide chain release factors (eRFs), homologues to these factors have been annotated in *Plasmodium* spp., however, as with elongation factors, functional characterisation of these proteins has yet to be carried out (Vembar, Droll and Scherf, 2016).

1.4.4 Translational repression

Although technically coming under the umbrella term of 'post-transcriptional' modifications, translational repression of mature mRNAs has been extensively characterised in *Plasmodium* spp., and occurs at a point *after* regulation of pre-mRNA by ribonucleases or cryptic RNAs, but *before* mRNA processing and translation in the ribosome. In *Plasmodium* spp., translational repression of mRNA transcripts has been shown to be carried out by a number of proteins, including, dihydrofolate reductase-thymidylate synthase (DHFR-TS) (inhibits translation of its cognate mRNA) (Zhang and Rathod, 2002); *P. falciparum* protein Alba1 (a homologue to archaeal Acetylation lowers binding affinity protein 1) (Chêne *et al.*, 2012; Vembar *et al.*, 2015); messenger ribonucleoprotein (mRNP) complexes containing the DDX6 DEAD-box RNA helicase, DOZI (Development of Zygote Inhibited) and the Sm-like factor, CITH (a homologue of the *Caenorhabditis elegans* factor, CAR-I and fly Trailer Hitch) (Mair *et al.*, 2006; Mair *et al.*, 2010); and *P. falciparum* RNA-binding proteins, *Pf*Puf1 and *Pf*Puf2 (Puf standing for

Pumilio and fem-3 binding factor homologue) (Müller, Matuschewski and Silvie, 2011; Shrestha *et al.*, 2016).

Unlike the bidirectional protein, *Pf*DHFR-TS, which binds its own cognate mRNA, therefore acting as a translational repressor of itself (Zhang and Rathod, 2002), the *Pf*Alba1 RNA-binding protein has been shown to bind 1193 RNA transcripts in replicative trophozoites, followed by carefully timed release of previously *Pf*Alba1-bound mRNAs encoding erythrocyte invasion components in later schizont stages (Vembar *et al.*, 2015). With these results, the authors propose that *Pf*Alba1 acts as a post-transcriptional regulator of mRNAs during the *P. falciparum* IDC, with a particularly important role in regulating erythrocyte invasion as the end of the asexual parasite stage nears.

The DOZI and CITH proteins (characterised in *P. berghei*) act at different parasite stages, exhibiting crucial roles as translational regulators of gametocytogenesis, sexual differentiation, and post-fertilisation development in the mosquito vector (Mair *et al.*, 2006; Mair *et al.*, 2010). Further studies into the roles of *Pb*DOZI and *Pb*CITH established that, without the presence of these repressors, mRNA transcripts of the mature female gametocyte are destabilised, abolishing *Plasmodium* transmission and zygote development (Guerreiro *et al.*, 2014).

In *P. falciparum*, integrated transcriptomic and proteomic analyses of male and female gametocytes identified a putative set of translationally repressed *P. falciparum* transcripts in female gametocytes, with 260 novel repressed transcripts. This study also highlighted the fact that some of these newly-identified, translationally-repressed transcripts in female gametocytes encoded for genes whose corresponding proteins have only ever been reported in oocyst- or sporozoite- stage parasites. Therefore, these mRNA transcripts may not only be repressed until after fertilisation in the mosquito, but may be translationally repressed in a more long-term manner (Lasonder *et al.*, 2016). Despite these findings, it remains to be determined whether these *P. falciparum* female gametocyte transcripts are subject to translational repression by DOZI, CITH, or an as-yet-unidentified mechanism of translational repression. In fact, to date, the functions of the DOZI-CITH complex in the human malaria parasite *P. falciparum* remain to be characterised (Shrestha *et al.*, 2016).

In the case of *P. falciparum* Puf proteins, the RNA-binding protein, *Pf*Puf1 has also been determined to regulate sexual development in the parasite. *Pf*Puf1 has been shown to be expressed in all gametocyte stages, with higher levels reported in female gametocytes in particular (Shrestha *et al.*, 2016). At present, *Pf*Puf1 is postulated to promote either translational repression by blocking transcription factor binding itself, by recruiting further repressor complexes, or by recruiting factors involved in mRNA degradation. In eukaryotes, two common mechanisms of repression by Puf proteins have been identified: inhibition of mRNA 5'-capping to block translation initiation, or interaction with mRNA 3'UTR *cis* regulatory elements to recruit the CCR4-POP2-NOT deadenylase complex which mediates polyA-tail removal (Miller and Olivas, 2011). In *P. falciparum*, it is quite possible that *Pf*Puf1 may act in concert with the DOZI-CITH complex to repress mRNA translation and regulate sexual development, although only with further experiments can this relationship be examined (Shrestha *et al.*, 2016).

In addition to *Plasmodium* Puf1, the Puf2 RNA-binding protein has also been examined in detail in both *P. berghei* and *P. falciparum*. In *P. falciparum,* Puf2 gene disruption resulted in an increase of gametocyte conversion and differentiation to male gametocytes, with an increase in *Pf*Puf2 repressing male gametocyte maturation and exflagellation (Miao *et al.*, 2010). Further studies revealed that *Pf*Puf2 is in fact an instigator of translational repression of a number of gametocyte mRNA transcripts, including mRNA of the ookinete surface proteins P25 and P28 (Miao *et al.*, 2013). Unlike *Pf*Puf1, the mechanism by which *Pf*Puf2 mediates translational repression in *P. falciparum* has been determined. In a study by Miao *et al.* (2013), *Pf*Puf2 was shown to bind to Puf-binding elements (PBEs) in the 3'UTR of the *P. falciparum* P28 mRNA transcript (*pfs28*) and the 5'UTR of the *P. falciparum* P25 mRNA transcript (*pfs25*). This study represented the first time that the mechanism of *Pf*Puf2-mediated translational repression was determined in the human malaria parasite, and shows its regulatory role in *P. falciparum* gametocytogenesis.

In *P. berghei, Pb*Puf2 was also shown to be a key regulator in the transition of sporozoites from the mosquito salivary glands to EEF parasite liver stages in the mammalian host (Gomes-Santos *et al.*, 2011; Müller, Matuschewski and Silvie, 2011). During sporozoite-to-EEF transition, *Pf*Puf2 regulates the translation of the

key sporozoite protein, UIS4 (Upregulated in Infectious Sporozoites protein 4) by binding to its mRNA transcript, directly repressing translation (Silva *et al.*, 2016). Perturbation of Puf2 expression in an additional rodent malaria model, *P. yoelii* (strain *P. yoelii* 17 XNL (non-lethal)) confirmed the role of this translational repressor in the correct timing of sporozoite-to-EEF transformation, with premature *Py*Puf2 overexpression resulting in deregulation of sporozoite maturation and infectivity (Lindner *et al.*, 2013).

The role of a third Puf protein, Puf3, has only recently been investigated in both *P. falciparum* and *P. yoelii,* though the function of this RNA-binding protein appears to be unrelated to gametocytogenesis (Liang *et al.*, 2018). In *P. falciparum*, Puf3 activity appeared localised to the nucleolus during the IDC. With subsequent analysis of *P. yoelii* Puf3 during mosquito and liver stages, a movement from the nucleolar compartment to cytosolic puncta were observed in oocysts, oocyst sporozoites, and early liver-stage sporozoites. Protein co-immunoprecipitation studies and RNA-immunoprecipitation with high-throughput sequencing (RIP-seq) further demonstrated an association of *Pf*Puf3 with ribosomal proteins and enrichment at rRNA internal-transcribed spacer 2 (ITS2) sequences. Taken together, these findings suggest that Puf3 is involved in ribosomal biogenesis at parasite transmission stages within the mosquito and early intra-hepatocytic development (Liang *et al.*, 2018).

The roles of DHFR-TS, Alba1, CITH, DOZI, and the Puf proteins mark only the beginning of investigations into mechanisms of translational repression in *Plasmodium* parasites. Experiments have only just begun to reveal the possible roles of the transcriptional regulators, SAP1 (sporozoite asparagine-rich protein 1), CELF 1 (CUG-BP, Elav-like family proteins 1), and PABP 1 and 2 (Poly(A)-binding proteins 1 and 2) in the translational repression of *Plasmodium* spp. RNA (Aly *et al.*, 2011; Wongsombat *et al.*, 2014; Minns *et al.*, 2018). In a comprehensive study using hidden Markov model (HMM) searches, nearly 1000 possible RNA-binding proteins were identified in *P. falciparum* (Bunnik *et al.*, 2016). Among these were a number of mRNA-binding proteins associated with polysomes throughout the *P. falciparum* IDC, some of which have been identified as translational repressors (Alba1 and PABP1 for example), and more with the potential for roles in

translational repression such as Alba3, CAF1, and CAF40 (Bunnik *et al.*, 2013, 2016) (**Figure 1.11**).

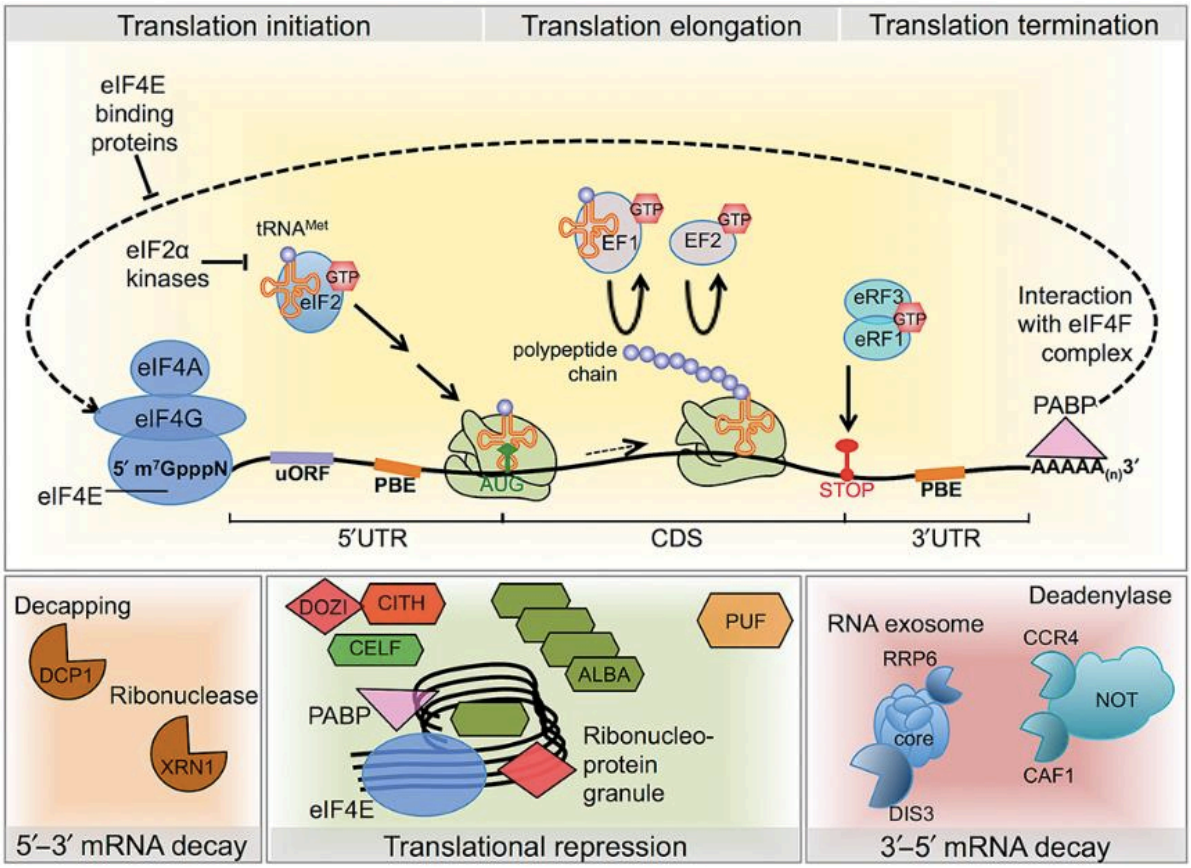


Figure 1.11: An overview of the regulation of translation in *Plasmodium* spp. This image is a cartoon representation of our current knowledge of translational regulation in *Plasmodium* spp. The upper panel of this image shows the stages of mRNA translation from the 5'untranslated region (5'UTR), through the coding sequence (CDS), to the 3'untranslated region (3'UTR), with the stages of initiation, elongation, and termination highlighted. Upon entering the cell cytoplasm, mature mRNA can be subject to translation initiation by eukaryotic initiation factors (eIFs), a number of which are shown in the above image (eIF4A, eIF4G and eIF2α). Initiation factors can be inhibited in their action by binding proteins, such as the eIF4E binding proteins shown. The factor eIF4E binds the 5'm^{7}GpppN cap of the mRNA molecule and forms part if the heterotrimeric translation initiation complex with the helicase eIF4A, and scaffold protein eIF4G. Eukaryotic translation usually begins with the assemblage of the 80S ribosome and a methionine-loaded transfer RNA (tRNA^{met}) at the mRNA start codon (AUG). The eIF2 protein is involved in this process but can be inhibited in its actions by phosphorylation of the eIF2α subunit. In *Plasmodium* eIF2α kinases (as shown above) can repress translation by phosphorylation of the eIF2α subunit. The polypeptide chain is then elongated by eukaryotic elongation factors (eEFs) and the final polypeptide released from the ribosome by eukaryotic translation release factors (eRFs). A core element of successful translation is the interaction of the eIF4F complex (made up of eIF4A, eIF4E and eIF4G) with a poly(A)-binding protein (PABP) at the 3' polyA-tail of the mRNA. This interaction forms the mRNA-

ribonucleoprotein (mRNP) structure that prevents the mRNA being translated from degradation by outside factors. The three panels beneath the main mRNA translation image depict methods by which *Plasmodium* spp. translation can be inhibited; again from the 5'UTR to the 3'UTR. A number of *Plasmodium* translation repressors are depicted in the centre image (ALBA, CITH, DOZI, PUF, CELF, and PABP). uORF: upstream open reading frame; PBE: Puf-binding element; DCP: decapping protein; XRN1: exoribonuclease 1; RRP6: exosome complex exonuclease RRP6; DIS3: exosome complex exonuclease RRP44; CCR4: carbon catabolite repressor 4; NOT: Negative on TATA; CAF1: CCR4-associated factor 1. Source: (Vembar, Droll and Scherf, 2016).

1.5 Transcription factors in *Plasmodium*: the AP2 protein family

Outside of the mechanisms of transcriptional, post-transcriptional, and translational control of gene expression as outlined in section **1.4**, and regulation by epigenetic mechanisms (detailed in section **1.6**), *Plasmodium* spp., as well as other members of the Apicomplexa genus, show an unexpected paucity of transcription factor families with homology to those found in other eukaryotes (Balaji *et al.*, 2005). Unlike other mechanisms of gene regulation that have been previously mentioned, transcription factors are proteins that bind to regulatory elements of a gene (such as a specific sequence motif) and regulate the baseline expression of that gene. Transcription factors can act as either activators or repressors and are composed of at least a family-designating DNA-binding domain (DBD) and a transcription regulatory domain that mediates activation or repression of the gene (Staby *et al.*, 2017). Transcription factor regulatory domains can also bind to co-activators of transcriptional machinery, or take part in remodelling or modification of adjacent chromatin (Näär, Lemon and Tjian, 2001).

In *Plasmodium* spp., the few identified transcription factors are: *P. falciparum* Myb1 and Myb2 ("Myb" from "myeloblastosis") (Gissot *et al.*, 2005; Bischoff and Vaquero, 2010); high-mobility-group box (HMGB) factors 1-4 (Kumar *et al.*, 2008; Tuteja, Ansari and Chauhan, 2011); the Apicomplexan Apetela2/ethylene response factor (AP2/ERF) family of DNA-binding proteins (Balaji *et al.*, 2005); and the novel *P. falciparum* Prx regulatory element binding protein, *Pf*PREBP (Komaki-Yasuda *et al.*, 2013). At present, it is the AP2 family of transcription factors that are subject to the most intensive investigations, with studies in *P. berghei* revealing roles for 11 out of 26 *P. berghei* AP2 family genes in transmission. An additional 14 genes in *P. berghei* were demonstrated as being

necessary for asexual parasite stage development, with one final gene of the 26 being incompatible with the knockout screen used and possibly requiring further elucidation (Modrzynska *et al.*, 2017). In *P. falciparum*, the AP2 transcription factor family consists of 27 members; the Apicomplexan parasite *Toxoplasma gondii* possesses 67 AP2 DNA-binding proteins; and *Cryptosporidium parvum* contains only 23 AP2 domains across 18 proteins (Oberstaller *et al.*, 2014; Hong, Radke and White, 2017; Modrzynska *et al.*, 2017). Despite the seemingly unconserved nature of the AP2 family of transcription factors in the Apicomplexa genus, in *Plasmodium* spp., these factors have been shown to be fundamental to developmental stage transitions in the parasite (Modrzynska *et al.*, 2017).

One of the first *Plasmodium* AP2 transcription factors to be functionally characterised was AP2-sp (AP2 in sporozoites), a major transcriptional regulator of gene expression in sporozoites (Yuda *et al.*, 2010). In *P. berghei*, disruption of the *pbap2-sp* gene resulted in parasite development up until the point of sporozoite formation within the mosquito oocyst. In *pbap2-sp*-disrupted parasites, oocysts were shown to have undergone nuclear division, but with no invagination of the oocyst plasma membrane, therefore hindering sporoblast development. This study also demonstrated that the AP2-sp DNA-binding domain recognised a specific sequence motif beginning with TGCATG, located in the proximal promoter region of all sporozoite-specific genes (Yuda *et al.*, 2010).

In *P. falciparum*, the AP2-sp orthologue (PF3D7_1466400) was shown to play another biological role (Martins *et al.*, 2017). In this human-infective parasite, Martins *et al.* demonstrated that the orthologous *P. falciparum* transcription factor, which was called 'AP2-exp' throughout this study, was in fact responsible for regulating a subset of clonally variant genes involved in virulence. Upon disruption of *pfap2-exp* in red blood stage parasites, 190 genes were upregulated and 47 genes downregulated, most belonging to groups of subtelomeric gene families that showed mutually-exclusive expression, including surface proteins of the *rif* and *stevor* gene families. Of considerable importance was the fact that this study demonstrated a role for the non-DNA-binding domain of AP2-exp in the expression of clonally-variant gene families (Martins *et al.*, 2017). However, it remains to be seen whether AP2-exp retains its function in sporozoite

development in *P. falciparum*, or whether *Pb*AP2-sp also acts as a regulator of clonally-variant gene families in blood stages of the rodent malaria model.

In a similar manner, a *P. berghei* AP2 gene necessary for the motility of sporozoites within oocysts, and subsequent infectivity of the parasite into the host liver (*pbap2-sp3*) (Modrzynska *et al.*, 2017) contained an orthologue in *P. falciparum* (PF3D7_0622900) that was subjected to investigation for a very different reason. This *P. falciparum* orthologue, denoted as *Pf*AP2Tel (for *P. falciparum* AP2-telomeric) was pulled down as a member of a telomere-binding protein complex during the intraerythrocytic parasite life cycle (Sierra-Miranda *et al.*, 2017). As a result, *Pf*AP2Tel was again implicated as having a role in the regulation of clonally-variant subtelomeric gene families such as *var*, *rif* and *stevor* gene families during blood stages, with a possible function in sporogony of *P. falciparum* that has yet to be examined.

In yet another study of *P. falciparum* chromosome end biology, the AP2 domain-containing protein, *Pf*SIP2 (*P. falciparum* SPE2-interacting protein) was also shown to interact with SPE2 motifs (two imperfect 6 bp repeats separated by 4 bp) at heterochromatic domains upstream of subtelomeric *var* genes (Flueck *et al.*, 2010). In the later *P. berghei* AP2 protein knockout screen, a *pbsip2* knockout was lethal to the parasite (Modrzynska *et al.*, 2017), suggesting that this transcription factor is indispensable for asexual growth. In future studies, an examination of the roles of *Pf*AP2-exp and *Pf*AP2Tel in sporozoite development and infectivity in the human malaria parasite would be enlightening, as these studies could corroborate results seen in the *P. berghei* rodent malaria model. Alternatively, examining transcriptional changes in knockouts of *Pb*AP2-sp and *Pb*AP2-sp3 may reveal as-yet-undetermined subtelomeric clonally-variant gene families in *P. berghei*. At the very least, a comparison of transcriptomic changes in *P. falciparum Pf*AP2-exp or *Pf*AP2Tel knockout lines (as the non-lethality of these knockouts has been demonstrated in *P. berghei*) may reveal gene families, or single genes, necessary for sporozoite development and invasion, these such proteins being targetable as vaccine candidates.

Outside of the characterisation of *Pb*AP2-sp (*Pf*AP2-exp), *Pb*AP2-sp3 (*Pf*AP2Tel), and SIP2, other members of the AP2 transcription factor family have been shown

to be critical in maintaining developmental stage transitions throughout the *Plasmodium* life cycle. In *P. berghei*, AP2-L (AP2-liver) has been shown to be necessary for intrahepatocytic parasite development (Iwanaga *et al.*, 2012), AP2-O (AP2-ookinete) is responsible for control of ookinete morphogenesis, locomotion, and mosquito midgut penetration (Kaneko *et al.*, 2015; Modrzynska *et al.*, 2017), AP2-ookinete proteins 2 to 4 (AP2-O2 to AP2-O4) and AP2-sp protein 2 (AP2-sp2) are also necessary for successful ookinete and sporozoite development respectively (Modrzynska *et al.*, 2017), and AP2-gametocyte proteins 1 and 2 (AP2-G and AP2-G2) are master regulatory factors of commitment to sexual development and gametocyte maturation (Sinha *et al.*, 2014; Yuda *et al.*, 2015; Modrzynska *et al.*, 2017) (**Figure 1.12**).

Despite this knowledge, the exact point at which commitment to sexual development is initiated is still unknown. We do know, however, that the mechanism for initiation of gametocytogenesis must lie upstream of AP2-G, AP2-G2, and GDV1 (in human-infective *Plasmodium* spp.), and that these transcription factors are essential for correct development of the *Plasmodium* gametocyte. Undoubtedly, future experiments will shed light on the roles of other AP2 family proteins in the *Plasmodium* life-cycle, and that novel methods for genetic manipulation may aid in the examination of these transcription factors in the human malaria parasite.

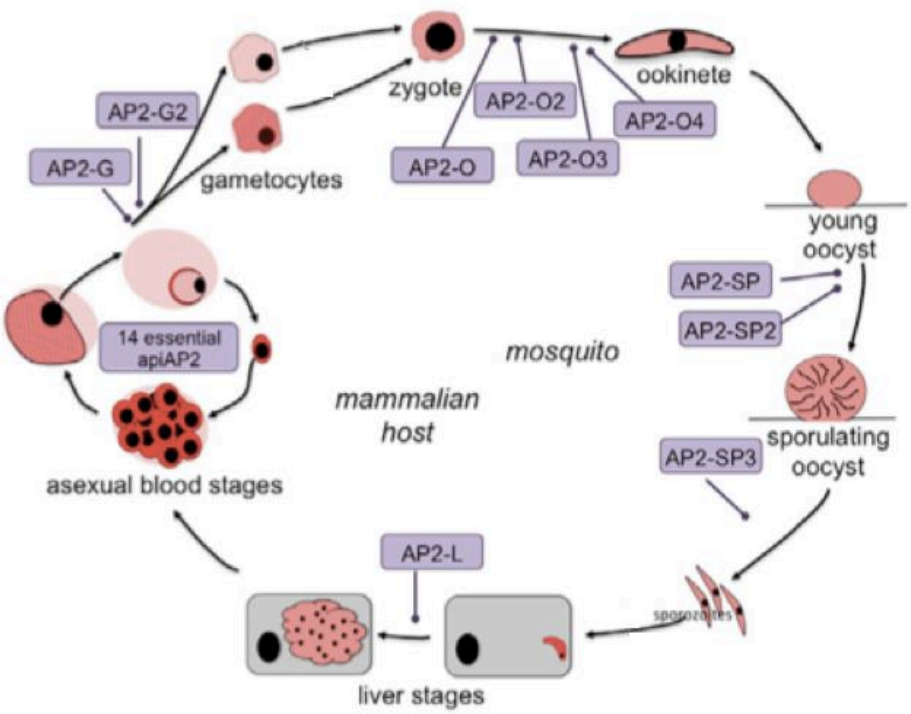


Figure 1.12: An overview of the involvement of AP2 family proteins in the *P. berghei* life-cycle. This image shows a simplified *P. berghei* parasite life-cycle with Apicomplexan AP2 (ApiAP2) transcription factors indicated at the point at which they have been shown to have a regulatory role. Beginning on the left of the image with the asexual blood stages in the host erythrocyte (merozoite invasion to intraerythrocytic schizont stage), 14 ApiAP2 protein knockouts failed to grow asexually and so it has been suggested that they play essential regulatory roles in asexual replication. Moving up to male and female gametocyte stages, AP2-G has been shown to be essential to gametocyte formation (both males and females), whereas AP2-G2 has shown to be essential for gametocyte maturation (Kafsack *et al.*, 2014; Sinha *et al.*, 2014; Yuda *et al.*, 2015; Modrzynska *et al.*, 2017). Modrzynska *et al.* (2017) also showed that AP2-G2 acts as a transcriptional repressor in both asexual and sexual stages, with liver- and sporozoite- stage genes dysregulated in asexual stages upon *pbap2-g2* knockout.

After zygote formation, the formation of the *P. berghei* ookinete is directed by the action of four AP2-ookinete transcription factors, AP2-O to AP2-O4. AP2-O proteins act in concert for the formation of an ookinete with motile capabilities and the capacity for oocyst development. The absence of any one of these transcription factors results in loss of ookinete formation, maturation or oocyst formation at a specific point in this developmental cycle. After oocyst formation, AP2-sporozoite proteins 1 to 3 (AP2-sp to AP2-sp3) regulate normal development of the sporozoite from its development in the oocyst to the release of motile and infective sporozoites in the salivary glands of the mosquito. Once sporozoites have entered the mammalian host and have traversed to a suitable hepatocyte for intrahepatocytic growth, the AP2 transcription factor, AP2-liver (AP2-L) regulates transcriptional changes necessary for exoerythrocytic form (EEF) development. Without AP2-L expression, the EEF developmental stage is lost and the parasite life-cycle cannot continue. Not shown are the roles of *Pb*SIP2 and an as-yet-unnamed AP2 domain-containing protein

(corresponding to gene PBANKA_1313200) for which no knockout vector could be generated. Source: Image adapted from the graphical abstract of Modrzynska *et al.*, 2017.

1.6 Epigenetics and higher order genome organisation in *Plasmodium*

1.6.1 Epigenetic regulation of *Plasmodium* virulence: telomere position effect

The study of epigenetics in malaria parasites has stemmed almost solely from investigations into the molecular mechanisms responsible for monoallelic *var* gene expression in *Plasmodium falciparum*. At the heart of this immune evasion process is the intricate organisation and regulation of the family of ~60 *var* genes from the point of initiation of gene transcription (Kaestli *et al.*, 2006; Kyes *et al.*, 2003; Scherf *et al.*, 1998; Su *et al.*, 1995).

Epigenetic regulation of *var* genes takes place via a number of different mechanisms, from the level of the gene itself to its location within the three-dimensional chromatin structure of the *Plasmodium* nucleus. Telomere position effect, i.e. the reversible transcriptional repression of genes as a result of their relative position to the chromosome end, has been demonstrated as playing an important role on the silencing of *var* genes in *Plasmodium* (Gottschling *et al.*, 1990; Duraisingh *et al.*, 2005). *Var* genes are predominantly located at subtelomeric regions of the genome and are clustered at the nuclear periphery, with transcriptional silencing being reversed upon physical repositioning of a *var* locus to a position that is permissive of transcriptional activation *(*Freitas-Junior *et al.*, 2000; Voss *et al.*, 2007; Marty *et al.*, 2006). In *P. falciparum*, the transcriptional repression of *var* genes occurs at both subtelomeric regions of the chromosome and at so-called 'chromosome-central' regions, with both loci demonstrating variable rates of transcriptional switching and stochastic gene expression patterns (Frank *et al.*, 2007). The lack of a pre-determined order of *var* gene transcription more than likely contributes to the heterogeneity of clonal parasite populations and the severity of *P. falciparum*-mediated malarial illness by contributing to the evasion of host immunity.

In addition to being localised to specific chromosome regions, *var* genes are also regulated by distinct *var* gene promoters that correlate strongly with their

chromosomal location. Following completion of the *P. falciparum* genome, three separate classes of *var* gene promoter have been determined, all being highly conserved and upstream of their respective genes (hence the *ups* nomenclature): *ups*A, *ups*B and *ups*C (Voss *et al.*, 2003). Variant *var* gene promoters mediate different *in situ* transcriptional effects through differential binding to nuclear proteins and different physical interactions between *var* introns and upstream elements within the gene sequence (Voss *et al.*, 2006; Calderwood *et al.*, 2003). In particular, the *ups*C-type *var* promoter has been implicated in epigenetic silencing through an interaction between AUG-rich, *cis*-acting promoter motifs and the silencing machinery of *P. falciparum* (Brancucci *et al.*, 2012).

Though positioning of genes within the nucleus and their position relative to telomeres is a decisive factor in the epigenetic regulation of virulence genes in *P. falciparum*, so too are factors that influence nucleosome structure and function: histone-modifying enzymes (Coleman *et al.*, 2014; Brancucci *et al.*, 2014; Chaal *et al.*, 2010; Miao *et al.*, 2010; Mancio-Silva *et al.*, 2013; Merrick *et al.*, 2015; Tonkin *et al.*, 2009); histone modifications (histone "marks") (Jiang *et al.*, 2013; Lopez-Rubio *et al.*, 2007; Chookajorn *et al.*, 2007; Cabral *et al.*, 2012; Karmodiya *et al.*, 2015; Miao *et al.*, 2006; Dastidar *et al.*, 2013; Gupta *et al.*, 2013; Sindikubwabo *et al.*, 2017); the interchange between canonical and variant histones (Petter *et al.*, 2013; Hoeijmakers *et al.*, 2013; Bártfai *et al.*, 2010); the activity of nucleosome assembly proteins (NAPs) that chaperone histone proteins to and from the nucleosome structures (Gill *et al.*, 2010); and dynamic nucleosome positioning relative to transcription start sites (Bunnik *et al.*, 2014; Kensche *et al.*, 2015) **(Figure 1.13).**

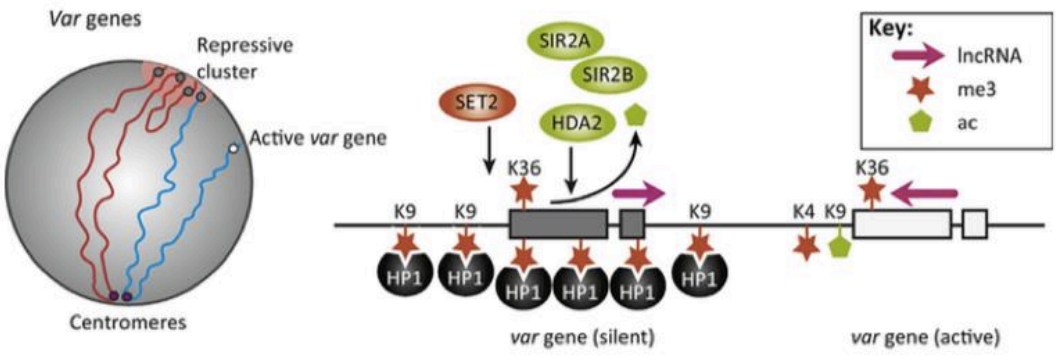


Figure 1.13: Epigenetic regulation of *var* genes. This image depicts a number of the known epigenetic mechanisms responsible for monoallelic *var* gene expression in the *P. falciparum*

parasite. *Var* genes are localised to the nuclear periphery of the parasite, with one active *var* gene isolated from a repressive cluster of silent *var* genes. Silenced *var* genes are associated with the repressive histone 3, lysine 9 trimethylation (H3K9me3) modification, which in turn recruits heterochromatin protein 1 (HP1) to form tightly compacted heterochromatin. Histone deacetylases such as histone deacetylase 2 ('HDA2' or 'HDAC2'), silent information regulator 2A (SIR2A), and SIR2B are responsible for removing acetylation modifications, therefore contributing to heterochromatin formation by facilitating the action of methyltransferases such as SET2 (Su(var)3–9, Enhancer of zeste, and Trithorax 2). H3 lysine 36 trimethylation (H3K36me3) is also associated with silenced *var* genes and is a product of the SET2/*PfSETvs* methyltransferase enzyme. The active *var* gene is marked by H3 lysine 4 trimethylation (H3K4me3) and H3K9 acetylation (H3K9Ac). In addition, long non-coding RNAs (lncRNAs) are transcribed from a bidirectional promoter in *var* introns and contribute to *var* gene transcription. Source: Adapted from (Batugedara *et al.*, 2017).

1.6.2 Epigenetic regulation of *Plasmodium* transcription: histone acetylation

In the malaria parasite, the discovery of the involvement of the histone deacetylase PfSIR2A in the transcriptional silencing of virulence-associated *var* and repetitive interspersed family (*rif*) genes marked the first instance that a histone-modifier was identified in *Plasmodium* gene regulation (Tonkin *et al.*, 2009). Since then, PfSIR2A and its paralogue, PfSIR2B, have been shown to mediate the silencing of subtelomeric *var* and *rif* genes by propagating the *cis*-spreading of silent chromatin from telomeres to adjacent regions (Mancio-Silva *et al.*, 2013; Merrick *et al.*, 2015).

In the past decade, a host of enzymes involved in modulating acetylation and deacetylation throughout the complex *Plasmodium* life cycle have been characterised. During the asexual intraerythrocytic parasite cycle, gene activation in *P. falciparum* is closely regulated by the activity of a single histone acetyltransferase of the MYST family of histone acetyltransferases, *Pf*MYST, which is essential for cell cycle control *(*Miao *et al.*, 2010). In addition, inhibiting global histone deacetylase (HDAC) activity by introducing a potent chemical inhibitor has profound detrimental effects on the intraerythrocytic transcriptional cascade; both disrupting the stage-specific histone modification profile, as well as producing a mass up-regulation of transcription factors that de-regulate normal parasite growth (Chaal *et al.*, 2010).

The disruption of HDAC activity not only effects the asexual intraerythrocytic developmental cycle of *Plasmodium*, but also the rate of gametocyte conversion. In a study by Coleman (Coleman *et al.*, 2014), epigenetic regulation of gametocytogenesis was demonstrated in an examination of the role of histone deacetylase 2 (HDAC2) on the level of gametocyte conversion in *P. falciparum*. HDAC2 depletion in *P. falciparum* results in a significant up-regulation in the expression of AP2-G, the transcriptional switch responsible for commitment to sexual development (Sinha *et al.*, 2014; Kafsack *et al.*, 2014). Peak HDAC2 expression also coincides with late trophozoite and schizont parasite stages in wild-type strains, consistent with the point of the asexual intraerythrocytic cycle at which parasites must commit to gametocytogenesis (Bruce *et al.*, 1990). In a complementary study, Brancucci *et al.* (2014) conditionally depleted the conserved regulator of gene silencing, heterochromatin protein 1 (HP1), which binds to H3K9Me3, stabilising this repressive mark (Zeng *et al.*, 2010). The result of this was a dysregulation of *var* gene monoallelic expression, and an increase in *Pf*AP2-G expression and gametocytogenesis (Brancucci *et al.*, 2014). Together, these studies testify to a relationship between enzymes that modulate acetylation of histones and regulation of both the asexual and sexual life cycles in the malaria parasite, *P. falciparum* **(Figure 1.14).**

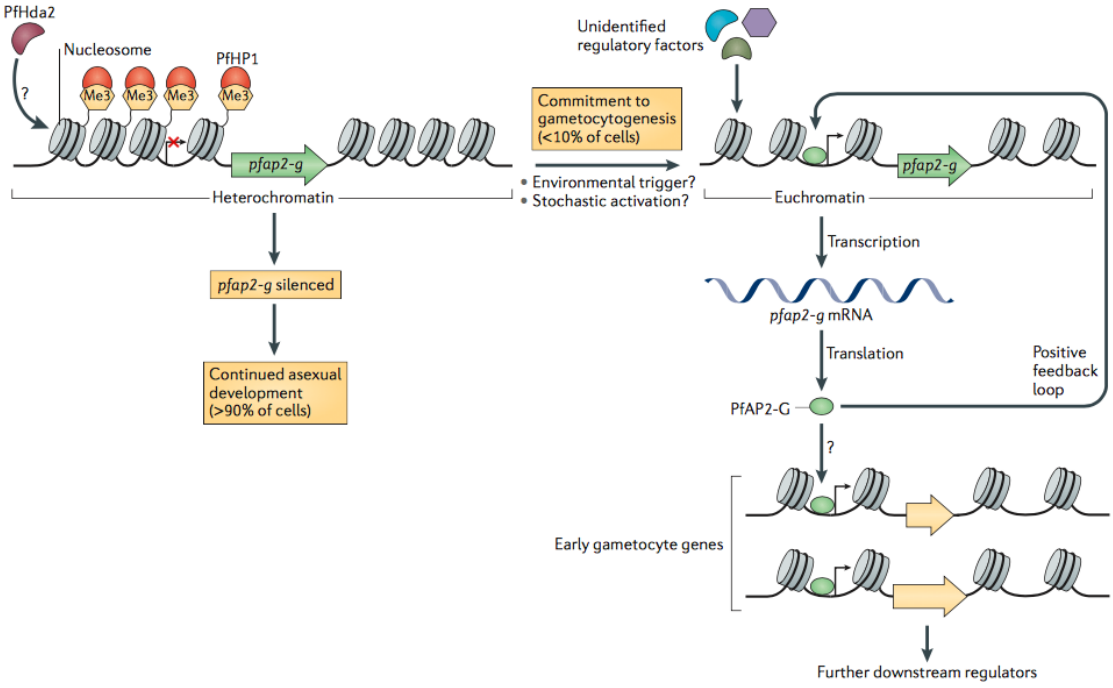


Figure 1.14: Epigenetic regulation of *Plasmodium* sexual development. In *P. falciparum* parasite populations, only 10% or less of all parasites commit to gametocytogenesis, a process under control of the 'master transcriptional regulator', *Pf*AP2-G (Kafsack *et al.*, 2014; Sinha *et*

al., 2014). In the majority of the parasite population, the *ap2-g* locus is located in a heterochromatic region of the nucleus that is marked by methylation of H3K9 (Lopez-Rubio *et al.*, 2009). Depletion of both *Pf*Hda2 (Coleman *et al.,* 2014) and heterochromatin protein 1 (HP1) (Brancucci *et al.*, 2014) have been shown to result in an increase in *Pfap2-g* transcription and gametocyte conversion, though the exact mechanism by which *Pf*Hda2 elicits its effect is currently unknown. In the case of *Pf*HP1, it is proposed that the opening of the *Pfap2-g* locus to transcription results in increased *Pf*AP2-G expression which in turn maintains a positive feedback loop facilitating more *Pf*AP2-G transcription, as well as activation and expression of its downstream effectors (Sinha *et al.,* 2014). In a natural *P. falciparum* infection, it is not yet known whether the initial signal responsible for sexual commitment is stochastic, sporadic, or resulting from environmental triggers. Despite our current knowledge, there is still much about the regulation of gametocyte commitment and development that we have yet to uncover, not least about possible role of further unidentified regulatory factors such as histone methyltransferases and chromatin remodeling complexes. Source: (Josling & Llinas, 2015).

Outside of *P. falciparum* histone acetyltransferases (HATs) and HDACs that have been characterised so far (*Pf*MYST, *Pf*SIR2A, *Pf*SIR2B, and *Pf*HDAC2), recent phylogenetic analyses of HATs and HDACs of the *P. falciparum* genome have revealed the presence of 10 putative HATs of the MYST and Gcn5-related *N*-acetyltransferases (GNATs) families, 6 HDACs of the Class I, II and III families of HDACs, and a novel unclassified HDAC with homology to the eukaryotic chromodomain helicase DNA-binding (CHD) family of chromatin modifiers (Marfella and Imbalzano, 2007; Kanyal *et al.*, 2018). A list of the genes encoding these potential *P. falciparum* HAT and HDAC proteins, alongside orthologous *P. berghei* genes, the enzyme family to which they belong, and the binding domains denoted by their sequence, can be found in **Table 1.1.**

***P. falciparum* gene identifier**	***P. berghei* gene identifier**	**Protein superfamily**	**PlasmoDB name (if given)**	**Conserved domain (s)**
PF3D7_1003300	PBANKA_1201700	GNAT	N-terminal acetyltransferase A complex catalytic subunit ARD1, putative (ARD1)	i) Acetyltransferase domain
PF3D7_0109500	PBANKA_0204100	GNAT	N-acetyltransferase, putative	i) Acetyltransferase domain ii) FR-47-like domain
PF3D7_0805400	PBANKA_1225200	GNAT	Acetyltransferase, putative	i) Acetyltransferase domain ii) RimL domain
PF3D7_0629000	PBANKA_1127600	GNAT	N-acetyltransferase, putative	i) Acetyltransferase domain
PF3D7_1437000	PBANKA_0611800	GNAT	N-acetyltransferase, putative	i) Acetyltransferase domain ii) NAT domain
PF3D7_0823300	PBANKA_0707300	GNAT	Histone acetyltransferase GCN5 (GCN5)	i) Acetyltransferase domain ii) Bromodomain
PF3D7_1323300	PBANKA_1338500	GNAT	Acetyltransferase, GNAT family, putative	i) Acetyltransferase domain
PF3D7_1227800	PBANKA_1442500	GNAT	Elongator complex protein 3, putative (ELP3)	i) Radical SAM/HAT domain

PF3D7_1020700	PBANKA_0504900	GNAT	Histone acetyltransferase, putative	i) Acetyltransferase domain ii) Helicase domain iii) tRNA-binding domain iv) tRNA(Met) cytidine acetyltransferase TmcA domain
PF3D7_1118600	PBANKA_0929500	MYST	Histone acetyltransferase (MYST) **$**	i) Acetyltransferase domain ii) Chromodomain-like domain iii) C2HC-type zinc finger domain
PF3D7_1472200	PBANKA_1335400	HDAC Class II	Histone deacetylase, putative (HDA1)	i) Histone deacetylase domain (x3) ii) Ankyrin repeats
PF3D7_0506600	PBANKA_1106200	HDAC Class II	Histone deacetylase, putative, pseudogene	i) Histone deacetylase domain
PF3D7_1008000	PBANKA_1206200	HDAC Class II	Histone deacetylase 2 (HDA2/HDAC2) **$**	i) Histone deacetylase domain (x3) ii) DUF2981 domain iii) Inositol polyphosphate kinase domain
PF3D7_1023900	PBANKA_0508100	Unclassified	Chromodomain-helicase-DNA-binding protein 1 homolog, putative (CHD1)	i) Chromo/chromo-shadow domain-like ii) Helicase family ATP-binding domain iii) SNF2-related, N-terminal domain iv) Helicase domain
PF3D7_0925700	PBANKA_0826500	HDAC Class I	Histone deacetylase 1 (HDAC1)	i) Histone deacetylase domain
PF3D7_1328800	PBANKA_1343800	HDAC Class III	Transcriptional regulatory protein SIR2A (SIR2A) **$**	i) DHS-like NAD/FAD binding/sirtuin domain

PF3D7_1451400	PBANKA_1315100	HDAC Class III	Transcriptional regulatory protein SIR2B (SIR2B) **$**	i) Histone deacetylase domain (x2)

Table 1.1: Histone acetyltransferases (HATs) and histone deacetylases (HDACs) identified in *P. falciparum* by Kanyal *et al*. (2018) and their orthologues in *P. berghei*. The table above details HATs and HDACs identified as such by Kanyal *et al.*, 2018 in the format *P. falciparum* gene identifier; *P. berghei* gene identifier; protein superfamily into which the protein has been categorised; PlasmoDB name for protein (if one is present); and protein domains encoded by the *P. falciparum* gene sequence (as identified by Kanyal *et al*. 2018). A "**$**" symbol next to the name of the protein in the fourth column indicates that the protein has been characterised in *Plasmodium* previously. Table generated using PlasmoDB [up-to-date as of the 23rd of January, 2018] (The EuPathDB Project Team, 2018).
GNAT: Gcn5-related *N*-acetyltransferase; MYST: MOZ, Ybf1/Sas3, Sas2, and Tip60; HDAC/HDA: Histone deacetylase; ELP: Elongator complex protein; SIR2: Silencing information regulator 2; CHD: Chromodomain helicase DNA-binding; SNF2: Sucrose non-fermentation 2; DUF2981: Domain of Unknown Function 2981; TmcA: tRNA(met) cytidine Acetyltransferase; NAT: *N*-acetyltransferase; NAD: nicotinamide adenine dinucleotide; FAD: flavin adenine dinucleotide; DHS: deoxyhypusine synthase.

1.6.3 Epigenetic regulation of *Plasmodium* transcription: histone methylation

Histone lysine methylation may play an equally important role in the regulation of stage transitions in *Plasmodium*. Bioinformatic analysis has identified nine SET-domain-containing genes ('SET' from the *Drosophila* chromatin-modifying enzymes Su(var)3–9, Enhancer of zeste, and Trithorax), each with putative histone lysine methyltransferase (HKMT) activity in *P. falciparum*, and two Jumonji C (JmjC)-domain-containing genes with putative histone lysine demethylase capabilities (Cui *et al.*, 2008, Jenuwein *et al.*, 2000).

One silencing SET gene, encoding the *P. falciparum* SET2/*PfSETvs* methyltransferase, has been subject to comprehensive functional experimentation (Jiang *et al.*, 2013). In an *in vitro* study of a SET2 knockout *P. falciparum* line, haemagglutinin (HA)-tagged SET2 specifically localised to transcription start sites and intronic promoters of *var* genes, and was associated with *var* family repression of these genes via H3K36 trimethylation. In contrast, *P. falciparum* SET10 (*Pf*SET10), has been shown to localise to H3K4 mono-, di- and tri- methylation marks (H3K4me1, H3K4me2, and H3K4me3) at a euchromatic region of the nuclear periphery corresponding to the site of active *var* gene expression (Volz *et al.*, 2010; Volz *et al.*, 2012). In these *Pf*SET10 studies, the authors propose that *Pf*SET10 is likely required for the maintenance of the transcriptionally permissive chromatin state at the active *var* gene promoter, and the maintenance of the open chromatin state during increased active *var* gene transcription at the *P. falciparum* ring stage (6-8 hours post invasion (hpi)). Though undoubtedly a component of the active *var* gene expression site and a methyltransferase of H3K4, the role of *Pf*SET10 will likely be elucidated further as more is learned about the roles of other SET-domain-containing proteins and their combined activities in regulating monoallelic *var* gene expression (Volz *et al.*, 2012).

The most recent *Plasmodium* methyltransferase to be enzymatically characterised has been *P. falciparum* SET7 (*Pf*SET7), an s-adenosyl methionine (AdoMet)-dependent methyltransferase of histone 3, with highest activity towards lysines 4 and 9 (Chen *et al.*, 2016). Steady state enzyme kinetics were recorded for a codon-

optimised version of wild-type full-length *Pf*SET7 (*Pf*SET7FL), expressed as a recombinant protein using the baculoviral expression system in Spodoptera frugiperda clone 9 (Sf9) insect cells. Recombinant *Pf*SET7 showed a specific substrate preference towards nucleosomes marked with H3 lysine 14 acetylation (H3K14ac), with fluorescently-tagged *P. falciparum* SET7 showing localisation to distinct foci adjacent to the cell nucleus in erythrocytic and liver stage parasites.

In salivary gland sporozoites, *Pf*SET7 expression was seen throughout the cell cytoplasm, suggesting an altered expression of this protein between the vertebrate host and insect vector, though the reasons behind this altered expression are not understood at present. As well as association with methylation of H3K9 and H3K4 (in the presence of H3K14ac), *Pf*SET7 was also shown to methylate lysines 27 and 36 of histone 3 (H3K27 and H3K36) to a lesser extent (Chen *et al.*, 2016). Considering that methylation of H3K9 is associated with silencing of *var* genes (Chookajorn *et al.,* 2007), while K3K4 methylation is associated with transcribed *var* promoters (Lopez-Rubio *et al.*, 2007) in *P. falciparum,* the role of *Pf*SET7 in parasite virulence is almost certainly specific to location within the parasite cytoplasm or nucleus, the open or repressed chromatin state of the surrounding nuclear environment, and the presence or absence of additional histone modifications or their modifying enzymes.

To add to the complexity of histone methyltransferases and their role in epigenetically regulating *Plasmodium* virulence and sexual development, an improved *P. cynomolgi* genome assembly has recently demonstrated a paralogous expansion of the *Plasmodium* interspersed repeat (*pir*) genes, genes encoding for merozoite surface proteins (MSPs), and an expansion of 36 subtelomeric methyltransferase pseudogenes (Pasini *et al.*, 2017). All 36 pseudogenes cluster to one full-length core *P. cynomolgi* gene (PcyM_0947500) on chromosome 9, a gene with a conserved orthologue across all *Plasmodium* species. Evidence of a duplication of methyltransferases was also discovered in *Plasmodium simiovale,* a closely related simian malaria parasite (Cogswell *et al.,* 1991; Pasini *et al.,* 2017). It remains to be seen whether these paralogous genes are functional in either *Plasmodium* species.

Although it is clear that future studies regarding *Plasmodium* HKMTs and JmjC-domain-containing genes will no doubt reveal further fascinating examples of epigenetic regulation (and perhaps a link between histone lysine methylation and regulation of gametocytogenesis), many more studies have endeavoured to map histone modifications across the *Plasmodium* life cycle irrespective of the enzymes that underlie their presence. For example, singularly active *var* genes are flanked by active marks H3K4me2 and H3K4me3 at their 5' coding region, as well as the repressive H3K9me3 modification (Chookajorn *et al.,* 2007; Lopez-Rubio *et al.,* 2007). Genome-wide profiles of histone modifications have identified H3K36me2 as a global repressive mark (Karmodiya *et al.,* 2015), the association of H3K9ac and H3K4me2 with active *rif* loci (Cabral *et al.,* 2012), steady acetylation of H4K5, H4K8, and H4K12 throughout the intraerythrocytic life cycle *(*Miao *et al.,* 2006), an association of H4K8ac with promoter regions (Gupta *et al.*, 2013), and the presence of monomethylated H4K31 at inactive, pericentromeric promoter regions (Sindikubwabo *et al.*, 2017), all in *P. falciparum*. Though known to be present as a histone modification, comprehensive genome-wide and stage-specific analyses of histone phosphorylation patterns in *P. falciparum* remain to be undertaken (Dastidar *et al.,* 2013).

1.6.4 Regulation of *Plasmodium* gene expression by histone variation

Another level of epigenetic control observed in *Plasmodium* parasites so far is the regulation of gene expression in a stage-specific manner via the exchange of canonical histones with alternative histone variants. As with other mechanisms of *P. falciparum* epigenetic regulation, the exchange of a canonical histone, in this case, *Pf*H2A, with the variant histone, *Pf*H2A.Z, was examined in the context of *var* gene antigenic variation (Petter *et al.*, 2013). During the asexual intraerythrocytic cycle, *Pf*H2A.Z associated with histone marks at transcriptionally active open chromatin regions, and in *var* genes, *Pf*H2A.Z is enriched only at transcription start sites of the active gene. Chromatin immunoprecipitation (ChIP) in tandem with high-throughput sequencing (ChIP-seq) has shown co-localisation of *Pf*H2A.Z with the active marks H3K9ac and H3K4me3 (Bártfai *et al.,* 2010). With the apicomplexan-specific histone variant *Pf*H2B.Z, *Pf*H2A.Z has also been shown to establish a H2A.Z/H2B.Z double-variant nucleosome subtype that is enriched

at AT-rich promoter regions, including that of the single active *var* gene, highlighting its correlation with open chromatin and active gene transcription in this malaria parasite (Hoeijmakers *et al.*, 2013).

Though chromatin architecture has been hitherto discussed within the context of *var* gene monoallelic transcription, recent studies have demonstrated that the dynamicity of the nucleosome landscape is fundamental to the regulation of global transcriptional activity in *P. falciparum* throughout the entirety of the life cycle. By taking advantage of the capabilities of micrococcal nuclease (MNase) and subsequent high-throughput sequencing (MNase-seq), Bunnik *et al.* demonstrated a relationship between the absence of nucleosome occupancy and active gene transcription, with a particular focus on gene promoter regions *(*Bunnik *et al.*, 2014). Furthermore, in a recent study by Kensche (Kensche *et al.*, 2015) it was shown that throughout the intraerythrocytic life cycle of *P. falciparum*, loss of nucleosome binding to transcription start sites and the structuring of the nucleosome landscape around TGCATGCA intergenic motifs could potentially be used to predict the functionality of regulatory DNA elements; an idea further developed by Lu *et al.* in their identification of putative novel genes based upon the nucleosome positioning landscape of *P. falciparum* (Lu *et al.*, 2015).

1.6.5. Epigenetic regulation by DNA methylation

No analysis of the epigenetic regulation of a species can be complete without discussing the possibility of regulation via DNA cytosine methylation. So far, we have discussed epigenetic regulation as dictated by changes in nucleosome architecture and control of the histones that make up this DNA-packaging structure. The regulation of DNA transcription by cytosine methylation at gene bodies, promoters, or enhancers is fundamental for normal mammalian development (Chen & Riggs, 2011). However, in *Plasmodium* parasites, little information has been gleaned regarding possible methylation profiles across the genome. In an early study of DNA methylation at the DHFR-TS gene in mixed blood stages of *P. falciparum* using restriction enzyme analysis, Pollack et al. (1991) highlighted a complete lack of methylation at adenine residues and partial

cytosine methylation at a putative CpG island within the gene (Pollack *et al.*, 1991).

A later study using an improved liquid chromatography/electrospray ionisation mass spectrometry (LC/ESI-MS) protocol for the detection of 5-methyl-2'-deoxycytidine (5mC) on genomic DNA (gDNA) (Friso *et al.*, 2002) disputed this previous claim, finding no evidence whatsoever for the presence of 5mC in *P. falciparum* mixed blood stage parasite DNA (Choi, Keyes and Horrocks, 2006). Choi *et al.* suggest that *Plasmodium* spp. parasites regulate expression of genes epigenetically by reversible histone modifications alone, and that the previous findings by Pollack *et al.* (1991) resulted from the inconsistency in the enzymatic activity of methyl-sensitive restriction enzymes.

Taking into consideration the methods of examining gDNA methylation used by Pollack *et al.* (1991) and Choi *et al.* (2006), and the fact that both groups derived their results from mixed *P. falciparum* cultures, it is no surprise that the question of DNA methylation as a mechanism for gene expression in *Plasmodium* was raised again in later studies. With the advent of bisulfite conversion of methylated cytosines coupled to high-throughput sequencing methods (bisulfite-sequencing or 'Bis-Seq'), Ponts *et al.* (2013) set out to map the genome-wide DNA methylation profile of *P. falciparum* using this new technique, in combination with mass spectrometry (LC-MS/MS) and *in vitro* cytosine methylation assays (Cokus *et al.*, 2008; Ponts *et al.*, 2013).

In their study, Ponts *et al.* demonstrated asymmetrical DNA methylation of the *P. falciparum* genome from an asynchronous culture, in a manner similar to undifferentiated mammal and plant cells. Moreover, hypomethylation at core gene promoters and sharp methylation changes at nucleosome and intron-exon boundaries suggest a role for this epigenetic regulation in the malaria parasite. This study also suggested differential total DNA methylation profiles between synchronous *P. falciparum* ring (0 hpi), trophozoite (16 hpi), and schizont (30 hpi) stages of the IDC, though these results were determined by hypomethylating drug assays only, with Bis-Seq and LC-MS/MS only carried out on asynchronous mixed blood stage cultures (Ponts *et al.*, 2013).

In the most recent study of quantitative detection and mapping of 5mC in *Plasmodium* gDNA, McInroy *et al.* (2016) highlighted some serious problems with the previously-utilised Bis-Seq protocol that throw the Ponts *et al.* (2013) results into question (McInroy *et al.*, 2016). McInroy *et al.* discuss the facts that the Bis-Seq protocol is incapable of distinguishing 5mC residues from 5-hydroxymethylcytosine (5hmC) residues, and that the harsh bisulfite chemical treatment can induce gDNA strand scission, leaving up to 99.9% of the DNA fragments in a Bis-Seq library preparation unsequencable (Tanaka and Okamoto, 2007; McInroy *et al.*, 2016). To generate a non-biased, more accurate methylation analysis of a *Plasmodium* parasite, McInroy *et al.* (2016) developed a novel ReBuilT (Recovery after Bisulfite Treatment) method for PCR-free Bis-Seq.

In a comparative analysis of PCR-based Bis-Seq DNA preparation, ReBuilT, and LC-MS/MS, McInroy *et al.* observed low-level asymmetric methylation throughout the *P. berghei* genome (taken from an asynchronous culture of intraerythrocytic parasites), with no preference for methylation of cytosines at CG-rich regions (a feature of mammalian genome methylation) (Li and Zhang, 2014). To conclude, McInroy *et al.* (2016) suggest that data generated from the ReBuilT protocol provides compelling evidence for the superiority of PCR-free methylation analysis as a method, but that in *P. berghei*, methylation is either not highly conserved across the IDC; present at only key active or silenced genes; or that cytosine methylation across the genome is different between the four life stages accounted for in the asynchronous culture used (ring stage, trophozoite, schizont, or gametocyte).

It is clear from these attempts at studying the *Plasmodium* spp. DNA methylome that progress has been hindered by shortcomings in DNA methylation analysis technology. The use of asynchronous cultures in all four studies (Pollack, Kogan and Golenser, 1991; Choi, Keyes and Horrocks, 2006; Ponts *et al.*, 2013; McInroy *et al.*, 2016), apart from a brief mention of *in vitro* cytosine methylation assays of synchronous *P. falciparum* ring stages, trophozoites, and schizonts by Ponts *et al.* (2013), also prevents 5mC changes from being compared between the parasite life-stages. The low level of cytosine methylation in the *Plasmodium* genome, and its asymmetric organisation, suggests that cytosine methylation of *Plasmodium*

DNA may not be as important a factor in the regulation of developmental stage transitions as, for example, transcription factors of the AP2 family (detailed in section **1.5**) or epigenetic regulation via changes to chromatin conformation (sections **1.6.2** to **1.6.4**). It remains to be seen whether direct methylation of cytosines in *Plasmodium* DNA correlates with changes to gene expression, or whether putative *Plasmodium* DNMTs show intrinsic methyltransferase activity, as was only just demonstrated in the related Apicomplexan parasite, *T. gondii* (Wei *et al.*, 2017).

1.7 Using a rodent model, *P. berghei*, to study epigenetic regulation of malaria parasites

1.7.1 Rodent models of malaria

This study was carried out using the *Plasmodium berghei* rodent model of malaria, the first rodent malaria parasite to be identified (Vincke & Lips, 1948), isolated, and maintained in a laboratory environment (Cox, 2010). Though sometimes controversial owing to the use of animal models in scientific research, *P. berghei* (in particular, the Antwerp/Kasapa (ANKA) strain) has been used extensively in malaria research to allow for more detailed investigations into multiple pathophysiological processes observed in human malaria patients (Craig *et al.*, 2012). Though ideal, studies of malaria pathogenicity using humans are limited by lack of access to relevant organs and tissue samples from infected patients; an inability to manipulate host immunity for mechanistic studies; the lethality of cerebral malaria to the host; and general ethical constraints (Langhorne *et al.*, 2011). At present, *P. berghei* infection of CBA or CB57BL/6 mice is a widely-used model of human cerebral malaria, despite different clinicopathological processes implicated in both human and murine malarial disease (White *et al.*, 2010).

In addition, studies of human-infective *Plasmodium* parasites using *in vitro* cultivation methods are limited. Continuous *in vitro* culturing systems differ substantially from the human host *in vivo* environment. *In vitro P. falciparum* cultures lack the temperature fluctuations, immune environment, shear force inside the host circulatory system, as well as additional metabolites and hormones

unique to the *in vivo* surroundings (LeRoux *et al.*, 2009). Appropriate *in vitro* cultivation techniques to accurately mimic the host microvasculature and to enable the preservation of *P. falciparum* multiplication, invasion and rosetting phenotypes (significant contributing factors in human malaria pathology) have only recently been optimized and necessitate the addition of human serum and maintenance of a controlled micro-aerophilic environment (Ribacke *et al.*, 2013). In the cases of the four remaining human malaria parasites, *P. vivax, P. knowlesi, P. malariae,* and *P. ovale* (both subspecies), continuous *in vitro* culturing conditions have been attempted with limited success (Schuster, 2002). The human malaria parasite with the highest morbidity and mortality outside of *P. falciparum*, *P. vivax*, has proven difficult to maintain consistently *in vitro* (Bermúdez *et al.*, 2018).

Rodent models of malaria also have the advantage of enabling immunological experiments as a result of the availability of inbred and congenic rodent strains with gene-targeted and well-defined natural immune deficits or conditions. Outside of *P. berghei* ANKA as a rodent model of the entire *Plasmodium* life-cycle, and as a model for human cerebral malaria, the three remaining rodent *Plasmodium* species (*P. chabaudi*, *P. yoelii*, *P. vinckei,* and their strains) display differences in parasite biology and pathogenicity in their hosts that can be suited to particular areas of malaria research. For example, *P. yoelii yoelii* strain YM is a lethal rodent parasite that infects both reticulocytes and mature red blood cells (RBCs), and is used to test vaccine candidates; whereas *P. chabaudi chabaudi* clone AS infections in BALB/c mice have been used as a model for severe malarial anaemia (Lamb & Langhorne, 2008; Wykes & Good, 2009).

1.7.2 Methods of *P. berghei* genetic manipulation used in the present study

The first genetic manipulation of a malaria parasite occurred with the expression of firefly luciferase in the bird malaria parasite, *P. gallinaceum,* following transfection of plasmid DNA into zygote and gamete parasites via electroporation (Goonewardene *et al.*, 1993). Transfection of DNA into the haploid genome of the rodent *P. berghei* parasite followed soon after, with the additional transfection of a gene locus encoding a drug-selectable marker (a drug-resistant form of the

dihydrofolate reductase-thymidylate synthase (DHFR-TS) enzyme) allowing genetically modified parasites to be selected based on their resistance to the drug, pyrimethamine (van Dijk, Waters & Janse, 1995). In contrast to *P. falciparum, P. berghei* parasites cannot be cultured long-term *in vitro,* and so drug selection must be carried out only after reinfection of a rodent with live, electroporated merozoites. Since this first transfection, which resulted in the presence of stably-expressed, though episomally-located, drug-resistant plasmids, many more tools for genetic manipulation of malaria parasites have become available, often developed in the *P. berghei* or *P. falciparum* malaria parasites, and sometimes in tandem (de Koning-Ward *et al.,* 2015).

After the development of transfection capabilities in *P. berghei,* came many more techniques for genetic manipulation of malaria parasites: targeted gene deletion (or 'knockout' (KO)) (Ménard *et al.,* 1997; Crabb *et al.*, 1997); stable expression of fluorescent reporters (de Koning-Ward *et al.,* 1998; de Koning-Ward *et al.,* 1999); complementation of gene KOs (Sultan *et al.,* 2001); further advances in drug selection cassettes and transfection (Janse *et al.,* 2006); and a host of techniques for conditional knockdown (KD) of genes and proteins (Carvalho *et al.*, 2004; Fonager *et al.*, 2011; Lin *et al.*, 2011; Pfander *et al.*, 2011; Pino *et al.*, 2012; Wagner *et al.*, 2014; Philip and Waters, 2015; Hughes and Waters, 2017). A selection of these methods were employed in the present study to investigate the function (or lack thereof) of a number of putative epigenetic regulators in the *P. berghei* genome. These methods are detailed below.

1.7.2.1 Reverse genetics in *P. berghei*: positive and negative drug selection

Much of our current knowledge of the molecular biology of the malaria parasite, *Plasmodium*, has been gleaned as a result of reverse genetics approaches, i.e. the investigation of a phenotype that arises *after* a known genetic change has been introduced, such as a complete gene knockout (KO). In contrast, forward genetics describes a study in which a known phenotype is examined in an attempt to determine the underlying genetic changes that arose *before* the phenotype was identified (Alonso and Ecker, 2006).

The finding that revolutionised the study of *Plasmodium* parasites was the discovery that mutations in the bifunctional dihydrofolate reductase (DHFR) and thymidylate synthase (TS) (DHFR-TS) gene in *P. falciparum* resulted in resistance of the parasite to the DHFR reductase inhibitor, Pyrimethamine (Bzik *et al.*, 1987; Cowman *et al.*, 1988; Peterson, Walliker and Wellems, 1988; Zolg *et al.*, 1989; Sirawaraporn *et al.*, 1990). The antifolate antimalarial drug, Pyrimethamine, acts by binding and inhibiting the function of the DHFR-TS enzyme, therefore depleting deoxythymidine monophosphate (dTMP) production and disrupting DNA synthesis (Basco *et al.*, 1995). A single amino acid substitution at position 108 of the DHFR-TS amino acid sequence, a serine-to-asparagine change, was all that was necessary to confer Pyrimethamine resistance to the parasite, with the addition of alanine-to-valine substitution at position 16 also conferring resistance to a second antifolate antimalarial, Proguanil (Foote, Galatis and Cowman, 1990). Further investigations also revealed that dual alanine-to-valine (position 16) and serine-to-threonine (position 108) substitutions conferred resistance in *P. falciparum* to Proguanil but not Pyrimethamine. However, significant cross-resistance to both drugs also occurred in the presence of dual serine-to-asparagine (position 108) and isoleucine-to-leucine (position 164) mutations (Peterson, Milhous and Wellems, 1990).

The ability to confer drug resistance to the *Plasmodium* parasite via single point mutations to the DHFR-TS gene raised the possibility that a modified DHFR-TS gene could be used as a drug-selectable marker in future genetic studies of the parasite, a feat accomplished by van Dijk, Waters, and Janse, a number of years later (van Dijk, Waters and Janse, 1995). From this point forward in *P. berghei*, stable transfection of circular plasmids containing a DHFR-TS gene with a single serine-to-asparagine mutation in position 110 of the codon sequence (orthologous to the *P. falciparum* position 108), resulting in resistance of the parasite to Pyrimethamine, was used for genetic manipulation of the parasite. Although integration of linear vectors into the *P. berghei* genome via homologous recombination was attempted, at this time Pyrimethamine resistance was conferred through episomal replication of the mutated DHFR-TS gene (van Dijk, Waters and Janse, 1995).

Since these studies, *Plasmodium* biology has benefitted greatly from further advances in both positive and negative drug selection protocols in *P. berghei*, and the ability to generate parasite lines in which drug selectable markers are integrated into the parasite genome (de Koning-Ward *et al.*, 2000). In a positive drug selection protocol, parasites are selected that are resistant to a drug, for example, through introduction of a resistance-conferring gene to the parasite genome. In a negative drug selection protocol, drug sensitivity is conferred to a parasite through genetic manipulation and so parasites are killed upon drug administration.

In *P. berghei*, three positive selectable markers are available: the mutated *P. berghei* DHFR-TS gene (*pbdhfr-ts*), conferring resistance to Pyrimethamine (van Dijk, Waters and Janse, 1995); the mutated *T. gondii* DHFR-TS gene (*tgdhfr-ts*), which also confers resistance to Pyrimethamine (Donald and Roos, 1993); and the human dihydrofolate reductase gene (*hdhfr*), that is innately resistant to both the antimalarial antifolate drug, Pyrimethamine, and a second antifolate inhibitor, WR99210 (de Koning-Ward *et al.*, 2000). For negative selection in *P. berghei,* only one option is available: transgenic expression of a bifunctional gene containing the fused yeast cytosine deaminase and uridyl-phosphoribosyltransferase enzymes (*yfcu*). With the expression of the yFCU bifunctional protein in *P. berghei*, administration of the prodrug 5-fluorocytosine (5-FC) to the *P. berghei*-infected mice results in the conversion of 5-FC into the cytotoxic compound 5-fluorouracil (5-FU) by cytosine deaminase, and the subsequent conversion of this 5-FU to 5-fluoro-uridine monophosphate (5-fluoro-UMP) by uridyl-phosphoribosyltransferase. The combined effects of these enzymes is to inhibit RNA synthesis (Braks *et al.*, 2006; Maier *et al.*, 2006).

Despite what appears to be a limited repertoire of drug selectable markers, strategies for positive- and negative- selectable marker 'recycling' have been developed in *P. berghei*. With the discovery that the *yfcu* negative selectable marker could be used in *P. berghei*, and that 5-FC was safe to administer to infected rodents in drinking water, came the development of positive-negative selectable marker systems (Braks *et al.*, 2006; Orr, Philip and Waters, 2012). By

introducing both a positive- and a negative- selectable drug marker together, a number of outcomes are possible.

Taking pyrimethamine (*hdhfr*) and yFCU (*yfcu*) as examples; a transfected parasite expressing hdhfr::yfcu can initially be selected for by Pyrimethamine. All non-transfected parasites are killed and only successfully-transfected parasites remain. Negative selection pressure is then applied by administering 5-FC to the parasites *in vivo* (Orr, Philip and Waters, 2012). All *hdhfr::yfcu*-containing parasites are then killed. The only surviving parasites following negative selection will be those in which the *hdhfr::yfcu* marker has been excised by homologous recombination around the drug selection cassette. This strategy made possible the integration of a modified locus (e.g. transgene, gene knockout, tagging epitope) in *P. berghei* that could be selected for by Pyrimethamine administration (positive selection) and then the subsequent generation of a marker-free, modified *P. berghei* line by negative selection with 5-FC. This ability to 'recycle' drug selectable markers meant that, for the first time, multiple sequential genetic manipulations could be made to this rodent malaria parasite, including complementation studies (Sultan *et al.*, 2001; Braks *et al.*, 2006; Orr, Philip and Waters, 2012). However, negative selection with 5-FC does not necessarily guarantee a population of parasites in which the *hdhfr::yfcu* cassette has been 100% excised (perhaps the *yfcu* sequence alone was excised). To confirm a purely marker-free *P. berghei* line, cloning by limiting dilution *in vivo* was subsequently carried out, followed by PCR or Southern blot analysis to demonstrate loss of the drug cassette (Orr, Philip and Waters, 2012; Manzoni *et al.*, 2015) (**Figure 1.15**).

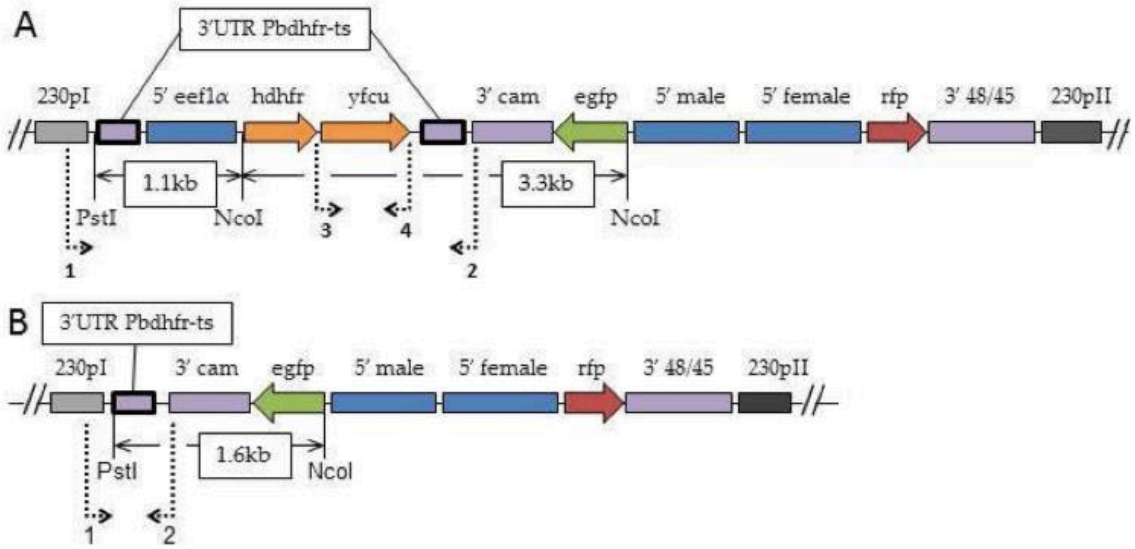


Figure 1.15: A modified *P. berghei* genetic locus before and after positive and negative drug selection. This image from Orr, Philip, and Waters (2012) shows the integration of plasmid pL1186 into the *P. berghei p230p* locus before (**A**) and after (**B**) both positive and negative drug selection. The *p230p* integration locus (an upstream paralogue to the P230 6-cysteine male gametocyte surface protein) is a *P. berghei* gene locus with a dispensable role during normal life cycle progression and is thus used for stable integration and expression of transgenes (Thompson, Janse and Waters, 2001; van Dijk *et al.*, 2010; Matz and Kooij, 2015). The *P. berghei p230p* locus is split into three sites, each one a possible site for vector integration: *p230p*I, *p230p*II, and *p230p*III. Sites *p230p*I (230pI) and *p230p*II (230pII) are labelled in the above image with the pL1186 plasmid inserted into the *p230p*I site (upstream of *p230p*II). In (A), the *hdhfr::yfcu* fusion gene is shown in orange and is flanked upstream by a 3'UTR of the *P. berghei* mutated *dhfr-ts* gene (3'UTR Pbdhfr-ts) and a 5' promoter of the constitutively-expressed *P. berghei* elongation factor 1 alpha (5'eef1α), and downstream by a second 3'UTR Pbdhfr-ts sequence (de Koning-Ward, Janse and Waters, 2000; Franke-Fayard *et al.*, 2004). Downstream of the drug selection cassette is the desired transgene to be integrated into *P. berghei*. This transgene contains a 3'UTR from the *P. berghei* calmodulin gene (3'cam), an enhanced green fluorescent protein gene (*egfp*) being driven by the male gametocyte-specific heavy dynein chain promoter (upstream of PBANKA_0416100) (5'male), a red fluorescent promoter (*rfp*) being driven by the female gametocyte-specific LCCL-domain-containing gene promoter (upstream of PBANKA_1319500) (5'female), and the 3'UTR of the *p48/45* gene (3'48/45) (Ponzi *et al.*, 2009). Fragment sizes (in kb) that resulted upon restriction digest are indicated (for enzymes PstI and NcoI) and the sites of primers for integration PCR analyses are indicated by the numbers 1 to 4. (**A**) depicts the *P. berghei* locus as it would appear in a successfully-transfected parasite before and after positive selection pressure by Pyrimethamine. If transfection had not been successful, the parasite would be killed under Pyrimethamine administration.

(**B**) depicts the same *P. berghei* locus after negative selection with 5-FC. In this depiction, the drug selection cassette has been excised from the parasite genome. Parasites with the integration locus in this conformation only (or with an *hdhfr* sequence only) would survive 5-FC negative selection. PCR and restriction fragment analyses could then be used to confirm the loss of the

hdhfr::yfcu fusion gene, with retention of the desired transgene. Source: Image adapted from (Orr, Philip and Waters, 2012).

1.7.2.2 Monitoring the production of male and female gametocytes in *P. berghei*: the '820' background line

As shown in figure **1.15** as an example of a construct containing both positive and negative drug-selectable markers, the construct used to create the '820cl1m1cl1' *P. berghei* line warrants further discussion as a result of its pivotal role in the investigation of mechanisms underlying commitment to gametocytogenesis, and differences between male and female gametocyte conversion and maturation (Ponzi *et al.*, 2009; Sinha *et al.*, 2014; Yuda *et al.*, 2015; Yeoh *et al.*, 2017).

Used extensively to study gametocytes in *P. berghei* ANKA, the 820cl1m1cl1 line was introduced originally by Ponzi *et al.* (2009) as a means of examining the role of the MDV1/PEG3 (male development protein 1/protein of early gametocyte 3) protein in egress of both male and female gametes from the host erythrocyte (Ponzi *et al.*, 2009). As depicted in figure **1.15** (**A**), an enhanced green fluorescent protein sequence (*egfp*) is driven in this construct by the promoter of the male-specific dynein heavy chain protein (encoded by PBANKA_0416100), discovered as highly male-specific in a proteomic study of separated male and female *P. berghei* gametocytes (Khan *et al.*, 2005). In this way, when commitment to male gametocytogenesis begins, the GFP fluorescent reporter is also expressed in addition to the native dynein heavy chain protein, thus separating male sexual-stage parasites from asexual parasites or female gametocytes by its green fluorescence.

In a similar manner to the male gametocyte fluorescence strategy, the 820cl1m1cl1 line construct contained a red fluorescent protein sequence (*rfp*) that is driven by the promoter of the predominantly female-specific Limulus factor C, Coch-5b2 and Lgl1 (LCCL) domain-containing protein, CCp2, (encoded by PBANKA_1319500), a female-specific protein again identified by proteomic analysis (Khan *et al.*, 2005). The newly transfected line with the '820' DNA construct integrated at the p230pl locus, as depicted in figure **1.15** (**A**), was

positively selected by Pyrimethamine and cloned by limiting dilution to generate the '820 clone 1' *P. berghei* ANKA line (shortened to '820cl1'). This line was then negatively selected by 5-FC in another mouse (mouse 1) and cloned once more by limiting dilution to give the '820cl1 mouse 1, clone 1' (called '820cl1m1cl1') *P. berghei* line, a drug-selectable marker-free parasite line in which male gametocytes express GFP, and female gametocytes express RFP (as depicted in figure **1.15** (**B**) (Ponzi *et al.*, 2009).

Though highly enriched in male and female gametocytes respectively, it is important to note that both dynein heavy chain protein and CCp2 *are* present in parasites of the opposite sex, and so one must be careful when sexing parasites at early stages of commitment, when fluorescence of GFP and RFP are faint (Khan *et al.*, 2005). In fact, fluorescence intensity is dependent upon gametocyte maturation, with early gametocytes having a fainter fluorescence than mature gametocytes (personal observation).

1.7.2.3 Generating conditional gene knockout and protein knockdown systems in *P. berghei* for examination of putative epigenetic regulators

1.7.2.3.1 Dimerisable Cre Recombinase (DiCre)-based inducible gene knockout or expression

Site-specific recombination in the *P. berghei* genome can be achieved by using either a Flp/FRT system (Falae *et al.*, 2010; Lacroix *et al.*, 2011), or the analogous Cre/*loxP* recombination system ('Cre' standing for 'creates recombination' and '*loxP*' derived from 'locus of crossover [x] in P1 bacteriophage') (Kühn and Torres, 2002). In both systems, a recombinase enzyme must be expressed and, in *P. berghei*, the recombinase must be expressed in background ('mother' or 'deleter') parasite lines before a target construct containing a FRTed or floxed (flanked by *loxP* sites) DNA construct is integrated into the genome by recombination. In the Flp/FRT system, Flp recombinase is generally stage-specific as constitutive recombinase expression would result in a target DNA sequence being excised or 'flipped', or otherwise manipulated, at every parasite stage.

To overcome the limitations of these stage-specific recombination strategies, a new approach was developed in which the Cre recombinase enzyme was split into two inactive moieties that were then fused to an FK506-binding protein (FKBP12) (Siekierka *et al.*, 1989) and the ligand-binding domain of FKBP12-rapamycin associated protein (FRB) (Chiu, Katz and Berlin, 1994) respectively. These fusions led to the production of a ligand-inducible dimerisable Cre (DiCre) recombinase that could be expressed at a desired time point upon rapamycin administration (Jullien *et al.*, 2007). It was not long before this inducible site-specific recombination system was adapted for use in Apicomplexa, including *Plasmodium* (Andenmatten *et al.*, 2013; Collins *et al.*, 2013).

To begin with, inducible DiCre expression was used as a means to conditionally knock out genes of interest in *P. falciparum* by excision of floxed sequences of interest (Collins *et al.*, 2013; Yap *et al.*, 2014). However, using the inducible DiCre site-specific recombination system, in combination with a pair of asymmetrical *loxP* sites (*lox66* and *lox71*) that highly favour unidirectional inversion of a targeted DNA sequence (Oberdoerffer, 2003), rapamycin-induced, DiCre-based overexpression of a gene of interest was undertaken in *P. berghei* (**Figure 1.16**). To be specific, the unidirectional DiCre-based recombination was used as a means to induce gametocytogenesis in the parasite by overexpression of the *pbap2-g* transcriptional regulator (Kent *et al.*, 2018).

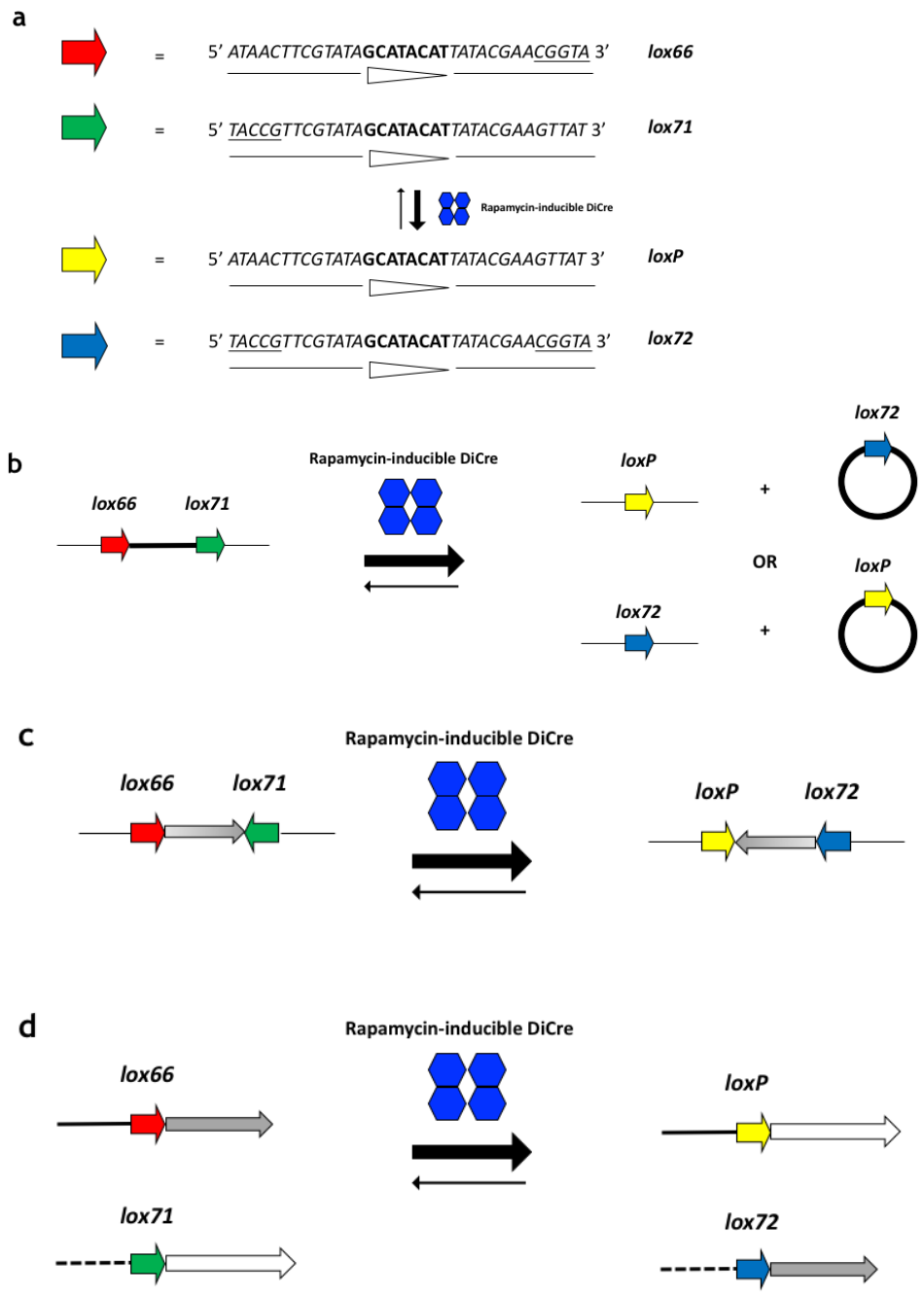


Figure 1.16: Strategies available for rapamycin-inducible genetic manipulation using a predominantly unidirectional DiCre-based recombination system. Drawn in the style of Lacroix *et al.*, 2011 (in their explanation of possibilities for gene manipulation using the Flp/FRT system in *Plasmodium*), the above image shows how a target gene (or genes) can be manipulated using the rapamycin-inducible DiCre system (Andenmatten *et al.*, 2013) in combination with the predominantly unidirectional (so far shown to be experimentally irreversible) pair of *loxP* sequences, *lox66* and *lox71* (Zhang and Lutz, 2002; Oberdoerffer, 2003). (**a**) shows the 5' to 3' 34 bp nucleotide sequences of *lox66*, *lox71*, *loxP*, and *lox72*, with corresponding coloured arrows that are used to indicate these sequences in (**b**) to (**d**). In these sequences, the 8 bp 'spacer' sequence is indicated in bold letters, with the underlying white arrowhead indicating the orientation of this sequence. The remaining semi-palindromic sequences flanking the 'spacer' sequence are italicised

and underscored by a straight line. Where present, a mutated 5 bp sequence is indicated in the semi-palindromic sequences by being underlined. A single mutated 5 bp sequence is present in both *lox66* and *lox71*. Following rapamycin-induced conglomeration of a complete dimerisable Cre recombinase (DiCre) (indicated by a blue tetramer), the semi-palindromic sequences of *lox66* and *lox71* are cut, with recombination resulting in the *loxP* and *lox72* sequences. DiCre-mediated recombination results in a *loxP* sequence with no 5 bp inverted repeat and a *lox72* sequence containing two mutated inverted repeats. In the presence of DiCre recombinase, the *lox66* and *lox71* to *loxP* and *lox72* sequence recombination event is highly favoured over the reverse recombination event (as indicated by bold and thin arrows). As a result, induction of an almost irreversible genetic manipulation is possible by flanking genes of interest with *lox66* and *lox71* sequences in a background cell line that expresses rapamycin-inducible DiCre.

(**b**) depicts the excision of a gene by the DiCre-*lox66*/*lox71* system. The target gene is flanked by *lox66* and *lox71*. Recombination after DiCre induction (with both 'spacer' sequences in the same orientation) will result in excision of a gene with a *loxP* or *lox72* sequence remaining in the genome.

In **(c),** unidirectional inversion of a sequence of interest is depicted. Inversion is mediated by flanking the target sequence with *lox66* and *lox71,* with 'spacer' sequences present in opposite orientations. DiCre recombination then results in an inverted sequence that is flanked by *loxP* and *lox72* sites.

Part (**d**) of the above image depicts a genetic manipulation that is possible, but has yet to be undertaken in a *Plasmodium* parasite. In this case, two different linear sequences, such as target sequences of adjacent chromosomes are flanked upstream by a *lox66* site (one sequence) and a *lox71* sequence (second target sequence). Rapamycin-induced DiCre recombination would then result in a predominantly unidirectional switching of *lox66* and *lox71* sites and their downstream sequences. The resulting linear sequences contain a *loxP* and a *lox72* site. All of the above manipulations are almost irreversible, but in each case large, bold arrows are accompanied by an inverted thin arrow to highlight the fact that, though unlikely, recombination in reverse is not impossible and should therefore be considered.

In *P. berghei* (ANKA strain), unidirectional recombination using the DiCre-*lox66*/*lox71* system was used as depicted in figure **1.16 (c)** to generate a parasite line called 'G1142' (Glasgow 1142) (Kent *et al.*, 2018). In this background parasite line, rapamycin-induced DiCre recombination resulted in inversion of a potent *P. berghei* heat shock protein 70 (*hsp70*) promoter (encoded for upstream of PBANKA_0711900) (Manzoni *et al.*, 2015) which had been placed upstream of the *pbap2-g* gene (**Figure 1.17).** As a result, rapamycin-induced DiCre recombination was used to overexpress *Pb*AP2-G, therefore having the overall phenotypic effect of inducing gametocytogenesis in *P. berghei* (Kent *et al.*, 2018).

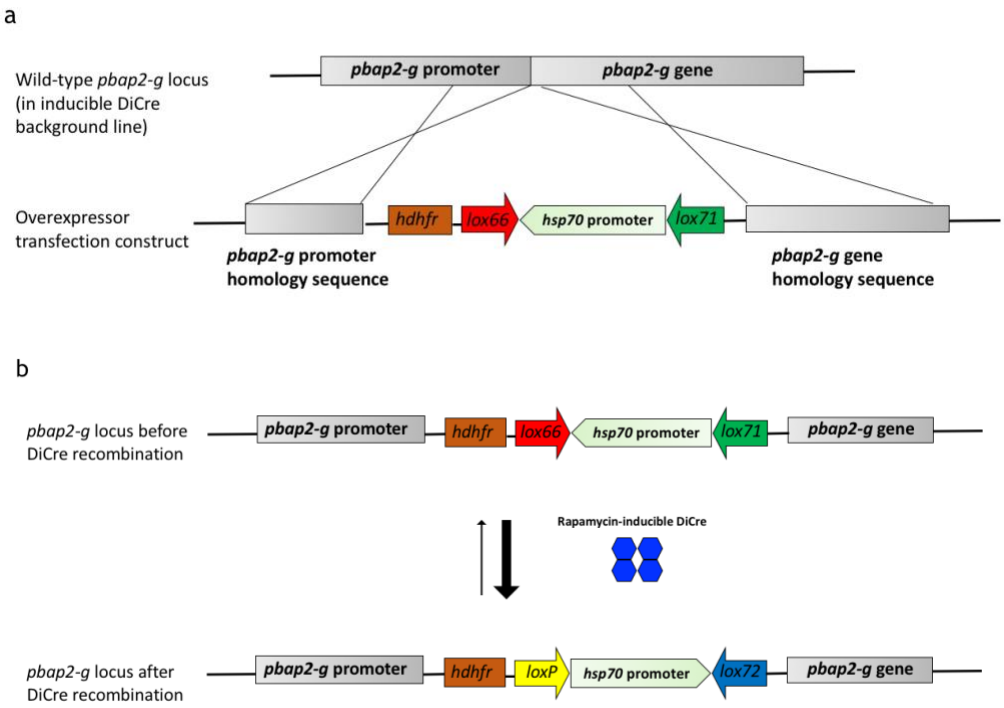


Figure 1.17: Inducible *pbap2-g* expression in *P. berghei* ANKA using the rapamycin-inducible DiCre-*lox66/lox71* recombination system. In this image, (**a**) depicts the *pbap2-g* locus in a background *P. berghei* ANKA line that already contains the sequences for both moieties of the rapamycin-inducible DiCre recombinase enzyme (at the *p230p* locus), in addition to the DNA construct used to introduce a drug-selectable marker and floxed *hsp70* promoter upstream of the *pbap2-g* gene. The 'overexpressor' construct is integrated into the background *P. berghei* DiCre line by homologous recombination with the wild-type *pbap2-g* locus. In this construct, a silent (inverted) *hsp70* promoter is flanked by *lox66* and *lox71* sequences in opposite orientations. (**b**) shows the *P. berghei* locus after the 'overexpressor' construct has been integrated (and selected for by Pyrimethamine). Upon rapamycin-induced DiCre recombination, the *lox66* and *lox71* sites are cut and recombined to generate an active *hsp70* promoter that is flanked by *loxP* and *lox72* sequences respectively. Thick and thin arrows in this image are used to depict the predominantly unidirectional recombination event. The likelihood after DiCre recombination therefore, is that the *hsp70* promoter drives downstream *pbap2-g* expression in the so-called 'G1142' *P. berghei* line, inducing gametocytogenesis.

1.7.2.3.2 The auxin-inducible degron (AID) system

Originally developed using a family of plant hormones (auxins) produced by *Arabidopsis thaliana*, the auxin inducible degron (AID) system was adapted for other eukaryotes by Nishimura *et al.* (2009) for targeted conditional depletion of proteins (Nishimura *et al.*, 2009). In plants, hormones of the auxin family, which

includes the endogenous indole-3-acetic acid (IAA), or synthetic auxins such as 1-naphthalenacetic acid (NAA), act by binding to the so-called 'F-box' transport inhibitor response protein 1 (TIR1) (Ruegger *et al.*, 1998). Auxin-binding to TIR1 promotes (or enhances) the interaction of the E3 ubiquitin ligase SCF-TIR1 (SCF standing for Skp1, Cullin and F-box complexes) with native plant auxin/IAA transcriptional repressors (Tan *et al.*, 2007). The SCF-TIR1 complex recruits an E2 ubiquitin-conjugating enzymes that polyubiquitylates the auxin/IAA transcriptional repressor, marking it for degradation by the 26S proteasome (Teale, Paponov and Palme, 2006).

In *P. berghei*, the AID system takes advantage of the presence of the highly conserved SCF-complex-mediated protein degradation pathway. In *P. berghei* ANKA parasites, the introduction of a constitutively expressed thermostable TIR1 protein from the rice plant *Oryza sativa* (*Os*TIR1), under control of the potent *hsp70* promoter, enables the targeting of proteins of interest (POIs) for degradation if they are tagged with an auxin-inducible degron (i.e. an auxin/IAA response protein). In *P. berghei,* the AID tag used was the sequence of the *A. thaliana* auxin-responsive protein, IAA17 (gene IAA17_ARATH) (Nishimura *et al.*, 2009; Philip and Waters, 2015). Therefore, when an auxin hormone is administered to *P. berghei in vitro,* the fused POI::IAA17 complex interacts with the SCF-TIR1 E3 ubiquitin ligase, resulting in proteosomal degradation of the target protein (**Figure 1.18**). The short half-life of an IAA17-fused protein (8-10 minutes in *A. thaliana* seedlings) (Dreher, 2006; Nishimura *et al.*, 2009), and the need for auxin addition to elicit the degradation response, means that the AID system in *P. berghei* is useful for rapid, and reversible, conditional degradation of proteins (Philip and Waters, 2015). The reversibility of the AID system is very useful for functional characterisation of *Plasmodium* proteins in that a response phenotype can be observed and then restoration of the target protein investigated in the

same experiment upon removal of auxin from parasite media.

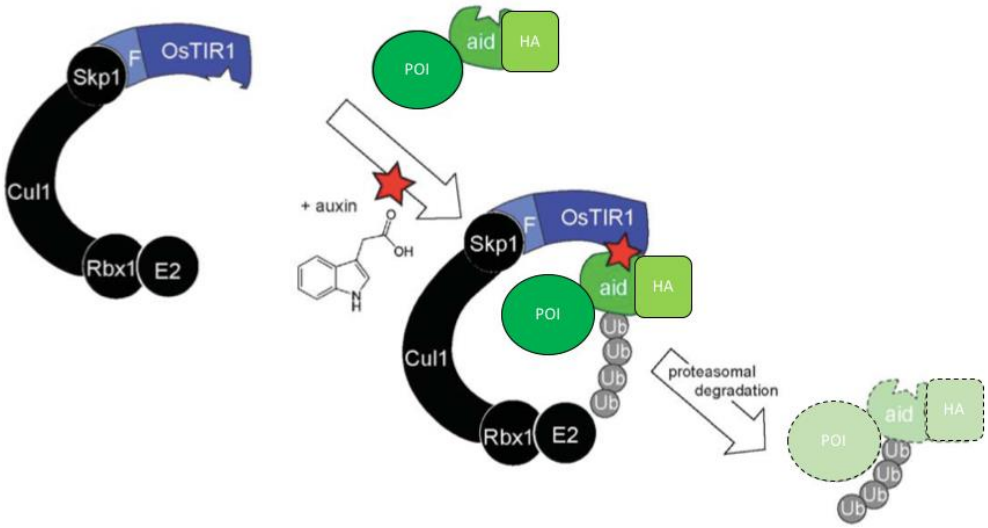


Figure 1.18: Conditional degradation of proteins of interest in *P. berghei* using the auxin-inducible degron (AID) system. This image shows the SCF (Skp1, Cullin, F-box)-TIR1 E3 ubiquitin ligase complex (containing a RING-box protein 1 (Rbx1) E3 ubiquitin ligase component), bound to an E2 ubiquitin-conjugating enzyme before auxin addition, and then after auxin addition (+ auxin). In the presence of an auxin hormone (indole-3-acetic acid is shown here), the SCF-TIR1-E2 complex binds to an auxin-inducible degron (aid)-tagged protein of interest (POI). In this image, the POI::aid complex is also attached to a haemagglutinin (HA) tag for further experiments and Western blot analyses. Auxin-induced interaction of both complexes results in the addition of ubiquitin (Ub) proteins to the tagged protein, resulting in its degradation in the proteasome. OsTIR1: *Oryza sativa* transport inhibitor response 1 protein; F: F-box protein; Cul1: Cullin 1 protein. Sourse: Image adapted from (Trost, Blattner and Lehner, 2016).

1.7.2.3.3 The *Plasmodium* Genetic Modification (*Plasmo*GEM) and Rodent Malaria Genetically Modified Database (RMgmBD) resources

A considerable number of methods for the generation of genetically modified *P. berghei* lines are available, and many have been applied to specific proteins of interest in this rodent malaria parasite. Such an array of mutated *P. berghei* lines necessitates a platform by which knockout or knockdown lines can be searched, or by which *P. berghei* proteins can be searched to prevent experiments being duplicated unnecessarily. Fortunately, a number of databases do exist for this very purpose: the online *Plasmodium* Genomics Resource, PlasmoDB (Bahl *et al.*, 2003); the Rodent Malaria genetically modified DataBase (RMgmDB) (Khan *et al.*, 2013); the malaria parasite phenotype database, PhenoPlasm (Sanderson and

Rayner, 2017); and the *Plasmodium* Genetic Modification (*Plasmo*GEM) database (Schwach *et al.*, 2015).

Of note among these databases are the RMgmDB and *Plasmo*GEM repositories, as both of these databases not only allow a search of *P. berghei* KO lines and phenotypes, but provide means by which vectors for genetic manipulation of the *P. berghei* parasite, and specific mutant *P. berghei* lines, can be ordered for further studies. In fact, the *Plasmo*GEM resource incorporates a high-throughput vector production pipeline that couples the generation of a bacteriophage N15-based genome-scale plasmid library (Pfander *et al.*, 2011) with Red recombinase-based engineering ('recombineering') (Pfander *et al.*, 2013) to produce barcoded *P. berghei* modification vectors on a large scale (Schwach *et al.*, 2015).

At present, the *Plasmo*GEM database includes nearly 2000 vectors for the complete or partial deletion of *P. berghei* genes, over 200 vectors for C-terminal-tagging of proteins of interest with a triple haemagluttinin (3xHA) tag, and a near genome-scale *P. berghei* artificial chromosome (*Pb*AC) vector resource (Schwach *et al.*, 2015; *Plasmo*GEM, 2018). Production of all of these DNA constructs begins with the selection of the genomic DNA (gDNA) insert of interest from a library of bacteriophage N15-based pJAZZ-OK vectors that cover over 95% of the *P. berghei* ANKA genome in a fragmented form (Godiska *et al.*, 2009; Pfander *et al.*, 2011). These linear kanamycin-resistant pJAZZ-OK vectors are capable of propagating the AT-rich *P. berghei* inserts stably in *Escherichia coli,* with the resultant arrayed gDNA library being termed *Pb*GO1 (Pfander *et al.*, 2011).

For the creation of a gene knockout vector, a PCR product containing a bicistronic bacterial drug-selectable cassette (*zeo-PheS*) and 50 bp homology regions at either side of the gene-of-interest (GOI) is transformed into the *E. coli Pb*GO1 library clone via *lambda* Red/ET recombineering (Wang *et al.*, 2006; Pfander *et al.*, 2013). At this point, positive drug selection selects for successfully-transformed *E. coli* containing the intermediate vector. The *zeo-PheS* cassette is flanked by *attR1* and *attR2* sites of the Gateway LR Clonase system which are used to direct replacement of the bacterial drug selection cassette with a larger *P. berghei* negative-positive drug selection cassette upon activation of the Gateway

recombinase reaction. To select for bacteria containing the correctly-transformed final knockout vector, negative drug selection with *p*-chlorophenylalanine is carried out (Pfander *et al.*, 2013) (**Figure 1.19**).

To carry out transfection of a *Plasmo*GEM knockout vector to a *P. berghei* parasite line, the transfection vector (contained within kanamycin-resistant *E. coli*) is grown until sufficient DNA for transfection is generated. Isolation and purification of DNA, followed by a *NotI* restriction digest, releases the final knockout vector from the remaining pJAZZ-OK vector backbone and renders it capable of integration into the *P. berghei* ANKA genome by homologous recombination (Pfander *et al.*, 2011; Pfander *et al.*, 2013) (**Figure 1.19**).

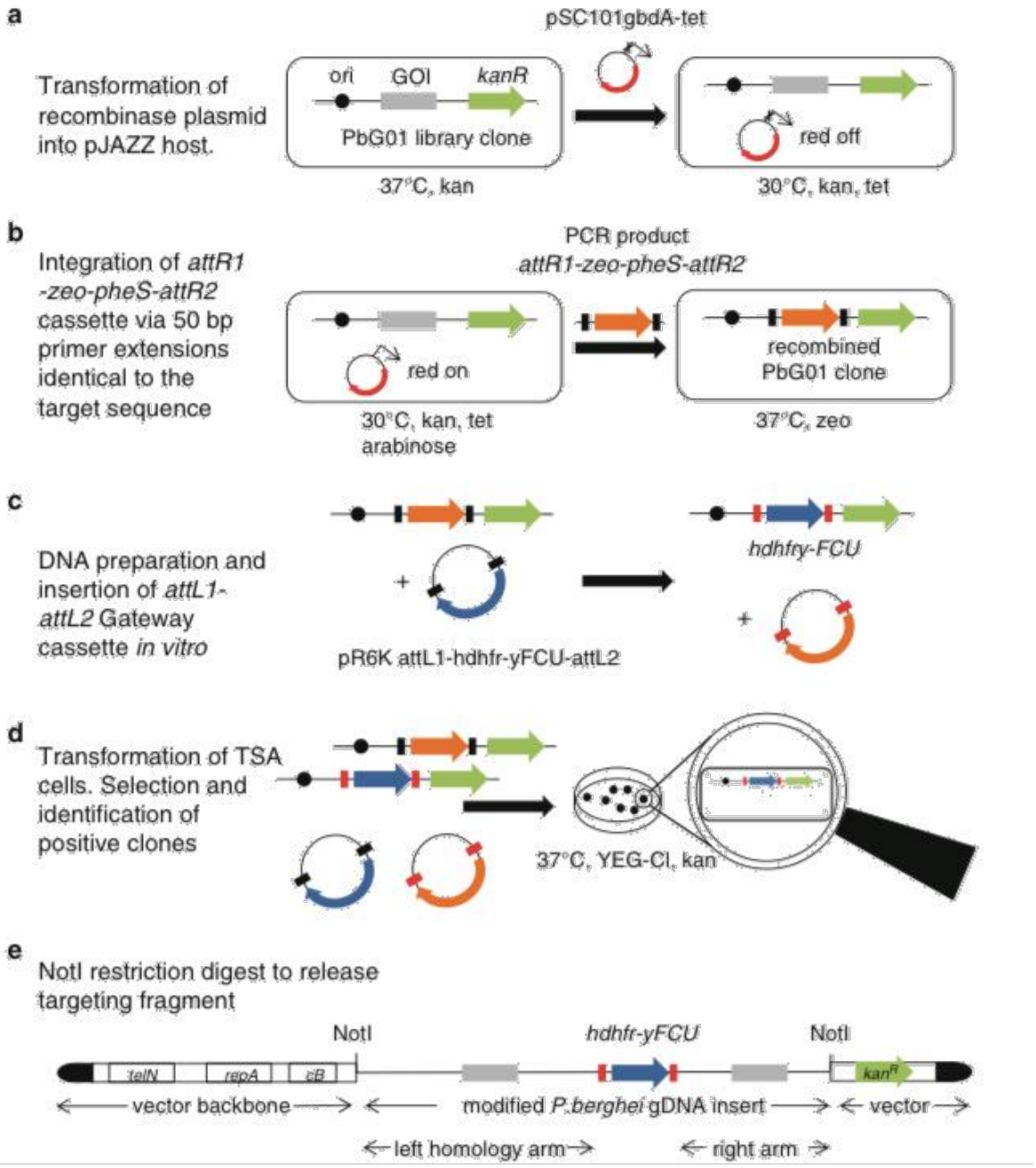


Figure 1.19: Large-scale, recombination-mediated genetic engineering of *Plasmo*GEM *P. berghei* ANKA knockout vectors. The above image, reproduced from Pfander *et al.*, 2013, depicts the process of creating a *Plasmo*GEM vector from selection of the *Pb*GO1 library clone (in a pJAZZ-OK vector) (**a**) to the final product that can be transfected into a *P. berghei* ANKA line after *NotI* restriction digest (**e**). (**a**) shows the gDNA library (*Pb*GO1) clone, a linear pJAZZ-OK vector (inside an *E. coli* cell) that contains the GOI insert, a kanamycin-resistance gene (*kan*R), and an *E. coli* origin of replication (ori) site. These cells are grown up at optimum temperature (37°C) and a pSC101gbdA-tet plasmid containing an arabinose-inducible recombinase (shown in red) is transformed into the *E. coli* cell by electroporation. *E. coli* colonies are then grown at the optimum pSC101gbdA-tet plasmid temperature (30°C) under both kanamycin and tetracycline drug selection. (**b**) shows propagation of the pSC101gbdA-tet-transformed *E. coli* once the recombinase is 'switched on' by the addition of arabinose. After arabinose induction of the recombinase

enzyme, the cells are transformed with a targeted *att*R1-*zeo-PheS-att*R2 PCR product, i.e. the bacterial positive-negative drug selection cassette flacked by Gateway LR Clonase *att* sites. Synthetic oligonucleotides with 5' 50 bp extensions homologous to the GOI flanking sequences are used to produce the *att*R1-*zeo-PheS-att*R2 PCR product so that it can be integrated into the pJAZZ-OK vector by homologous recombination, replacing a large section of the GOI insert. Recombined *E. coli* clones are then positively-selected under zeocin pressure at 37°C (zeocin resistance conferred by the '*zeo*' fragment). At (**c**), the linear pJAZZ-OK vector of interest is isolated from *E. coli* and combined with the Gateway circular plasmid, pR6K *att*R1-*hdhfr-yfcu-att*R2. This pR6K plasmid contains the *P. berghei* positive-negative *hdhfr::yfcu* drug cassette which is recombined into the pJAZZ-OK vector at *att*R1 and *att*R2 sites, replacing the bacterial drug cassette. In (**d**), both pJAZZ-OK vectors (containing the bacterial and *P. berghei* drug selection cassettes) are transformed into Lucigen TSA electrocompetent *E. coli* cells. These TSA *E. coli* cells are then plated onto agar containing both kanamycin and *p*-chlorophenylalanine (YEG-Cl) and grown at 37°C. This procedure removes any linear vectors that still harbor the *zeo-PheS* bacterial drug selection cassette, as the mutant phenylalanine tRNA synthase (*PheS*) sequence confers sensitivity to YEG-Cl. The final panel, (**e**) shows the linear pJAZZ-OK vector which now contains a *P. berghei* insert in which the GOI has been replaced by a *hdhfr::yfcu* drug cassette (in blue), and which contains long homology arms for recombination into the wild-type *P. berghei* ANKA genome for gene knockout. The pJAZZ-OK vector backbone must be cut from the linear plasmid of interest by a *NotI* digest that is carried out before transfection. Source: Pfander *et al.,* 2013.

This large-scale recombineering method produced *Plasmo*GEM vectors that are capable of disrupting or deleting approximately one third of genes in the *P. berghei* ANKA genome (Schwach *et al.*, 2015). As such, the *Plasmo*GEM resource is invaluable in providing vectors for genetic manipulation (free-of-charge for non-commercial institutions) for smaller studies in which time or resources are limited.

1.8 Summary and study context

As outlined in sections **1.4** to **1.6,** the complex developmental changes required by *Plasmodium* spp. parasites to survive both mammalian and dipteran host environments (detailed in section **1.3**), and to transmit successfully between these two hosts, necessitates strict genome regulation by at least twelve different mechanisms:

1. Regulation of transcription initiation by 'TATA-box'-dependent or independent mechanisms, transcription factor binding (such as by AP2 proteins), and pre-initiation complex (PIC) assembly.
2. Regulation of transcription by RNA polymerase II and transcription-associated proteins (TAPs).
3. Regulation of nascent mRNA and cryptic short-lived mRNAs by exoribonuclease-mediated degradation, such as *Pf*RNAse II.
4. Regulation of gene expression by non-coding regulatory elements such as antisense long non-coding RNAs (lncRNAs) or circular RNAs (circRNAs).
5. Alternative splicing of mRNA transcripts.
6. Regulation of mRNA maturation by 5' m7GpppN capping, and 3' polyA-tailing enzymes.
7. Nuclear export of mature mRNA and transport to specific organelles for translation.
8. Translational repression of mRNA transcripts by direct mRNA binding proteins such as poly(A)-binding proteins (PABPs), messenger ribonucleoprotein (mRNP) complexes containing DOZI, or the CCR4-POP2-NOT deadenylase complex (resulting in polyA-tail removal).
9. Translational repression of mRNA by proteins/factors that interact with core translational machinery, such as the proposed role of Puf3 (by associating with ribosomal proteins).
10. Epigenetic regulation of gene accessibility by post-translational modifications on histone residues.
11. Epigenetic regulation of gene accessibility by histone variation.
12. Epigenetic regulation of gene accessibility by modifiers of higher order chromatin organisation, such as nucleosome assembly proteins (NAPs), an example of which are *Pf*NAP-S and *Pf*NAP-L in *P. falciparum* (Chandra *et al.*, 2005; Navadgi *et al.*, 2006).

In light of the discovery of AP2-G (Kafsack *et al.*, 2014; Sinha *et al.*, 2014) as a transcriptional regulator of gametocytogenesis in *Plasmodium,* and the role of both HP1 and HDA2 in gametocyte conversion in *P. falciparum* (Brancucci *et al.*, 2014; Coleman *et al.*, 2014), the present study was designed to further elucidate

the role of epigenetics in gametocytogenesis using the rodent malaria model, *P. berghei* ANKA and experimental techniques as described in section **1.7.**

1.9 Research hypothesis and aims:

1.9.1 Hypothesis:

Expression of the AP2-G transcription factor and commitment to sexual development in the rodent malaria model, *P. berghei* ANKA, is regulated by epigenetic mechanisms.

1.9.2 Aim 1:

A number of known and putative epigenetic regulatory proteins in *P. berghei* ANKA will be subject to both complete gene knockout by double crossover (DXO) homologous recombination and conditional protein degradation by the auxin-inducible degron (AID) system. These experiments will be undertaken in a high-gametocyte producer (HP) *P. berghei* ANKA line (820cl1m1cl1) expressing both green fluorescent protein (GFP) in male gametocytes and red fluorescent protein (RFP) in female gametocytes that will enable the subsequent effect of knockout/knockdown on gametocytogenesis to be examined and quantitated in a sex-specific manner by fluorescence microscopy and/or fluorescence-activated cell sorting (FACS) (Ponzi et al., 2009) (detailed in section **1.7.2.2**). Where possible, cellular localization of tagged proteins will be assessed by fluorescence microscopy and interacting factors will be determined by co-immunoprecipitation and mass spectrometry **(Chapter 3).**

1.9.3 Aim 2

At present, histone modifications have been investigated only in *P. falciparum,* with no studies reported for *P. berghei* (Westenberger *et al.*, 2009; Lopez-Rubio *et al.*, 2009; Miao *et al.,* 2006; Salcedo-Amaya *et al.,* 2009; Ponts *et al.*, 2010; Petter *et al.,* 2011; Coetzee *et al.,* 2017). To remedy this situation, histones will be purified from both asexual and sexual stages of the *P. berghei* ANKA parasite line (namely: mature schizont and mature gametocyte stages) by acid extraction

from chromatin and mass spectrometry to determine the repertoire of histone modifications and their stage-specificity in *P. berghei* (Longhurst & Holder, 1997; Miao *et al.*, 2006; Shechter *et al.*, 2007). Where possible, these modifications will be confirmed using commercial antibodies (e.g. Life Technologies). Common modifications that have not yet been tested nor seen in these analyses, and for which commercial antibodies exist, will also be sought. The outcome of this study will be a comprehensive catalogue of the type and stage-specificity of histone modifications in *P. berghei* **(Chapter 4).**

1.9.4 Aim 3

Histone modifications that appear to correlate with sex-specific expression (from **Aim 2**) will be examined further to determine their distribution across the *P. berghei* genome using chromatin immunoprecipitation-sequencing (ChIP-seq) and ChIP-seq in combination with Tn5 transposase library preparation (ChIPmentation), paying particular attention to *ap2-g*. A number of histone modifications that have previously been examined in *P. falciparum* and for which ChIP-grade commercial antibodies are available will also be examined. Optimisation of the ChIPmentation procedure will be carried out in *P. berghei* ANKA asexual-stage parasites (mature schizonts) to establish a standard experimental and bioinformatic analysis pipeline by which further *P. berghei* developmental stages can be investigated in future **(Chapter 5).**

2. Materials and Methods

2.1 *Plasmodium berghei* methods

2.1.1 Experimental animals: ethical considerations

All animal experimentation carried out during the present project was done so in accordance with the Animals (Scientific Procedures) Act 1986 (ASPA) and European Union Directive 2010/63/EU. This Directive is based on the Three 'R's principle: to replace, reduce, and refine the use of animal for scientific purposes. All animal experiments were carried out at the University of Glasgow Joint Research Facility (JRF) under Procedure Project Licence (PPL) No. 60/4443 and Procedure Individual Licence (PIL) No. I3F0B9C21. All animal experiments carried out were deemed as 'moderate' according to Annex 8 criteria of ASPA.

2.1.2 Infection of laboratory animals with *P. berghei* parasites

Infection of laboratory animals (Theiler's Original (TO) female mice, 26-30 g or Wistar female rats, 150-175 g) with *P. berghei* parasites was performed using cryopreserved parasite stocks, purified parasites from synchronous schizont (asexual parasite) culture, or, directly from heart or tail-vein blood from previously-infected animals. Taking into account the preference of *P. berghei* for immature red blood cells (RBCs), reticulocyte production was increased in experimental animals 2 days prior to infection using 100 μl of phenylhydrazine-hydrochloride at 12.5 mg/ml in solution administered via intraperitoneal injection. Injection of 100 μl of phenylhydrazine-hydrochloride at 12.5 mg/ml in solution was not undertaken in the event of *P. berghei* samples to be used for fluorescence-activated cell sorting (FACS).

2.1.3 Infection of animals with *P. berghei* parasites via intravenous injection

Intravenous injection of *P. berghei* parasites was undertaken using restrained, un-anaesthetised TO mice (female, 26-30 g). Intravenous injection of parasites was used for infection of animals with purified synchronous schizonts from culture; transfected merozoites following electroporation of plasmid DNA (see below); or when diluting parasite lines for cloning of a *P. berghei* line. Where possible, selection of fluorescently-labelled parasite populations using FACS was carried out

to produce isogenic parasite lines. This procedure reduced the number of animals used for cloning of *P. berghei* lines. Animals were placed in a temperature-regulated transparent Perspex chamber for 15-20 min at 37°C prior to intravenous injection. After sufficient venodilation, 200 µl of parasites in 'rich' phosphate-buffered saline (PBS) (10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.7 mM KCl (1× PBS), supplemented with 20 mM HEPES, 20 mM glucose, 4 mM $NaHCO_3$, 0.1% bovine serum albumin (BSA)) was injected using insulin syringes (Becton-Dickinson, Micro-fine +, 0.5 ml; 0.30 mm (30G) x 8 mm, cat. No. 324870). Following all animal infections, parasitaemia of each animal was monitored daily by tail drop (venepuncture) and Giemsa-staining on a glass slide. Parasitaemia was counted as a percentage of infected red blood cells divided by the total red blood cell population.

2.1.4 Cervical dislocation of terminally anaesthetised animals following cardiac puncture

Upon induction of anaesthesia with vapourised isofluorane and collection of *P. berghei*-infected blood via cardiac puncture, anaesthetised animals were killed by cervical dislocation. For cervical dislocation, the mouse was turned over to the prone position and the neck manually broken by pinching of the cervical vertebrae just below the skull with a thumb and index finger, and the base of the tail pulled until the skull and vertebrae had completely separated. Once the neck was broken, the carcass was then deposited in a designated biohazard plastic bag and placed in the appropriate freezer for future disposal. All surfaces were disinfected upon completion of all animal procedures.

2.1.5 Transfection of plasmid DNA to *P. berghei* parasites via electroporation

2.1.5.1 Preparing a plasmid DNA sample for transfection

For the present study, all plasmids were linearised before transfection into *P. berghei* merozoites. No circular plasmids were transfected, with no experiments resulting in episomal expression of plasmid DNA in parasites. Throughout this study, one or more restriction enzymes were used to cut out the Ampicillin resistance fragment of a circular plasmid from the remaining DNA, leaving the rest

of the plasmid in linear form **(Figure 2.1).** To ensure that all fragments were linear, the digested plasmid was run on a gel (1% agarose) and the desired linear fragment extracted using a QIAquick Gel Extraction Kit (QIAGEN). To ensure a correct cut, a small sample (~5 μl) of the original cut plasmid was kept and run on a 1% agarose gel next to the gel-extracted fragment. The rest of the gel-extracted linear plasmid was then used for transfection. For each transfection, ~15-20 μg of circular plasmid DNA was digested with the appropriate restriction enzymes for a final gel-extracted concentration of 1-10 μg in sterile water. For transfection, 10-15 μl of linear plasmid DNA at 5-10 μg was ideal.

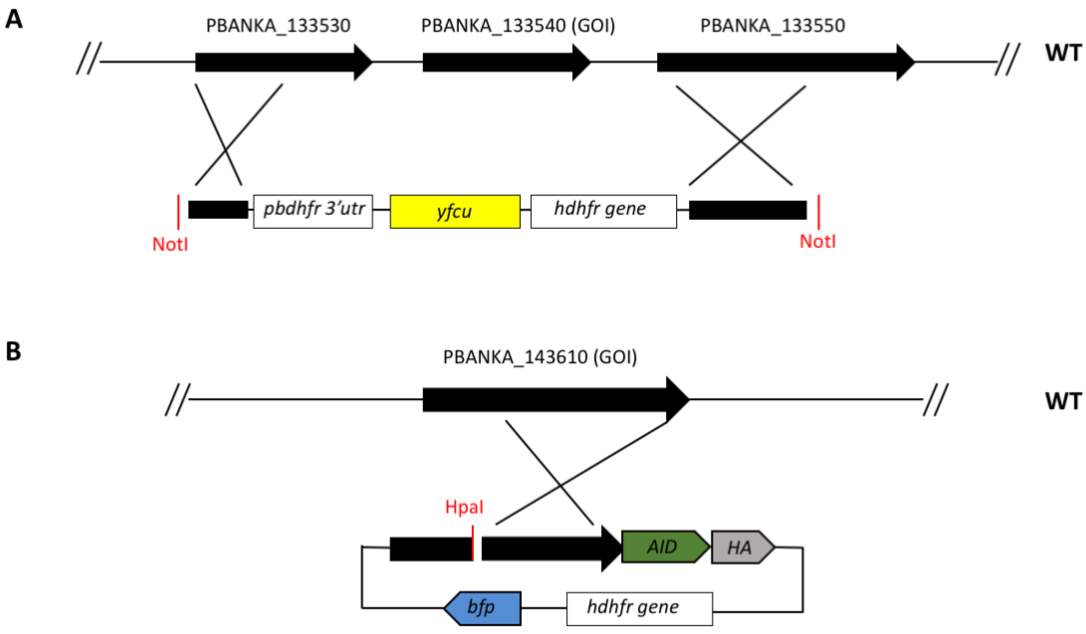


Figure 2.1: Example restriction enzyme sites used for the linearization of circular plasmids for double crossover and single crossover homologous recombination. Image **A** is a schematic representation of a wild type gene of interest (PBANKA_133540; histone deacetylase 1 (HDA1)) and adjacent genes. Underneath is a schematic representation of a linearised plasmid to be integrated into the *P. berghei* wild-type genome. The location of a digested NotI restriction site is shown. The digestion of the circular plasmid with NotI linearised the sequence for transfection. Image **B** is a schematic shows a similar schematic representation of a gene of interest, in this case PBANKA_143610 (Heterochromatin protein 1 (HP1)), and the circular plasmid used to tag *pbhp1* with HA, AID, and BFP (as well as to incorporate drug resistance sequences (only the *hdhfr* gene shown here)). The restriction site used for linearization of the cicular plasmid was HpaI and its location is indicated in image **B.** More detailed descriptions of these plasmids and their design are located in Chapters 3 and 4 of the present study.

2.1.5.2 Cultivating *P. berghei* merozoites for transfection

Once a plasmid was prepared for transfection, a donor animal was treated with 100 µl of phenylhydrazine-hydrochloride at 12.5 mg/ml in solution to facilitate the growth of host reticulocytes for infection with a *P. berghei* parasite line. After 48 hours, the donor animal was infected with *P. berghei* parasites of the parasite line into which the linear plasmid was to be integrated, e.g. the 820 GFP/RFP reporter *P. berghei* line (Ponzi *et al.*, 2009). Following infection of the donor animal with the *P. berghei* line, the animal is monitored daily by venepuncture and Giemsa staining until the *P. berghei* infection reaches a parasitaemia of 1-5%. At this point, the *P. berghei*-infected donor blood was harvested via cardiac puncture of the donor animal under terminal anaesthesia. Parasites were then cultured *in vitro* for 24 h in 'schizont culture medium' (RPMI-1640 medium with L-glutamine, supplemented with 20% heat inactivated fetal bovine serum (FBS), 20 mM HEPES buffer solution, 7.3 mM $NaHCO_3$, 0.007% HT supplement (×100), 0.002% Pen-Strep, pH 7.3), gassed with a 5% CO_2/5% O_2/90% N_2 mixture (0.5 bar), and shaking (50 rpm) at 37 °C to produce a synchronous *P. berghei* schizont culture. Synchronous *P. berghei* schizonts were then purified with a 55% Nycodenz/'rich' PBS density gradient, the resulting parasites of which were carefully extracted and kept at 37°C before being subjected to electroporation using Amaxa™ Nucleofector™ Technology (Lonza). For *P. berghei* parasites, the Nucleofector™ 'Human T-cell' protocol was applied (Janse *et al.*, 2006).

2.1.5.3 General *in vitro* cultivation of mature schizonts

One to 8 ml of infected blood was collected from an animal by cardiac puncture (using a 10 ml syringe in the case of a rat; 2 ml syringe in the case of a mouse). The blood was then transferred to 5-8 ml of complete 'schizont' culture medium to which 0.3 ml stock-solution of heparin had been added. After being transported back to the main laboratory, cells were then pelleted by centrifugation (8 min at 450 ×g) at 21°C. The supernatant was then discarded from the cell pellet. Cells were then re-suspended in complete culture medium and transferred to a 150 cm^2 cell culture flask. For a successful schizont culture, every 1 ml of blood was added to 120 ml of complete culture medium. 120 ml was the maximum amount of media added to each 150 cm^2 culture flask. The culture flask was then placed in a sterile fume hood/cabinet and each culture gassed with a mixture of 5% CO_2/5% O_2/90%

N_2. Each flask was gassed for exactly 60 seconds and then the lid closed and the flask immediately placed in a 37°C shaking incubator at 50 rpm for 24 h for synchronisation of *P. berghei* schizonts.

After ~ 20 - 22 h, and again at 23.5 – 24 h, a small sample (0.5 ml) of parasite culture was taken and placed in a sterile microtube. Cells were then pelleted by centrifugation (maximum speed; 5s) and the supernatant discarded. A thin blood smear of the cells was then made on a glass slide, fixed with methanol, and stained with Giemsa to check the quality of the parasite culture and to ensure that development of schizonts has occurred. Schizont smears were examined using a light-microscope at a 1000× magnification (100× objective; immersion oil). Viable schizonts were distinguished by the presence of 12-24 'free' merozoites within each infected red blood cell, usually with one or more clusters of 'malaria pigment' (haemozoin). Smearing of schizonts on the microscope slides often damaged the host red blood cell membranes and merozoites were visible as single, or clustered, free parasites. A pink/red-coloured nucleus and purple cytoplasm was characteristic of viable merozoites. Mature gametocytes were also seen in these smears. Degenerate schizonts often showed a compact morphology in which the separate merozoites were difficult to recognise. Once it had been determined that the cultured schizonts had reached maturity and that there was no contamination of the parasite culture, purification of synchronous schizonts from other parasite developmental stages and host cells was undertaken.

2.1.5.4 Purification of mature schizonts

Prior to transfection, schizonts (~1-3% of the total cell population) were separated from uninfected erythrocytes and other host cell debris by density gradient. For density gradients, Nycodenz was used as it does not affect the viability of *P. berghei* parasites after separation. Approximately 3×10^8 - 1×10^9 schizonts (equivalent to 3×10^9 – 1×10^{10} merozoites) could be separated from 5 - 6 ml of heart blood (1 rat cardiac puncture). A Nycodenz stock solution was made by dissolving 27.6 g of Nycodenz powder (Progen Biotechnik; Cat. No. 1002424) in Nycodenz-buffered medium (for a density at 20°C of 1.15 g/ml). This solution is autoclaved and stored at 4°C. Nycodenz-buffered medium consisted of sterile water with 5

mM Tris-HCl, pH7.5, 3 mM KCl, and 0.3 mM of calcium disodium versenate ($CaNa_2EDTA$). For separation of mature schizonts, a 55% Nycodenz/rich PBS solution (v/v) was prepared. In general, a total volume of 50 ml Nycodenz stock solution/22.5 ml rich PBS was required for a culture suspension of 150 ml. To begin, the culture suspension containing the schizonts was separated into 50 ml centrifuge tubes (35 ml per tube). Then, using a 10 ml pipette, 10 ml of 55% Nycodenz-solution was gently added to the bottom of each tube, i.e. **underneath** the culture suspension, so that a sharp contrasting division is visible between the two suspensions **(Figure 2.2)**. The suspension was then gently placed in a centrifuge and centrifugation carried out at 21°C for 30 min at 500 ×g using a swing out rotor with no brake.

Following centrifugation, the 'brown' layer at the interface between the two suspensions was collected using a 5 ml Pasteur/Dropping pipette **(Figure 2.2)**. Uninfected cells and debris formed a pellet at the bottom of each tubes after centrifugation. Schizonts were then pelleted by centrifugation (450 ×g for 8 min at 21°C). The resulting supernatant could then be discarded and the parasites taken to the next phase of experimentation.

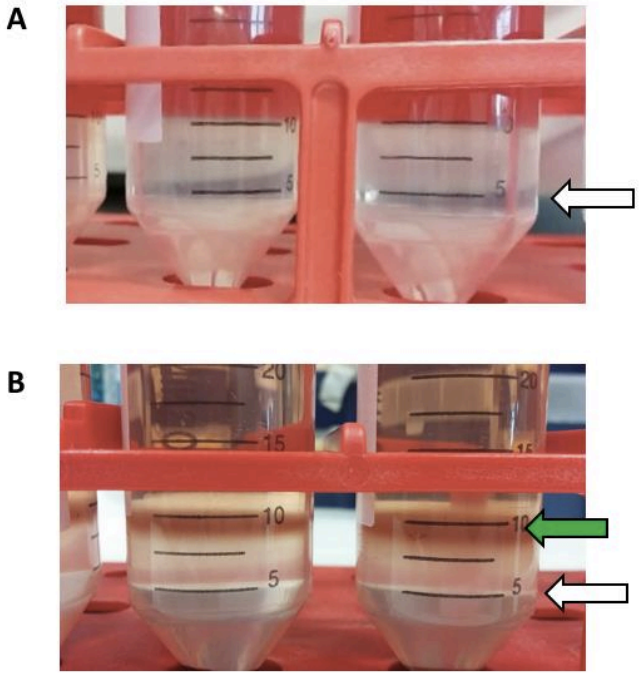


Figure 2.2: Purifying mature schizonts from a mixed culture using a Nycodenz density gradient. This image shows the same samples of synchronous schizont culture (a mixture of infected blood and 'schizont' complete culture medium) before and after separation of parasites

using a 55% Nycodenz density gradient. Image **A** shows the blood/medium mixture (35 ml) above a clear 10 ml volume of 55% Nycodenz solution (indicated by a white arrow). Image **B** shows the same sample as in **A** after centrifugation. The white arrow here indicated the clear Nycodenz solution with a brown layer of schizonts directly above it (indicated by the green arrow). Above the schizont layer is a layer of clear (yellow/orange in colour) complete culture medium from which blood and parasites had been separated out. In **B**, there was also blood cell debris in the form of dark red pellets in the base of the centrifuge tubes (not shown).

2.1.5.5 Transfection of plasmid DNA into *P. berghei* merozoites via electroporation

The parasite pellet obtained following Nycodenz-mediated separation of schizonts from culture medium was then re-suspended in fresh culture medium (warmed to 37°C) and transferred to microtubes at a concentration of ~1×10^7 - 3×10^7 schizonts per tube. 1×10^7 - 3×10^7 schizonts were sufficient for one transfection. 100 µl of the 'Human T-cell Nucleofactor™' solution was then added to the (~10 µl) of re-suspended plasmid DNA. *P. berghei* schizonts were then pelleted by centrifugation (5 s, maximum speed, in Eppendorf™ microcentrifuge) and the supernatant discarded. Each parasite pellet was gently re-suspended in the ~110 µl plasmid DNA/Nucleofactor™' solution mix and this mixture then transferred to an Amaxa™ 0.2 µl cuvette, being careful to avoid the production of any bubbles. Plasmid DNA was then transfected into parasite merozoites with the Amaxa™ gene pulser, using protocol U33. Immediately after transfection, 50 µl of culture medium was added to the cuvette.

2.1.5.6 Intravenous injection of transfected merozoites to uninfected animals

Animals to be injected intravenously were placed in a transparent temperature-controlled Perspex chamber at 37°C for 15-20 min prior to injection of transfected *P. berghei* parasites. Rodent tail veins venodilated at this temperature, making the intravenous injection procedure easier to undertake. Transfected parasites were injected into tail veins of mice using insulin syringes (Becton-Dickinson, Micro-fine +, 0.5 ml; 0.30 mm (30 G) x 8 mm, cat. No. 324870). The ~160 µl suspension containing parasites transfected with the Amaxa™ gene pulser was injected immediately into one animal following transfection.

2.1.5.7 Selection of transfected parasites with pyrimethamine

Animals infected with transfected parasites (in which the appropriate drug selection cassette was present) were provided with drinking water containing pyrimethamine 24 h after transfected parasites had been injected. Pyrimethamine was provided in drinking water for a period of 4-7 days, or until the parasitaemia had reached a percentage at which the blood could be harvested from the infected animal.

2.1.6 Preparation of pyrimethamine for drug selection

Pyrimethamine powder was dissolved in dimethyl sulfoxide (DMSO) to a final concentration of 7 mg/ml for a 100× stock solution. 10 ml of 100× stock solution was then diluted 100 times with tap water in a 1 litre glass bottle and the pH adjusted to between pH 3.5 and pH 5.0 using 1M HCl solution. 1× pyrimethamine (pH 3.5 - 5.0) was then immediately stored at 4°C and protected from light for future use.

2.1.7 Preparation of 5-fluorocytosine for drug selection

5-fluorocytosine (5-FC) is stored in white powder form at 4°C (Sigma-Aldrich; Cat. No. F7129). A disposable mask was worn when handling and measuring powdered 5-FC as this substance is teratogenic. 5-FC powder is dissolved in tap water to a concentration of 1.5 mg/ml in solution. Usually 400 ml of tap water was sufficient for selection of resistant *P. berghei* in 1 animal (0.6 g of 5-fluorocytosine powder in 400 ml for a 1.5 mg/ml solution). 5-FC was stored at 4°C and the glass bottle container wrapped in tin foil to protect the sensitive solution from light. Animal drinking bottles were also covered in tin foil to maintain 5-FC activity for the duration of the drug selection process.

2.1.8 Preparation of sulfadiazine in animal drinking water for gametocyte enrichment

During the present study, sulfadiazine was used to enrich gametocyte populations from asynchronous *P. berghei* infections. Sulfadiazine selectively kills asexual

replicating parasites in a *Plasmodium* infection and so was used to enrich sexual-stage (gametocyte) populations in this study (Beetsma *et al.*, 1998). To prepare sulfadiazine, 1 litre of fresh tap water was added to a sterile 1 litre glass 'Duran' flask and 25 mg of sulfadiazine powder (Sigma-Aldrich; Cat. No. S8626-25G) added to the water. This solution was then mixed using a magnetic stirrer and pill for 8-12 h continuously. A low heat was also applied to the stirring solution if the powder was not readily dissolving in solution. After 8-12 h, the solution turned clear and light yellow in colour. This indicated that the sulfadiazine powder had fully dissolved and the solution could then be removed from the stirrer and placed in a 4°C fridge for storage and future use.

2.1.9 Preparation of rapamycin for *in vivo* and *in vitro* experimentation using an inducible DiCre gametocyte overexpression *P. berghei* line

A lab stock of rapamycin was made up to 4 mg/ml solution, dissolved in DMSO (Sigma-Aldrich; R0395 SIGMA; Rapamycin from *Streptomyces hygroscopicus*; >95% (HPLC); powder; $C_{51}H_{79}NO_{13}$; Molecular weight 914.17; 1 mg per vial). When bought, this rapamycin stock was soluble in DMSO or 2 mM ethanol and was stored continuously at -20°C when not in use.

2.1.9.1 *In vivo* rapamycin procedure

For rapamycin-induction of gametocytogenesis in the DiCre *ap2-g* overexpression line (*P. berghei* line G1142), animals were first infected with the *P. berghei* line (G1142) and monitored until a parasitaemia of ~10% was observable in the test animal. This was to ensure that enough parasites were present to facilitate a high gametocyte conversion rate, which in turn produced a sufficient number of gametocytes for subsequent experiments. Animals were then weighed just prior to injection with rapamycin. Rapamycin (4 mg/ml stock) was first defrosted in a waterbath at 37°C and then a solution of 1 mg/ml rapamycin prepared by dissolving stock rapamycin solution in 'rich' PBS that had been warmed to 37°C. Solution volumes were prepared after the animal had been weighed. Each animal (TO mouse or Wistar rat) was injected with 1 mg/kg rapamycin from the 1 mg/ml solution in 'rich' PBS. The solution was kept at a temperature of 37°C until the moment of injection. Rapamycin was administered intraperitoneally in all cases.

2.1.9.2 *In vitro* rapamycin procedure

Following induction of gametocyte conversion *in* vivo, animals infected with the *ap2-g* inducible overexpression line were monitored by Giemsa staining of blood tail drops from approximately 12 - 24 h after induction with rapamycin. Upon observation of a high gametocytaemia in infected animals at ~24 h post-induction, blood was harvested via cardiac puncture and transferred to a 150 cm^2 culture flask at 1 ml of blood to 120 ml of complete culture medium. This culture was gassed with a mixture of 5% CO_2/5% O_2/90% N_2 and placed in a shaking incubator at 37°C (shaking at 50 rpm). To keep this induced culture under rapamycin pressure, 100 nM rapamycin was added from a 100 mM stock solution and the culture incubated overnight (12–16 h in culture) to ensure complete gametocyte commitment and maturation of rapamycin-induced *P. berghei* parasites.
To make a 100 mM stock solution of rapamycin to be added to cultures *in vitro*, 100 mM rapamycin was made up from the 4 mg/ml solution in sterile water. 0.27 µl of 4 mg/ml rapamycin was added to 9999.73 µl of sterile H_2O to give 10 ml of 100 mM rapamycin solution. Ten 1 ml aliquots of 100 mM rapamycin stock were then added to microtubes and frozen at -20°C until further use. Therefore, for a 100 nM concentration of rapamycin in a culture of 120 ml, 120 µl of a 100 mM aliquot would be added. This culture was then monitored by Giemsa staining and gametocytes purified using a 53% Nycodenz/'rich' PBS solution when appropriate. Nycodenz-purification of gametocytes from culture was carried out in the same manner as purification of mature schizonts from an *in vitro* culture (section **2.1.5.4**).

2.1.10 Preparation of 3-indoleacetic acid (auxin) for conditional degradation using the auxin-inducible degron (AID) system in *P. berghei*

Conditional degradation of proteins using the auxin-inducible degron (AID) system in *P. berghei* is carried out by addition of auxin (Sigma-Aldrich; Cat. No. 13750; 3-indoleacetic acid 98% (IAA); $C_{10}H_9NO_2$; Molecular Weight 175.18; 100 g) to parasite cultures *in vitro*. Stock solutions of auxin at different micromolar concentrations (50 mM, 100 mM, 250 mM, and 500 mM concentrations) for testing were made up by dissolving auxin powder in 95% ethanol and were then stored in 1 ml aliquots at -20°C. The most effective concentration of auxin *in vitro* was 500 µM (as

demonstrated in (Philip and Waters, 2015)). A 500 µM concentration *in vitro* was generated by addition of microlitre amounts of a 500 mM stock solution to a *P. berghei* culture *in vitro*. Cultures to which auxin had been added were incubated at 37°C in a shaking incubator at 50 rpm.

2.1.11 *In vitro* auxin-induced degradation (AID) procedure

A protocol for *in vitro* auxin-induced degradation of proteins was developed for use with synchronised schizont cultures of parasites transfected with *Os*TIR1/AID/HA sequences. Samples from parasite cultures were taken prior to addition of auxin to check parasite morphology via Giemsa stain and to check for the presence of complete protein prior to degradation by a HA-antibody-stained Western blot. Taking into account the number of microlitres of media in the culture flask, parasite cultures were incubated with 500 µM of auxin at 37°C in an incubator with mild shaking (50 rpm). Samples were then taken to test for protein concentration (via Western blot) at time-points such as every 15 min; or at the points: 30 min, 1 hour, 4 hours, 16 hours and 24 hours incubation with 500 µM auxin. The incubation process was carried out until maximal protein depletion was seen.

2.1.12 Giemsa stain preparation for monitoring of *P. berghei* parasitaemia

To prepare fresh (12.5%) Giemsa stain, 5 ml of Giemsa stain, modified solution (Fluka Analytical, Sigma-Aldrich; Cat. No. 48900-500ML-F) was added to 35 ml Giemsa buffer (Sörensen staining buffer) to give a total of 40 ml of fresh 12.5% Giemsa solution. Giemsa stain was then stored at room temperature for future use. To prepare a stock of 5 l of Giemsa buffer (Sörensen staining buffer), 2.541 g of KH_2PO_4 and 0.5507 g of $Na_2HPO_4.2$ were dissolved in 5 l of de-ionised water. The pH of the solution was then adjusted to 7.2 with NaOH. Prior to staining, slides on which parasites are present were fixed in methanol for 5 s before staining in 12.5% Giemsa stain for 10 min. Slides were then allowed to air-dry at room temperature.

2.1.13 Freezing *P. berghei*-infected blood stabilates for long-term storage

To make an undiluted stabilate, a 2 ml cryotube (Cryogenic vial, graduated, internally threaded with silicone O-ring, Alpha Laboratories Limited; Product No. LW3534) was labelled and 250 µl of blood added to an equal volume of 30% glycerol/'rich' PBS solution, supplemented with 10 I.U./ml heparin. Stabilates were frozen slowly at -80°C by wrapping the sample tube in tissue and then placing this at -80°C overnight. Once the initial slow-freeze had taken place, the stabilate was then moved to a regular -80°C Cryovial storage box (Alpha Laboratories Limited; Product No. LW3440B) or transferred to liquid nitrogen. To make a diluted stabilate if the starting *P. berghei* parasitaemia was high (e.g. >5%), 125 µl of blood was diluted in 125 µl of 'rich' PBS, and then 250 µl of sterile freezing solution (30% glycerol/'rich' PBS solution, supplemented with 10 I.U./ml heparin) added to the diluted blood. These stabilates were then slow-frozen as described for undiluted stabilates. Each stabilate was made up to 500 µl as this volume provided enough material for infection of 2 mice via intraperitoneal injection (200 µl each).

2.1.14 Removal of contaminating host leukocytes from *P. berghei* infected blood using Plasmodipur filters

In advance of all histone extraction and sequencing experiments, contaminating host leukocytes were removed from *P. berghei*-infected red blood cells immediately after harvesting of blood by cardiac puncture and before infected blood was cultured *in vitro*. At present, removal of leukocytes by Plasmodipur filters (EuroProxima BV) has been shown to be effective in removing contaminating cells from the host, while retaining a good yield of *Plasmodium*-infected red blood cells for further analyses (Auburn *et al.*, 2011; Li *et al.*, 2017).

To begin with, the Plasmodipur filter was mounted on a 20 ml syringe with the plunger removed. The Plasmodipur filter was then pre-wet with 5 ml of 'rich' PBS that had been warmed to 37°C. The pre-warmed 'rich' PBS was then passed through the filter at a gentle pressure. Each 1 ml of *P. berghei*-infected blood was diluted in 20 ml of complete culture medium. This diluted blood sample was then gently passed through the Plasmodipur filter at constant pressure, with the filtered material collected in a 50 ml centrifuge tube. Each Plasmodipur filter was used only once and then discarded, as was the syringe. Each diluted 1 ml of whole

blood was purified by passing through 2 clean Plasmodipur filters in sequence to ensure that all possible contaminating leukocytes were removed (or as much as possible using this method). Purified samples were then pelleted by centrifugation before further experimentation (8 min at 450 ×g; 21°C).

2.1.15 Purification of *P. berghei* parasites by magnetic separation

In some cases, where purification by Nycodenz of *P. berghei* parasites from culture or directly from cardiac puncture had resulted in a low parasite yield, diluted whole blood containing *P. berghei*-infected red blood cells was passed through a magnetic MACS® LD column (Miltenyi Biotec; 25 LD columns; Order No. 130-042-901) held within a strong magnetic stand. These magnetic separation columns were effective in separating intraerythrocytic haemozoin-producing *P. berghei* parasite stages (schizonts or gametocytes) from solution as a result of their high iron content (Ribaut *et al.*, 2008).

To begin with, MACS® LD columns were placed in a very strong custom-built magnetic stand within a fume cupboard (to prevent any contamination). Infected host blood was diluted (1 ml of blood to 20 ml of complete culture medium) before being added to the magnetic separation columns in 10 ml increments. Magnetic columns were suspended above a plastic waste tray (as shown in **Figure 2.3**) and the parasites suspended in culture medium allowed to pass through the column over time (without the use of the provided plunger).

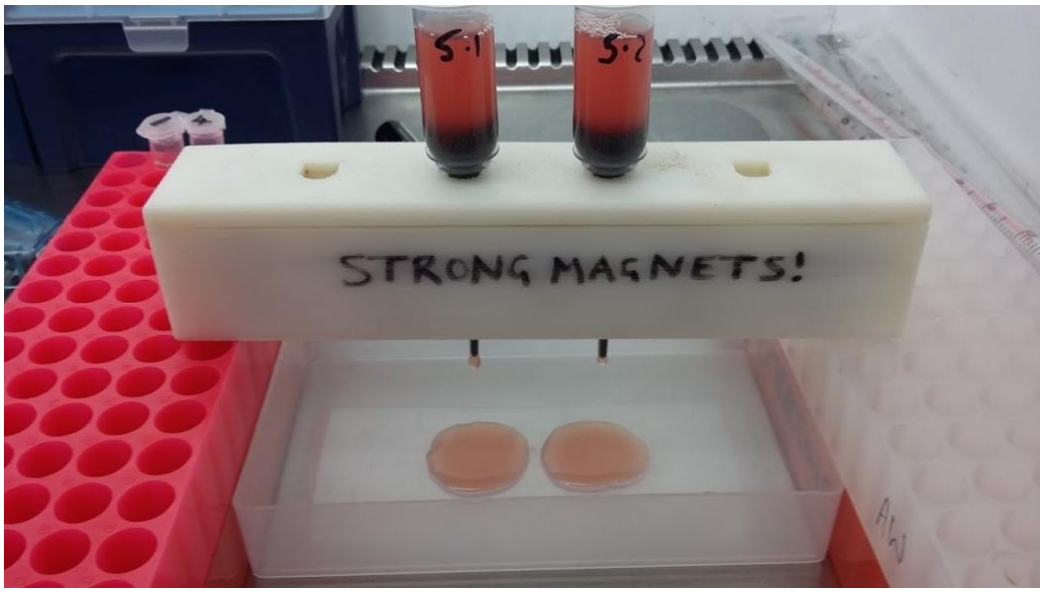


Figure 2.3: Purification of *P. berghei*-infected red blood cells from host blood by magnetic separation. The image above was taken during the present study and shows the magnetic separation of synchronous schizonts from a blood/culture medium mixture using MACS® LD columns placed inside a strong magnet. The solution (without parasites) is dripping into a plastic waste tray, leaving the schizonts adhered to the magnetic column. These would then be washed out with 'rich' PBS into a clean centrifuge tube when the columns were removed from the magnetic stand.

Each column was used to purify approximately 50 ml of parasites in suspension (each column can hold up to 1×10^8 cells as per manufacturers recommendations). After all of the parasite suspension had passed through the column, the column was removed from the magnetic stand and placed over a sterile 15 ml or 50 ml centrifuge tube (depending on the volume into which the magnetically-bound parasites were to be eluted). The parasites attached to the sides of the column were then eluted by pipetting 10 ml of 'rich' PBS to the column and allowing it to drip through until the solution coming through was clear. If it was thought that more parasites were still adhered to the sides of the magnetic column, a further 10 ml of 'rich' PBS was added and allowed to drip through. After this purification stage was complete, parasites were pelleted by centrifugation at 450 ×g for 8 min (21°C) before further analyses were carried out.

2.1.16 Preparation of live *P. berghei* parasites from tail blood drops (mice or rats) for fluorescence microscopy

Hoechst stain (Hoechst 33258, pentahydrate (bis-benzimide), 10 mg/ml in water, 16.0 mM solution; Thermo Scientific™; Cat. No. H3569) was prepared before staining of nuclei by adding 3.4 µl of Hoechst stain to 5 ml of 'rich' PBS to make a 10 µmol solution. Hoechst solution was kept away from light at all times and stock solution stored at 4°C. For each separate blood sample, 100 µl of 10 µmol Hoechst was added to a microtube and a tail drop taken via venepuncture. Blood drops were collected via sterile glass capillary tube (Brand® Micro-haematocrit Capillary tubes; Thermo Scientific™; Cat. No. 11383994) and then blood drops pushed from the capillary tube into the Hoechst solution using a 1 ml dropper. Blood drop samples in Hoechst solution were then transported to the main laboratory and incubated at 37°C for 15 min. Following incubation, blood drops were pelleted by centrifugation at 9500 ×g for 2 min at room temperature. The supernatant was discarded and the blood pellet re-suspended in 10 µl of 'rich' PBS. Approximately

6 µl of this re-suspension was then dropped onto a clean glass slide and a cover slip placed on top. The cover slip was then sealed to the glass slide using clear nail varnish before parasites were viewed under a fluorescence microscope.

2.1.17 Preparation of *P. berghei* samples for fluorescence-activated cell sorting (FACS)

Before analysis of *P. berghei* samples by FACS, Hoechst stain was prepared in 'rich' PBS and warmed to 37°C in an incubator (kept away from light at all times). The purpose of the Hoechst stain in this case was to dye the nuclei of *P. berghei*-infected red blood cells. Host uninfected red blood cells do not contain nuclei, and so any stained nuclei are indicative of *P. berghei*-infected red blood cells and can be identified by FACS. Prior to taking blood drops from infected animals, 6.8 µl of Hoechst stain solution was added to 10 ml of 'rich' PBS and 1 ml aliquots of this diluted Hoechst solution placed in microtubes. Once heated to 37°C, one tail drop from an infected animal with the parasite line of interest was added to 1 ml of diluted Hoechst solution. Samples were incubated in Hoechst stain at 37°C in a shaking incubator (50 rpm) for 30 minutes. Following incubation, 500 µl of stained sample was passed through Nitex mesh (45 micron Nitex microfibre Nylon cloth; Cadisch Precision Meshes) to filter out any contaminating leukocytes or debris. Samples were filtered into 5 ml round-bottom Falcon® tubes (Corning®; 12 × 75 mm Polystyrene test tubes, Product No. 352058) containing 500 µl of FACS buffer (1× Phosphate-buffered saline (1× PBS) with 5mM EDTA, pH 8.0). All FACS buffer was sterile and had been passed through a 0.22 µm pore vacuum filter unit (Merck; Stericup-GP, 0.22 µm, polyethersulfone, Cat. No. SCGPU05RE) prior to use. These filtered samples were then transported to the FACS laboratory (protected from light at all times).

2.1.18 Protocol for FACS using a BD FACSCelesta™ flow cytometer

Once *P. berghei* samples had been prepared as described in section **2.1.17,** FACS analysis of many lines was carried out using a BD FACSCelesta™ flow cytometer (BD Biosciences). For this machine, the FACS facility room was kept at a temperature of 21-22ºC at all times. FACS data was analysed at the time of running

samples using BD FACSDiva Software (BD Biosciences) and then afterwards using FlowJo® software (FlowJo®, LLC).

To begin, samples were run at a maximum of 10,000 events per second and the forward scatter area (FSC-A) (x-axis), which determines cell size, plotted against side scatter area (SSC-A) (y-axis), which determines the granularity of cells. From this plot, the parasite population was gated, excluding any very large or highly granular cells. Next, the FSC-A (x-axis) was plotted against the forward scatter height (FSC-H) (y-axis). This plot was used to gate singlet cells (single cells) from doublet cells (aggregated cells or multiply-infected cells). Infected red blood cells were then gated by selecting only cells that were positive for Hoechst nuclear stain (corresponding to BD Horizon Brilliant™ Violet 421 (BV421) on the FACSCelesta™ cytometer). Infected red blood cells could then be analysed/gated according to their fluorescence spectra. For example, parasites transfected with sequences encoding mCherry or TagBFP could be gated from background non-fluorescent parasites by selecting for emission spectra corresponding to BD Horizon™ PE-CF594 and BD Horizon Brilliant™ Violet 421 (BV421) respectively. In the case of TagBFP fluorescence, no Hoechst staining was carried out prior to FACS; successfully transfected parasites were identified based on their 421 nm emission spectra. To examine parasite lines emitting RFP and GFP, cell populations were gated according to emission spectra that corresponded to R-phycoerythrin (PE) (578 nm maximum emission) and BD Horizon Brilliant™ Blue 515 (BB515) (515 nm maximum emission) respectively **(Figure 2.4).**

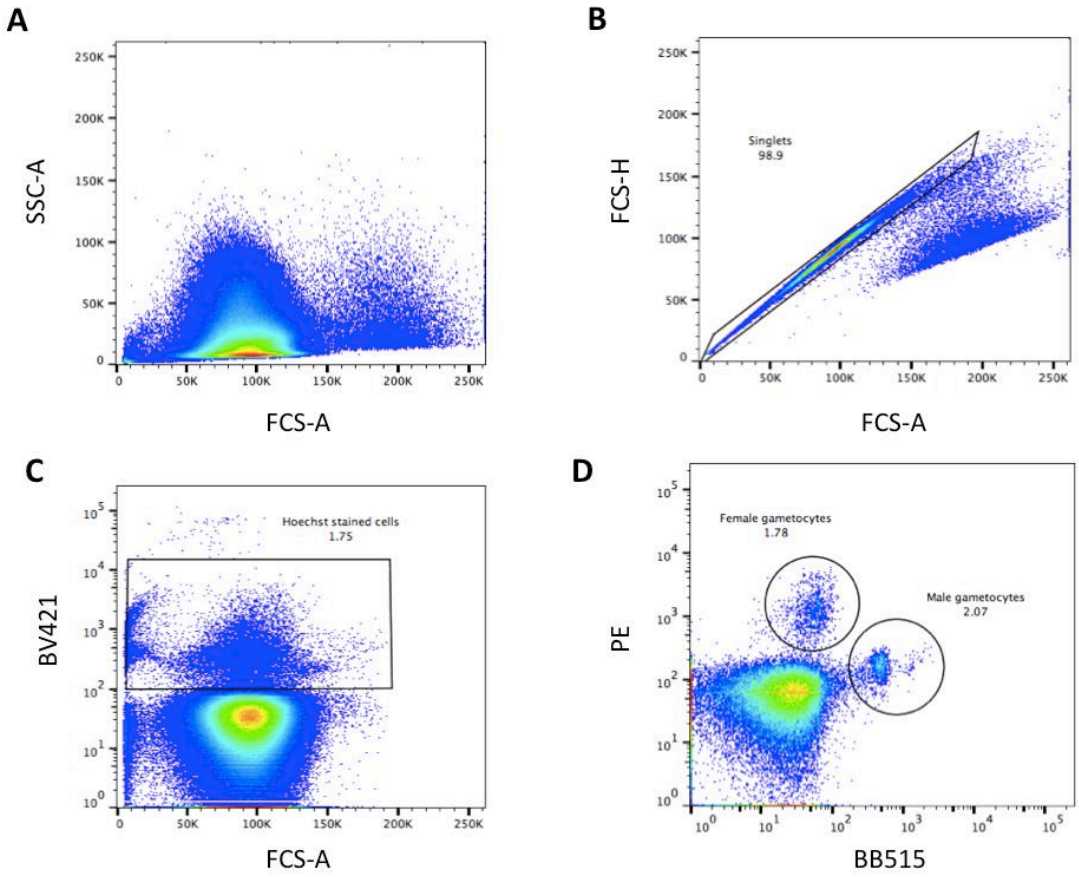


Figure 2.4: Gating strategy used when examining the *P. berghei* 820 TBB parental line. This image shows an example of the gating strategy applied to an 820 TBB *P. berghei* parental line 4 days post infection (dpi). 1 million events were recorded in this instance. In female gametocytes of this parasite line, RFP is driven by a female-specific promoter (PBANKA_131950), while GFP is present in male gametocytes, as the GFP sequence is driven by a male-specific promoter (PBANKA_041610). Image **A** depicts all cells (SSC-A versus FCS-A). Image **B** then shows the gating of singlets, in this case, the majority of cells sorted were singlets. Singlets are then gated further in Image **C**: only infected red blood cells are chosen based on the staining of their nuclei with Hoechst. Image **D** shows the final gating of male and female gametocytes from the infected red blood cell population. Using the BD FACSCelesta™ flow cytometer, GFP fluorescence corresponds to BD Horizon Brilliant™ Blue 515 (BB515) (515 nm maximum emission) and RFP fluorescence corresponds to R-phycoerythrin (PE) (578 nm maximum emission). Further FACS analysis and gating strategies can be found in Chapters 3 and 4.

2.1.19 *P. berghei* competition growth curve protocol (for analysis by FACS)

To begin a competitive growth curve analysis of *P. berghei* lines by FACS, 2 animals were infected with *P. berghei*: one with the parasite line of interest in which a fluorescent tag has been incorporated via transfection (in the present study: Tag-

BFP or mCherry fluorescent proteins), and the second animal with a 507 TBB *P. berghei* parasite line in which GFP is expressed constitutively. After 48 - 72 h of parasite growth *in vivo*, the parasitaemia of each animal was carefully counted (usually an average parasitaemia generated after 10 separate counts was used). Three clean TO female mice were then infected with an equal number of parasites from both lines. Mice were infected with low numbers of parasites to begin with to maintain a low starting parasitaemia. This produced an infection that could be monitored by FACS over the course of 4 - 5 days post-infection (dpi) before the animals needed to be killed.

Each animal in a competitive growth curve experiment was inoculated intravenously with a 0.01% total parasitaemia (0.005% parasitaemia from each parasite line), assuming a volume of 2.3 ml of blood in each mouse. Tail drops of live *P. berghei* parasites from donor animals were taken via venepuncture and added to 1 ml of 'rich' PBS warmed to 37°C. Parasites were counted and then the correct number of microlitres of each *P. berghei* line in solution added to warmed 'rich' PBS to make a total volume of 200 µl for intravenous injection into each mouse. Enough of the mixture of *P. berghei* lines in 'rich' PBS was made up to infect n+2 animals to account for any loss of solution while preparing syringes. The 3 mice to be monitored by FACS were then inoculated with the *P. berghei* mix from this solution. Infection with a 0.01% total parasitaemia resulted in a subsequent infection that was just about detectable by FACS after 24 h and could be monitored by FACS for a period of 4 -5 days. The ratio of GFP-fluorescent parasites to mCherry/Tag-BFP-fluorescent parasites was then recorded every 24 h. The parasite population could also be followed over a longer period of time if the infection was inoculated again via mechanical passage to clean animals.

After infection, samples were collected after every 24 h and analysed by FACS where 500,000 – 1 million events were recorded for GFP, mCherry/Tag-BFP and Hoechst. Gating was carried out on Hoechst positive cells (nuclei indicated *P.* berghei-infected red blood cells) and then GFP-positive and mCherry/Tag-BFP-positive parasites counted. Some double-positive parasites where sometimes collected, resulting from multiply-infected host red blood cells. Mechanical passage or cull of animals was carried out once a parasitamia reached ~5%. In the

event that FACS could not be carried out on the same day as sample collection; the samples were stored in 'rich' PBS at 4ºC and analysed the following day.

2.1.20 Generating a clonal *P. berghei* line by limiting dilution

Following transfection and drug selection of a *P. berghei* parasite line with an integrated genetic construct, cloning by limiting dilution was carried out to ensure a clonal population of parasites without contamination with wild-type parasites that may have survived transfection and drug treatment. After confirming the presence of wild-type genes and integration of the desired sequence by PCR, one phenylhydrazine-treated mouse was infected with a stabilate of the parasite line to be cloned via intraperitoneal injection. After 24 h, a parasite line had usually grown to a red blood cell parasitaemia of 0.1-0.5%. Limiting dilution was always carried out when a donor mouse had reached a low parasitaemia without any multiply-infected cells.

Using a light microscope, 25 fields of view with ~300 parasites in each field were counted, and an average parasitaemia calculated in this manner. Ten mice were then placed in a transparent temperature-controlled Perspex chamber at 37°C for 15-20 min prior to intravenous injection. From the donor mouse, ~5 µl of blood was taken by venepuncture and placed in 1 ml of 'rich' PBS (warmed to 37°C to maintain parasite viability). Approximately 10 µl of this blood/'rich' PBS mixture was placed on a haemocytometer and 10 counts of red blood cells taken (usually a new 4 nl section was counted each time). From these counts, an accurate measure of red blood cells per nl was calculated. Red blood cell concentration per µl was then calculated and, using the previously obtained parasitaemia percentage, an approximate number of infected red blood cells per µl was determined. A limiting dilution of the blood mixture in warm 'rich' PBS was carried out until a sample with ~0.8 parasites per 200 µl was obtained. Each of the 10 mice were then injected with 200 µl of solution as detailed in section **2.1.3**, to infect approximately 0.8 parasites into each mouse.

Infected mice were monitored from ~7 days post-infection by Giemsa staining of blood from tail drops. On average, 3 mice out of 10 that had been used for cloning became infected with a *P. berghei* parasite. Parasites were harvested via cardiac

puncture under terminal anaesthesia. PCR was then carried out to confirm the presence of a clonal parasite line.

2.1.21 Generating an isogenic *P. berghei* line by FACS

In an effort to reduce the number of animals used in scientific experiments, attempts to produce isogenic *P. berghei* lines were made by selecting fluorescently-active parasite lines using a BD FACSAria II Cell Sorter (BD Biosciences), followed by immediate infection of selected parasites into mice via intravenous injection. This method was used in the production of isogenic cell lines when a genetically-modified *P. berghei* parasite line expressed a fluorescent protein that distinguished it from background *P. berghei* wild-type parasites. When preparing parasites for selection followed by re-injection into new animals, nuclear staining with Hoechst stain was not carried out as this could be detrimental to cell viability. Parasites were only selected based on their transfected fluorescent marker (e.g. mCherry or TagBFP). On the cell sorter, singlet red blood cells were identified by their size (forward scatter) and granularity (side scatter) and gated according to these cytometric events. *P. berghei* parasites with the desired genetic background were then identified by their fluorescence and this population gated and sorted from background cells. Selected parasites (usually 100 single parasites) were sorted into 1.5 ml microtubes containing 250 µl of 'rich' PBS (sterile) that had been warmed to 37°C. These tubes were then placed into a 37°C heat block for transport to the animals to be infected. Unlike cloning by limiting dilution, only 2 mice would be infected intravenously with sorted parasites and often, 100% of infected animals developed a parasite infection after this procedure was used. However, this method was not as successful as limiting dilution in producing lines devoid of wild-type *P. berghei* contamination.

2.2 Molecular biology methods

2.2.1 Genomic DNA isolation from *P. berghei* parasites

After cardiac puncture, 200 µl of blood was placed directly into 12 ml of 1× erythrocyte lysis (eLysis) buffer (1.5M NH_4Cl, 0.1M $KHCO_3$, 0.01M EDTA) on ice.

This mixture of blood was then allowed to lyse on ice for 5-10 min until the solution had become clear. The mixture was then pelleted by centrifugation at 400 ×g for 10 min and the resulting supernatant discarded. The parasite pellet was then re-suspended in 0.5 ml eLysis buffer and transferred to a clean microtube. This solution was then centrifuged at maximum speed for 1 min and the supernatant removed and discarded. This parasite pellet was then stored at -20 ºC for later DNA isolation or the DNA isolated immediately.

To isolate DNA, the parasite pellet was re-suspended in 350 µl of TNE buffer (10 mM Tris pH 8.0, 5 mM EDTA pH 8.0, 100 mM NaCl) and then 100 µg of RNase A (10 µl of a 10 mg/ml solution; acquired from Sigma-Aldrich; Cat. No. LSKPMRN30) and 1% (v/v) sodium dodecyl sulfate (SDS) (50 µl of a 10% solution) was added. This solution was then made up to 500 µl with demineralised water. These parasites in solution were then incubated for 10 minutes at 37ºC before 200 µg of Proteinase K from *Tritirachium album* (10 µl of a 10 mg/ml solution; Sigma-Aldrich; Cat. No. P6556) was added and the mixture incubated for 1 h at 37ºC. Following incubation, 0.5 ml of buffered phenol (Sigma-Aldrich; Cat. No. P4557) was added and the tube inverted several times. Centrifugation for 3 min at 10,000 ×g was carried out at room temperature and the resultant aqueous upper phase transferred to a clean microtube. 0.5 ml of buffered phenol:chloroform:isoamylalcohol (25:24:1) was then added to this aqueous phase and the tube inverted several times. This tube was then centrifuged for 3 min at 10,000 ×g. Again, the resulting aqueous upper phase was transferred to a clean microtube. To this solution, 0.5 ml chloroform:isoamylalcohol (24:1) was added and the tube inverted several times before being centrifuged for 3 minutes at 10,000 ×g. The resulting aqueous upper phase was transferred to a clean microtube and 0.1 volume (~50 µl) of 3 M NaAc, pH 5.2, and 2 volumes (~1 ml) of ice cold 96% ethanol added to precipitate the DNA.

The solution was inverted several times and the DNA precipitated at –20ºC for a couple of hours to overnight. To isolate the DNA from the NaAc/ethanol solution after this period, the DNA microtube was centrifuged for 20 min at maximum speed at 4ºC. The DNA pellet was then washed with 150 µl ice cold 70% ethanol and centrifuged for 5 min at maximum speed at 4ºC. The supernatant was then

discarded and the DNA pellet air-dried for 5-10 min. Parasite DNA was then re-suspended in 50 µl demineralised water if the procedure started with a small parasite pellet, and up to 250 µl water if started with a large parasite pellet. DNA was dissolved at 37ºC for 30 min and then the concentration calculated using a NanoDrop spectrophotometer. If the DNA solution was very viscous or if the concentration was greater than 1 µg/µl, then the sample could be diluted further. Diluted DNA stocks at a concentration of 50-100 ng/µl were created for analysis by PCR.

2.2.2 Whole protein extraction from *P. berghei* parasites

After blood was extracted by cardiac puncture, 200 µl of blood (or more if necessary) was placed directly into 12 ml of 1× erythrocyte lysis (eLysis) buffer on ice. Once the blood/buffer mixture had become clear, the mixture was centrifuged at 400 ×g for 10 min. The parasite pellet was then re-suspended in 0.5 ml eLysis buffer and transferred to a clean microtube before being pelleted once more at maximum speed in a centrifuge for 1 min. At this point, the pellet was either frozen and stored at -20ºC or protein extraction undertaken.

To complete whole protein extraction, the parasite pellet was re-suspended in 50 µl of RIPA buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS) with 1 mM DTT and 1× complete protease inhibitor tablet in solution. This sample was then incubated on ice for 30 min with frequent re-suspension. The sample was then centrifuged at 14,000 ×g for 15 min and the supernatant transferred to a clean microtube. To this supernatant (containing extracted whole protein), 2× SDS-PAGE protein sample buffer (80 mM, 2% SDS, 10% glycerol, 0.0006% Bromophenol blue, 0.1 M dithiothreitol (DTT)) or 2× Laemmli protein sample buffer (4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% Bromophenol blue, 0.125 mM Tris-HCl pH 6.8) was added. The suspended parasite pellet was then boiled for 10 min at 95ºC. After boiling, the parasite suspension was placed on ice to cool. Supernatant from this pellet could then be run on an SDS-PAGE gel.

2.2.3 Nuclear fractionation of proteins from *P. berghei* parasites

All steps of a nuclear fractionation procedure were carried out at 4ºC. To begin, blood extracted from an infected animal was placed into 1× eLysis buffer and incubated on ice for 5 min (or until solution was clear). Parasites were then pelleted by centrifugation (450 ×g for 8 min; Acc. 6, Dec. 4) and 5 volumes of nuclear fractionation buffer ('nuclear lysis buffer') added to the blood pellet. This was then incubated on ice for 5 min. The solution was then centrifuged at 2500 ×g for 5 min at 4ºC. The supernatant from this step was kept as the cytosolic protein fraction.

To extract nuclear proteins, 2 volumes of 'nuclear extraction buffer' were added to the parasite pellet and this suspension incubated on ice for 1 h. Following 1 h incubation on ice, the suspension was centrifuged at full speed for 30 min at 4ºC. The supernatant from this step was retained as the 'nuclear protein fraction'. The remaining pellet was re-suspended in 1× PBS and kept as the 'insoluble fraction'.

Buffer name	Composition
'Nuclear lysis buffer'	20 mM HEPES, 10 mM KCl, 1 mM EDTA, 0.65% Tergitol® type NP-40, 1 mM DTT and 1× Complete Protease Inhibitor cocktail added just before use
'Nuclear extraction buffer'	10 mM Tris-HCl pH 8.0, 2% SDS, 5 U/ml Benzonase® endonuclease, 1 mM DTT and 1× Complete Protease Inhibitor cocktail added just before use

Table 2.1: Composition of buffers for nuclear fractionation from whole proteins.

2.2.4 Acid extraction of *P. berghei* histones

This protocol is adapted from acid extraction methods tested in *P. falciparum* for the maximal retention of native histone post-translational modifications (as

published in (Miao *et al.*, 2006) and (Moll *et al.*, 2013)). All step were performed at 4ºC.

P. berghei parasites of the desired developmental stage were extracted by Nycodenz purification (discribed in section **2.1.5.4**) or magnetic separation (see section **2.1.15**) following the removal of contaminating host leukocytes from blood using Plasmodipur filtration (section **2.1.14**). Ideally, 1-3 $\times$ 10^9 parasites (Miao *et al.*, 2006) at either synchronized asexual stage or mature gametocyte stage were harvested for acid extraction of histones to obtain enough material for identification of proteins and histone modifications via mass spectrometry. After centrifugation of parasites following Nycodenz separation or magnetic extraction from host blood, parasite pellets were washed twice with cold 1$\times$ PBS. Re-suspended pellets were centrifuged at 450 $\times$g for 8 min at 4ºC after each wash. Parasite pellets were then washed twice with a wash buffer (25 mM Tris-HCl, pH 7.8, 1 mM EDTA, 0.2% IGEPAL CA-630) to remove haemoglobin and re-pelleted at 450 $\times$g for 8 min at 4ºC. Parasite pellets were then washed three times with 0.8% NaCl to remove excess haemoglobin and IGEPAL CA-630 detergent from the previous wash steps. Parasites were re-pelleted at 450 $\times$g for 8 min at 4ºC each time.

Histones were then extracted twice with 0.25 M HCl. To do this, 10 volumes of 0.25 M HCl was added to the parasite pellet and incubated on ice for 1 h with frequent re-suspension. After 1 h, this sample was centrifuged at 12,000 $\times$g for 2 min to pellet acid-insoluble contaminants. The acid-soluble fraction (supernatant) was then transferred to a clean 1.5 ml microtube. This fraction was kept on ice while a second acid extraction of the remaining pellet was carried out. To the remaining pellet, 10 volumes of 0.25 M HCl was again added and the sample incubated on ice for 1 hour with frequent re-suspension. This parasite pellet was then centrifuged at 12,000 $\times$g for 2 min (or longer if needed) and remove acid-soluble fraction (supernatant). This supernatant was then pooled with the previous acid-soluble fraction on ice. This acid-soluble supernatant pool contained the extracted *P. berghei* histones. Pooled acid-soluble fractions were then mixed with an equal volume of 25% TCA and incubated on ice for 1 hour. This step precipitated the parasite histones. Acid extracted samples were then centrifuged

at 12,000 ×g for 15 min at 4ºC. When finished, precipitated histones were pelleted at the bottom of the centrifuge tube and the supernatant was discarded.

Parasite pellets were then washed with acetone and allowed to dry completely on ice overnight at -20ºC. To facilitate drying at -20ºC, parafilm was placed over the open cap of the microtube and a hole pierced in this with a needle. This allowed the pellet to dry overnight at -20ºC without ice formation. Dried histones could then be analysed by 15% SDS-polyacrylamide gel electrophoresis (PAGE) the following day after boiling the pellet in 2× SDS protein sample buffer or 2× Laemmli sample buffer at 95ºC for 10 min (as described when extracting whole protein in section **2.2.2**). Acid extracted histones could subsequently be examined by SDS-PAGE, Western blotting, or mass spectrometry. Histone gel bands could also be excised individually and digested with trypsin for analysis of individual histones by mass spectrometry.

2.2.5 Preparation of *P. berghei* samples for immunofluorescence microscopy

To observe fluorescent proteins or tagged proteins of interest by fluorescence microscopy, blood or parasite culture samples were smeared onto glass microscope slides and fixed in 4% paraformaldehyde (made up in sterile water this study, 16% paraformaldehyde single-use ampoules were used in each case (Paraformaldehyde 16% solution, EM grade; Electron Microscopy Sciences; Cat. No. 15710). In the case of *P. berghei* parasites in whole blood; addition of 4% paraformaldehyde lysed red blood cells and so the smear appeared clear after fixation. However, parasites had been successfully fixed to the glass slide. After 10 min staining, the 4% paraformaldehyde was poured from the slides and disposed of appropriately. Slides were then rinsed once in 1× PBS and permealised by incubating in a 0.1% Triton in 1× PBS (1× PBST) solution for 5 min with gentle agitation. Slides were then rinsed once more in 1× PBS. Blocking of non-specific antibody binding was then achieved by incubating the slides for 45 min in 1% BSA/1× PBS on an orbital shaker. After this, slides were incubated with a primary antibody diluted in 1% BSA/1× PBS for 1 h at room temperature (with slides in a slide box in the dark). For staining with primary and secondary antibodies, just 100 µl to 200 µl of antibody solution/dilution was pipetted onto each slide and

covered with a long cover slide. Primary and secondary antibodies were often diluted 1:400 or 1:500 in 1% BSA/1× PBS solution.

Following primary antibody staining, the slides were again rinsed in clean 1× PBS before being incubated with a fluorescently-labelled secondary antibody in 1% BSA/1× PBS solution for 45 min at room temperature. After a final rinse in 1× PBS, slides were then mounted on the slide with 1-2 drops of Vectashield mounting medium with 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI) (Vector Laboratories, Inc.; Cat. No. H01200). A cover slip was then placed onto each glass slide and the outer edges of the cover slips sealed to the slides with clear nail varnish. Slides were then visualised using a fluorescence microscope or stored at 4ºC until use.

2.2.6 Co-immunoprecipitation of *P. berghei* proteins using antibody-bound magnetic beads

2.2.6.1 Cell lysis procedure

In the present study, co-IPs were often carried out at schizont stages of *P. berghei* lines to maximise the amount of tagged protein available. For a large (120 ml) culture, or more than one large culture, material from culture flasks was poured into 50 ml centrifuge tubes and centrifuged at 450 ×g for 8 min at room temperature. Centrifugation was repeated in this manner until all samples were pooled in 10 ml of complete culture medium in a single 50 ml centrifuge tube. Cold 1× eLysis buffer was then added to the sample to bring to a volume of 50 ml. This suspension was then incubated on ice for 10 min, or until the solution was clear. Once clear, the 50 ml of suspended parasites were spun down at 450 ×g for 8 min at 4ºC. The supernatant was then discarded and the pellet washed in 50 ml 1× PBS, with re-pelleting by centrifugation at 450 ×g for 8 min at 4ºC (6 acceleration; 4 deceleration). The supernatant was then discarded and the pellet re-suspended in 1 ml 1× PBS and transferred to a clean microtube. Parasites were then pelleted by centrifugation at 14,000 ×g for 1 min. The supernatant was discarded and the pellet re-suspended in 500 µl ice-cold 'Co-IP Lysis buffer' (20 mM Tris-HCl pH 8.0, 137 mM NaCl, 10% Glycerol, 1% Tergitol® Type NP-40, 2 mM

EDTA, 1× Complete Protease Inhibitor cocktail added just before use). This re-suspended parasite pellet was then incubated on ice for 30 min.

The amount of ice-cold 'Co-IP lysis buffer' added to each sample was adjusted in cases where there was a very large parasite pellet. During incubation on ice, the tube was inverted several times to mix the sample. Samples were never vortexed at this stage: this prevented disruption of any protein-protein interactions. After 30 min incubation, the sample was centrifuged for 15 min at 14,000 ×g (4°C). The supernatant was then transferred to a clean microtube. This lysate contained the proteins to be immunoprecipitated. An aliquot of lysate was kept for future Western blot analysis, with the remaining lysate proceeding to the Co-IP procedure. Lysate samples were never frozen for storage at this point; again to prevent the disruption of protein-protein interaction.

Buffer name	**Composition**
'Co-IP lysis buffer'	20 mM Tris-HCl pH 8.0, 137 mM NaCl, 10% Glycerol, 1% Tergitol® Type NP-40, 2 mM EDTA, 1× Complete Protease Inhibitor cocktail added just before use

Table 2.2: Chemical composition of *P. berghei* co-IP lysis buffer.

2.2.6.2 Co-immunoprecipitation procedure

To prepare anti-HA magnetic beads (Pierce™ Anti-HA Magnetic Beads; Thermo Scientific; Cat. No. 88836), 175 µl of 1× PBST was placed in a 1.5 ml microtube to which 25 µl of re-suspended anti-HA magnetic beads were then added (Anti-HA magnetic beads were vortexed well before use). Magnetic beads were then placed in a magnetic microtube stand and the supernatant discarded after beads had separated to the side of the tube. Beads were then re-suspended in 200 µl 1× PBST, vortexed, and placed back in the magnetic stand. After separation of beads to the side of the tube, beads were once more washed in 200 µl of 'Co-IP lysis buffer' and the supernatant discarded. These magnetic anti-HA beads were then re-suspended in the protein lysate from section **2.2.6.1**. Beads were re-suspended

in the lysate by gentle pipetting and this mixture incubated for 4 h to overnight at 4°C in a rotator.

Following incubation, the tube was placed in the magnetic stand, the beads separated, and the supernatant removed and stored for future Western blot analysis. Anti-HA beads were then washed 4 times with 400 µl of 'Co-IP lysis buffer'. Washed fractions were stored for western blot analysis. Anti-HA beads (now bound to proteins containing a HA-tag) were re-suspended in 100 µl of 'Co-IP lysis buffer' and transferred to a new tube (this step helped to prevent carry over of contaminating proteins from previous incubation steps). The tube was placed in a magnetic stand, supernatant removed, and the bound proteins eluted by incubation of beads with 50-100 µl 2× SDS sample buffer or 2× Laemmli sample buffer for 10 min at 95°C. The tube was then cooled on ice. Magnetic beads were then separated in the magnetic stand and the eluted proteins transferred to a clean microtube. 1% 2-mercaptoethanol was then added to the sample to reduce disulphide bond formation. Co-immunoprecipitated proteins were then analysed or stored at -20°C for future analyses.

2.2.7 Western blotting (immunoblotting)

2.2.7.1 SDS-PAGE protocol for separation of proteins

SDS-PAGE was used in the present study to separate protein mixtures following acid extraction of histones from *P. berghei,* and to visualise proteins following whole protein extraction, nuclear fractionation, and co-immunoprecipitation of tagged proteins. Throughout this study, all SDS-PAGE and subsequent immunoblotting/Western blotting procedures were carried out using the Bio-Rad Mini-PROTEAN® Tetra Vertical Electrophoresis Cell and accessories for this module. 15% SDS-PAGE resolving gels were prepared for analysis of histones in this study. For standard protein SDS-PAGE (e.g. tagged protein analyses), pre-cast 4-20% Mini- PROTEAN® TGX™ gels were used (Bio-Rad Laboratories, Inc.; Cat. No. 4561093). All reagents used for casting of a 15% SDS-PAGE gel are listed in **Table 2.3** below.

Reagent/Buffer name	Composition (volumes)
15% resolving gel (for one 1.5 mm gel)	3.75 ml Bis-acrylamide (30%), 1.9 ml 1.5 M Tris-HCl pH 8.8, 37.5 µl 20% SDS, 1.8 ml H_2O, 100 µl 10% ammonium persulfate (APS), 10 µl *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED)
5% stacking gel (for four 1.5 mm gels)	650 µl Bis-acrylamide (30%), 1.25 ml 0.5 M Tris-HCl pH 6.8, 25 µl 20% SDS, 3 ml H_2O, 100 µl 10% ammonium persulfate (APS), 10 µl *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED)

Table 2.3: Reagents used for casting a 15% SDS-PAGE gel for analysis of *P. berghei* ANKA histones.

Protein samples to be run on an SDS-PAGE gel were denatured by boiling at 95°C for 10 min in 2× SDS protein sample buffer or 2× Laemmli sample buffer. The inner chamber of the vertical electrophoresis cell was then filled with SDS running buffer (1× Tris/Glycine/SDS buffer from a 10x stock) until the buffer level reached over the top of the gels (but making sure not to overfill the inner chamber). The 'lower' chamber, or the electrophoresis tank as a whole, was then filled with ~200 ml of SDS running buffer or until the tank was ~1/3 full. Protein samples were then loaded into the wells in the stacking gel next to a protein ladder. Gels were then run at 90 V for ~1 h, or until samples had run to the base of the gel. The chemical composition of buffers used for SDS-PAGE in the present study are listed below in **Table 2.4.**

Reagent/Buffer name	Composition
1× TAE buffer	40 mM Tris (pH 7.6), 20 mM acetic acid, 1 mM EDTA
2× SDS-PAGE protein sample buffer	80 mM, 2% SDS, 10% glycerol, 0.0006% Bromophenol blue, 0.1 M dithiothreitol (DTT)

2× Laemmli protein sample buffer	4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% Bromophenol blue, 0.125 mM Tris-HCl pH 6.8

Table 2.4: Buffers used during SDS-PAGE experiments in the present study.

2.2.7.2 Transfer of proteins from SDS-PAGE gel to membrane and antibody binding

In the present study, protein samples were solubilised in 2× SDS protein sample buffer or 2× Laemmli sample buffer before being separated on an SDS-PAGE gel (section **2.2.7.1**). For immunodetection purposes, proteins on an SDS-PAGE gel were first transferred to polyvinylidene fluoride (PVDF) membrane (Amersham™ Hybond™ P0.45 PVDF; GE Healthcare Life Sciences; Cat. No. 10600023) using a Mini-PROTEAN® Tetra Cell tank (Mini Trans-Blot® Module; Bio-Rad Laboratories, Inc; Cat. No. 1658033).

To begin, transfer buffer (1× Tris/Glycine buffer from a 10x stock) was cooled on ice and 4 pieces of blotting paper (Whatman™ 3MM Chr Chromatography Paper; ThermoFisher Scientific) cut to 9 × 7 cm and soaked in pre-cooled transfer buffer. Blotting fibre pads were then soaked in cold transfer buffer with the blotting paper. A section of PVDF membrane (large enough to cover the surface of the gel) was then cut and activated by placing in methanol for 10 s. The PVDF membrane was then washed in distilled water for 5 min on an orbital shaker before being placed in pre-cooled transfer buffer for 10 min (again, on an orbital shaker). The SDS-PAGE gel was then removed from the electrophoresis tank and placed in transfer buffer for 10 minutes (with shaking).

Immunoblotting transfer cassettes were then placed in the transfer tank with an ice pack inserted to maintain a cold temperature, and then the tank filled to the top with ice-cold transfer buffer. Transfer of proteins was then carried out at 90 V for 1 h or at 20 V overnight. Following transfer, the PVDF membrane (now containing proteins of interest) was removed from the cassette and placed in blocking solution for 1-2 h on an orbital shaker at 4ºC in a cold room. Once the

membrane had been blocked, blocking buffer was poured off and the membrane washed three times (5 min each) with 1× TBST. The membrane was then incubated with a primary antibody diluted in 5 ml or 10 ml of antibody dilution buffer with agitation on an orbital shaker overnight at 4ºC. After incubation with the primary antibody, the membrane was washed 3 times with 1× TBST (5 min each). A secondary horseradish peroxidase (HRP)-conjugated antibody was then diluted in 5-10 ml dilution buffer and the membrane incubated with this at room temperature for 1-2 h with gentle agitation. Following the final incubation, the membrane was washed 3 times with 1× TBST (5 min each) before being imaged using a commercial chemiluminescence kit and X-ray detection (Medical X-ray Blue/MXRE Film 18 × 24 cm; Carestream Healthcare, Inc.; Product No. 814 3059). Solutions used for transfer of proteins from SDS-PAGE gels to PVDF membranes are listed in **Table 2.5.**

Reagent/Buffer name	**Composition**
10× Tris-buffered saline (TBS)	0.2 M Tris, 1.5M NaCl
1× Tris-buffered saline, 0.1% Tween-20 (1× TBST)	20 mM Tris, 150 mM NaCl, 0.1% Tween-20
1× SDS-PAGE Transfer buffer	190 mM glycine, 25 mM Tris, 20% methanol
Western blot blocking/Antibody dilution buffer	20 mM Tris, 150 mM NaCl, 0.1% Tween-20, 5% (w/v) non-fat milk powder

Table 2.5: Solution compositions for transfer of proteins to a PVDF membrane prior to immunoblotting.

2.2.7.3 Detection of antibody-bound proteins by chemiluminescence

To proceed with the detection step of the Western blot, the peroxidase solution (with luminol) and Acridan solution, from a commercial chemiluminescence kit were mixed together according to the manufacturers guidelines. This mixture of solutions was then pipetted onto the membrane that had previously been probed with a HRP-conjugated antibody. The solution mixture was kept on the membrane for the length of time indicated in the kit guidelines. The membrane was then

wrapped in plastic ('cling film'. This blot was then placed in an autoradiography cassette (Amersham™ Hypercassette™; GE Healthcare Life Sciences; Product No. RPN11629) and brought to a darkroom for development. To detect proteins using X-ray film, the X-ray film was placed over the chemiluminescent Western blot for increments of 10 s, 30 s, 1 min, 5 min, 10 min, or longer if necessary; until protein bands were observed. Western blot membranes were then discarded after detection. X-rays were scanned and stored in a dark place.

2.2.8 Protein sample preparation for LC-MS using a FASP protein digestion kit

In the present study, all samples to be analysed by liquid chromatography-mass spectrometry (LC-MS) were prepared using a filter-aided sample preparation (FASP) digestion kit (FASP Protein Digestion Kit; Expedeon Ltd.; Cat. No. 44250) with a maximum capacity of 0.4 mg of protein in 30 µl of solution.

2.2.8.1 Preparation of urea sample solution

Fresh urea sample solution was prepared before each new protein sample digestion. 1 ml of Tris-HCl was added to one tube of urea (provided in the FASP kit). This tube was then vortexed until all urea powder had dissolved.

2.2.8.2 Preparation of 10× Iodoacetamide solution

10× iodoacetamide solution was prepared fresh before each protein sample digestion. To begin, 100 µl of the urea sample solution was added to one tube of iodoacetamide from the FASP kit. Iodoacetamide powder was dissolved in solution by pipetting up and down approximately 15 times (or until powder had fully dissolved). The fresh solution was then transferred to a clean microtube.

2.2.8.3 Preparation of digestion solution

Digestion solution was prepared fresh for each sample (75 µl each time) by dissolving 4 µg of trypsin powder in 75 µl of 50 mM ammonium bicarbonate solution to a final concentration of 0.05 µg/µl.

2.2.8.4 FASP digestion procedure

To begin digestion, up to 30 µl (at 0.4 mg) of protein extract was mixed with 200 µl of urea sample solution in a spin filter and centrifuged at 14,000 ×g for 15 min at room temperature. An additional 200 µl of urea sample solution was then added to the spin filter and centrifuged at 14,000 ×g for 15 min once more. The flow-through was then discarded (from the collection tube) and 10 µl 10× iodoacetamide solution and 90 µl urea sample solution added to the filter. This was vortexed for 1 min and then the spin filter incubated for 20 min at room temperature in the dark (without mixing). After incubation, the spin filter was centrifuged at 14,000 ×g for 10 min at room temperature. 100 µl of urea sample solution was then added to the spin filter and centrifuged 14,000 ×g for 15 min.

This step was then repeated twice more and the flow-through discarded. After this, 100 µl of 50 mM ammonium bicarbonate solution was added to the spin filter and centrifuged at 14,000 ×g for 10 min at room temperature. Again, this step was repeated twice more and the flow-through discarded. 75 µl of digestion solution (enzyme-to-protein ratio: 1:100) was added to the spin filter and the filter vortexed for 1 min. Parafilm was then wrapped around the top of the tube to prevent evaporation from the spin filter. The spin filter was then incubated at 37°C for 4-18 h. Following incubation, the spin transfer was then transferred to a clean collection tube and 40 µl of 50 mM ammonium bicarbonate solution added. The spin filter was then centrifuged at 14,000 ×g for 10 min at room temperature. A further 40 µl of 50 mM ammonium bicarbonate solution was then added and the filter centrifuged at 14,000 ×g for 10 min at room temperature once more and the flow-through discarded. Digested protein extract was then eluted by the addition of 50 µl 10% acetonitrile solution to the spin filter and centrifugation of this spin filter at 14,000 ×g for 10 min at room temperature. The eluted filtrate was then acidified by adding 1.0% trifluoroacetic acid (TFA) to the desired pH and desalting.

2.2.9 LC-MS/MS procedure for histone analysis

All samples in the present study were analysed on an Orbitrap™ Elite™ mass analyser coupled with an UltiMate™ 3000 RSLCnano ultra high performance liquid chromatography (UHPLC) system. Samples (5 µl) were loaded onto an Acclaim™

PepMap100™ column (75 μm i.d. x 15 cm, 3 μm C18; Thermo Scientific™) at a flow rate of 5 μl min^{-1} by the UltiMate™ 3000 RSLCnano autosampler (Thermo Scientific™). Loading solution composition was 0.1% formic acid and acetonitrile (98:2). Loaded sample was then washed into an Acclaim™ PepMap100™ trap column (100 μm i.d. x 2 cm, 5 μm, 100Å C18; Thermo Scientific™) at a flow rate of 0.3 μm min^{-1}. One-hour linear HPLC gradients were performed from 0-35% solvent B to 35% solvent B (A = 0.1% formic acid; B = 80% acetonitrile 0.08% formic acid). Nano flow and trap columns were maintained at 35°C.

Column eluate was directed to a Triversa NanoMate® electrospray ionisation (ESI) source (Advion™) operating in positive ion mode (2+) and then into the Orbitrap™ Elite™ mass analyser. Ionisation voltage was 2.5 kV and the capillary temperature was 230°C. Tandem mass spectrometry (MS/MS) was carried out from 300-2000 atomic mass units (amu) and the top 20 multiply-charged ions were selected from each full scan for MS/MS analysis. Fragmentation was by collision-induced dissociation (CID) at 35% collision energy. Ions were selected for MS^2 using a data-dependent method with a repeat count of 1 s and an exclusion time of 60 s. Precursor ions with unassigned charged states or a charge state (z) of 1 were rejected. The resolution of ions in the first stage (MS^1) was 60,000 and 7500 for the second stage (CID MS^2). Data were acquired using XCalibur™ software (version 2.1) (Thermo Scientific™).

2.2.10 LC-MS/MS data analysis

Raw LC-MS/MS data files ('.RAW' files from XCalibur™) were processed using Proteome Discoverer™ software (version 2.0; Thermo Scientific™). Raw files were converted to Mascot Generic Format (MGF) files in Proteome Discoverer™ and probability-based identification of histone peptides and PTMs from MS/MS data was carried out using the Mascot search engine (version 2.6 with patch release 2.6.2) (Perkins *et al.*, 1999). Peptide information for *P. berghei* ANKA was derived from the *Plasmo*DB database (version 35). Searches were performed with a peptide tolerance of 5 parts per million (ppm) (#^{13}C = 1; monoisotropic mass values), a fragment mass tolerance of ±0.3 Da, and a maximum of 5 missed cleavages by trypsin. Mascot search settings were for the fixed modification,

carbamidomethyl (C), and variable modifications of N-terminal acetylation (N-term), lysine acetylation (K), lysine mono-, di-, and tri- methylation (K), arginine mono-, di-, and tri- methylation (R), and phosphorylation (ST, Y). Error tolerant searches for all significant (P < 0.05) protein hits were displayed as shown in **Figure 2.5.** Comparative analysis of significant (P < 0.05) histone PTMs between *P. berghei* ANKA mature asexual and sexual stage parasites is described in **Chapter 4.**

Re-Search ◉ All queries ○ Unassigned ○ Below homology threshold ○ Below identity threshold

1. PBANKA_1420600.1-p1 **Mass:** 13741 **Score:** 5911 **Matches:** 228(228) **Sequences:** 12(12) **emPAI:** 852.41
histone H2B variant, putative

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
106	408.7321	815.4497	815.4501	-0.44	0	49	0.0084	1	U	R.EIQTAVR.L 87
198	414.7139	827.4132	827.4137	-0.58	0	52	0.00049	1		K.HAVSEGTK.A 189 190 191 192
1261	471.7715	941.5285	941.5294	-0.94	1	73	1.2e-05	1	U	K.TAGKTLGPR.H 1255 1256 1257
1677	485.3028	968.5910	968.5906	0.38	0	44	0.0059	1	U	R.LLLPGELSK.H 1668 1672 1675
2726	535.8192	1069.6238	1069.6244	-0.58	2	(44)	0.021	1	U	K.KTAGKTLGPR.H 2725
3219	556.8246	1111.6346	1111.6349	-0.28	2	61	0.00036	1	U	K.KTAGKTLGPR.H 3205 3206 320
4288	617.8207	1233.6268	1233.6281	-1.07	0	58	0.0017	1	U	R.TESFSLYIFK.V
5919	464.2502	1389.7286	1389.7292	-0.43	1	(46)	0.038	1	U	R.RTESFSLYIFK.V 5918
5921	695.8719	1389.7293	1389.7292	0.08	1	50	0.015	1	U	R.RTESFSLYIFK.V 5920
7247	516.2837	1545.8292	1545.8303	-0.71	2	47	0.048	1	U	K.RRTESFSLYIFK.V
7956	820.4597	1638.9049	1638.9053	-0.24	3	96	6.3e-07	1	U	K.SQAAKKTAGKTLGPR.H 7930 793
8920	908.9212	1815.8279	1815.8284	-0.28	0	(93)	1.2e-06	1	U	K.SMNIMNSFINDIFDR.L 8921
8928	909.4119	1816.8093	1816.8124	-1.69	0	(64)	0.00096	1	U	K.SMNIMNSFINDIFDR.L
8929	909.4128	1816.8111	1816.8124	-0.69	0	(75)	6.9e-05	1	U	K.SMNIMNSFINDIFDR.L
8977	916.9150	1831.8154	1831.8233	-4.30	0	112	1.6e-08	1	U	K.SMNIMNSFINDIFDR.L 8979
8978	916.9183	1831.8221	1831.8233	-0.64	0	(88)	4.1e-06	1	U	K.SMNIMNSFINDIFDR.L

Figure 2.5: Example Mascot output following LC-MS/MS of acid-extracted histone samples. This image is a screenshot of the Mascot (version 2.6) search engine output for a sample of G1142 mature gametocyte histones following LC-MS/MS as described in section **2.2.8.** The display is showing a selection of MS/MS spectra that have mapped to the putative *P. berghei* ANKA histone 2B (H2B) variant (H2B.Z) (PBANKA_1420600.1-p1) as described in *Plasmo*DB (version 35). In this instance, H2B.Z was the highest scoring protein (according to the Mascot scoring algorithm) with a score of 5911.

In the example experiment shown in **Figure 2.5,** the expected histone H2B.Z mass was 13741 atomic mass units/Daltons (amu/Da). The number of 'matches' describes the number of MS/MS spectra that matched to this protein, with the number of significant matches in brackets. In this case, all 228 matched spectra were significant. The 'sequences' value describes a count of matches to distinct peptide sequences, with significant matches in brackets. Again, in this example protein (H2B.Z), all 12 distinct peptide sequences had significant matches. To the very right is the emPAI (Exponentially Modified Protein Abundance Index) value, an approximate quantitation of this protein in the sample mixture based on protein coverage by peptide matches (Ishihama *et al.*, 2005). An emPAI value is

given only if 100 MS/MS spectra or more are present within the sample. In the example above, the 852.41 value was derived from $10(N^{observed}/N^{observable})$-1, where '$N^{observed}$' was the number of experimentally-derived peptides for H2B.Z and '$N^{observable}$' was the calculated number of observable peptides for the H2B.Z protein.

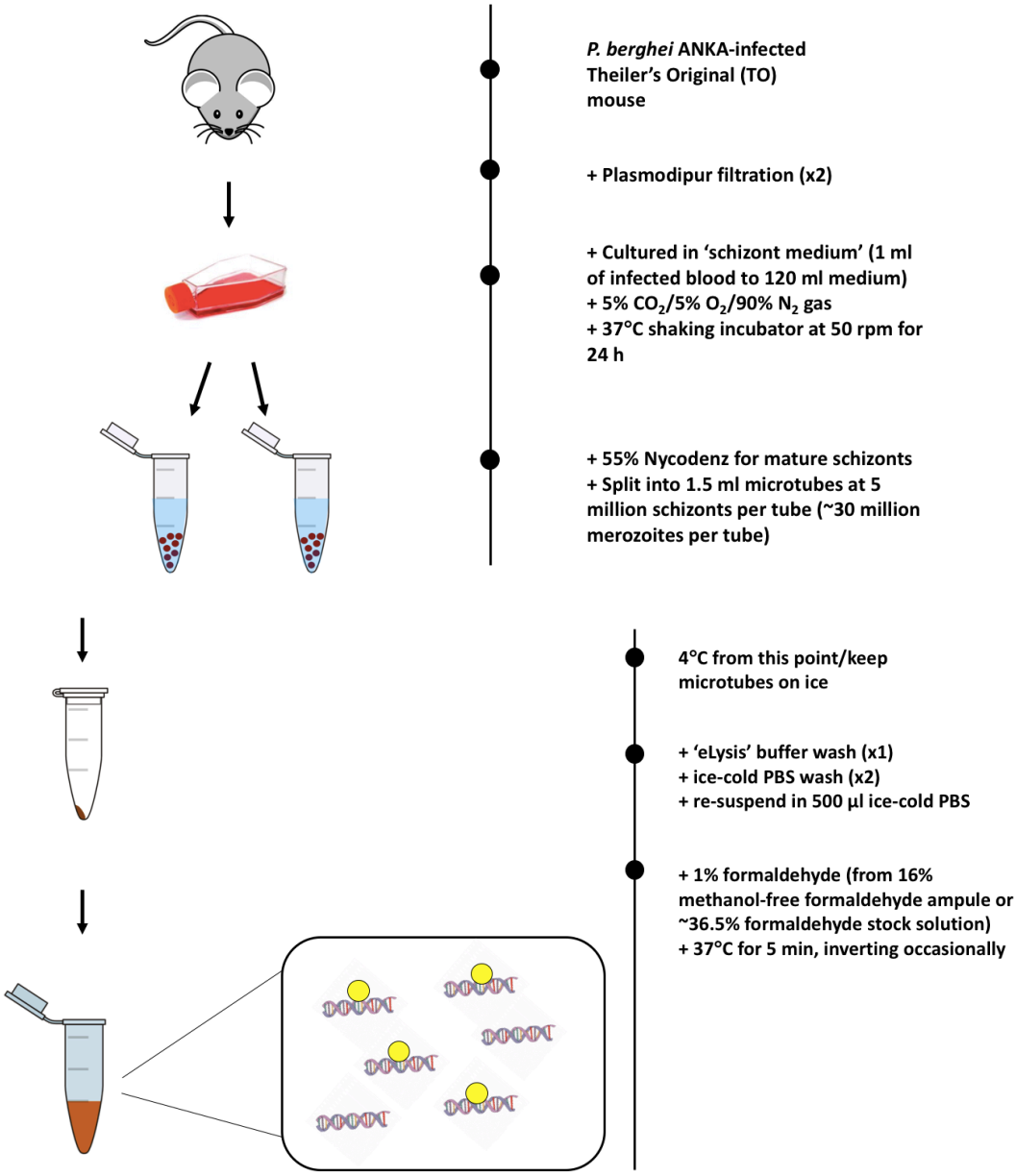
Below the protein descriptor from *Plasmo*DB in **Figure 2.5**, all matched queries are shown, with each query (an MS/MS spectrum 'queried' against *Plasmo*DB known sequences) given a unique identifying number (on the left in blue). The experimentally-derived mass/charge (m/z) value for each peptide is labelled in the 'observed' column, followed by values showing the experimental m/z value transformed to a relative molecular mass (Mr(expt)) and the calculated relative molecular mass (Mr(calc)) based on the matched peptide sequence. The parts per million ('ppm') value shows the difference between (Mr(expt)) and (Mr(calc)). In the 'miss' column, the number of missed protease cleavage sites are shown, with the 'score' value showing the Mascot-assigned 'ion score' (a measure of how well the observed MS/MS spectrum matches to the peptide). 'Score' values in brackets are scores for duplicate MS/MS spectra, but ones with lower ion scores than their duplicate.

In **Figure 2.5,** the expectation value ('Expect') is a measure of the likelihood that the given score would be seen by chance, with the 'rank' column showing a rank for the ion match from 1-10 (1 being the best match). A letter 'U' is assigned in the 'Unique' column if a query is unique to the protein hit. The final column then shows the amino acid sequence, with a period indicating a cleavage site, and a dash indicating the end of the protein sequence (a carboxyl or amino terminus). The Mascot output (**Figure 2.5)** also provides links to variable modifications on each peptide. These are shown in blue and in brackets to the right of the peptide single-letter sequence. Variable modifications in each sequence are underlined and those that correspond to PTMs are underlined in the main 'peptide' sequence. Red text indicates that the peptide shown is the highest-scoring match for a given MS/MS spectrum, with bold text indicating the highest scoring protein to which the peptide sequence appears (out of all proteins identified in the sample run) (Perkins *et al.*, 1999). Results such as those seen in **Figure 2.5** were obtained for

all mature schizont and gametocyte histone samples run in the present study, with results described in **Chapter 4.**

2.2.11 Chromatin immunoprecipitation coupled with high-throughput DNA sequencing (ChIP-seq) protocol for *P. berghei* parasites

A standard ChIP procedure for the examination of DNA regions bound to specifically-modified histones in *P. berghei* ANKA cells is depicted in **Figure 2.6** (original ChIP protocol provided by Dr Agnieszka Religa at the University of Glasgow, in collaboration with Dr Valerie M. Crowley of the University Health Network, Toronto General Hospital, Canada).



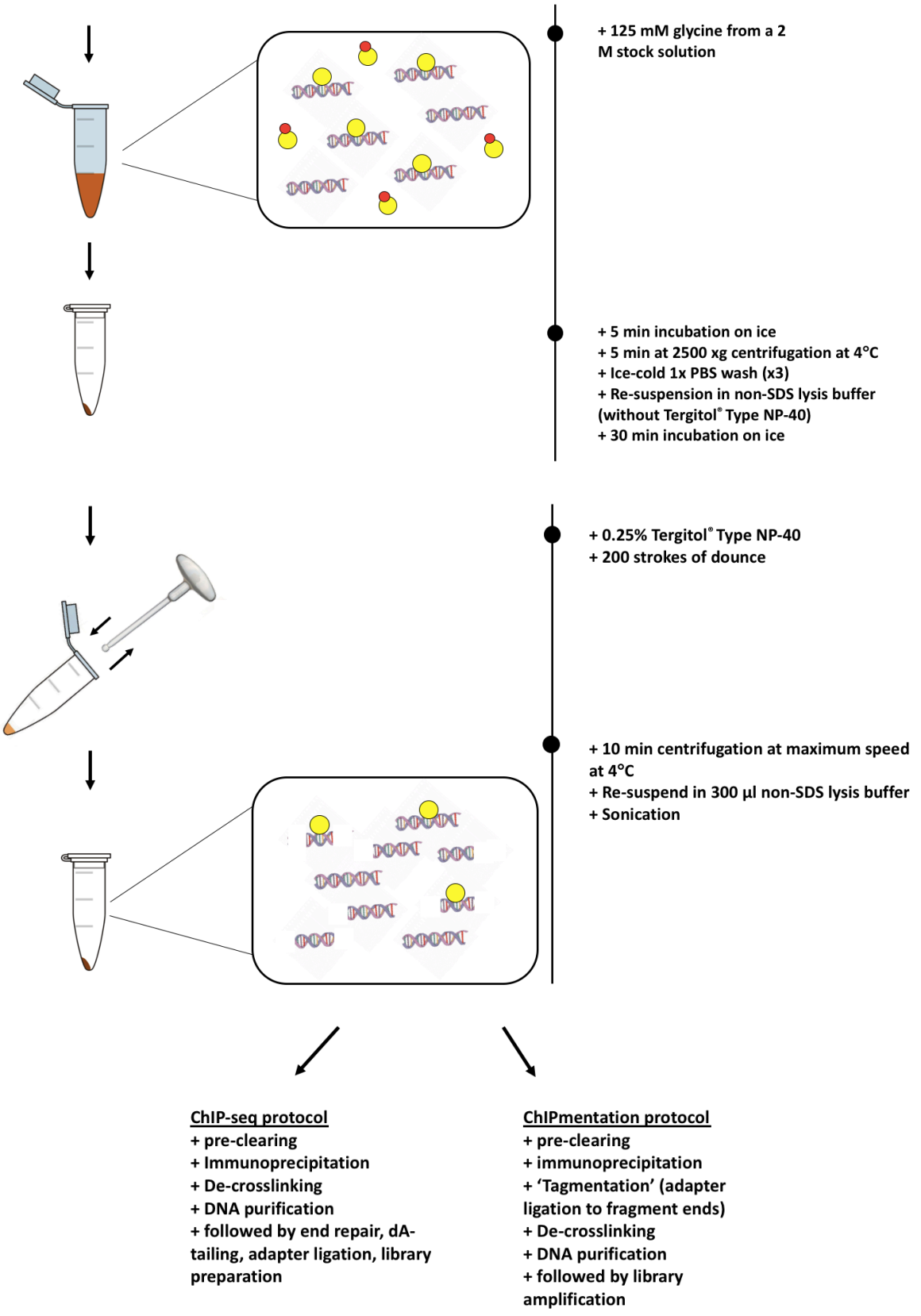
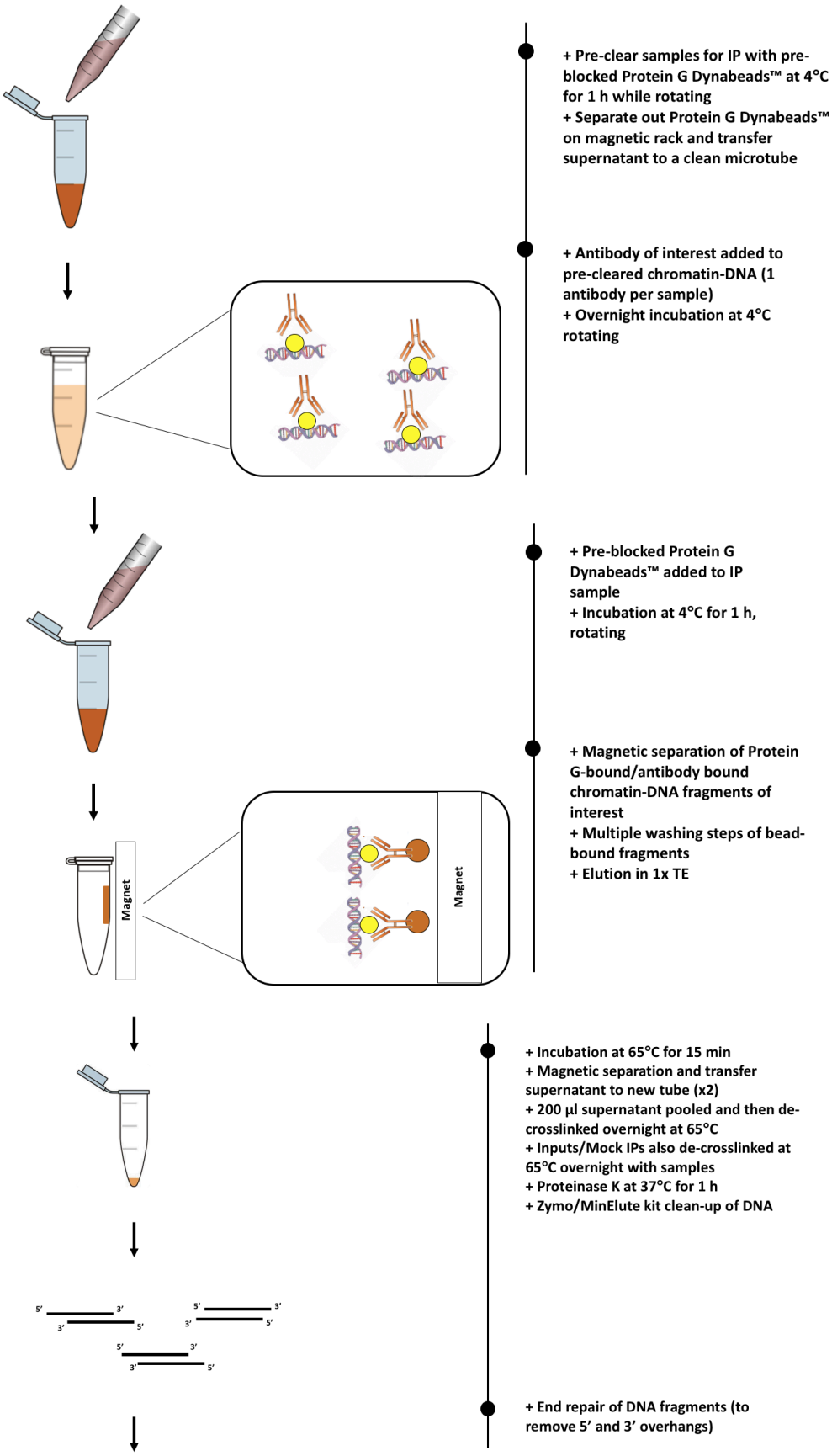


Figure 2.6: Chromatin immunoprecipitation (ChIP) of *P. berghei* ANKA parasites: from cell culture to library preparation. The stages of chromatin immunoprecipitation (ChIP) from collection of infected blood from a TO mouse to the point at which ChIP-seq and ChIPmentation procedures diverge (the point just after immunoprecipitation (IP) of fragmented chromatin-DNA)

are shown in the images above. In the present study, only mature asexual *P. berghei* ANKA parasites were examined by ChIP-seq or ChIPmentation to maximise genetic material with which to work with, and to optimise both ChIP-seq and ChIPmentation protocols for future ChIP studies in *P. berghei*. To begin with, infected blood was collected from animals via cardiac puncture and filtered through Plasmodipur filters as described in section **2.1.14.** Mature asexual parasites (schizonts) were then purified from culture using a 55% Nycodenz gradient, and samples split into microtubes containing 5 million parasites each (equating to ~30 million merozoites, assuming each *P. berghei* ANKA schizont contained 12 segmented merozoites). These tubes were then placed on ice and remained at this cooled temperature (or roughly 4ºC) for the remainder of the ChIP procedure. Erythrocyte lysis buffer and ice-cold 1x PBS washes were then used to remove host red blood cells contaminants and any remaining culture medium contaminants. Crosslinking, quenching and fragmentation of chromatin-DNA was then carried out as described in sections **2.2.11.1** and **2.2.11.2.** In the cartoon above, free proteins/chromatin (in yellow) are shown being bound by glycine molecules (represented by red circles). In each case, some unbound DNA strands are shown, as it is always possible that not all chromatin-DNA interactions are preserved, even after formaldehyde treatment.

In **Figure 2.6**, the initial ChIP procedure is shown up to the pre-clearing and immunoprecipitation steps. At this point, i.e. just following sonication, input samples are chosen that are not subject to immunoprecipitation. For ChIP analyses of histone modifications and their associated DNA regions, mock antibody controls or a histone 3 (H3) antibody pulldown could also be used as a control for ChIP samples (Flensburg *et al.*, 2014). In the present study, an anti-GFP mock antibody control was used in an initial pilot study of ChIPmentation in *P. berghei* schizonts. All remaining experiments were carried out with input ChIP-DNA controls, i.e. sonicated chromatin-DNA after de-crosslinking, with no immunoprecipitation step.

Though the 'ChIPmentation' protocol described in section **2.2.12** also incorporates a ChIP component (as depicted in **Figure 2.6**) and culminates in DNA sequencing of amplified ChIP-DNA libraries, a standard ChIP-seq protocol is defined by its library preparation method, incorporating end-repair of de-crosslinked DNA fragments, dA-tailing, adapter ligation, and library amplifications, size selections, and purifications (Furey, 2012) (**Figure 2.7**).


+ Pre-clear samples for IP with pre-blocked Protein G Dynabeads™ at 4°C for 1 h while rotating
+ Separate out Protein G Dynabeads™ on magnetic rack and transfer supernatant to a clean microtube
+ Antibody of interest added to pre-cleared chromatin-DNA (1 antibody per sample)
+ Overnight incubation at 4°C rotating
+ Pre-blocked Protein G Dynabeads™ added to IP sample
+ Incubation at 4°C for 1 h, rotating
Magnet
Magnet
+ Magnetic separation of Protein G-bound/antibody bound chromatin-DNA fragments of interest
+ Multiple washing steps of bead-bound fragments
+ Elution in 1x TE
+ Incubation at 65°C for 15 min
+ Magnetic separation and transfer supernatant to new tube (x2)
+ 200 µl supernatant pooled and then de-crosslinked overnight at 65°C
+ Inputs/Mock IPs also de-crosslinked at 65°C overnight with samples
+ Proteinase K at 37°C for 1 h
+ Zymo/MinElute kit clean-up of DNA
+ End repair of DNA fragments (to remove 5' and 3' overhangs)

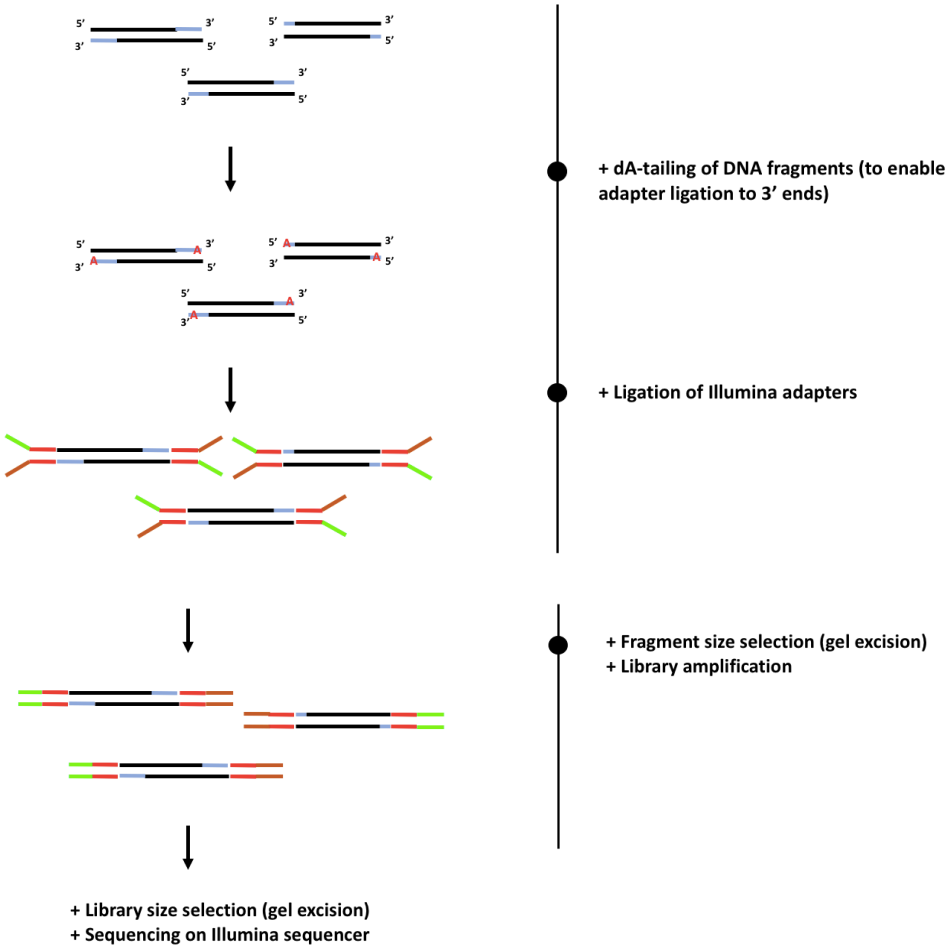


Figure 2.7: Standard ChIP-seq procedure in *P. berghei* ANKA from immunoprecipitation to sequencing on an Illumina platform. In the sequence of images above, the complete protocol for a standard *P. berghei* ANKA ChIP-seq is shown. The complete materials and methods for ChIP-sequencing using standard library preparation procedures are described from section **2.2.11.1** to section **2.2.11.12.**

The ChIP-seq library preparation method detailed below is modified from a protocol provided by Dr David Vetrie as part of the 'Chromatin Structure and Function' Wellcome Trust Advanced Course, held at the Wellcome Genome Campus, Hinxton, Cambridge, UK: 'PROTOCOL 3: Preparation of ChIP DNAs for sequencing on the Illumina Genome Analyzer IIe'.

Buffer name	**Buffer components**
Lysis Buffer (non-SDS) (kept cold)	10 mM HEPES (pH 7.9), 10 mM KCl, 0.1 mM EDTA (pH 8.0), 0.1 mM EGTA (pH 8.0), 1 mM DTT (just before use), 1×

	protease inhibitor cocktail (from 25× stock- added just before use), 0.25% Tergitol® Type NP-40 (added just before use directly into sample)
1% SDS Lysis Buffer	1% SDS, 10 mM EDTA (pH 8.0), 50 mM Tris-HCl (pH 8.1)
Dilution buffer	0.01% SDS, 1.1% Triton-X100, 1.2 mM EDTA (pH 8.0), 16.7 mM Tris-HCl (pH 8.1), 167 mM NaCl, 1× protease cocktail inhibitor
Low Salt Wash Buffer	0.10% SDS, 1% Triton X-100, 2 mM EDTA (pH 8.0), 20 mM Tris-HCl (pH 8.1), 150 mM NaCl
Salty Wash Buffer	0.10% SDS, 1% Triton X-100, 2 mM EDTA (pH 8.0), 20 mM Tris-HCl (pH 8.1), 500 mM NaCl
LiCl Wash Buffer	250 mM LiCl, 1% Tergitol® Type NP-40, 1% sodium deoxycholate (from fresh 5% stock), 1 mM EDTA (pH 8.0), 10 mM Tris-HCl (pH 8.0)
Elution Buffer	1% SDS, 0.1 M $NaHCO_3$ (from single use aliquots)
TE Buffer	10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)

Table 2.6: Composition of buffers for a standard ChIP-seq procedure in *P. berghei*.

2.2.11.1 Crosslinking of chromatin to DNA

The procedure began by chilling glass dounce tissue grinder sets (Sigma-Aldrich; Cat. No. D8938) on ice and pre-cooling a centrifuge to 4°C. All stages of this protocol were carried out on ice unless otherwise stated. Parasites of a synchronised asexual or sexual stage population (Plasmodipur filtered and Nycodenz/magnetically separated) were washed twice in 1 ml of ice-cold 1× PBS and then re-suspended in 150 µl to 500 µl of ice-cold 1× PBS, depending upon

parasite numbers (cell number guidelines were taken from (Schmidl *et al.*, 2015)), with small cell numbers being up to 5×10^5 cells and large cell numbers from 5×10^5 cells up to 10×10^6 cells. If parasite numbers were larger, samples were split into smaller cell numbers and the ChIP DNA pooled after the procedure. Re-suspended parasites were crosslinked with 1% formaldehyde (from a 16% stock of methanol-free formaldehyde). The samples were mixed by inversion of the microtube and incubated at 37°C for 5 min, inverting occasionally. After incubation, the samples were briefly spun in a microfuge to pool any evaporated material to the base of the microtube. The crosslinking reaction was then quenched with 125 mM glycine from a 2 M glycine stock. The samples were mixed well by gentle inversion of the tubes and then incubated on ice for 5 min. After this incubation period, samples were then centrifuged at 2500 ×g for 5 min at 4°C and the supernatant discarded.

Samples were then washed 3 times with 1 ml 1× PBS (each time re-pelleting the sample at 2500 ×g for 5 min at 4°C). The pellet was then re-suspended in cold **non-SDS** lysis buffer (without Tergitol® Type NP-40) to a concentration of ~3.33×10^5 parasites/µl, or fewer in the case of small cell numbers. Samples were incubated on ice for 30 min before Tergitol® Type NP-40 was added to a final concentration of 0.25%. One sample at a time was then transferred to a dounce (up to 2 ml) on ice and the sample homogenised with 200 strokes of the dounce (repeated with more of the sample if the sample was greater than 2 ml). For very small cell numbers, samples were retained within the microtubes on ice and homogenised using single-use disposable pestles (Fisherbrand™ RNase-Free Disposable Pellet Pestles; Fisher Scientific; Cat. No. 13236679). The samples were pipetted from the dounce (if glass dounce was used) and placed into clean microtubes. Samples greater than 2 ml were pooled together once more if separated for douncing. Tubes were then centrifuged at maximum speed for 10 min at 4°C to pellet any insoluble material. The supernatant was then transferred to a clean microtube.

Note: In the case of cell types that were difficult to lyse, or if the 'insoluble' pellet appears quite large, the process of douncing in non-SDS, cold lysis buffer was repeated to obtain more sample. At this point, SDS, or any other

denaturing detergent was not used, as it may have disrupted protein-protein or protein-DNA interactions. Additional supernatant could then be added to the previous eluate.

Samples were topped up to a total of 300 µl per microtube in cold **non-SDS** lysis buffer prior to sonication.

2.2.11.2 Sonication of crosslinked chromatin/DNA

Aliquots of 300 µl per microtube of cross-linked chromatin/DNA were sonicated using a Bioruptor® Plus sonication device (Diagenode; Cat. No. B01020001) at the 'high power' setting in ice water. The following sonication sequence was used each time: 20 sec ON, 2 min OFF, 20 sec ON, 3 min OFF, 30 sec ON, 15 sec OFF, 15 sec ON, 2 min OFF, 30 sec ON.

2.2.11.3 Determination of DNA fragment size and concentration

After sonication, samples were centrifuged at 10,000 ×g for 10 min at 4°C. The supernatant (sonicated chromatin/DNA) was then transferred to a clean microtube and kept on ice. Chromatin/DNA concentration was checked using a NanoDrop at OD_{260}. The concentration of DNA in µg/ml is $OD_{260} \times 10{,}000$. This method was used to calculate the DNA concentration of the chromatin preparation. A 5-10 µl sample of chromatin/DNA was also run on a 1% agarose gel to check DNA fragment size after sonication. A majority of chromatin/DNA at ~200 bp is ideal but a smear on an agarose gel may stretch from 150 bp to 2000 bp. If fragments were too large, samples were sonicated once more according to the sequence in section **2.2.11.2.** After determination of fragment size and concentration, ~1 µg of DNA (or as close to 1 µg as possible) per sample was used for immunoprecipitation.

2.2.11.4 Chromatin/DNA pre-clearing

Prior to pre-clearing, fresh dilution buffer was made up and kept on ice. Samples of chromatin/DNA were then diluted in dilution buffer to a concentration of 1 µg of DNA in 500 µl. Magnetic Protein G Dynabeads™ (Thermo Scientific™; Cat. No. 10003D) were vortexed until mixed well and 50 µl of beads per sample separated

from solution using a magnetic rack. Each 50 µl of beads was then re-suspended in 250 µl blocking buffer (1× PBS with 5% BSA) and placed at 4°C rotating for 1 h. After blocking, tubes containing magnetic Protein G beads were placed in the magnetic tube rack and the beads separated from solution. The blocking buffer was then discarded and beads re-suspended in the 500 µl chromatin/DNA sample. Chromatin/DNA samples were then pre-cleared by rotating at 4°C for 1 h. After pre-clearing, the magnetic beads were separated from the chromatin/DNA solution using a magnetic tube rack. Pre-cleared samples were then pipetted from this tube into a clean microtube for immunoprecipitation.

2.2.11.5 Immunoprecipitation

To each sample, the antibody of interest was added at a 1:100 antibody dilution. All samples were incubated overnight at 4°C on a rotator. A non-antibody control (input) was also incubated overnight in the same manner.

2.2.11.6 De-crosslinking and DNA purification

Following overnight incubation with antibodies (or no antibody in the case of the control), fresh wash buffers were made up and kept on ice. Fresh elution buffer (TE buffer) was then made up and kept at room temperature. A heating block was pre-warmed to 65°C. Samples were then taken from the rotator at 4°C and briefly spun in a microfuge to ensure that no solution remains in the lid of the microtubes. 40 µl of fresh Protein G magnetic beads were then added to each sample and control tube. These tubes were then rotated once more at 4°C for 1 h. After this incubation, samples were placed in a magnetic microtube rack and each sample washed with 1 ml of the following solutions in sequence:

1) Once in low salt wash buffer
2) Once in salty wash buffer
3) Once in LiCl wash buffer
4) Twice in TE buffer

The chromatin/DNA complexes were then eluted by adding 100 µl of the fresh TE buffer to each sample and vortexed briefly at low speed. Tubes were then spun briefly in a microfuge and the samples tapped to re-suspend the beads in TE solution. All samples were then placed in the 65°C heat block for 15 min. After 15 min, the samples were vortexed at low speed and spun in a microfuge briefly. The samples were tapped to re-suspend beads in solution and then the samples incubated at room temperature for 1 min. Samples were then placed in the magnetic tube rack and the beads separated from solution. Chromatin/DNA that was eluted in the 100 µl TE buffer was removed from these tubes and placed in clean microtubes (kept on ice). The microtubes containing the beads were then filled with another 100 µl of fresh TE buffer and the samples vortexed at low speed, spun on a microfuge briefly and the tubes tapped to re-suspend beads. These beads in solution were then once again incubated at 65°C, this time for 5 min.

After incubation, the samples were vortexed, spun on a microfuge and the beads separated from solution using the magnetic tube rack as had been done previously. The new 100 µl of chromatin/DNA in solution was then pooled with the previous 100 µl, creating 200 µl samples of eluted chromatin/DNA. Inputs were also brought to 200 µl with the appropriate amount of dilution buffer. All 200 µl samples and inputs were then placed at 65°C overnight for de-crosslinking. After de-crosslinking, tubes were removed from 65°C heat block and 1 µl of Proteinase K (10 mg/ml) added to each tube (including inputs). Samples were then placed at 37°C for 1 h. Following this final digestion step, DNA was purified using the Zymo DNA Clean and Concentrator™-5 kit. DNA was eluted in 100 µl of distilled water. Samples were stored at -20°C at this point or immediately taken forward to the library preparation procedure for sequencing (preparation for Illumina sequencing in this study).

2.2.11.7 ChIP DNA fragment end repair

To begin, Klenow DNA polymerase (DNA polymerase I, Large (Klenow) Fragment; New England BioLabs, Inc.; Cat. No. M0210S) was diluted 1:5 in distilled water for a final concentration of 1 U/µl. T4 DNA Polymerase (New England BioLabs, Inc.;

Cat. No. M0203S) was then diluted 1:3 in distilled water to give a final concentration of 1 U/µl also. A T4 polynucleotide kinase (New England BioLabs, Inc.; Cat. No. M0201S) and T4 DNA ligase buffer with 10 mM ATP (10× T4 DNA Ligase Reaction Buffer; New England BioLabs, Inc.; Cat. No. B0202S) were also used in the end repair procedure. The following reaction was then prepared:

Reagent	**Volume (for 50 µl total reaction)**
ChIP DNA	20 µl
Distilled water	20 µl
T4 DNA ligase buffer with 10 mM ATP	5 µl
dNTP mix	2 µl
T4 DNA polymerase (1:3)	1 µl
Klenow DNA polymerase (1:5)	1 µl
T4 polynucleotide kinase (T4 PNK)	1 µl

Table 2.7: Reagents necessary for one ChIP DNA end repair reaction.

Each 50 µl reaction was then placed in a thermal cycler and incubated at 20ºC for 30 min. For DNA clean-up after end repair, the QIAquick® PCR Purification Kit (QIAGEN; Cat. No. 28106) was used. 250 µl of Buffer PB was added to 1 volume of the end repair sample and mixed. A QIAquick® spin column was then placed in a provided 2 ml collection tube and the sample added to this column. Samples were then centrifuged at 13,000 rpm (as per manual) for 60 s. The flow-through was discarded and then the QIAquick® spin column washed with 750 µl of Buffer PE. Samples were then spun once more at 13,000 rpm for 60 s. The flow-through was discarded and the column spun again at 13,000 rpm for 60 s in the collection tube to remove residual buffer or ethanol. The QIAquick® spin column was then placed in a clean 1.5 ml microtube and 34 µl of Buffer EB added to elute DNA. The sample was then spun at 13,000 rpm for 60 s and the QIAquick® spin column discarded.

2.2.11.8 ChIP DNA fragment dA-tailing

For preparation of Illumina libraries for sequencing, a deoxyadenosine 5'-monophosphate (dAMP) was incorporated onto the 3' end of blunted DNA fragments, a process known as dA-tailing. This process enabled the ligation of

adapters with complementary dT-overhangs to the ChIP DNA fragments and prevented concatamer formation at downstream ligation steps. In the present study, a truncated DNA Polymerase I (Klenow Fragment 3'->5' exo-; New England BioLabs, Inc.; Cat. No. M0212S) was used. The dA-tailing reaction was prepared as follows:

Reagent	**Volume (for 50 µl total reaction)**
End-repaired DNA sample	34 µl
Klenow buffer	5 µl
dATP	10 µl
Klenow Fragment 3'->5' exo-	1 µl

Table 2.8: Reagents for one ChIP DNA fragment dA-tailing reaction.

Samples were then incubated at 37ºC for 30 min. Purification of dA-tailed ChIP DNA was then undertaken using a MinElute® PCR Purification Kit (QIAGEN; Cat. No. 28004). For this clean-up step, 250 µl of Buffer PB was added to 1 volume of dA-tailed DNA and the sample mixed. A MinElute® spin column was then placed in a provided 2 ml collection tube and the dA-tailed DNA added to the column. To bind DNA, the MinElute® spin column was centrifuged at 13,000 rpm for 60 s (according to manual recommendations). The flow-through was discarded and 750 µl of Buffer PE buffer added and then the sample spun at 13,000 rpm for 60 s once more. The flow-through was discarded and the MinElute® spin column placed back into the empty collection tube and centrifuged at 13,000 rpm for 60 s to remove residual buffer and ethanol. The MinElute® spin column was then placed in a clean 1.5 ml microtube. dA-tailed ChIP DNA was then eluted by addition of 10 µl Buffer EB to the column and the column spun at 13,000 rpm for 60 s. The column was then discarded.

2.2.11.9 Ligation of adapters to ChIP DNA

Adapters and reagents for this step were provided in the Illumina TruSeq ChIP Library Preparation Kit (Illumina; Cat. No. IP-202-1012) in the present study as all sequencing was undertaken using Illumina platforms. If a non-ChIP Illumina DNA preparation kit was used instead of a ChIP Library Preparation kit (e.g. a standard

NextSeq 500/550 High Output v2 kit), the adapter solution was diluted 1:10 to account for the lower DNA amount present in ChIP samples. Once an adapter set/kit is chosen, following reaction was prepared:

Reagent	Volume (for 30 µl reaction)
ChIP DNA sample	10 µl
DNA ligase buffer	15 µl
Adapter oligonucleotide mix (possibly 1:10)	1 µl
DNA ligase	4 µl

Table 2.9: Reagents required for ligation of adapters to dA-tailed ChIP DNA.

These samples were then incubated for 15 min at room temperature. Purification of DNA following adapter ligation was carried out using a MinElute® PCR Purification Kit (QIAGEN; Cat. No. 28004). To begin, 150 µl of Buffer PB was added to 1 volume of the ligation sample and mixed well. A MinElute® spin column was then placed in a provided 2 ml collection tube and the adapter-ligated DNA added to the column. To bind DNA, the MinElute® spin column was centrifuged at 13,000 rpm for 60 s. The flow-through was then discarded and 750 µl of Buffer PE buffer was added to the sample. The sample was spun at 13,000 rpm for 60 s and the resulting flow-through discarded. The MinElute® spin column was then placed back into the empty collection tube and centrifuged at 13,000 rpm for 60 s to remove residual buffer and ethanol. The MinElute® spin column was then placed in a clean 1.5 ml microtube and the adapter-ligated ChIP DNA eluted by adding 10 µl Buffer EB to the column and the column spun at 13,000 rpm for 60 s. The column was then discarded.

2.2.11.10 Agarose gel size selection of ChIP library

To begin, a 50 ml 2% agarose gel was prepared using 1× TAE buffer and ~5 µl of Sybr Safe. Loading dye was then added to the 10 µl of adapter-ligated ChIP DNA. Approximately 500 ng of DNA ladder (1 kb Plus DNA Ladder) was added in lanes at either side of the wells to contain DNA samples and samples the loaded to the 2% gel. Gaps of 1-2 wells were left between each sample and between each sample

and ladder well. The gel was then placed in an electrophoresis tank and run for ~120 min (or more if necessary) at 60-70 V. The gel was then viewed on a UV transilluminator and bands of ~200 bp in size excised from the gel using GeneCatcher disposable gel excision tips (6.5 mm × 1.0 mm; Gel Company; Cat. No. PKB6.5) attached to a P1000 pipette. New disposable tips were used for each sample to avoid contamination. Each excised piece of gel was then placed in a clean 1.5 ml microfuge tube. The gel was imaged before being discarded (after excision of all gel slices).

Purification of DNA from gel slices was then undertaken using a QIAquick® Gel Extraction Kit (QIAGEN; Cat. No. 28706). The gel slice was weighed in a microfuge tube prior to beginning purification and then 3 volumes of Buffer QG was added to 1 volume of gel (assume 100 mg = 100 µl). The samples were then incubated at 50ºC for 10 min with frequent vortexing until the gel slice had completely dissolved. If the sample was not yellow in colour, 10 µl of 3 M sodium acetate was added to adjust the pH. 1 gel volume of isopropanol (assume 100 mg = 100 µl) was then added to each sample and the sample then placed into a QIAquick® spin column that had been placed inside a 2 ml collection tube. The samples were then centrifuged at 13 000 rpm for 60 s and the flow-through discarded. The column was then placed back into the tube and 0.5 ml of Buffer QG added to the column. Samples were then centrifuged at 13 000 rpm for 60 s and the flow-through discarded. To wash samples, 750 µl of Buffer PE was added and then the sample spun at 13,000 rpm for 60 s once more. The flow-through was discarded and the QIAquick® spin column placed back into the empty collection tube and centrifuged at 13,000 rpm for 60 s to remove residual buffer and ethanol. The QIAquick® spin column was then placed in a clean 1.5 ml microtube and 36 µl of Buffer EB added to the column to elute DNA. The sample was eluted by centrifugation at 13,000 rpm for 60 s.

2.2.11.11 PCR amplification of the ChIP library

For ChIP library amplification, Phusion® High-Fidelity DNA Polymerase (New England BioLabs, Inc.; Cat. No. M0530S) was used. The reaction was set up as follows:

Reagent	Volume (for 50 µl total reaction)
ChIP DNA	36 µl
5× Phusion buffer	10 µl
dNTP mix	1.5 µl
PCR primer 1.1 (index primer of choice from Illumina kit)	1 µl
PCR primer 2.1 (index primer of choice from Illumina kit)	1 µl
Phusion polymerase	0.5 µl

Table 2.10: Reagents required for PCR amplification of ChIP DNA library following adapter ligation.

The sample reactions were amplified according to the following PCR protocol:

30 s at 98ºC
18 cycles of:
10 s at 98ºC
30 s at 65ºC
30 s at 72ºC
5 min at 72ºC
Hold at 4ºC

2.2.11.12 Agarose gel size-selection of amplified ChIP library

A 50 ml 2% agarose gel was prepared with 1× TAE buffer and ~5 µl of Sybr Safe. Loading dye was then added to the 50 µl samples of amplified ChIP DNA. Approximately 500 ng of DNA ladder (1 kb Plus DNA Ladder) was added in lanes at either side of the wells to contain DNA samples and samples the loaded to the 2% gel. Gaps of 1-2 wells were left between each sample and between each sample and ladder well. The gel was then placed in an electrophoresis tank and run for a minimum of 120 min at 60-70 V. The gel was then viewed on a UV transilluminator and sample lanes corresponding to DNA of ~175 +/- 25 bp in size excised from the gel. New disposable tips/razors were used for each sample excision to avoid

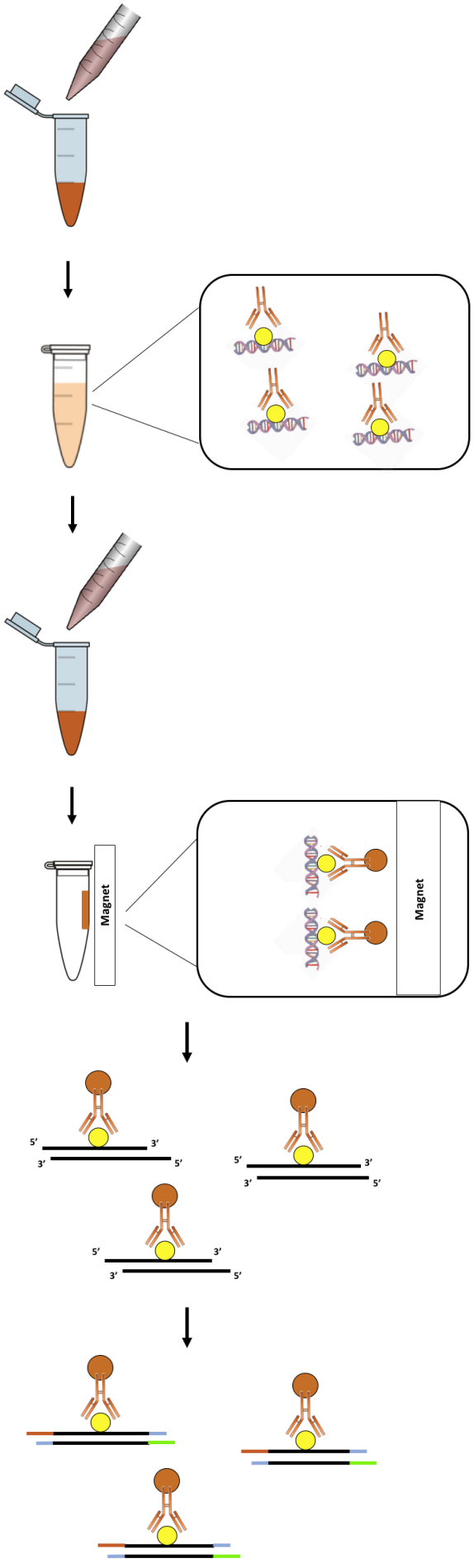
contamination. Each excised piece of gel was then placed in a clean 1.5 ml microfuge tube. The gel was imaged before being discarded.

Amplified ChIP DNA in excised gel slices was then extracted and purified using a QIAquick® Gel Extraction Kit. After cleaning, the amplified ChIP DNA library was eluted in 30 µl Buffer EB. The 30 µl Buffer EB was added to the QIAquick® spin column and let to soak into the membrane for ~60 s. The samples were then centrifuged for 60 s at 13 000 rpm. The eluted DNA was then quantified and assessed using a Qubit 3.0 fluorimeter (Thermo Scientific™) and 2100 Bioanalyzer Instrument (Agilent Technologies) before being sequenced on an Illumina platform.

2.2.12 Chromatin immunoprecipitation coupled with high-throughput DNA sequencing and library preparation by Tn5 transposase 'tagmentation' ('ChIPmentation') protocol for *P. berghei* parasites

As shown in **Figure 2.7,** the process of creating ChIP DNA libraries from *P. berghei* samples using a standard ChIP-seq protocol required a considerable number of clean-up steps and size selections, all with fewer cells than a conventional ChIP-seq protocol (Furey, 2012). In an effort to reduce the loss of genetic material through the multiple DNA purification steps required for standard ChIP-seq library preparation, a novel ChIP-seq method was employed in which immunoprecipitation of chromatin-DNA fragments of interest was immediately followed by the addition of a hyperactive Tn5 transposase that allowed simultaneous *in vitro* DNA fragmentation and adapter tagging ('tagmentation'): a process termed 'ChIPmentation' (Adey *et al.*, 2010; Schmidl *et al.*, 2015).

Unlike the standard ChIP-seq protocol shown in **Figure 2.7**, ChIPmentation of *P. berghei* ANKA parasites (in this study, mature asexual schizonts), is a shorter procedure, with the end-repair to adapter ligation stages of the library preparation protocol substituted with a single Tn5 transposase-mediated reaction (Schmidl *et al.*, 2015) (**Figure 2.8**).



+ Pre-clear samples for IP with pre-blocked Protein G Dynabeads™ at 4°C for 1 h while rotating
+ Separate out Protein G Dynabeads™ on magnetic rack and transfer supernatant to a clean microtube

+ Antibody of interest added to pre-cleared chromatin-DNA (1 antibody per sample)
+ Overnight incubation at 4°C rotating

+ Pre-blocked Protein G Dynabeads™ added to IP sample
+ Incubation at 4°C for 2 h, rotating

+ Magnetic separation of Protein G-bound/antibody bound chromatin-DNA fragments of interest
+ Multiple washing steps of bead- bound fragments
+ Elution in 10 mM Tris-HCl, pH 8.0

+ Illumina adapter-loaded Tn5 transposase addition ('tagmentation' step)

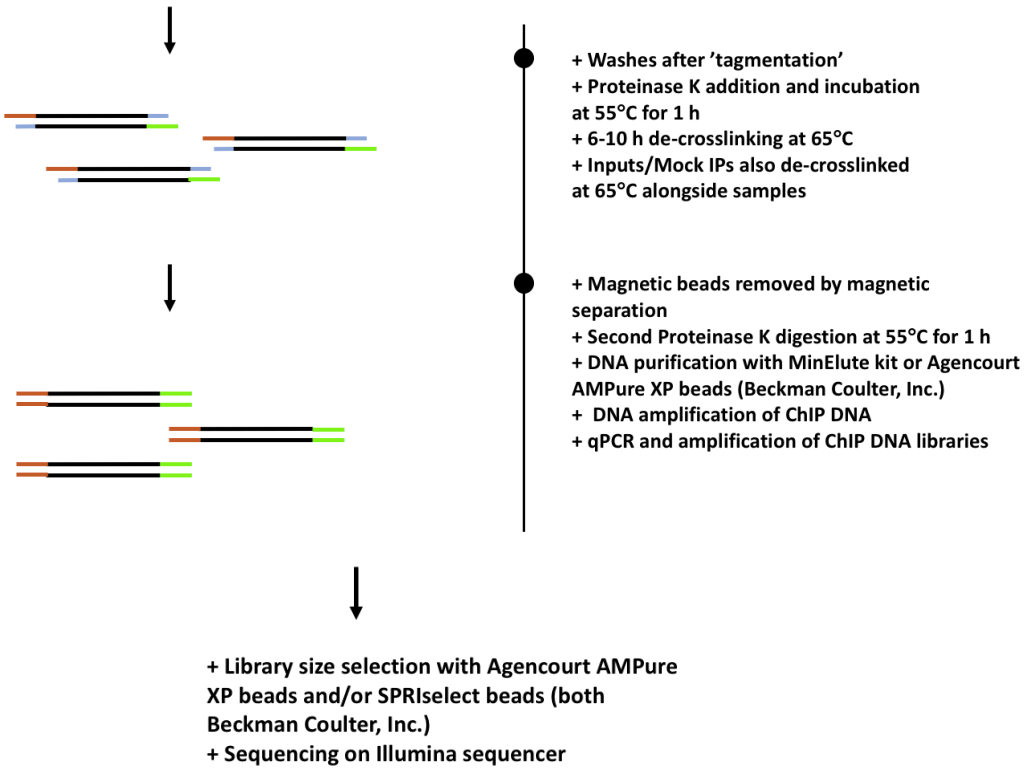


Figure 2.8: Depiction of a 'ChIPmentation' protocol in *P. berghei* ANKA parasites from immunoprecipitation to DNA sequencing on an Illumina platform. The sequence of images above show the ChIPmentation protocol that follows on from the ChIP preparation depicted in **Figure 2.6.** Though pre-clearing with Protein G Dynabeads™ and immunoprecipitation are identical to that of a standard ChIP-seq procedure (**Figure 2.6),** Protein G-bound chromatin-DNA fragments are not de-crosslinked before addition of adapter-loaded Tn5 transposase. DNA fragments attached to chromatin are both fragmented, and then ligated at 5' ends with Illumina adapters in a single transposase-mediated *in vitro* reaction (Adey *et al.*, 2010). Only following this 'tagmentation' procedure, are all samples digested with Proteinase K and de-crosslinked for 6-10 h (or overnight) at 65°C alongside input samples or mock IP controls. A complete list of steps for a ChIPmentation procedure are described from section **2.2.12.1** to **2.2.12.9.**

Buffer name	**Buffer components**
Lysis Buffer (non-SDS) (store at 4°C)	10 mM HEPES (pH 7.9), 10 mM KCl, 0.1 mM EDTA (pH 8.0), 0.1 mM EGTA (pH 8.0), 1 mM DTT (just before use), 1× protease inhibitor cocktail (from 25× stock- added just before use), 0.25% Tergitol® Type NP-40 (added just before use directly into sample)

1% SDS Lysis Buffer	1% SDS, 10 mM EDTA (pH 8.0), 50 mM Tris-HCl (pH 8.1)
5% (w/v) sodium deoxycholate stock (store at room temperature)	5% (w/v) sodium deoxycholate, powder
1× Tris-EDTA (TE) buffer (store at 4ºC)	10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0)
RIPA-LS buffer (store at 4ºC)	10 mM Tris-HCl (pH 8.0), 140 mM NaCl, 1 mM EDTA (pH 8.0), 0.1% SDS, 0.1% sodium deoxycholate, 1% Triton X-100
RIPA-HS buffer (store at 4ºC)	10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), 500 mM NaCl, 1% Triton X-100, 0.1% SDS, 0.1% sodium deoxycholate
RIPA-LiCl buffer (store at 4ºC)	10 mM Tris-HCl (pH 8.0), 1 mM EDTA (pH 8.0), 250 mM LiCl, 0.5% Tergitol® Type NP-40, 0.5% sodium deoxycholate
ChIP elution buffer (store at room temperature)	10 mM Tris-HCl (pH 8.0), 5 mM EDTA (pH 8.0), 300 mM NaCl, 0.4% SDS

Table 2.11: Composition of buffers for a 'ChIPmentation' procedure in *P. berghei*.

2.2.12.1 Crosslinking of chromatin to DNA

The procedure began by chilling glass dounce tissue grinder sets (Sigma-Aldrich; Cat. No. D8938) on ice and pre-cooling a centrifuge to 4°C. All stages of this protocol were carried out on ice unless otherwise stated. Parasites of a synchronised asexual or sexual stage population (Plasmodipur filtered and Nycodenz/magnetically separated) were washed twice in 1 ml of ice-cold 1× PBS and then re-suspended in 150 µl to 500 µl of ice-cold 1× PBS, depending upon parasite numbers, with small cell numbers being up to 5×10^5 cells and large cell numbers from 5×10^5 cells up to 10×10^6 cells. If parasite numbers were larger, samples were split into smaller cell numbers and the ChIP DNA pooled after the

procedure. Re-suspended parasites were crosslinked with 1% formaldehyde from a 16% stock of methanol-free formaldehyde. Samples were mixed by inversion of the microtube and incubated at 37°C for 5 min, inverting occasionally. After incubation, the samples were briefly spun in a microfuge to pool any evaporated material to the base of the microtube. The crosslinking reaction was then quenched with 125 mM glycine from a 2 M glycine stock. The samples were mixed well by gentle inversion of the tubes and then incubated on ice for 5 min. After this incubation period, samples were then centrifuged at 2500 ×g for 5 min at 4°C and the supernatant discarded.

Pellets were then washed 3 times with 1 ml 1× PBS (each time re-pelleting the sample at 2500 ×g for 5 min at 4°C). The pellet was then re-suspended in **non-SDS** cold lysis buffer (without Tergitol® Type NP-40) to a concentration of ~3.33×10^5 parasites/µl, or fewer in the case of small cell numbers. Samples were incubated on ice for 30 min before Tergitol® Type NP-40 was added to a final concentration of 0.25%. One sample at a time was then transferred to a dounce (up to 2 ml) on ice and the sample homogenised with 200 strokes of the dounce (repeated with more of the sample if the sample was greater than 2 ml). For very small cell numbers, samples were retained within the microtubes on ice and homogenised using single-use disposable pestles (Fisherbrand™ RNase-Free Disposable Pellet Pestles; Thermo Scientific™; Cat. No. 12-141-364). The samples were pipetted from the dounce (if glass dounce was used) and placed into clean microtubes. Samples greater than 2 ml were pooled together once more if separated for douncing. Tubes were then centrifuged at maximum speed for 10 min at 4°C to pellet insoluble cell debris. The supernatant was then transferred to a clean microtube.

Note: In the case of cell types that were difficult to lyse, or if the 'insoluble' pellet appears quite large, the process of douncing in non-SDS, cold lysis buffer was repeated to obtain more sample. At this point, SDS, or any other denaturing detergent was not used, as it may have disrupted protein-protein or protein-DNA interactions. Additional supernatant could then be added to the previous eluate.

Samples were topped up to a total of 300 µl per microtube in cold **non-SDS** lysis buffer prior to sonication.

2.2.12.2 Sonication of crosslinked chromatin/DNA

Aliquots of 300 µl per microtube of cross-linked chromatin/DNA were sonicated using a Bioruptor® Plus sonication device (Diagenode; Cat. No. B01020001) at the 'high power' setting in ice water. The following sonication sequence was used each time: 20 sec ON, 2 min OFF, 20 sec ON, 3 min OFF, 30 sec ON, 15 sec OFF, 15 sec ON, 2 min OFF, 30 sec ON.

Following sonication, inputs and samples were chosen and the chromatin/DNA topped up with **non-SDS** cold lysis buffer to volumes of 500-1000 µl for large cell numbers and 300-500 µl for small cell numbers per IP. To each sample, the antibody of interest was added to the desired dilution (e.g. a 1:100 or 1:500 antibody dilution). No antibody was added to inputs. All samples were incubated overnight at 4°C on a rotator.

2.2.12.3 Preparation of magnetic beads for ChIP

For each IP, 10 µl of Magnetic Protein G Dynabeads™ (Thermo Scientific™; Cat. No. 10003D) and 25 µl of Protein G Dynabeads™ (for small or large cell numbers respectively) were added to 175 µl of 0.1% BSA/RIPA buffer solution and incubated at 4ºC rotating overnight.

2.2.12.4 Immunoprecipitation

After an overnight incubation with a blocking buffer, Protein G beads were separated from blocking solution by placing in a magnetic microtube rack. Using 10 µl of beads for small cell numbers and 25 µl of beads for large cell numbers; antibody-bound chromatin/DNA was added to the beads and the tubes inverted to mix samples well. These samples were then incubated at 4ºC for 2 h on a rotator. During the incubation of antibody-bound chromatin/DNA and beads, a magnetic rack was cooled on ice and protease inhibitors were added to fresh buffers (also kept on ice). After the 2 h incubation, beads were washed (150 µl per small cell

number and 1000 µl per large cell number) with wash buffers in the following sequence:

1) Twice in RIPA-LS
2) Twice in RIPA-HS
3) Twice in RIPA-LiCl
4) Once in 10 mM Tris-HCl, pH 8.0

After these washing steps, the beads were gently re-suspended in 150 µl 10 mM Tris-HCl, pH 8.0 and transferred to a clean tube. These tubes were then placed in a magnetic microtube rack and, after separation of beads from solution, the supernatant was discarded.

2.2.12.5 'Tagmentation' reaction with chromatin/DNA ('ChIPmentation')

For each ChIP sample, the following reaction was prepared using components from a Nextera® DNA Library Preparation Kit (Illumina; Cat. No. FC-121-1030):

Reagent	Volume
5× 'Tagmentation' buffer	5 µl
Nuclease-free water	19 µl
'Tagment' DNA Enzyme	1 µl

Table 2.12: Reagents required for a single 'tagmentation' reaction.

Each sample (bound to magnetic beads) was re-suspended in the 25 µl reaction mixture and incubated for 10 min at 37ºC (with a re-suspension by gentle pipetting at 5 min). Once incubation had been completed, the reaction tubes were placed on ice to cool. After this, 150 µl of RIPA-LS (for small cell numbers) or 1000 µl of RIPA-LS (for large cell numbers) was added to samples (ice-cold) and the samples washed in the following sequence:

1) Twice in RIPA-LS
2) Twice in 1× TE buffer

2.2.12.6 De-crosslinking and DNA purification

After the final 1× TE buffer wash, beads were then re-suspended in 48 µl ChIP elution buffer and 2 µl of Proteinase K (10 mg/ml) added to each sample. These samples were then incubated at 65ºC for 6-10 h to de-crosslink DNA from chromatin. Input samples were also brought to 48 µl with 0.4% SDS, 300 mM NaCl solution (final concentrations) and then 2 µl Proteinase K (10 mg/ml) added. The inputs were then incubated at 65ºC for 6-10 h. After de-crosslinking at 65ºC, the beads were separated from solution using a magnetic microtube rack and the supernatant transferred to a clean 1.5 ml DNA Lo-Bind tube. A further 19 µl of ChIP elution buffer and 1 µl of Proteinase K were added to each sample of beads and these beads were once more incubated at 55ºC for 1 h. After this incubation, the beads were separated on the magnetic microtube rack and the supernatant removed and pooled with the previous eluate. All samples were then cleaned using a MinElute® PCR Purification Kit (QIAGEN; Cat. No. 28004) with elution of DNA in 22 µl of EB buffer or 10 mM Tris-HCl, pH 8.0. An additional clean-up with Agencourt AMPure XP Beads (Beckman Coulter, Inc.; Cat. No. A36880) at 1.8× DNA volume was then carried out.

For the Ampure XP bead clean-up, 1.8× DNA volume of beads was added to each reaction and gently mixed by pipetting. Samples were then incubated at room temperature for 10 min. The beads were then separated from solution using a magnetic microtube rack (leaving 10 min to allow all beads to separate) and the supernatant discarded. Beads were then washed twice with 100 µl of 80% ethanol (5 s washes) and then the beads air-dried at room temperature on the magnet. Sample tubes were then removed from the magnet and 22 µl of distilled water added to the beads. Beads and water were mixed gently by pipetting before the samples were left to incubate at room temperature for 10 min. The samples were then added to the magnetic rack, and the beads separated out for 5 min. The supernatant (containing clean ChIP DNA) was then transferred to a new clean microtube.

2.2.12.7 ChIP DNA library amplification

To begin, a PCR master mix (n+1 reactions) was made up using NEBNext® High-Fidelity 2X PCR Master Mix (New England BioLabs, Inc.; Cat. No. M0541S) as follows:

Component	Volume for 35 µl master mix
NEBNext® High-Fidelity 2X PCR Master Mix	25 µl
Distilled water	10 µl

Table 2.13: PCR master mix composition for ChIP DNA library amplification.

For each sample, the following was combined in one PCR tube:

Component	Volume for 50 µl reaction
NEBNext® High-Fidelity 2X PCR Master Mix diluted in water	35 µl
PCR primer 1.1 (index primer of choice from Illumina Nextera kit) (25 µM)	2.5 µl
PCR primer 2.1 (index primer of choice from Illumina Nextera kit) (25 µM)	2.5 µl
Purified ChIP DNA	10 µl

Table 2.14: Complete reagent list for PCR ChIP DNA library amplification after 'tagmentation'.

These samples were then incubated in a thermal cycler according to the following sequence:

68ºC for 5 min
98ºC for 30 s
5 cycles of:
98ºC for 10 s
63ºC for 30 s
68ºC for 1 min

Hold at 4ºC

Samples were then removed from the cycler and either stored at 4ºC or immediately used for qPCR. A qPCR reaction master mix (for n+1 reactions) was set up for each sample as follows:

Component	**Volume for 9.5 µl master mix**
NEBNext® High-Fidelity 2X PCR Master Mix	5 µl
Distilled water	4.32 µl
50× Sybr Green	0.18 µl

Table 2.15: qPCR master mix composition for ChIP DNA library amplification.

Each sample was then combined with qPCR master mix in a 48 well plate (with a H_2O control) with the final 15 µl reactions made up as shown below:

Component	**Volume for 15 µl reaction**
qPCR master mix	9.5 µl
PCR primer 1.1 (index primer of choice from Illumina Nextera kit) (25 µM)	0.25 µl
PCR primer 2.1 (index primer of choice from Illumina Nextera kit) (25 µM)	0.25 µl
Amplified ChIP DNA (from previous PCR reaction)	5 µl

Table 2.16: Complete qPCR reaction mixture for determining optimal amplification of individual ChIP DNA libraries.

The qPCR reaction was then carried out in the following sequence:

98ºC for 30 s

40 cycles of:

98ºC for 10 s

63ºC for 30 s

68ºC for 1 min
Hold at 4ºC

For ChIP DNA, 40 cycles were always necessary to reach a plateau in the qPCR amplification curve. For an Assay for Transposase-Accessible Chromatin with high throughput sequencing (ATAC-seq) protocol, 20 cycles were sufficient (from collaborative work with Dr Sebastian Kirchner). qPCR data was then used to inform the number of cycles to be used for ChIP DNA library amplification (using remaining pre-amplified DNA). The minimal and maximal intensity for each sample amplification curve was noted and the signal corresponding to one third of maximal fluorescence was used to infer the number of cycles with which to amplify the remaining pre-amplified ChIP DNA samples.

The remaining pre-amplified ChIP DNA (that had been amplified for only 5 cycles previously) was then placed back in the thermal cycler and amplified for the specific number of cycles obtained for each sample by qPCR:

98ºC for 30 s
'X' number of cycles of:
98ºC for 10 s
63ºC for 30 s
68ºC for 1 min
Hold at 4ºC

2.2.12.8 Final DNA clean-up before high throughput sequencing

The amplified ChIP DNA library obtained following the procedures carried out in section **2.2.12.7** was immediately cleaned once more with Agencourt AMPure XP Beads (Beckman Coulter, Inc.; Cat. No. A36880) as detailed in section **2.2.12.6** with DNA libraries eluted in 35 µl of 10 mM Tris-HCl, pH 8.0. These samples were then stored at -20ºC for long-term storage, or immediately examined using a Qubit™ 3.0 fluorimeter (with Qubit™ dsDNA High Sensitivity (HS) Assay Kit; Thermo Scientific™; Cat. No. Q32851) and 2100 Bioanalyzer (Agilent Technologies). 1 µl of each sample was examined using each machine and the sample then sent for sequencing on an Illumina platform, or cleaned once more if necessary. In the

case of adapter dimer peaks from the Bioanlyzer analysis, a further DNA clean-up with SPRISelect beads (Beckman Coulter, Inc.; Cat. No. B23317) was carried out and then the DNA re-examined by Qubit and bioanalyzer to assess ChIP DNA library quantity and quality. Samples to be sent for sequencing had at least 10 ng of DNA suspended in 10-30 µl 10 mM Tris-HCl, pH 8.0.

2.2.12.9 Illumina index adapter sequences and primers

The following primers (de-salted) with Illumina adapter sequences were used in the present study. Primer sequence construction was based on Illumina guidelines and methods detailed in (Schmidl *et al.*, 2015) and (Buenrostro *et al.*, 2013).

Primer Type	Index name	Index sequence	Waters Lab No.	Primer sequence
Index 1 (i7)	N701	TCGCCTTA	GU4107	CAAGCAGAAGACGGCATACGAGATTCGC CTTAGTCTCGTGGGCTCGGAGATGT
Index 1 (i7)	N702	CTAGTACG	GU4338	CAAGCAGAAGACGGCATACGAGATC TAGTACGGTCTCGTGGGCTCGGAGATGT
Index 1 (i7)	N703	TTCTGCCT	GU4403	CAAGCAGAAGACGGCATACGAGATTT CTGCCTGTCTCGTGGGCTCGGAGATGT
Index 1 (i7)	N704	GCTCAGGA	GU4339	CAAGCAGAAGACGGCATACGAGATG CTCAGGAGTCTCGTGGGCTCGGAGATGT
Index 1 (i7)	N705	AGGAGTCC	GU4340	CAAGCAGAAGACGGCATACGAGAT AGGAGTCCGTCTCGTGGGCTCGGAGATGT
Index 1 (i7)	N706	CATGCCTA	GU4341	CAAGCAGAAGACGGCATACGAGAT CATGCCTAGTCTCGTGGGCTCGGAGATGT
Index 2 (i5)	N501	TAGATCGC	GU4342	AATGATACGGCGACCACCGAGATCTACAC TAGATCGCTCGTCGGCAGCGTCAGATGTG
Index 2 (i5)	N502	CTCTCTAT	GU4343	AATGATACGGCGACCACCGAGATCTACAC CTCTCTATTCGTCGGCAGCGTCAGATGTG
Index 2 (i5)	N503	TATCCTCT	GU4404	AATGATACGGCGACCACCGAGATCTACACT ATCCTCTTCGTCGGCAGCGTCAGATGTG
Index 2 (i5)	N504	AGAGTAGA	GU4405	AATGATACGGCGACCACCGAGATCTACAC AGAGTAGATCGTCGGCAGCGTCAGATGTG

Index 2 (i5)	No index	--------	GU4106	AATGATACGGCGACCACCGAGATCTACACT CGTCGGCAGCGTCAGATGTG

Table 2.17: List of Illumina index adapter sequences used for *P. berghei* ChIPmentation sample sequencing.

2.2.13 ChIP-seq and ChIPmentation analysis: sequencers, read number, and software

Following library preparation by ChIP-seq (depicted in **Figure 2.7**), or ChIPmentation (depicted in **Figure 2.8**) methods, both DNA library quantity and fragment length were analysed using a Qubit® 3.0 Fluorometer with double-stranded DNA high-sensitivity (HS) reagents (Thermo Scientific™) and an Agilent 2100 Bioanalyzer with Agilent High Sensitivity DNA kit reagents (Agilent Technologies). No samples prepared using a standard ChIP-seq protocol **(2.2.10)** were suitable for high-throughput sequencing. However, successfully amplified ChIPmentation libraries were sent for whole genome paired-end sequencing to the Glasgow Polyomics Facility, specifying 25 M, 75 bp paired-end reads using the NextSeq 500 system (Illumina, Inc.). Output was provided in FASTQ format.

All ChIPmentation analysis was carried out with help from Dr Kathryn Crouch and Sejal Modha of the bioinformatics group within the Wellcome Centre for Molecular Parasitology. Raw data quality (in FASTQ format) for all samples was assessed using FastQC (www.bioinformatics.babraham.ac.uk). Sequencing adapters were removed from sample data (in FASTQ format) using Trim Galore! (www.bioinformatics.babraham.ac.uk) and trimmed reads were aligned to the *P. berghei* ANKA genome (version 34; *Plasmo*DB) using Bowtie2 (Langmead and Salzberg, 2012). Bowtie2-generated BAM (binary sequence alignment/map) files were then normalised to a ChIP input sample using 'bamCompare' (Ramírez *et al.*, 2016). The 'bamCompare' tool generated input-normalised ChIP coverage files in 'bigWig' format which could be visualised using the IGV or Artemis genome browsers (Robinson *et al.*, 2011; Carver *et al.*, 2012).

To find *P. berghei* ANKA genome regions to which paired-end fragments aligned, i.e. genome regions at which a 'peak' of reads was present, specifically-designed

peak-calling software was used. To optimise peak-calling within the AT-rich *P. berghei* ANKA genome, and for the relatively 'broad' peaks generated by ChIP antibodies to histone modifications, comparative analyses of four peak-calling tools was carried out by Sejal Modha at the Wellcome Centre for Molecular Parasitology (Heinig *et al.*, 2015). As peak-calling software required input in '.bed' format, BAM files were converted to BED files using the 'bamtobed' script within SAMtools (Li *et al.*, 2009). Peaks were called using MACS2 ('--broad' parameter selected) (Zhang *et al.*, 2008), HOMER (findPeaks with '--style histone' parameter) (Heinz *et al.*, 2010), SICER (Xu *et al.*, 2014), and Epic (a diffuse domain ChIP-seq caller based on SICER) (Xu *et al.*, 2014). The complete bioinformatic pipeline used for ChIPmentation analyses in this study is depicted in **Figure 2.9**.

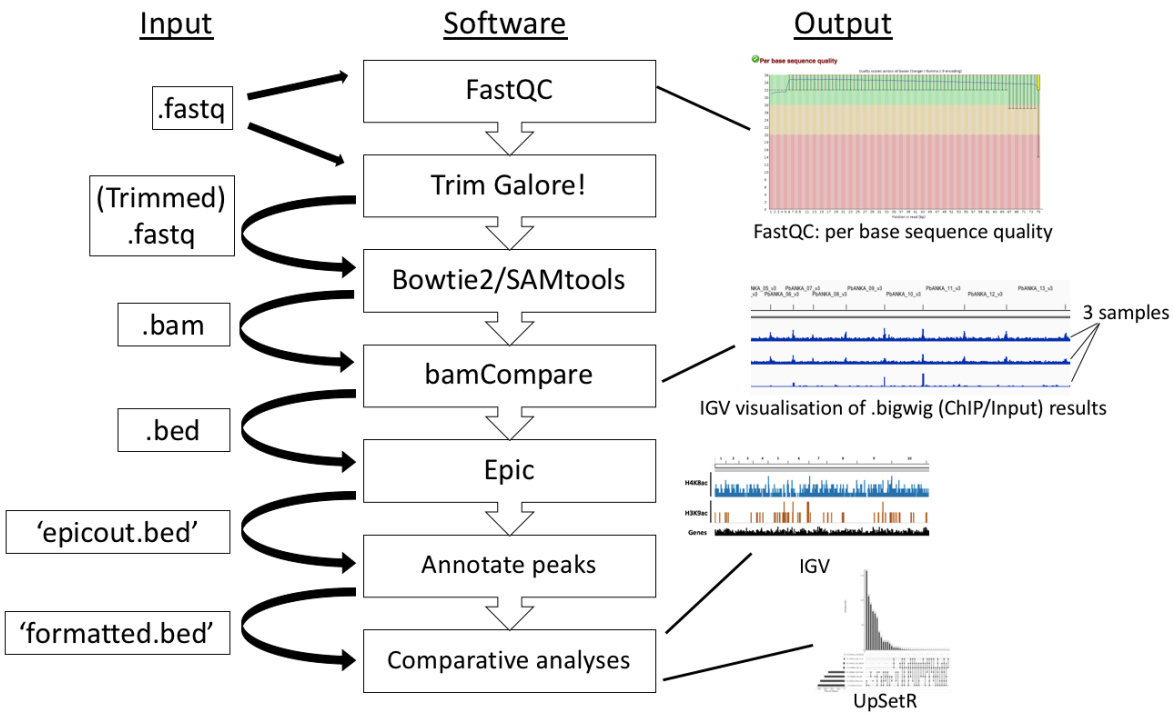


Figure 2.9: Sequencing analysis of ChIPmentation data from an Illumina platform. This image shows the bioinformatics software used to process and then compare ChIPmentation results in the present study. In the centre, a flow chart shows the names of software used and to the left are the names of the file formats that are required as input to generate data using the next program in the sequence. To the right of the image, a number of examples are provided for the type of results generated at a number of stages in the analysis. To the right of the FastQC step is an example of the FastQC (Babraham Bioinformatics) output, in this case, a chart depicting the quality of each base (A,T,G, or C) corresponding to sequence reads in the sample (by Illumina sequencing). In the example shown (one sample from this study), almost all reads have bases with high quality scores (shown in green). The blue line that runs through each sample represents the

mean quality score, with yellow and red areas of the graph indicating reasonable quality or poor quality bases respectively. Below this is a screenshot of the IGV browser showing 3 separate 'bigwig' files (ChIP/input) being compared in individual tracks. Tracks in IGV show BAM or 'bigwig' files aligned to the reference genome. In this image (from the present study), peaks are already visible by IGV visualisation alone. The final two screenshots in this image, stemming from the 'Comparative Analyses' step of the flowchart, show two methods by which ChIPmentation data can be compared: firstly, by visualising the Epic peaks (with or without corresponding BAM or 'bigwig' files) in IGV, and, secondly, by using the 'UpSetR' R-based package which generates a chart comparing the number of peaks called using each number of cells, sequencing conditions, or peak-calling software used.

Peaks called using three of these tools (SICER had been superseded by Epic and so Epic was used for comparative analysis) were visualised in comparison to one another using the 'UpSetR' R-based package (Conway, Lex and Gehlenborg, 2017) based on the 'UpSet' technique showing intersections of multiple sets and their sizes (Lex *et al.*, 2014) **(Figure 2.9) (Chapter 5)**. Following the choice of Epic as primary peak-calling software for this study (as a result of most shared peaks between samples), shared peaks between two or more Epic output files were identified using 'epic-merge'. Epic-derived BED files were used as the input for these comparisons. Annotation of ChIP peaks for both individual and merged files was undertaken using a modified HOMER (Hypergeometric Optimization of Motif Enrichment) script. The original HOMER annotation script was downloaded from http://homer.ucsd.edu/homer/ngs/annotation.html. The original script was modified for use with Epic parameters in Perl with the resulting file labelled 'annotatePeaks.pl'. Epic software provided output files for both individual and merged peak results as BED files which could be visualised using the IGV genome browser. All programming scripts (in 'Bash') used to run bioinformatic analysis of ChIPmentation data in the present study are provided in **Supplemental File 2 (S2).**

2.2.14 RNA-sequencing analysis for comparison with ChIPmentation results

To correlate ChIPmentation data to gene expression, analysis of RNA-sequencing results from a previous study was carried out to compare these with ChIPmentation results from the present study. To facilitate this, duplicate RNA-seq results for both 16 h *P. berghei* ANKA trophozoites and 22 h *P. berghei* ANKA schizonts were downloaded from NCBI

(https://www.ncbi.nlm.nih.gov/sra/SRX325688) (Otto, Böhme, *et al.*, 2014). In each case, sequencing adapters were trimmed and reads that became shorter than 20 bp were discarded. Cleaned and trimmed reads were aligned to the *P. berghei* ANKA reference genome (version 34) using HISAT2 which was downloaded from https://ccb.jhu.edu/software/hisat2/index.shtml (Kim *et al.*, 2016).

Differentially expressed genes between 22 h schizonts and 16 h trophozoites were identified using edgeR (Robinson, Mccarthy and Smyth, 2010). Raw edgeR results are listed in **Electronic Supplemental Material 4 (ES4),** with more detailed differential expression analyses displayed in the **Supplemental Material RNA-seq glimma-plots** folder as interactive graphics (Su *et al.*, 2017).

A multi-dimensional scaling (MDS) analysis was also undertaken to investigate the similarity between both 22 h schizont and 16 h trophozoite RNA-seq samples and their replicates. By default, the top 500 genes in each group are used to generate each representative MDS display point. Distances between two display points reflects the leading log-fold-change (logFC) separating samples (Su *et al.*, 2017). Final differential RNA-seq results identified by paired edgeR analysis are then displayed as interactive MA (log ratio/mean average) and volcano plots for easy visualisation of data **(Supplemental Material RNA-seq glimma-plots).** All RNA-seq analysis was carried out by Sejal Modha at the Wellcome Centre for Molecular Parasitology.

2.2.15 Multi-dimensional scaling (MDS) analysis of ChIPmentation results

Upon completion of all ChIPmentation analysis and peak calling, MDS analysis of the similarity between all samples used in the present study was carried out using edgeR (Robinson, McCarthy and Smyth, 2010). This analysis was carried out to inform as to which ChIPmention datasets were consistent between replicates of 2.5 million *P. berghei* schizonts, and to identify any outlying datasets that might skew findings incorrectly. The 'plotMDS' function of edgeR was applied which produced a two-dimensional scatter plot in which distances between samples represented the leading $\log_2$ fold-change ($\log_2$FC) between samples (for the top 500 genes that distinguish those samples). When generating the MDS plot, sequence reads were counted for each bin in the genome with a genomic bin size

of 1000. In this case, the gene selection method was 'pairwise', meaning that a different set of top genes is selected for each pair of samples (as opposed to a 'common' gene selection approach in which the top genes are always chosen as those with the largest standard deviation between samples). (Ritchie *et al.*, 2015). The resulting MDS plot for all triplicate ChIPmentation samples carried out in the present study (2.5 million *P. berghei* ANKA schizonts in triplicate) is shown in **Chapter 5** alongside comprehensive analysis of ChIPmentation results.

2.2.16 Generation of *P. berghei* constructs for double-crossover or single-crossover homologous recombination

2.2.16.1 Restriction digest of background DNA vector and DNA fragments

Restriction enzymes were used in the present study for the generation of cloning vectors and to then linearise vectors prior to transfection. Restriction enzymes for cloning were carefully chosen using vector maps and genomic DNA maps in the CLC Genomics Workbench, version 7 (QIAGEN). High fidelity restriction enzymes were sourced from New England BioLabs, Inc. (NEB). In most double-crossover (DXO) constructs, two identical restriction sites were placed at either side of an Ampicillin sequence to facilitate fast removal of the bacterial drug selection sequence before plasmid linearisation. If this was not possible, separate restriction sites and enzymes were used. In all reactions, the restriction enzyme volume was <10% of the total reaction volume, with reaction solutions gently mixed by pipetting. Restriction sites were incorporated into primer design to create fragments with compatible ends for ligation into a cut plasmid. Restriction digests were undertaken on both plasmids and PCR fragments to make sequence ends compatible for ligation.

After amplification of a sequence of interest from *P. berghei* genomic DNA, e.g. fragments homologous to sequences at either side of a gene (for DXO recombination), PCR products were cleaned up using a QIAquick® PCR Purification Kit (QIAGEN; Cat. No. 28106) before restriction digest. After PCR fragment purification, the fragment DNA was eluted in 50 µl of distilled water. 5 µl of the reaction was checked on a gel to ensure that the PCR had worked before digesting the rest of the PCR product. 40 µl of this 50 µl volume was used for cloning (corresponding to 2-5 µg of DNA). Volumes were adjusted as necessary. Restriction

enzyme digestion reactions were usually 50 µl in volume, containing 2 µl of each enzyme. Example restriction digest reactions for both plasmids and PCR fragments are shown below.

Reagent	**Volume (for a 50 µl total reaction)**
Distilled water	33 µl
Buffer (10×)	5 µl
Plasmid DNA (~4.5 µg DNA)	10 µl
Restriction enzyme	2 µl

Table 2.18: General reagent mixture for a single plasmid vector restriction digest.

Reagent	**Volume (for a 50 µl total reaction)**
Distilled water	3 µl
Buffer (10×)	5 µl
PCR product DNA (2-5 µg DNA)	40 µl
Restriction enzyme	2 µl

Table 2.19: General reagent mixture for a single PCR product restriction digest.

Samples were digested at the appropriate restriction enzyme temperature for a minimum of 1 h (~1-2 h). Digestion reactions were checked before ligation by running ~5 µl (~350 ng of DNA) from each reaction on a 1% agarose gel alongside a 1 kb Plus DNA ladder. Approximately 1 µl of uncut plasmid was run alongside the digested plasmid to determine if digestion had occurred. After 2 h of incubation, a top-up of restriction enzyme in buffer was added to each 50 µl reaction: a 10 µl top-up of 1 µl buffer, 0.5 µl restriction enzyme, and 8.5 µl distilled water. After the enzyme top-up, the reaction was incubated for a further 30 min. Once the incubation was completed, the enzyme was heat inactivated (if required) and the PCR product or plasmid fragment purified using the QIAquick® PCR Purification Kit or the QIAquick® Gel Extraction Kit (QIAGEN; Cat. No. 28706).

Often, if a single double-cutting enzyme was used, the linearised vector was dephosphorylated with calf intestinal phosphatase (CIP) (New England BioLabs, Inc.; Cat.No. M0290S) to prevent self-ligation. To achieve this, 1 µl CIP was added

to the 50 µl digest reaction (after heat-inactivation of enzymes) and the reaction incubated at the appropriate temperature for an additional 30 min. Before ligation, the dephosphorylated plasmid fragment was purified using a PCR purification kit or gel extraction. At the end of the purification process, digested fragments were re-suspended in 30 µl of sterile, distilled water.

2.2.16.2 Ligation of DNA fragments

Ligation reactions were carried out in this study to form circularised plasmids from a linearised plasmid backbone and a PCR product. Experiments were planned to ensure that compatible sticky ends were present on the insert and vector by using the same restriction enzyme(s). Alternatively, the restriction sites would need to be blunted i.e. using Klenow (DNA polymerase large fragment) to fill in 5' overhangs. 3' overhangs could be digested back with T4 polymerase. However, this latter method was not necessary for plasmid construction in the present study. After, plasmid and PCR fragment clean-up, DNA concentrations were checked using a NanoDrop. For the ligation reaction, vector: insert ratios were calculated using an online ligation calculator (*in silico*; Universität Düsseldorf) or according to the table below.

Insert Size	500bp	1kb	2kb	3kb	4kb	5kb
Vector (ng)	150ng	150ng	150ng	150ng	150ng	150ng
Insert (ng)	22.5ng	45ng	90ng	135ng	180ng	225ng

Table 2.20: Standard grid of examples for 3:1 ratio of PCR product to plasmid for ligation reactions.

After choosing the correct vector:insert ratio, 30 µl ligation reactions were carried out using the Rapid DNA ligation kit (Sigma-Aldrich (formerly Roche); Cat. No. 11635379001), with reactions set up as follows:

Example	**Amount**	**Generic Amounts**
Vector (10kb, 150ng at ~20ng/µl)	8 µl	Vector (150ng)

Insert (500bp, 22.5ng at ~12ng/µl)	2 µl	Insert (calculated for a 3:1 ratio)
Buffer (2×)	12.5 µl	Buffer (2×): 12.5 µl
Ligase	1 µl	DNA ligase: 1 µl
Distilled water	6.5 µl	Distilled water (to make up to 30 µl)

Table 2.21: Components for a standard 30 microlitre DNA ligation reaction.

When prepared, the 30 µl reaction was incubated at room temperature for 10 min to 1 h. This reaction was then transformed into competent *E. coli* cells for propagation of the new plasmid.

2.2.16.3 Transformation of ligated constructs into *E. coli*

In the present study, transformations were carried out using Fusion-Blue Competent *E. coli* (Clontech Laboratories, Inc.). A water bath was warmed to 42ºC and then microtubes containing the ligation reaction and *E. coli* were placed on ice (*E. coli* stocks were stored at -80ºC in aliquots when not in use). 65 µl of *E. coli* was then added to each ligation reaction and the reactions incubated on ice for 10 min. Each tube was then transferred to the 42ºC water bath for 1 min exactly (heat shock). After this, tubes were then incubated on ice for 5 min. 400 µl of sterile LB broth (or Super Optimal Broth with catabolite repression (SOC medium)) was then added to each reaction tube and the samples incubated for 45 min at 37ºC in a shaking incubator. After incubation, samples were then centrifuged at 8500 rpm for 2 min and 350 µl of supernatant discarded. The remaining supernatant was then used to re-suspend the pellet, and this *E. coli* sample spread on an Ampicillin agar plate using sterile techniques. Agar plates were then placed at 37ºC overnight. A control Ampicillin agar plate was also incubated with each batch of transformations.

2.2.16.4 *E. coli* colony PCR

To check for successfully transformed *E. coli* colonies, PCRs were carried out on samples taken from colonies that had grown on agar plates after transformation.

Usually, 8-10 bacterial colonies were picked from an Ampicillin/Kanamycin agar plate using clean P2 pipette tips and these colonies then placed in microtubes containing 25 µl sterile water. The microtube lids were closed and the samples incubated at room temperature for 30 min. Controls were taken from the control agar plate (i.e. a fragment of LB agar) and treated in the same manner as samples. A PCR master mix was then prepared (in the present study, the KAPA HiFi PCR kit was used; KAPA Biosystems; Cat. No. KR0368). For a 200 µl master mix, the following solutions were added:

Reagent	**Volume (for a 200 µl total reaction)**
Distilled water	120 µl
KAPA HiFi Buffer (5×)	40 µl
Forward primer	6 µl
Reverse primer	6 µl
dNTP mix	6 µl
KAPA HiFi DNA polymerase	2 µl

Table 2.22: Standard components for an *E. coli* colony PCR master mix.

This reaction was then split over 10 PCR tubes (18 µl per tube) and 2 µl of each colony/water mixture then added to one tube. The PCR was then run according to the following sequence:

- 95ºC: 4 min
- 95ºC: 30 s
- XºC: 30 s (primer T_m)
- 68ºC: 30 s per kb
- GO TO 2: 29 times
- 68ºC: 4 min
- 10ºC: Hold

PCR products were then run on a 1% agarose gel after loading dye added. Positive *E. coli* colonies were then grown up in larger LB broth cultures and the plasmids purified using a HiSpeed® Plasmid Midi Kit (QIAGEN; Cat. No. 12643) or QIAprep Spin Miniprep Kit (QIAGEN; Cat. No. 27106), depending on the culture volume. To

confirm the correct integration of restriction sites in the new plasmid, a restriction digest was sometimes also performed and fragment size observed on an agarose gel.

2.2.17 Storing transformed *E. coli* as glycerol stocks

To store *E. coli* samples for future propagation of a plasmid of interest, glycerol stocks were made and stored at -80ºC. To make a glycerol stock, 700 µl of bacterial culture was added in a cryotube to 200 µl 100% glycerol. The sample and glycerol were mixed and then the cryotubes stored in a -80ºC freezer for long-term storage.

2.2.18 Concentration of DNA by ethanol precipitation

To concentrate DNA samples, 10 volumes of 100% ethanol was added to a DNA sample along with 0.1 volume of 3 M NaAc, pH 5.2. The sample was then left to precipitate overnight at -20ºC. After precipitation overnight, a centrifuge was then pre-cooled to 4ºC and the samples centrifuged at this temperature at maximum speed for 20-30 min. Samples were then taken from the centrifuge and the supernatant discarded, being careful to avoid dislodging the DNA pellet at the base of the tube. The pellet was then washed once with 150 µl ice-cold 70% ethanol and allowed to air-dry for ~10-15 min. The DNA pellet was then re-suspended in distilled water.

2.2.19 Preparation of Ampicillin for selection of resistant *E. coli*

To make up an Ampicillin stock (100 mg/ml), 1 g of Ampicillin sodium salt (Sigma-Aldrich; Cat. No. A0166) was dissolved in 10 ml of sterile, distilled water (within a fume cupboard to maintain sterility). The dissolved Ampicillin solution was then passed through a single-use filter (Minisart filters, 0.2 µm pore size; Sigma-Aldrich; Cat. No. 16534K) using a 10 ml syringe. Aliquots of 1 ml 100 mg/ml Ampicillin were then made up and stored at -20ºC for up to one year.

3. Determining the role of putative epigenetic regulators throughout the *P. berghei* life cycle

3.1 Chapter Aim

The aim of this chapter was to knockout or conditionally knockdown a number of known and putative epigenetic regulatory proteins in *P. berghei* to examine their potential role in gametocytogenesis.

3.2 Introduction

Plasmodium epigenetic research has focussed largely on enzymes involved in the post-translational modification of histones, such as histone lysine methyltransferases (HKMTs), histone demethylases (HDMs), histone acetyltransferases (HATs), and histone deacetylases (HDACs) (Duffy *et al.*, 2013). The full complement of epigenetic regulators so far identified in *Plasmodium* spp. is listed in **Supplemental Table S1.**

In the present study, *P. berghei* ANKA histone deacetylases (HDACs) and histone acetyltransferases (HATs) were targeted based upon knowledge at the outset of the study; such as the fact that HDAC2/IPK1 conditional degradation, and that of HP1, resulted on an increase in gametocyte conversion relative to background *P. falciparum* lines (Brancucci *et al.*, 2014; Coleman *et al.*, 2014). As such, it was postulated that histone acetylation and deacetylation may play a role in commitment to gametocytogenesis in *P. berghei*.

Putative epigenetic regulators to be subject to complete gene knockout or conditional protein degradation via the auxin-inducible degron (AID) system are listed in **Table 3.1.** In addition to histone acetyltransferases (HATs) and histone deacetylases (HDACs) described in **Supplemental Table S1**, an examination of genes encoding for *P. berghei* ANKA heterochromatin protein 1 (HP1), chromatin-associated exoribonuclease, RNAseII (Zhang *et al.*, 2014), and inositol polyphosphate kinase 2 (IPK2), were also carried out. A study of IPK2 was chosen owing to this protein's similarity to the IPK1 region of the HDAC2/IPK1 combined gene (Coleman *et al.*, 2014) (**Table 3.1**).

Present study			Studied in tandem (Drs Kirchner and Vaikkinen)		
P. berghei Gene ID	**Protein name**	**KO or KD vectors**	**P. berghei Gene ID**	**Protein name**	**KO or KD vectors**
PBANKA_1436100	Heterochromatin protein 1 (HP1), putative	KO + KD	PBANKA_1128500	SET domain protein (SET1), putative	KO
PBANKA_0415300	Exoribonuclease II (RNAseII), putative	KO +KD	PBANKA_0702900	SET domain protein (SET3), putative	KO + KD
PBANKA_1414600	Inositol phosphate kinase 2 (IPK2), putative	KO only	PBANKA_0811200	SET domain protein (SET4), putative	KO + KD
PBANKA_0826500	Histone deacetylase 1 (HDAC1), putative	KO + KD	PBANKA_1430000	Histone-lysine N-methyltransferase (SET5), putative	No vectors available
PBANKA_1206200	Histone deacetylase 2/Inositol phosphate kinase 1 (HDAC2/IPK1), putative	KO + KD	PBANKA_1131800	Histone-lysine N-methyltransferase (SET6), putative	KO + KD
PBANKA_1335400	Histone deacetylase 1 (HDA1), putative	KO + KD	PBANKA_0932500	Histone-lysine N-methyltransferase SET7, putative	KO + KD
PBANKA_1106200	Histone deacetylase, putative (HDAP)	KD only	PBANKA_1001600	SET domain protein (SET8), putative	KO + KD
PBANKA_0718400	Histone acetyltransferase 1 (HAT1), putative	KO + KD	PBANKA_1107700	SET domain protein (SET9), putative	KO + KD
PBANKA_1343800	Transcriptional regulatory protein SIR2A	KD only	PBANKA_1436200	Histone-lysine N-methyltransferase, H3 lysine-4 specific (SET10), putative	KO + KD
PBANKA_1315100	Transcriptional regulatory protein SIR2B	KD only	PBANKA_0610100	Lysine-specific histone demethylase 1 (LSD1), putative	KO + KD
			PBANKA_0101600	JmjC domain-containing protein (JmjC2), putative	KO + KD

Table 3.1: Complete list of putative epigenetic modifiers in *P. berghei* ANKA that were subject to investigation in this study. This table lists the Gene identifier (ID) for each putative *P. berghei* epigenetic modifier examined in the present study, alongside those investigated in tandem by Drs Kirchner and Vaikkinen. The plasmid availability for each protein chosen, either as a knockout (KO) or auxin-inducible degron (AID) tagging/knockdown (KD) DNA construct (or both), is indicated next to the protein name in columns 3 and 6. 'No vector available', as is shown for protein SET5, indicates that no KO vector or tagging/KD vector was available for this protein. SET: Su(var)3-9, Enhancer of zeste, and Trithorax; JmjC: Jumonji C-terminal.

To completely knock out the gene of a putative epigenetic modifier from *P. berghei*, and to assess the effect of this gene knockout on both asexual growth and gametocytogenesis, KO vectors were transfected into parasites of two background lines: the '820' line (previously discussed in section **1.7.2.2**), and a novel 'early gametocyte marker' line, termed 'PbEGAM', in which green fluorescent protein (GFP) was constitutively expressed, and a red fluorescent protein (RFP) was expressed under control of a previously unexploited early gametocyte gene promoter (promoter of gene PBANKA_1018700) (Lee, Waters and Brewer, 2018).

For conditional degradation of an epigenetic modifier of interest, two *P. berghei* ANKA background lines were generated for integration of targeted auxin-inducible degron (AID)-tagged proteins. Conditional degradation of a protein of interest by endogenous or synthetic auxins requires the constitutive expression of an 'F-box' transport inhibitor response protein (TIR1) within the host cell. Therefore, to facilitate auxin-induced proteasomal degradation in the *P. berghei* ANKA parasite, which does not express a TIR1 protein endogenously, a TIR1-encoding sequence from the rice plant, *Oryza sativa* (*ostir1*) was inserted into the *P. berghei* ANKA high-gametocyte-producer (HP) genome at the *p230pI* locus (Nishimura and Kanemaki, 2014; Philip and Waters, 2015). Using conventional positive and negative drug selection, and cloning of the *Os*TIR1-expressing *P. berghei* ANKA line by limiting dilution, the 'G615cl1m1cl1' *P. berghei* background line was produced in which an *ostir1* sequence (tagged by a *9xmyc* sequence) was constitutively expressed under control of the 5' promoter of the HSP70 gene (PBANKA_0711900) (Manzoni *et al.*, 2015).

A second *Os*TIR1-expressing *P. berghei* background line was also generated (the G1448cl1m0cl1, or 'AID 820' line) in which the male- and female- specific fluorescent reporter sequences of the 820 line were expressed upstream of the *ostir1::9xmyc* sequence. With the creation of this background line, the effects of conditional degradation of a protein of interest could be monitored in both male and female gametocytes, as well as asexual growth stages.

All *P. berghei* background lines used or generated for complete gene knockout or conditional degradation via the AID system are listed in **Table 3.2.** Detailed linear plasmid maps are provided in **Supplemental Images S1A to S1O.**

***P. berghei* ANKA line name**	**Gene ID (if applicable)**	**Integration shown (Y/N)**	**Isogenic (FACS sorted) and/or clonal line (Y/N)**	**Linear plasmid map location**
P. berghei ANKA high-producer **(HP)**	N/A	N/A	Y	N/A
820 (820cl1m1cl1)	N/A	Y	Y	Suppl. 1A
PbEGAM (G1137cl2m0cl4)	N/A	Y	Y	Suppl. 1B
AID background (G615cl1m1cl1)	N/A	Y	Y	Suppl. 1C
AID 820 (G1448cl1m0cl1)	N/A	Y	Y	Suppl. 1D
G1512: HP1 KD in G615cl1m1cl1	PBANKA_143610	N*	N	Suppl. 1E
G1513: HDAC1 KD in G615cl1m1cl1	PBANKA_082650	Y	N	Suppl. 1F
G1514: HDAC2/IPK1 KD in G615cl1m1cl1	PBANKA_120620	Y	Y	Suppl. 1G
G1516: HDA1 KD into G615cl1m1cl1	PBANKA_133540	Y	N	Suppl. 1H
G1517: HDA, putative KD in G615cl1m1cl1	PBANKA_110620	Y	Y	Suppl. 1I
G1623: HP1 KD in G1448cl1m0cl1	PBANKA_143610	Y	N	Suppl. 1J
G1624: HDAC1 KD in G1448cl1m0cl1	PBANKA_082650	Y	N	Suppl. 1K
G1529: HAT1 KO, 1st transfection into 820	PBANKA_071840	Y	Y	Suppl. 1L
G1561: HAT1 KO, 2nd transfection into 820	PBANKA_071840	Y	N	Suppl. 1L
G1689: HAT1 KO, 3rd transfection into 820	PBANKA_071840	Y	N	Suppl. 1L
G1693: HAT1 KO, 1st transfection into PbEGAM	PBANKA_071840	Y	Y	Suppl. 1M
G1694: HAT1 KO, 2nd transfection into PbEGAM	PBANKA_071840	Y	Y	Suppl. 1M
G1511: HDA1 KO, 1st transfection into 820	PBANKA_133540	N	N	Suppl. 1N
G1560: HDA1 KO, 2nd transfection into 820	PBANKA_133540	N	N	Suppl. 1N
G1613: HDA1 KO, 3rd transfection into 820	PBANKA_133540	Y	N	Suppl. 1N
G1687: HDA1 KO, 4th transfection into 820	PBANKA_133540	Y	N	Suppl. 1N
G1688: HDA1 KO, 5th transfection into 820	PBANKA_133540	Y	N	Suppl. 1N
G1579: HDA1 KO, 1st transection into PbEGAM	PBANKA_133540	Y	N	Suppl. 1O

Table 3.2: Complete list of *P. berghei* ANKA parasite lines used and generated in the present study. This table provides an overview of all complete gene knockout (KO) or conditional gene knockdown (KD) *P. berghei* ANKA lines created during the present study. The background lines used, into which linear plasmid vectors were integrated, are also listed. The first column lists the common number or name given to the *P. berghei* line, alongside a more detailed description in the case of KO and KD lines. The second column lists the gene identifier (ID) of the gene to be targeted (in the case of a KO or KD line). The third column indicates if successful integration of the linear KO or KD vector into the genome of the chosen background line was demonstrated by PCR after transfection (all integration primers are listed in **Appendix, Table 7.1**). In the case of the heterochromatin protein 1 (HP1) background line, an asterisk in the third column (**N***) indicates that integration was only ever shown at the 3' UTR of the gene and not the 5' UTR, despite multiple primer designs being attempted. The fourth column indicates whether a successful clonal or isogenic line was generated for the KO or KD parasite line in question (either by limiting dilution or FACS). The fifth and final column listed the location of a more detailed plasmid map for the plasmid insertions made to the genome of each respective *P. berghei* ANKA line.

3.3 Results

3.3.1 Creation of KO vectors and confirmation of restriction site integrity

At the outset of this study, an attempt was made to generate KO plasmids for genes of interest (GOIs) using a pL0035 background plasmid (containing a *hdhfr::yfcu* drug selection sequence) and homology arms to facilitate double-crossover homologous recombination with the endogenous gene. Upstream and downstream homology arms for the final KO vectors were amplified from native genes (from the *P. berghei* ANKA high-gametocyte producer (HP) background line genomic DNA). All experimental methods used for the generation of KO and KD plasmids are detailed in section **2.2.15.** To confirm the presence of the correct GOI-specific homology arms in each pL0035-based KO construct, both confirmatory PCR and restriction digests were carried out. Restriction digest results for five such KO vectors are shown in **Figure 3.1.**

a

Plasmid name	Target protein	Size (minus *amp* backbone)
pL0035	None (plasmid backbone)	3286 bp
pG421	HP1	4730 bp
pG422	RNAseII	4849 bp
pG423	IPK2	4713 bp
pG424	HDAC1	4086 bp
pG425	HDAC2/IPK1	4658 bp

b

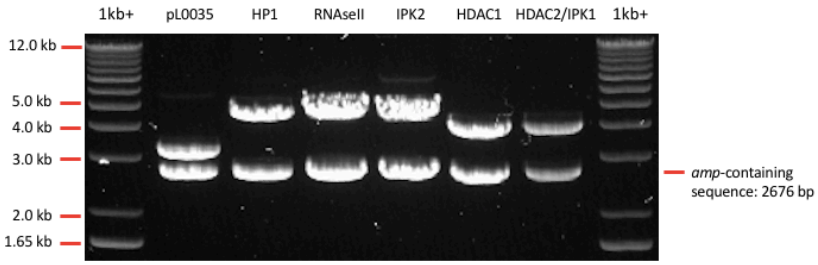


Figure 3.1: Confirmatory restriction digest of five pL0035-based KO constructs for *P. berghei* ANKA transfection. This figure shows five pL0035-based KO vectors and their expected product sizes after linearisation of the circular plasmid with XhoI and HindIII **(a).** Upon restriction digestion with these two enzymes, a background pL0035 sequence containing the bacterial Ampicillin resistance gene (*amp*) was excised, leaving behind a linearised vector in which the *hdhfr::yfcu* drug selection cassette, alongside its *5'eef1α* promoter and *3'pbdhfr-ts* UTRs, was flanked by homology arms that mapped to upstream and downstream regions adjacent to the GOI. In **(b),** the results of these restriction enzyme cuts are shown on an agarose gel. An untargeted pL0035 vector is shown next to the linearised pL0035-based vectors for comparison. All XhoI/HindIII restriction digests result in two bands: the *amp*-containing background plasmid sequence of 2676 bp, and the

larger fragment to be gel extracted for transfection experiments. All linearised plasmids in **(b)** produced products of the correct size with the exception of the pG425 vector for targeting of HDAC2/IPK1 in *P. berghei*. The linearised vector in this case seemed slightly smaller than the expected 4658 bp.

In a different manner to pL0035-based DXO KO constructs, the two *Plasmo*GEM KO constructs for HDA1 (PbGEM-250987) and HAT1 (PbGEM-236544) needed only to be digested with NotI to remove the bacteria-specific kanamycin resistance gene and remnants of the pJAZZ-OK background vector prior to transfection (Pfander *et al.*, 2013; Schwach *et al.*, 2015). Both *Plasmo*GEM vectors were NotI digested and transfected into *P. berghei* ANKA background lines to create new KO lines as depicted in **Supplemental images S1L to S1O**.

3.3.2 Creation of KD vectors and confirmation of restriction site integrity

In this study, 7 knockdown (KD) DNA constructs were successfully created using the pG364 background plasmid (six of which were transfected and the modified genetic loci of resultant lines are depicted in **Supplemental images S1E to S1I**). Three more DNA constructs were created in the pG472 BFP-containing plasmid for examination of the effects of three of these same proteins on both male and female gametocytogenesis after conditional degradation (two of which were transfected and the modified genetic loci of these lines are shown in **Supplemental images S1J and S1K**).

For three of the genes-of-interest targeted for degradation via the AID system (heterochromatin protein 1 (HP1), histone deacetylase 1 (HDAC1), and the joint histone deacetylase 2/inositol phosphate kinase 1 (HDAC2/IPK1) gene), endogenous unique restriction sites were used for linearisation of the final circular DNA construct (restriction sites HpaI, PmlI, and BglII respectively **(Supplemental images S1E-S1G**)). For all other genes-of-interest, a SalI restriction site was inserted at the centre of the chosen single-crossover (SXO) homology sequence (**Supplemental images S1H-S1I)**. Following PCR amplification of SXO homology sequences for all 7 successfully-generated pG364-based KD plasmids, the unique restriction sites of all 7 final plasmids were confirmed by sequencing and restriction digest (sequencing results not shown).

Following successful creation of 7 DNA constructs that were capable of targeting 7 separate putative epigenetic modifiers via auxin-induced conditional degradation, transfection of linearised plasmids into the *P. berghei* 'AID' background line were carried out. At a later stage in this study, the first three SXO inserts created (targeting genes for HP1, HDAC1, and HDAC2/IPK1) were amplified again and ligated into the pG472 BFP-expressing plasmid (with 2 successful results, as depicted in **Supplemental images S1J** and **S1K)**. These ligation reactions were carried out by Dr Heli Vaikkinen following creation of the 'AID 820' background line (**Supplemental image S1D)**.

3.3.3 Integration of KO constructs into the *P. berghei* ANKA genome: essential and dispensable gene identification

Using both *PlasmoGEM* and in-house KO DNA constructs (as shown in **Figure 3.1**), complete knockout of five putative epigenetic modifiers (HP1, HDAC1, HDAC2/IPK1, HDA1 and HAT1) and two related proteins (RNAseII and IPK2) was attempted in *P. berghei* ANKA. Methods for the transfection of linearised constructs and drug selection of positively-transfected KO parasites are described in section **2.1.5.** Upon confirmation of positive integration of a KO construct, the positively drug-selected KO *P. berghei* line was cloned by limiting dilution (detailed in section **2.1.20**).

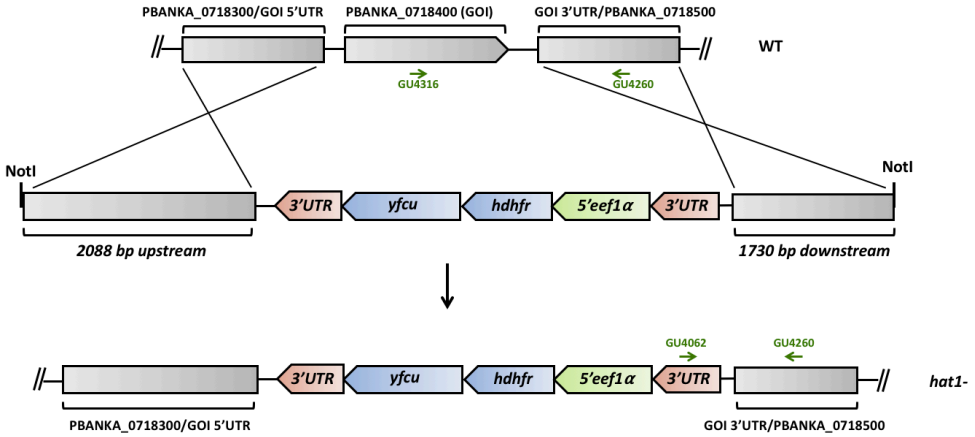
In the cases of *P. berghei* genes encoding for HP1, RNAseII, IPK2, HDAC1, and HDAC2/IPK1, three attempts each to transfect these KO constructs into the *P. berghei* ANKA 820 line resulted in either no parasites being produced after transfection, or generation of wild-type *P. berghei* ANKA parasites only (data not shown). In the cases of HP1, RNAseII, HDAC1, and HDAC2/IPK1, the failure to completely knock out these genes correlated with previous attempts in *P. falciparum* and *P. berghei* ANKA (Brancucci *et al.*, 2014; Coleman *et al.*, 2014; Zhang *et al.*, 2014).

However, in the case of IPK2, failure to knockout this gene did not correlate with the observed dispensable phenotype previously identified during *Plasmo*GEM screening (Bushell *et al.*, 2017). It is possible however, that the pL0035-based IPK2 KO construct was not as successful at integrating into the *P. berghei* ANKA genome

as a result of the shorter gene homology arms, which give the pJAZZ-OK-based *Plasmo*GEM constructs greater advantage when attempting to mediate gene knockout via double-crossover homologous recombination (Schwach *et al.*, 2015). The necessity for longer homology arms when attempting a complete KO of a *P. berghei* protein kinase, such as those of a pJAZZ-OK-based *Plasmo*GEM construct, has been observed previously (Tewari *et al.*, 2010).

In the cases of HDA1 and HAT1 complete knockout using *Plasmo*GEM vectors, transfection into both the 820 and PbEGAM background lines yielded positive integration PCR results, and generated *P. berghei* ANKA KO lines in which the role of both HDA1 and HAT1 in asexual growth and gametocytogenesis could be investigated. Details of *P. berghei* lines in which the HAT1 *Plasmo*GEM KO construct was integrated are shown in **Figure 3.2.** Details of HDA1 *Plasmo*GEM KO construct integration into both *P. berghei* ANKA background lines are similarly shown in **Figure 3.3.**

a



b

Primers	Product	Size
GU4260 + GU4316	Wild-type	2185 bp
GU4260 + GU4062	3' integration	2118 bp

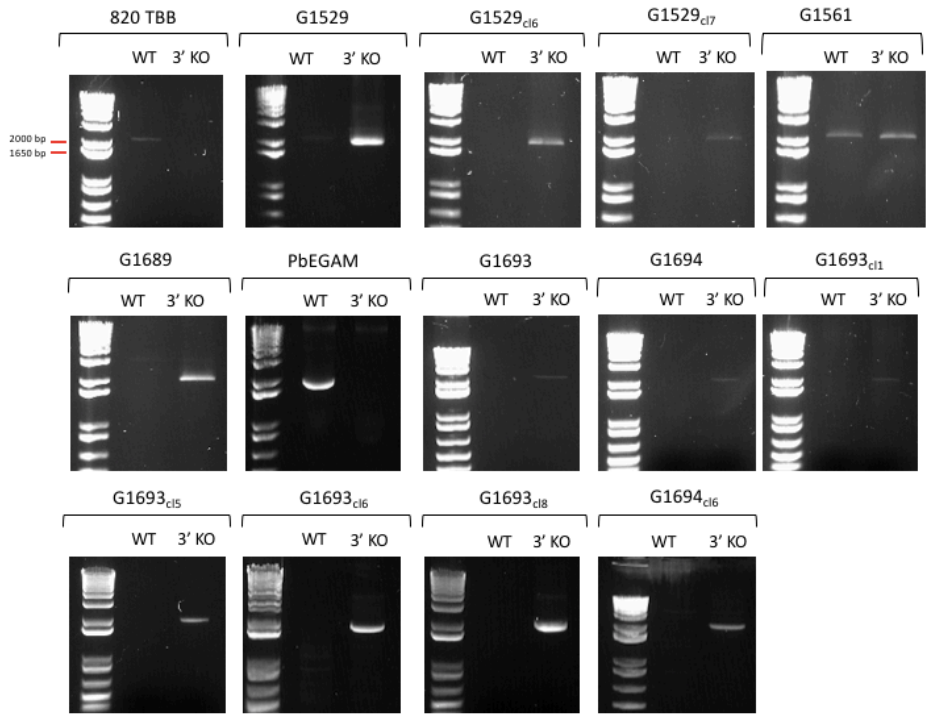


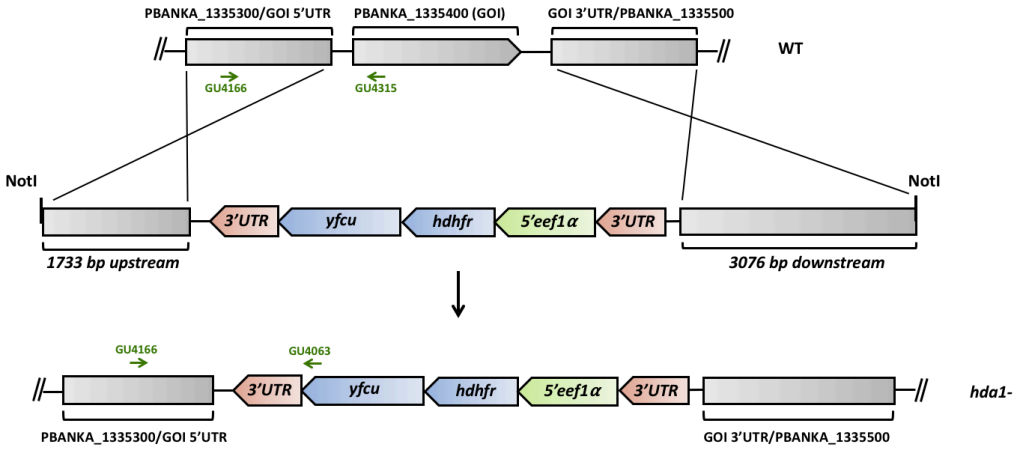
Figure 3.2: Integration of the HAT1 KO construct into *P. berghei* ANKA 820 and PbEGAM background lines. Images (**a**) and (**b**) together describe the homologous recombination of the HAT1 *Plasmo*GEM KO construct (PbGEM-236544) into the genome of two *P. berghei* ANKA background lines, 820 and PbEGAM. (**a**) is a schematic representation of the wild-type *P. berghei* HAT1 gene (PBANKA_0718400) before and after integration of the KO construct. Regions flanking the GOI that are targeted for homologous recombination include part of the upstream and downstream genes (PBANKA_0718300 and PBANKA_0718500 respectively), and part of the 5' and 3' untranslated regions (5' and 3' UTRs) of the HAT1 gene itself. The sites of NotI restriction digest are indicated in the linear plasmid and primer sites are represented by small green arrows and the 'GU' number

given to the primer. All primers are listed in **Table 7.1** as an appendix and the size of the resulting fragments in base pairs (bp) are listed in part (**b**) of this image.

(**b**) **T**able listing the primer products used to assess successful integration of the KO construct, with the PCR results shown below for all of the lines used, including controls and clones of the positively-transfected *P. berghei* KO lines. Transfection of the HAT1 KO construct was attempted three times into the 820 TBB line, generating lines G1529, G1561, and G1689. As G1529 yielded the strongest KO to WT ratio of all three lines, this line was cloned by limiting dilution to produce lines G1529 clone 6 ($G1529_{cl6}$) and G1529 clone 7 ($G1529_{cl7}$). Transfection of the HAT1 KO construct into PbEGAM was attempted twice, with both attempts resulting in successful integration, generating *P. berghei* lines G1693 and G1694. Both of these lines were cloned to give clonal HAT1 KO lines G1694 clone 6 ($G1694_{cl6}$) and 4 clones of G1693 (clones 1, 5, 6, and 8). Generating clonal lines in both *P. berghei* ANKA backgrounds enabled confirmation of the HAT1 KO phenotype with regard to both asexual growth, and the production of male and female gametocytes.

As shown in **Figure 3.2,** complete KO of the HAT1 gene produced *P. berghei* ANKA lines that were capable of asexual propagation and cloning via limiting dilution in mice. In **Figure 3.3**, repeated attempts to generate HDA1 KO lines in both 820 and PbEGAM did result in some success. However, the apparent slow asexual multiplication rate of recombinant parasites produced in HDA1 KO lines prevented cloning of the lines by limiting dilution in mice (7 attempts). The possible reasons for this are discussed further at the end of this chapter.

a



b

Primers	Product	Size
GU4166 + GU4315	Wild-type	2256 bp
GU4166 + GU4063	5' integration	2288 bp

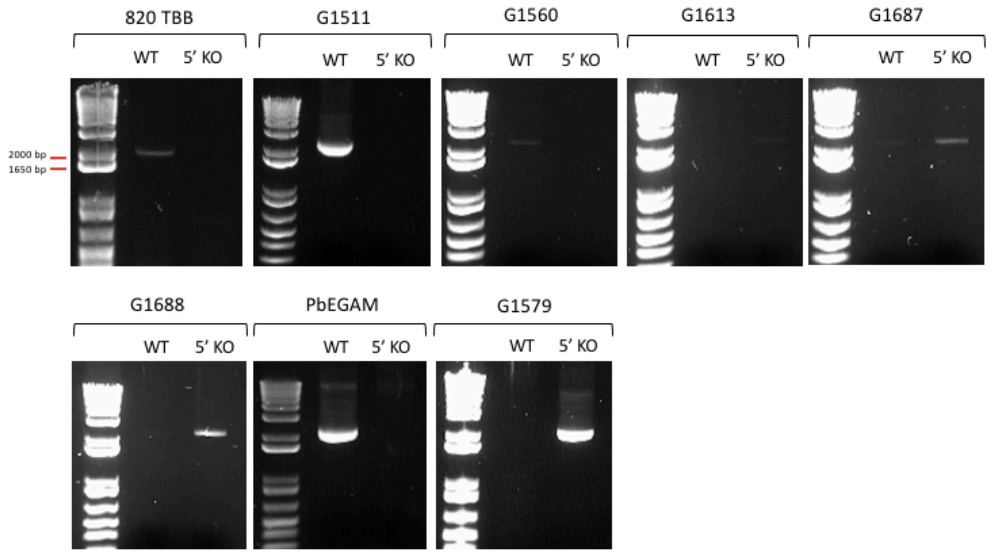


Figure 3.3: Integration of the HDA1 KO construct into *P. berghei* ANKA 820 and PbEGAM background lines. Images (**a**) and (**b**) detail the homologous recombination of the HDA1 *Plasmo*GEM KO construct (PbGEM-250987) into the genome of two *P. berghei* ANKA background lines, 820 and PbEGAM. (**a**) is a schematic representation of the wild-type *P. berghei* HDA1 gene (PBANKA_1335400) before and after integration of the KO construct. Regions flanking the GOI that are targeted for homologous recombination include part of the upstream and downstream genes (PBANKA_1335300 and PBANKA_1335500 respectively), and part of the 5' and 3' untranslated regions (5' and 3' UTRs) of the HDA1 gene itself. NotI and primer sites are indicated in black and green. The sizes of PCR fragments in base pairs (bp) are listed in part (**b**) of this image.

(**b**) lists the primer products used to assess successful integration of the KO construct, with the PCR results shown below for all of the lines used, including 820 and PbEGAM controls. Unlike the *Plasmo*GEM HAT1 KO construct, the HDA1 KO vector has a much longer downstream homology arm

and so integration was assessed at the 5' end of the construct. As seen in (**b**), transfection of the HDA1 KO construct was attempted five times into the 820 line, generating lines G1511, G1560, G1613, G1687 and G1688. From G1511 to G1688, the ratio of wild-type (WT) *P. berghei* parasites to HDA1 KO parasites decreases, with G1687 and G1688 having the highest number of HDA1 KO parasites. These WT:*hda1*$^{-}$ ratio differences reflect the point at which parasites were harvested from mice after transfection. As a result of continuously low parasitaemia counts in mouse blood smears, transfected lines G1511, G1560 and G1613 were only harvested from mice from 10-14 dpi (all while under positive Pyrimethamine selection pressure). It became apparent from these transfection attempts that HDA1 KO parasites were being outgrown by WT parasites over time. To combat this, two more transfection attempts into 820 were made, with G1687 and G1688 lines being harvested from mice at 7 dpi (after Giemsa staining from 3 dpi). In these lines, integration of the HDA1 KO construct was much clearer. One transfection of the HDA1 KO construct resulted in line G1579 which showed positive integration of the KO vector.

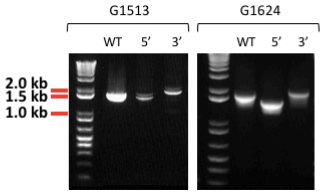
It is noticeable in **Figure 3.3** that the HDA1 KO construct could be integrated into the 820 and PbEGAM background lines, but no successful clonal lines were generated. In repeated attempts to create clonal lines of G1687, G688, and G1579, identical methods to grow synchronous schizonts for cloning by limiting dilution were carried out as detailed in sections **2.1.5.4** and **2.1.20.** In all cases, very few mature asexual (schizont) parasites could be cultivated, and 2 attempts each to clone lines G1687, G1688, and G1579 failed to yield any positive results. The combination of the outgrowth of transfected HDA1 KO parasites by WT parasites after transfection, and the failure to produce mature asexual parasites that could propagate *in vivo* after limiting dilution, pointed to a deficiency of asexual growth in the HDA1 KO *P. berghei* lines. Further analyses of non-clonal HDA1 KO lines are shown below.

3.3.4 Integration of KD constructs into the *P. berghei* ANKA genome

In the present study, KD vectors were linearised and transfected into the OsTIR1-expressing background line, G615cl1m1cl1 (the AID background line; as depicted in **Supplemental image S1C**). Following successful integration of a number of KD constructs, cloning via limiting dilution was attempted in some cases, with generation of isogenic lines by FACS also attempted to reduce animal usage (detailed in section **2.1.21**). All lines generated from these experiments in the present study are shown in **Figure 3.4.**

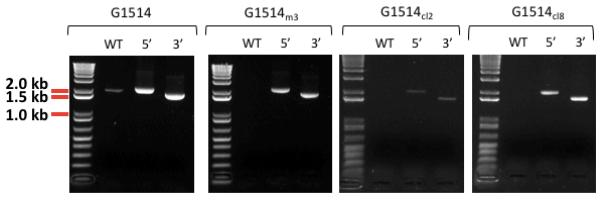
a

Name	Primers	Product	Size
HDAC1 (G1513)	GU4154 + GU4156	Wild-type	1767 bp
	GU4154 + GU4155	5′ integration	1567 bp
	GU4152 + GU4156	3′ integration	1929 bp



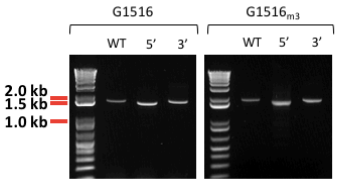
b

Name	Primers	Product	Size
HDAC2/IPK1 (G1514)	GU4631 + GU4160	Wild-type	1987 bp
	GU4631 + GU4158	5′ integration	2197 bp
	GU4159 + GU4160	3′ integration	1843 bp



c

Name	Primers	Product	Size
HDA1 (G1516)	GU4164 + GU4175	Wild-type	1900 bp
	GU4164 + GU4165	5′ integration	1775 bp
	GU4159 + GU4175	3′ integration	1948 bp



d

Name	Primers	Product	Size
HDAP (G1517)	GU4167 + GU4170	Wild-type	1409 bp
	GU4167 + GU4168	5′ integration	1336 bp
	GU4169 + GU4170	3′ integration	1481 bp

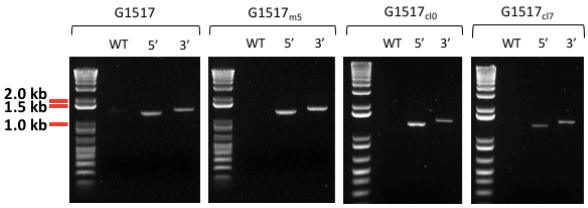


Figure 3.4: *P. berghei* ANKA conditional knockdown (KD) lines generated in the present study. Images (**a**) to (**d**) show expected and obtained PCR fragment sizes for integration of 4 KD plasmids into background lines: the pG439 vector targeting HDAC1; the pG440 vector targeting histone HDAC2/IPK1; the pG474 vector targeting HDA1; and the p475 vector targeting another putative histone deacetylase (termed 'HDAP'). Image (**a**) shows the successful integration of the HDAC1 KD plasmid into both the AID and AID 820 background lines, creating *P. berghei* ANKA lines G1513 and G1624 respectively. In each case, a WT gene fragment is shown next to fragments confirming both 5' and 3' integration of the KD construct into the parasite genome. In (**b**), integration of the HDAC2/IPK1 KD plasmid into the AID background line is shown, creating line G1514. Adjacent to agarose gel images confirming correct PCR fragment sizes in this line are results of both FACS-aided G1514 enrichment and cloning by limiting dilution. In this case, FACS-aided selection of G1514 parasites only (without WT background parasites) produced isogenic line $G1514_{m3}$ (m_3 indicating 'mouse 3') in which no WT PCR fragment could be detected. At the same time, cloning of this line was undertaken by limiting dilution, and clonal lines $G1514_{cl2}$ and $G1514_{cl8}$ were generated.

In (**c**), integration of the HDA1 KD plasmid into the AID background line is confirmed in the first gel image (creating line G1516), with an attempted FACS-selected isogenic line ($G1516_{m3}$) shown in the adjacent gel image. Unlike in the case of $G1514_{m3}$, $G1516_{m3}$ did not remain isogenic over time, with some WT parasites growing back, even under Pyrimethamine drug pressure. In (**d**), the HDAP KD plasmid is shown to be integrated into the AID background line also, generating line

G1517. As with G1514, both an isogenic line ($G1517_{m5}$), and two clonal lines ($G1517_{cl0}$ and $G1517_{cl7}$) were produced through FACS-aided selection of mCherry-fluorescent transfected parasites and cloning by limiting dilution in mice.

In addition to the *P. berghei* ANKA lines shown in **Figure 3.4**, attempts were made to transfect three more KD plasmids (for HP1 (shown in **Supplemental image S1E**), SIR2A, and HAT1 (latter two not shown) into the AID and AID 820 background lines, but to no avail. In the case of HP1, transfection of the KD vector into *P. berghei* lines 507, AID, and AID 820 background *P. berghei* ANKA lines were undertaken, creating parasite lines G1304, G1512, and G1623. In each case, 3' integration was observed following transfection. However, repeated primer redesign and PCR amplification under various conditions failed to confirm both 5' and 3' integration of this plasmid and so further experiments were undertaken with the successfully-transfected parasite lines shown in **Figure 3.4.**

3.3.5 Characterising *P. berghei* ANKA *hat1*- and *hda1*- KO lines by FACS

With each experiment of a *P. berghei* KO line by FACS, the equivalent background line (820 or PbEGAM) was run in parallel as a control. Under FACS analysis, the 820 background line produced three populations: GFP/RFP double-negative asexual-stage parasites and uninfected erythrocytes, GFP^+ male gametocytes, and RFP^+ female gametocytes (**Figure 3.5**). When examining the PbEGAM background line, four cell populations can be observed: GFP/RFP double-negative uninfected erythrocytes, GFP^+ asexual-stage parasites, GFP^+/RFP^+ male gametocytes, and GFP^{++}/RFP^+ female gametocytes. The enhanced GFP signal in the female gametocyte population is a result of the increased *5'eef1*α promoter expression in female gametocytes, in comparison to male gametocytes (Franke-Fayard *et al.*, 2004; Lee, Waters and Brewer, 2018). Therefore, in both the 820 and PbEGAM *P. berghei* lines, male-to-female gametocyte ratios could be characterised (**Figure 3.5**).

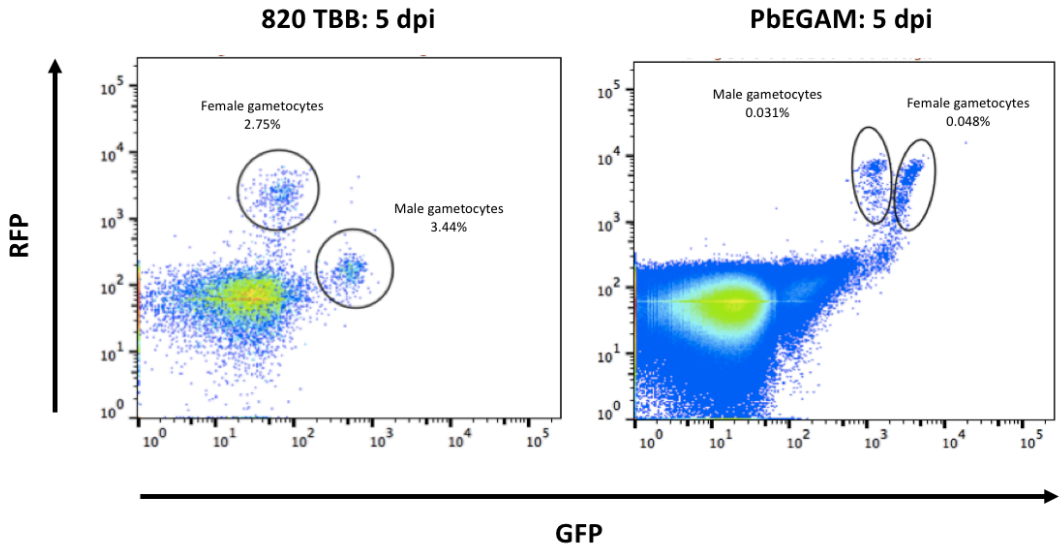


Figure 3.5: FACS analysis of *P. berghei* 820 and PbEGAM background lines. This image shows *P. berghei* ANKA background lines 820 and PbEGAM by FACS, 5 days after inoculation of the animal with 0.01% asynchronous parasites. Overall parasitaemias were 2.20% (820) and 3.24% (PbEGAM). Intensity of red fluorescent protein (RFP) is shown on the y-axis in both plots, with intensity of green fluorescent protein (GFP) shown on the x-axis of both FACS plots. GFP fluorescence corresponded to BD Horizon Brilliant™ Blue 515 (BB515) (515 nm maximum emission) and RFP fluorescence corresponded to R-phycoerythrin (PE) (578 nm maximum emission). Each plot represents 1 million events (cells) after gating. With regard to the 820 FACS plot, all cells shown represent Hoechst-stained singlets (to separate infected from uninfected erythrocytes; detailed in section **2.1.18**). In the PbEGAM FACS plot, all cells shown represent singlets alone, with infected erythrocytes identified by the constitutive expression of GFP in this background line. Population percentages of male and female gametocytes in these FACS plots are percentages of the gated population shown in the plot, and do not represent the percentages relative to all 1 million events.

In the 820 line, female gametocyte presence is based upon RFP fluorescence, which is under the control of the female gametocyte-specific LCCL-domain-containing gene 5' promoter (upstream of PBANKA_1319500) (Ponzi *et al.*, 2009). Both the 820 female- and male- specific promoters are based upon a proteomic study which showed that the LCCL-domain-containing protein, CCP2, is heavily female-specific, with the heavy dynein chain protein (encoded for by gene PBANKA_0416100) being heavily expressed in male gametocytes in *P. berghei* (Khan *et al.*, 2005). In the PbEGAM background line, gametocytes are differentiated from asexual parasites based upon RFP expression, driven by an 'early gametocyte marker', the 5' promoter of PBANKA_1018700, a conserved *Plasmodium* gene of as-yet unknown function (Lee, Waters and Brewer, 2018). Transcription of the 'early gametocyte' gene itself (PBANKA_1018700) was initially

noted during a study of differentially expressed genes between sexually-committed and asexual *P. berghei* ANKA parasites, which then led to the discovery of the AP2-G transcriptional regulator (Sinha *et al.*, 2014).

In the PbEGAM parasite line, RFP expression could be detected from 8-12 hours post-invasion (Lee, Waters and Brewer, 2018), with male and female gametocytes detectable as shown in **Figure 3.6.** When assessing total gametocytaemia in the PbEGAM background line by FACS, total gametocytaemia was determined to be all GFP^+/RFP^+ cells. When two separate GFP^+/RFP^+ populations became distinguishable, the GFP^{++}/RFP^+ population was counted as the female gametocyte population, and the separate GFP^+/RFP^+ population was labelled as the male gametocyte population. Because not all gametocytes were clearly distinguishable as either male or female, three gametocytaemia counts were made when using the PbEGAM background *P. berghei* line: total gametocytaemia; male gametocytaemia; and female gametocytaemia (as percentages of the overall infected-erythrocyte population) (**Figure 3.6**).

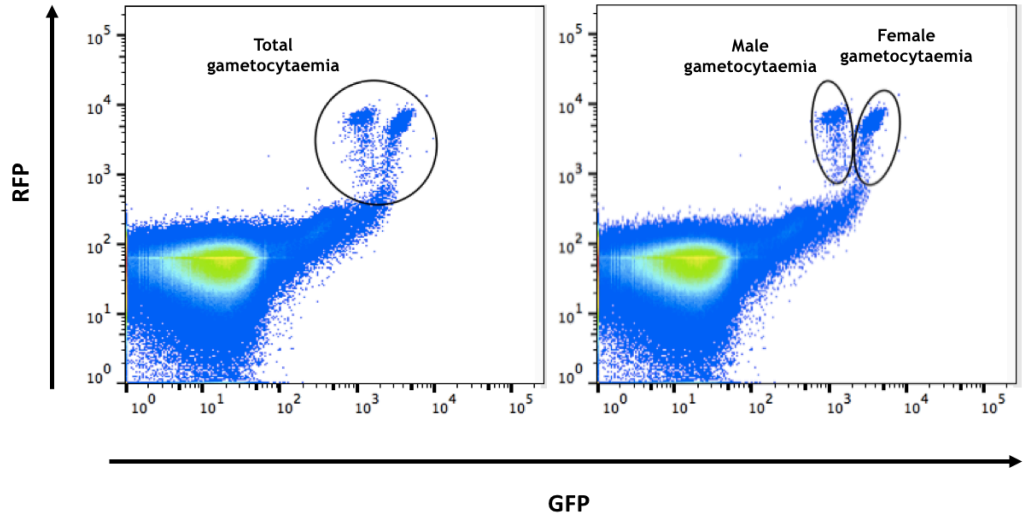


Figure 3.6: Determining total gametocytaemia, male gametocytaemia, and female gametocytaemia in the PbEGAM *P. berghei* line. This image shows an identical FACS plot of GFP^+ cells in a PbEGAM infection after 4 days (4 dpi). As in all cases in the present study, infection began at 0.01% parasitaemia in a female Theiler's Original (TO) mouse. In the left-hand panel, a population of GFP^+/RFP^+ cells is highlighted inside a circle. This population of GFP^+/RFP^+ cells as a whole is labelled as the total gametocytaemia in this particular sample. In the right-hand FACS plot, the same population of GFP^+ cells is shown, but in this case, the distinguishable male and female gametocyte populations are indicated. Male gametocytes are GFP^+/RFP^+ in the PbEGAM line. Female gametocytes are distinguished as the GFP^{++}/RFP^+ cells (Lee, Waters and Brewer, 2018), with the 5' promoter of the constitutively-expressed *P. berghei* ANKA elongation factor 1

alpha protein (*5'eef1α*) (encoded for by PBANKA_1133300 and PBANKA_1133400) being expressed to a greater degree in female gametocytes. However, not all GFP^+/RFP^+ gametocytes can be separated into male or female categories, and so the 'total gametocytaemia 'counts in the present study account for male, female, and early gametocyte populations as a whole.

3.3.6 *hat1- P. berghei* ANKA lines show reduced asexual growth and gametocytogenesis

Following repeated FACS analysis of clonal *hat1- P. berghei* lines $G1529_{cl6}$ (HAT1 KO in 820), $G1693_{cl6}$ (HAT1 KO in PbEGAM) and $G1694_{cl6}$ (HAT1 KO in PbEGAM from second transfection), a phenotype for *P. berghei* ANKA lines in which HAT1 had been knocked out became apparent. After 4 days, the population of male gametocytes makes up 3.86% of the overall Hoechst-stained singlet population in 820. RFP-fluorescing female gametocytes represent 2.75% of the overall 820-infected erythrocyte population. In the adjacent plot showing the same populations after HAT1 KO, the percentage of male gametocytes has fallen to 0.9%; with female gametocytes making up only 0.63% of the overall population. Taking these representative FACS samples alone, the HAT1 KO line has 76.7% fewer male gametocytes, and 77% fewer female gametocytes, than its HAT1-expressing background line (**Figure 3.7 (a), (c)** and **(d)).**

When characterising HAT1 KO in the 820 *P. berghei* ANKA line, 10 independent infections with 0.01% 820 over four days were examined to determine wild-type parasitaemia and gametocytaemia over time. Five independent infections with the HAT1 KO, $G1529_{cl6}$, again at a starting parasitaemia of 0.01%, were then examined. A comparison of wild-type *P. berghei* ANKA 820 parasitaemia and gametocytaemia are shown alongside HAT1 KO results in **Figure 3.7 (c)** and **(d).**

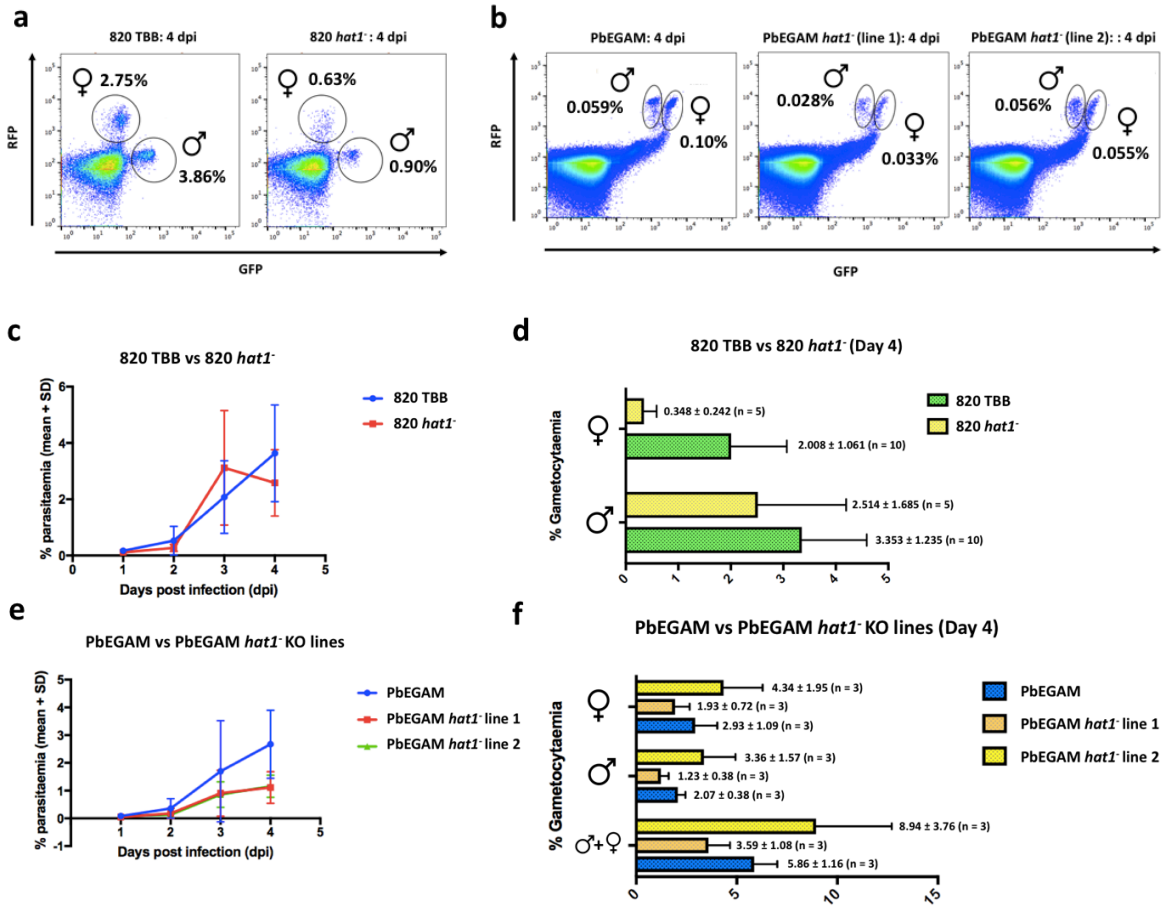


Figure 3.7: The *hat1*⁻ phenotype in *P. berghei* ANKA 820 and PbEGAM lines. (**a**) and (**b**) show representative examples of FACS results from all HAT1 KO *P. berghei* ANKA lines. RFP fluorescence is shown on the y-axis in all plots, with GFP fluorescence displayed at the x-axis of all FACS plots. In (**a**), the 820 line is shown at 4 dpi next to a clonal HAT1 KO line in this background, G1529cl6, again at 4 dpi. Each of these FACS plots represents a population of Hoechst-stained singlet cells, with male and female gametocyte populations represented by GFP fluorescence and RFP fluorescence respectively. (**b**) shows a FACS plot of the PbEGAM line at 4 dpi next to two examples of 820 *hat1*⁻ in the same line at 4 dpi also. **(c)** and **(d)** show a comparison of mean asexual growth and gametocytaemia in an 820 HAT1 KO line over four days. (**c**) shows the average percentage parasitaemia of both 820 and 820 *hat1*⁻ *P. berghei* lines over four days and image **(d)** shows the percentage of male and female gametocytes within these populations over the same four-day period. In (**c**), values used to generate a four-day growth curve of percentage parasitaemia in the 820 line were mean percentages (and standard deviations) from 10 independent infections starting at 0.01% parasitaemia. On the same graph, the mean percentage parasitaemia from 5 independent 820 *hat1*⁻ infections (including standard deviations) is shown. In **(d),** the percentage of infected erythrocytes that can be designated as either male or female gametocytes are shown. Bars shown in green represent gametocyte percentages for the 820 background line, while yellow bars represent the percentage male and female gametocytaemias for 820 *hat1*⁻. The percentage gametocytaemias at 4 dpi are shown (mean values and standard deviations). As with (**c**), mean values are taken from 10 independent 820 experiments and 5 independent 820 *hat1*⁻ experiments.

Finally, **(e)** and **(f)** show a comparison of mean asexual growth and gametocytaemia in two PbEGAM HAT1 KO lines over four days. Similarly to (**c**), **(e)** shows the average percentage parasitaemia of PbEGAM and both HAT1 KO clonal lines, $G1693_{cl6}$ (line 1) and $G1694_{cl6}$ (line 2), over four days. Counts of the total, male, and female gametocyte populations over the same four-day period are shown in (**f**). In both **(e)** and **(f),** percentage parasitaemias over 4 days of growth, and percentage gametocytaemias at 4 dpi are shown with mean values and standard deviations indicated also.

At days 1, 2, and 4 in **Figure 3.7** (**c**), the mean percentage parasitaemia is higher in the wild-type 820 line when compared to the HAT1 KO line. At day 3, the mean parasitaemia is 1.042% greater in the HAT1 KO line when compared to the wild-type, though standard deviations account for this difference. Further experiments may shed more light on the $G1529_{cl6}$ growth curve and explain this anomaly. At present asexual growth changes shown in **Figure 3.7 (c)** are not statistically significant (unpaired, two-tailed t-test; $P<0.05$). With regard to percentage gametocytaemias, the decrease in female gametocytes after HAT1 KO in this line was significant ($P=0.0048$; unpaired two-tailed t-test). However, the decrease in male gametocytes observed in **Figure 3.7** (**d**) is not significant ($P=0.29$; unpaired two-tailed t-test).

With repeated infections and FACS analysis of 820 and $G1529_{cl6}$ asexual and sexual growth over repeated four-day periods, it became apparent that there was attenuated asexual growth in the HAT1 KO clonal line, $G1529_{cl6}$, as was determined previously by *Plasmo*GEM screening (data not shown). However, more striking than this non-statistically significant growth attenuation was the significant decrease in female gametocytes over four days of growth, when compared to the wild-type line (**Figure 3.7 (c)** and **(d)**).

In the PbEGAM background line, the male and female gametocyte ($GFP^{+}RFP^{+}$ and $GFP^{++}RFP^{+}$) populations make up 0.059% and 0.1% of the total GFP^{+} singlet population. In the $G1693_{cl6}$ FACS plot, these populations have fallen to 0.028% male gametocytes and 0.033% female gametocytes. In the $G1694_{cl6}$ line at 4 dpi, male and female gametocyte populations are also lower than the background *P. berghei* line, at 0.056% and 0.055% of overall GFP^{+} singlets respectively. Taking these FACS plots alone, the $G1693_{cl6}$ line shows 53% fewer male gametocytes and 67% fewer female gametocytes than the background line. The $G1694_{cl6}$ line shows populations of 5% fewer male gametocytes and 45% fewer female gametocytes.

These examples alone indicated a decreased gametocytaemia in *hat-* *P. berghei* lines relative to background levels. These experiments were then replicated to confirm the existence of a HAT1 KO-induced, gametocyte-depleted phenotype (**Figure 3.7 (b), (e),** and **(f)).**

Using the PbEGAM *P. berghei* line, and counting gametocyte populations as shown in **Figure 3.6**, two HAT1 KO lines were examined: *P. berghei* ANKA clonal lines $G1693_{cl6}$ and $G1694_{cl6}$. Following an infection of four days, three independent biological replicates were studied. All results from the HAT1 KO lines are shown in **Figure 3.7 (e)** and **(f).** As in line $G1529_{cl6}$ (the 820 *hat1-* line), an attenuated growth phenotype over 4 days is seen in **Figure 3.7** (**e)** and (**f**), though the decreases in growth in both cases are not statistically significant (unpaired, two-tailed t-test; $P<0.05$; n=3 each). In **Figure 3.7** (**e**), taking gametocyte counts at 4 dpi, clonal line $G1693_{cl6}$ had a decreased overall gametocytaemia, though this was not statistically significant (P=0.0675; unpaired, two-tailed t-test).

The male gametocyte population was also decreased in a non-statistically significant manner (P=0.0550; unpaired, two-tailed t-test). Unlike in the 820 HAT1 KO, the decrease in female gametocytaemia in clonal line $G1693_{cl6}$ was not statistically significant (P=0.27; unpaired, two-tailed t-test). With regard to clonal line $G1694_{cl6}$, the HAT1 KO line showed an increase in overall gametocytaemia (in all populations), which differed to previous KO lines $G1529_{cl6}$ and $G1693_{cl6}$. This line showed an increase in gametocytaemia in all cases, though no change was statistically significant (unpaired, two-tailed t-test; $P<0.05$; n=3).

To summarise, HAT1 KO in both 820 and PbEGAM background lines resulted in a decreased rate of growth compared to wild-type *P. berghei* lines (**Figure 3.7**), though this attenuated growth rate was not statistically significant (using unpaired, two-tailed t-tests; $P<0.05$). In the 820 KO, $G1529_{cl6}$, repeated independent experiments concluded that the only statistically significant decrease in growth or gametocyte conversion regarded the generation of female gametocytes (**Figure 3.7 (d)**). After 5 independent experiments, a significant decrease in female gametocytes was observed (P=0.0048; unpaired two-tailed t-test).

Although these results were not observed when repeated in PbEGAM ($G1693_{cl6}$ and $G1694_{cl6}$), further experiments using these early-gametocyte marker lines may prove useful in providing a clearer picture of the effects of HAT1 KO on *P. berghei* sexual commitment. It may also be possible that a second mutation had occurred in the PbEGAM HAT1 KO line, $G1694_{cl6,}$ which would account for the inconsistency of results when comparing this line to both HAT1 KO lines, $G1529_{cl6}$ and $G1693_{cl6}$. Further discussion of the potential effects of HAT1 KO on *P. berghei* asexual and sexual stage maturation can be found in section **3.4** below.

3.3.7 Characterising asexual growth and gametocytogenesis in *hda1- P. berghei* ANKA lines

As mentioned previously in section **3.3.3**, integration of the HDA1 KO construct into both the 820 and PbEGAM background lines was successful, though repeated attempts to clone these lines proved futile. As demonstrated in **Figure 3.3**, failure to harvest peripheral blood immediately upon observation of a positive transfection experiment (as in lines G1511 and G1560), resulted in a greater ratio of wild-type 820 parasites when compared to the HDA1 KO line. As this asexual parasite growth deficiency was realised, later transfections were harvested from the moment parasites were observed in peripheral blood under Giemsa staining. These considerations resulted in lines G1613, G1687, and G1688, the latter two of which were chosen for cloning by limiting dilution.

After three attempts to clone line G1579, and two attempts each to clone lines G1687 and G1688, one mouse each was infected with an asynchronous 0.01% parasitaemia from non-clonal lines, G1687 and G1688, in an attempt to discern some information about the possible asexual growth defect effecting these HDA1 KO lines (**Figure 3.8**).

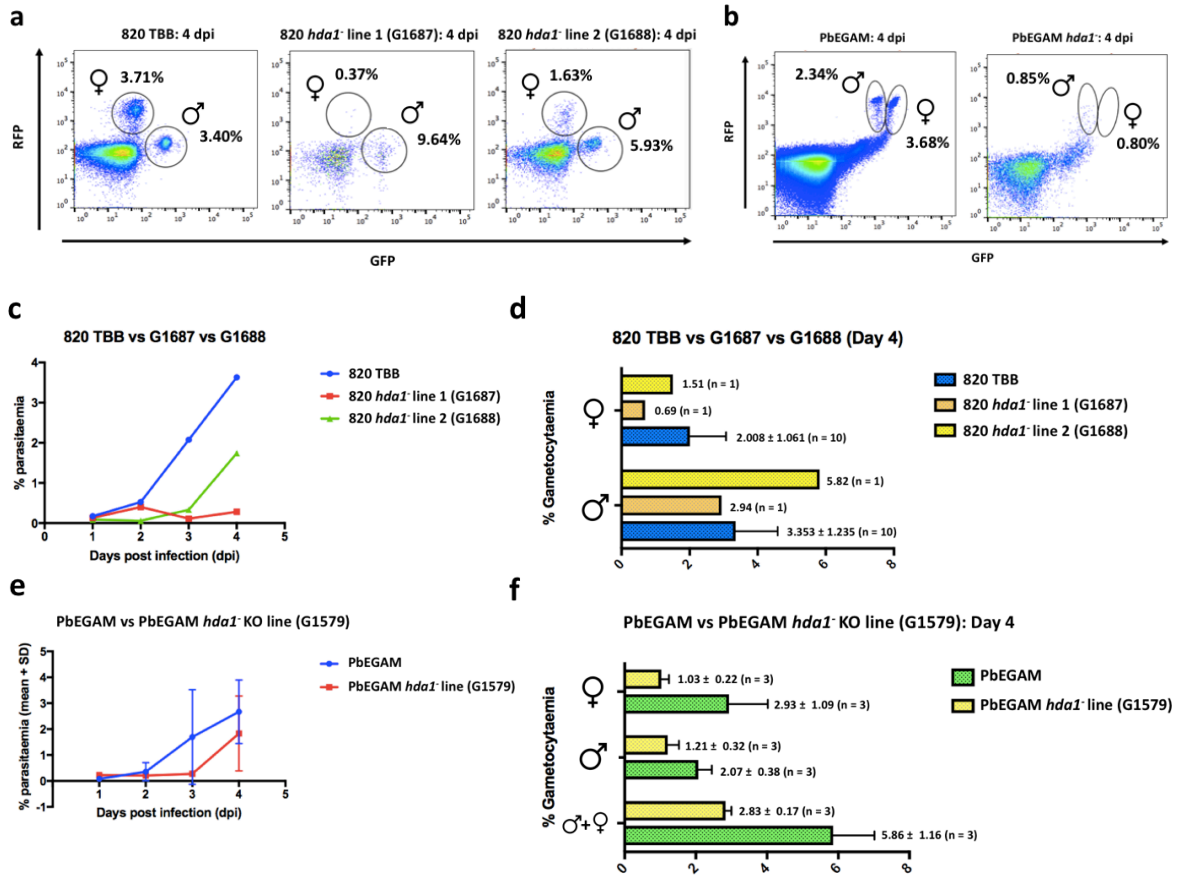


Figure 3.8: The *hda1*⁻ phenotype in *P. berghei* ANKA 820 and PbEGAM lines. In both (**a**) and (**b**), representative examples of FACS results from all HDA1 KO *P. berghei* ANKA lines are shown. Intensity of RFP fluorescence is shown on the y-axis in all plots, with intensity of GFP fluorescence shown on the x-axis. (**a**) shows an example of an 820 line is shown at 4 dpi next to FACS plots from two non-clonal HDA1 KO lines in this background: G1687 and G1688 (both at 4 dpi). Each of these FACS plots represents a population of Hoechst-stained singlet cells, with male and female gametocyte populations represented by GFP fluorescence and RFP fluorescence respectively. **(b)** shows representative FACS plots after 4 dpi from a WT PbEGAM sample and a G1579 HDA1 KO line (in the PbEGAM background). **(c)** and **(d)** show a comparison of asexual growth and gametocytaemia in two 820 HDA1 KO lines over four days. (**c**) depicts the average percentage parasitaemia of 10 820 experiments over four days of growth (the same 820 results used for comparison of 820 *hat1*⁻ KO parasite lines (**Figure 3.7**), alongside asexual growth of two *P. berghei* HDA1 KO lines over four days. (**d**) shows the percentage of male and female gametocytes within these populations over the same four-day period. In (**c**), values used to generate a four-day growth curve of percentage parasitaemia in the 820 line were mean percentages (and standard deviations) from 10 independent infections starting at 0.01% parasitaemia. On the same graph, the percentage parasitaemia from two single G1687 and G1688 infections are shown, again having started with a 0.01% parasitaemia.

In (**c**), a clear attenuated growth phenotype can be observed in both HDA1 KO lines when compared to the 820 background line. At 4 dpi, decreases in total parasitaemia seen in both HDA1 KO lines are striking, though only with repeated experiments can the significance of this attenuated growth

phenotype be assessed. In (**d**), again showing the mean values of 10 independent 820 experiments versus both single HDA1 KO samples, female gametocyte populations are decreased at 4 dpi in both HDA1 KO samples when compared to the background line, with the male gametocyte populations in both cases showing little decrease in population (G1687 KO) or even an increase in population (G1688 KO). **(e)** and **(f)** show a comparison of mean asexual growth and gametocytaemia in a PbEGAM HDA1 KO line over four days. Similarly to **Figure 3.8 (c)** and **(d),** (**e**) and (**f**) show a comparison of asexual growth and percentage gametocytaemia in a HDA1 KO (G1579) line versus its background line (PbEGAM) over 4 days infection, beginning at a parasitaemia of 0.01%. Taking the mean and standard deviations acquired over 3 independent experiments in each line, a decrease in overall parasitaemia over 4 days can be observed, in line with the attenuated growth phenotype seen in 820 HDA1 KO lines (**Figure 3.8 (a) and (c)**).

In the 820 FACS plot in **Figure 3.8** (**a**), two very distinguishable RFP^+ and GFP^+ populations can be observed. In the 820 background line, RFP^+ female gametocytes account for 3.71% of the total infected-erythrocyte population, with GFP^+ male gametocytes making up 3.40% of the total infected-erythrocyte population (**Figure 3.8 (a)**). Next to this FACS plot are corresponding plots from HDA1 KO lines, G1687 and G1688, showing drastically different results to those of the 820 line.

In addition to a considerable decrease in total infected-erythrocyte cell numbers, the RFP^+ female gametocyte populations at 4 dpi in these lines were 0.37% and 1.63% respectively. In complete contrast to this, the GFP^+ male gametocyte population made up 9.64% and 5.93% of the total infected-erythrocyte population in G1687 and G1688 respectively. Taking these representative FACS plots alone, the G1687 and G1688 HDA1 KO lines harboured female gametocyte populations that were 90% and 56% decreased in comparison to the 820 background *P. berghei* line. With regard to male gametocyte populations, the G1687 and G1688 KO lines showed male gametocytaemias that were 65% and 43% greater than that of WT 820 respectively.

Although attempts to clone all HDA1 KO lines were unsuccessful, 3 independent FACS analyses of G1579-infected erythrocytes were carried out as a result of this particular transfection showing no observable WT contamination upon PCR integration analysis (**Figure 3.3**). As shown in **Figure 3.8 (b),** PbEGAM GFP^+/RFP^+ male gametocytes and GFP^{++}/RFP^+ female gametocytes make up 2.34% and 3.68% of the total GFP^+ infected-erythrocyte population respectively. In the G1579 KO sample, taking the same populations, GFP^+/RFP^+ male gametocytes and

GFP^{++}/RFP^{+} female gametocytes make up only 0.85% and 0.80% of the overall GFP^{+} infected-erythrocyte population respectively, though this population as a whole is greatly reduced in comparison to the background line. Though the decrease in the GFP^{++}/RFP^{+} female gametocyte population is conserved among all HDA1 KO lines, the increased GFP^{+}/RFP^{+} male gametocyte population observed in the HDA1 KO in 820 is not replicated in the PbEGAM background line. Further analyses of these changes to asexual and sexual growth in HDA1 KO lines are shown in **Figure 3.8 (c)** to **(f).**

Initial observation of HDA1 KO lines by FACS, alongside monitoring of peripheral blood parasitaemia by Giemsa staining, revealed considerably attenuated asexual growth in these lines when compared to their wild-type background lines (**Figure 3.8 (c) and (e)**). Though experiments require replication (preferably with a clonal *P. berghei* line), initial comparison of asexual growth from 10 independent 820 experiments and single samples of both G1687 and G1688 HDA1 KO lines are shown in **Figure 3.8 (c).**

As a result of an inability to clone either the G1687 or G1688 KO lines, it is difficult to ascertain whether the growth seen at 4 dpi in both G1687 and G1688 (**Figure 3.8 (c)**) was a result of growth of the HDA1 KO line, or overgrowth of the background 820 asexual population. In either case, the lowered *hda1*$^{-}$ results indicate that the growth of a *P. berghei* ANKA line with this particular gene deletion is attenuated. However, to determine such a growth phenotype, a competitive growth assay with a *P. berghei* ANKA background line that constitutively expresses GFP throughout the entire life cycle should be carried out (detailed in section **2.1.20**). However, to undertake such an assay of asexual growth, a clonal HDA1 KO line would again be required, and no such line could be generated in the present study.

Despite a failure to successfully clone 820 *P. berghei* ANKA lines in which HDA1 was completely knocked out, transfection into the PbEGAM background line yielded slightly more promising results. Though cloning by limiting dilution was again attempted three times with no success, the lack of a visible wild-type 820 PCR fragment when checking integration of the HDA1 KO construct (**Figure 3.3**) warranted repeated investigation of the non-clonal G1579 line by FACS. The

results of 3 independent infections of PbEGAM in comparison to 3 independent infections of a PbEGAM hat1$^-$ line (G1579) are shown in **Figure 3.8 (e) and (f)**.

In the case of HDA1 KO in the PbEGAM background line, the total percentage parasitaemia at 4 dpi is not significantly different to that of the wild-type (unpaired, two-tailed t-test; P<0.05; n=3**)** (**Figure 3.8 (e)).** In **Figure 3.8** (**f**), total gametocytaemia (taking into account the entire GFP$^+$/RFP$^+$ population of immature and mature male and female gametocytes) at 4 dpi in G1579 are significantly lower than that of the wild-type line (P=0.0109; unpaired, two-tailed t-test). Female percentage gametocytaemia is significantly lower in G1579 when compared to the PbEGAM background line also (P=0.0415; unpaired, two-tailed t-test), with male percentage gametocytaemia also lowered significantly (P=0.0402; unpaired, two-tailed t-test).

However, without a clonal HDA1 KO in either the 820 or PbEGAM background lines, it is difficult to attribute these changes to a complete knockout of the HDA1 gene. However, with the observable attenuated growth phenotype after transfection (**Figures 3.8 (c) and (e)**), and an inability to produce a clonal line after seven independent attempts at cloning by limiting dilution, it should be suggested that complete KO of HDA1 is indeed affecting asexual replication in *P. berghei,* with a drop in gametocytaemia perhaps being attributable to this overall decrease in parasitaemia. To study this gene further, a HDA1 conditional knockdown (KD) line was generated using the auxin-inducible degron (AID) system, with the findings from these experiments detailed below.

3.3.8 Monitoring auxin degradation of epigenetic regulatory proteins by Western blot

All successfully-created *P. berghei* KD plasmids shown as **Supplemental Figures S1E-S1K** contained sequences for the expression of a Pyrimethamine resistance gene (*hdhfr*), a fluorescent protein (*mCherry* or *BFP*), and a hemagglutinin tag (*HA*). The main aim of these KD *P. berghei* lines was to assess the effect of putative epigenetic regulators on asexual growth and gametocytogenesis by conditional protein degradation. In previous studies in *P. berghei* (Philip and Waters, 2015), *P. falciparum* (Kreidenweiss, Hopkins and Mordmüller, 2013), and *Toxoplasma*

gondii (Brown, Long and Sibley, 2017), degradation of abundant cytoplasmic proteins (calcineurin, yellow fluorescent protein (YFP), and protein kinase G (PKG)) was assessed by a combination of fluorescence microscopy and Western blot analysis, and so this process was attempted in this study.

To begin with, Western blot analyses of each parasite line shown in **Figure 3.4** were carried out using whole protein lysates (extraction described as in section **2.2.2**), and later with separate nuclear, cytosolic, and insoluble fractions being run alongside one another (nuclear fractionation method described in section **2.2.3**). Protein detection on Western blots was carried out using both standard chemiluminescence reagents and Ponceau-staining. Samples from each line were run before and after the addition of auxin (3-indolacetic acid; Sigma-Aldrich) as detailed in section **2.1.11**.

In the present study, repeated attempts to conditionally degrade proteins of interest by auxin degradation failed. Instead, the fact that KD constructs contained a hemagglutinin (HA) tag, was exploited, and both fluorescence microscopy and co-immunoprecipitation (Co-IP) experiments were undertaken to determine the localisation of a tagged-protein within the parasite cell, and to discover proteins with which the putative epigenetic regulators interacted.

3.3.9 Organellar localisation of tagged epigenetic modifiers

With repeated failure to conditionally degrade putative epigenetic regulatory proteins by auxin degradation, HA-tagged proteins were subjected to fluorescence microscopy to determine whether they were localised to certain foci surrounding the parasite nucleus, as was the case with HP1, HDAC2/IPK1, and gametocyte development protein 1 (GDV1) in *P. falciparum* (Brancucci *et al.*, 2014; Coleman *et al.*, 2014; Filarsky *et al.*, 2018). In this study, *P. berghei* samples from asynchronous infections *in vivo* and synchronous mature schizonts *in vitro* were prepared for fluorescence microscopy as described in section **2.2.5.**

Taking all three proteins for which two tagged *P. berghei* ANKA lines existed (HDAC2/IPK1, HDA1, and 'HDAP'), and combining a primary rabbit anti-HA antibody (Sigma-Aldrich) with either an AlexaFluor® 488 or AlexaFluor® 594

secondary antibody, clear and reproducible results were seen for all three putative epigenetic modifiers at trophozoite and schizont asexual stages (**Figure 3.9).** As demonstrated in *P. falciparum* 3D7 (Coleman *et al.*, 2014), HDAC2/IPK1 was concentrated in foci at the nuclear periphery, a largely heterochromatic subcompartment of the parasite cell. These results were reproduced in both HDAC2/IPK1 *P. berghei* lines from *in vivo* asynchronous infections and *in vitro* synchronous schizonts (**Figure 3.9 (a) and (b)).**

With *P. berghei* histone deacetylases, HDA1 and HDAP, clear localisation of proteins to the nuclear subcompartment in the same manner as HDAC2/IPK1 was not observed. Under fluorescence microscopy, HDA1-tagged parasites from asynchronous *in vivo* infections showed no intracellular fluorescence, but rather, faint fluorescence at the cell membrane of trophozoites in each sample (**Figure 3.9 (d)).** Upon examination of HDA1-tagged synchronous schizonts, this faint membrane fluorescence was lost, with no striking fluorescence seen in any particular cell compartment **(Figure 3.9 (c)).** In the image shown in **Figure 3.9 (c)**, the HDA1-tagged schizont cell contains some faint fluorescence in one merozoite. This observation was made on more than one occasion but was not constant, and the significance of such a result could not be speculated upon at such an early stage in the characterisation of HDA1.

With regard to HDAP *P. berghei* lines ($G1517_{cl7}$), clearest images were produced when analysing the clonal G1517 line in which HDAP was tagged and integrated into the AID background line (**Figure 3.9 (e) and (f))**. In parasites from both asynchronous *in vivo* infections and synchronous *in vitro* schizont cultures, faint fluorescence was observed diffusely within the parasitophorous vacuole of trophozoites, and diffusely within the cytosol of merozoites (**Figure 3.9 (e) and (f)** respectively).

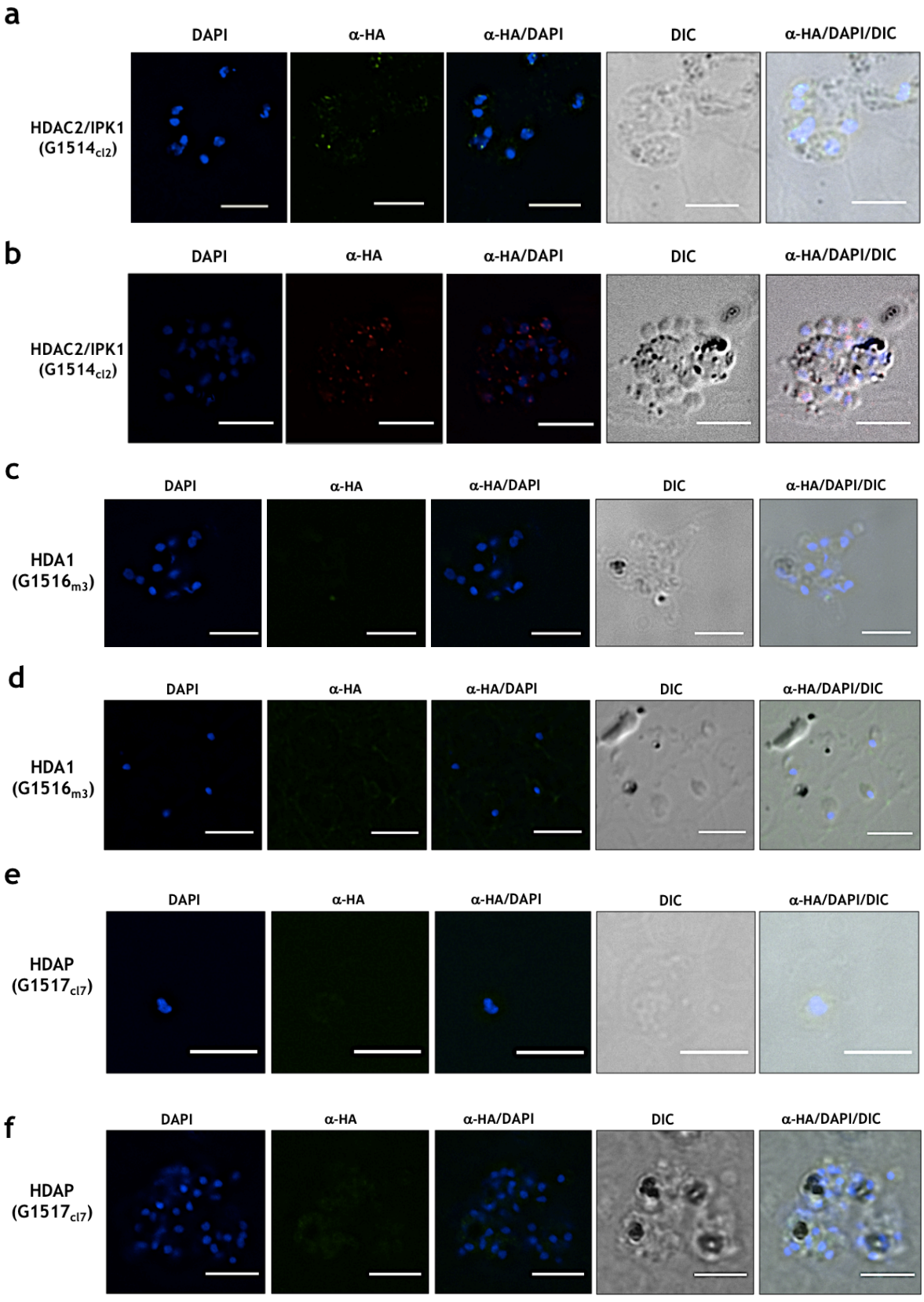


Figure 3.9: Fluorescence microscopy of putative epigenetic regulatory proteins HDAC2/IPK1, HDA1, and 'HDAP' in *P. berghei*. Clear representative images taken following fluorescence microscopy of three HA-tagged putative epigenetic proteins in *P. berghei* are shown in images **(a)** to **(f)** above. In (**a**), parasites from an asynchronous *in vivo* infection with clonal line G1514$_{cl2}$ (AID::HA-tagged HDAC2/IPK1) are shown after primary binding of an anti-HA antibody and

secondary incubation with an AlexaFluor® 488 anti-rabbit IgG secondary antibody. Parasites have also been incubated with a DAPI (4',6-diamidino-2-phenylindole) nuclear stain. In each image, a white scale bar indicates 5 μm in length. Single images with DAPI and anti-HA fluorescence are shown, followed by a merged image. A differential interference contrast (DIC) image of the same microscopic plane is then shown, followed by a merged image of all separate staining and DIC results. In this image, HDAC2/IPK1-staining is located in concentrated foci around the nuclear periphery. In **(b)**, the same HA-tagged protein is shown (HDAC2/IPK1), this time at a later stage of asexual development. In $G1514_{cl2}$, stained synchronous schizonts from an *in vitro* culture showed striking arrangement of multiple HDAC2/IPK1 foci around the nuclear periphery in all merozoites. In this image, an anti-HA primary antibody was bound to a secondary AlexaFluor® 594 anti-rabbit IgG antibody. In all cases, HDAC2/IPK1 appeared as concentrated foci around the parasite nuclear periphery.

Images **(c)** and (**d**) are representative examples of synchronous schizonts (**c**) and asynchronous trophozoites (**d**) from a *P. berghei* parasite line in which HDA1 has been AID::HA-tagged (line $G1516_{m3}$). In both cases, anti-HA antibody staining has been followed by a staining with a secondary AlexaFluor® 488 antibody. As in (**a**) and (**b**), all scale bars represent 5 μm. In schizonts (**c**), HA-tagged HDA1 appeared only very faintly in the cytosol of an occasional merozoite, and it remains to be seen whether this result was significant, or just background fluorescence. In (**d**), faint antibody-staining at the outer membranes of trophozoites can be observed.

Images **(e)** and (**f**) are representative examples of asynchronous trophozoites (**e**) and synchronous schizonts (**f**), taken from the same clonal *P. berghei* parasite line in which HDAP had been AID::HA-tagged and cloned by limiting dilution ($G1517_{cl7}$). In (**e**), only faint, if any, fluorescence can be detected within the trophozoite parasitophorous vacuole after anti-HA and AlexaFluor® 488 antibody staining. In parasites from an *in vitro* synchronous schizont culture (**f**), diffuse fluorescence can be seen in the cytosol of merozoites within the whole schizont, indicating a diffuse spread of HA-tagged HDAP within the cell cytosol.

As seen in **Figure 3.9**, fluorescence microscopy confirmed the localisation of HDAC2/IPK1 to the nuclear periphery of asexual parasite cells, as had been previously observed in *P. falciparum* 3D7 (Coleman *et al.*, 2014). Fluorescence imaging also determined that both HDA1 and HDAP in *P. berghei* ANKA do not localise in this manner, suggesting that these proteins are not associated with HP1 or GDV1 at asexual stages, certainly not in the same manner as HDAC2/IPK1 (Brancucci *et al.*, 2014; Coleman *et al.*, 2014; Filarsky *et al.*, 2018). The possible localisation of HDA1 to trophozoite cell membranes may be indicative of a role at this area of the parasite cell. The loss of this membrane localisation in the mature schizont (**Figure 3.9 (c)**) may also indicate that its function has ceased at this stage in the parasites development, with the protein moving to the cell membrane only at earlier asexual stages. Only further experiments will determine in future

the exact role of this putative epigenetic regulator. With regard to HDAP, the diffuse spread of this protein simply suggested that this protein is cytosolic at asexual parasite stages (**Figure 3.9 (e) and (f)).**

Though further examination is necessary, additional imaging of these AID::HA-tagged *P. berghei* lines in a timed fashion may lead to observation of patterns of protein movement within such life-cycle stages as rings, gametocytes or sporozoites. The successful tagging of these proteins, alongside HDAC1 (**Figure 3.4**), could also provide a useful starting point for studies into each epigenetic modifier's protein-protein interactions or enzyme activity by FRET (fluorescence resonance energy transfer) analysis (Bajar *et al.*, 2016).

3.3.10 Determining protein-protein interactions by co-immunoprecipitation and LC-MS/MS

With the failure of the conditional degradation system (section **3.3.8**), and the limited capacity to which fluorescence microscopy can add to our understanding of a protein's biology (section **3.3.9**), co-immunoprecipitation experiments were carried out to shed more light on the possible mechanism of action of putative epigenetic regulators HDA1, and HDAP. To maximise protein content, all co-immunoprecipitation (co-IP) experiments were carried out on purified, synchronous schizonts**,** with preparation for LC-MS/MS carried out as described in sections **2.2.6** and **2.2.8** to **2.2.10.** One exception came with the analysis of HDA1-interacting proteins, as co-IP was also carried out on gametocytes derived from *in vivo* sulfadiazine treatment, with gametocytes additionally purified by magnetic separation (method described in section **2.1.15**).

All co-IP results were compared to two background *P. berghei* ANKA high gametocyte producer (HP) line controls: one control co-IP undertaken with the standard procedure as described in section **2.2.6,** and one control co-IP in which an additional Protein G dynabead (ThermoFisher) pre-clear had been carried out. All co-IP LC-MS/MS results are listed in **Electronic Supplemental Material 1 (ES1).** These results will guide future co-IP experiments, along with additional experiments to characterise the possible interactions between epigenetic regulatory proteins examined in this study, and their interacting protein partners.

In the case of HDA1, promising KO results and the suggestion that this putative epigenetic regulator might be involved in segmentation of sexually-committed schizonts in *P. falciparum* (Poran *et al.*, 2017), led to investigation of both schizont and gametocyte stages of an HA-tagged HDA1 *P. berghei* ANKA line ($G1516_{m3}$) by co-IP. With this protein, shared unique results were seen at both schizont and gametocyte stages in comparison with control samples (**Table 3.3**). Two independent infections were carried out at each life-cycle stage, with all experiments carried out as described from sections **2.2.6** and **2.2.8** to **2.2.10,** but without the additional analysis of PTMs.

Proteins of Interest (HDA1 co-IP)					
P. berghei Gene ID	**Protein name**	**G1516m3 schizont (1)**	**G1516m3 schizont (2)**	**G1516m3 gametocyte (1)**	**G1516m3 gametocyte (2)**
PBANKA_1456100.1-p1	40S ribosomal protein S17, putative	✓	✓	✓	✓
PBANKA_1231000.1-p1	40S ribosomal protein S11, putative	✓	✓	✓	✓
PBANKA_0601900.1-p1	high mobility group protein B1, putative		✓	✓	✓
PBANKA_0210000.1-p1	vacuolar protein sorting-associated protein 51, putative		✓	✓	✓
PBANKA_0712900.1-p1	high mobility group protein B2		✓	✓	
PBANKA_1031100.1-p1	plastid replication-repair enzyme, putative		✓	✓	✓
PBANKA_1102000.1-p1	protein Mpv17, putative		✓	✓	✓
PBANKA_1340500.1-p1	conserved Plasmodium protein, unknown function		✓	✓	✓
PBANKA_1348600.1-p1	conserved Plasmodium protein, unknown function		✓	✓	✓
PBANKA_1131300.1-p1	conserved Plasmodium protein, unknown function			✓	✓
PBANKA_0803000.1-p1	origin recognition complex subunit 2, putative		✓	✓	✓
PBANKA_1106100.1-p1	conserved Plasmodium protein, unknown function			✓	✓
PBANKA_1432800.1-p1	multiple RNA-binding domain-containing protein 1, putative		✓	✓	
PBANKA_1108800.1-p1	leucine-rich repeat protein			✓	✓
PBANKA_0935900.1-p1	conserved Plasmodium protein, unknown function	✓		✓	
PBANKA_0315500.1-p1	conserved Plasmodium protein, unknown function, pseudogene	✓	✓		

Table 3.3: Unique protein identification from co-immunoprecipitation of HA-tagged HDA1 schizonts and gametocytes. This table lists 16 unique proteins that were identified in two or more co-IP experiments of HA-tagged HDA1 in schizonts or gametocytes. All *P. berghei* samples in these experiments came from the FACS-enriched $G1516_{m3}$ line (HA-tagged HDA1 in the AID background line). Sample type is listed at the head of each column and a check-mark indicates if the protein was found in this co-IP experiment.

When examining unique proteins co-immunoprecipitated with HA-tagged HDA1, 6 were conserved *Plasmodium* proteins and 2 were 40S ribosomal proteins that were also immunoprecipitated with HDAP (**Table 3.4**). Of the remaining 8 unique proteins, 2 were high mobility group proteins (B1 and B2), non-histone architectural chromosomal proteins that are highly conserved among eukaryotes and have been shown to bind DNA in *P. falciparum* (Kumar *et al.*, 2008; Reeves, 2015). High mobility group protein B2 (PBANKA_0712900) in particular has been implicated in transcriptional and translational control of sexual stage gene expression and oocyst formation (Gissot *et al.*, 2008). On the other hand, little is known at present about the role of vacuolar protein sorting-associated protein 51 (VSP51) in *Plasmodium*, though it may be involved in the transport of vesicles from endosomes to the trans-Golgi network (Jimenez-Ruiz *et al.*, 2016).

With regard to plastid replication-repair enzyme (PBANKA_1031100) and putative protein mpv17 (PBANKA_1102000), the former is associated with replication of the malarial plastid genome (Seow *et al.*, 2005), while the latter was found to be upregulated in *ap2-g2*-negative *P. berghei* 820 parasites and was suggested as a target of AP2-G2 (Yuda *et al.*, 2015). In fact, 7 of the 16 unique proteins from HDA1 co-IP experiments were identified as potential AP2-G2 targets by repeated ChIP-seq analysis: both 40S ribosomal proteins, S117 and S11, high-mobility group protein B2, putative mpv17, and three of the conserved *Plasmodium* proteins (PBANKA_1340500, PBANKA_1131300, and PBANKA_0315500) (Yuda *et al.*, 2015).

Of the remaining unique proteins identified, origin recognition complex subunit 2 (ORC2) may play a role in centrosome copy number control, and has been shown to bind tightly to heterochromatin protein 1 (HP1) in humans (Prasanth *et al.*, 2004). With regard to multiple RNA-binding domain-containing protein 1 (MRD1), no information has been gleaned so far about this protein's role in *Plasmodium*, though it has been marked as refractory to complete knockout in *P. berghei* (PlasmoGEM; 24/03/18), and so may be essential for asexual development. Conversely, a complete gene knockout of leucine-rich repeat protein 2 (LRR2) (encoded for by PBANKA_1108800) in *P. berghei* showed no difference in phenotype at asexual stages (PlasmoGEM; 24/03/18). However, LRR2 has been shown as having high nucleosome occupancy in ring stage *P. falciparum* 3D7 parasites, with low nucleosome occupancy at trophozoite and schizont stage

(Westenberger *et al.*, 2009). If a role for HDA1 in schizont segmentation is to be believed (Poran *et al.*, 2017), an association of HDA1 with LRR2 would be feasible, as low nucleosome occupancy at trophozoite and schizont stage, but not in ring-stage parasites, indicates that this protein could function at the same point in the *Plasmodium* life cycle as HDA1.

In the second group of co-IP experiments, carried out on a clonal line of G1517 in which the putative histone deacetylase, 'HDAP', has been HA-tagged, results listed in **Table 3.4** show unique proteins that were shared between two independent co-IP experiments. In the case of HDAP, only synchronous schizonts were examined to maximise protein content for co-IP samples.

Proteins of Interest (HDA, putative (HDAP) co-IP) G1517cl7	
P. berghei Gene ID	**Protein name**
PBANKA_1231000.1-p1	40S ribosomal protein S11, putative
PBANKA_1456100.1-p1	40S ribosomal protein S17, putative
PBANKA_0831000.1-p1	merozoite surface protein 1
PBANKA_1243400.1-p1	nucleolar preribosomal GTPase, putative
PBANKA_1341700.1-p1	conserved Plasmodium protein, unknown function
PBANKA_1135200.1-p1	GTP-binding protein, putative
PBANKA_0601200.1-p1	dynein heavy chain, putative
PBANKA_1129100.1-p1	DNA polymerase epsilon catalytic subunit A, putative
PBANKA_1200800.1-p1	RNA-binding protein, putative
PBANKA_1356800.1-p1	conserved Plasmodium protein, unknown function

Table 3.4: Unique protein identification from co-immunoprecipitation of HA-tagged HDAP from synchronous schizonts. In the table above, unique proteins are listed that were shared between two independent co-IP experiments of line $G1517_{cl7}$ schizonts. As was the case with HDA1, the top two unique proteins identified were 40S ribosomal proteins S17 and S11. All 10 proteins listed were shared between both co-IP experiments. Complete lists of proteins pulled down and their Mascot (Matrix Science) scores are listed in **Electronic Supplemental Material 1 (ES1)**.

As noted previously, both 40S ribosomal proteins S17 and S11 were pulled down from all HDA1 and HDAP co-IPs (**Electronic Supplemental Material 1 (ES1)**). Taking into consideration that these proteins may be contaminants, it is also possible that they both interact with HDA1 and HDAP. Both of these ribosomal proteins were also suggested as possible targets of AP2-G2 (Yuda *et al.*, 2015).

With this putative epigenetic regulator however, co-IP experiments pulled down merozoite surface protein 1 (MSP1), an early protein of merozoite invasion into host erythrocytes (Kadekoppala and Holder, 2010), and a dynein heavy chain protein (encoded for by PBANKA_0601200), the complete knockout of which resulted in a slow growth phenotype in asexual stages (PlasmoGEM; 24/03/18). In addition, genes transcribing the essential proteins, nucleolar preribosomal GTPase (PBANKA_1243400) and GTP-binding protein (PBANKA_1135200) are expressed to the greatest degree in 4 h ring-stage parasites and 16 h trophozoites (Otto *et al.*, 2014), further indicating that the function of HDAP is at early asexual stages in *Plasmodium* development.

The remaining two unique proteins of which there is a putative function, DNA polymerase epsilon catalytic subunit A (PBANKA_1129100) and RNA-binding protein (PBANKA_1200800) are expressed in both asexual and sexual stages, with the DNA polymerase epsilon subunit A (encoded for by PBANKA_1129100) particularly highly expressed in male gametocytes (Otto *et al.*, 2014; Yeoh *et al.*, 2017). Taken together, proteins co-immunoprecipitated with HDAP primarily indicate that this protein function in early asexual development, though its role may extend throughout the parasite life cycle, rendering it essential to parasite development.

3.4 Discussion

The key findings from this chapter are:

- Histone acetyltransferase 1 (HAT1; PBANKA_0718400) is dispensable for *P. berghei* asexual, intraerythrocytic growth.
- Complete HAT1 KO in the *P. berghei* 820 line resulted in a statistically significant decrease in mature female gametocytes ($P = 0.048$; $n = 5$).
- Histone deacetylase 1 (HDA1; PBANKA_1335400) knockout was possible in both *P. berghei* 820 and PbEGAM lines, but these lines were refractory to cloning by limiting dilution.
- Complete HDA1 knockout in a PbEGAM line resulted in a significant decrease in overall gametocytaemia ($P = 0.0109$; $n = 3$); a decrease in female

gametocytes from 2.93% ± 1.09 % to 1.03% ± 0.22% of total parasitaemia (P = 0.0415), and a decrease in male gametocytes from 2.07% ± 0.38% to 1.21% ± 0.32% of total parasitaemia (P = 0.0402).

- Four putative *P. berghei* histone deacetylases were successfully AID::HA-tagged: histone deacetylase 1 (HDA1; PBANKA_1335400), a second so-called 'histone deacetylase 1' (HDAC1; PBANKA_0826500), histone deacetylase 2/IPK1 combined gene (HDAC2/IPK1; PBANKA_1206200), and histone deacetylase, putative ('HDAP'; PBANKA_1106200).
- Auxin degradation of AID::HA-tagged lines was not possible for any of the proteins examined.
- HDAC2/IPK1 showed nuclear localisation in *P. berghei* asexual-stage parasites, while HDA1 and HDAP showed no discernible nuclear localisation.
- HDA1 co-immunoprecipitated with 16 unique proteins at mature schizont and gametocyte stages compared to control samples.
- HDAP co-immunoprecipitated with 10 unique proteins at mature schizont stage.

3.4.1 In context: epigenetic regulatory proteins and sexual development

It has recently been shown unequivocally that the AP2-G transcriptional regulator, responsible for *Plasmodium* gametocytogenesis, is controlled epigenetically, by the binding of heterochromatin protein 1 (HP1) to the promoter and 3' end of the *ap2-g* gene (Fraschka *et al.*, 2018). In addition, in *P. falciparum*, HP1 binding to *pfap2-g* is antagonised by the action of a protein known as gametocyte development protein 1 (GDV1), possibly forming a regulatory complex with chromodomain-helicase-DNA-binding protein 1 (CHD1) and a protein of unknown function encoded for by gene PF3D7_1451200 (homologous to PBANKA_1314900) (Filarsky *et al.*, 2018).

In a separate study in which single-cell transcriptomic analysis of a *P. falciparum* line producing *ap2-g,* in comparison to a line in which *ap2-g* was conditionally degraded, it was shown that four putative chromatin-remodelling proteins were expressed alongside *pfap2-g*, with all upstream regions of these genes containing an average of 2.6 AP2-G-binding motifs. These proteins were two SNF2 family

helicases (ISWI and SNF2L), the putative histone deacetylase, HDA1, and a putative histone lysine-specific demethylase, LSD2 (Poran *et al.*, 2017). Despite RNA co-expression, co-immunoprecipitation of tagged-HDA1 in *P. berghei* during this study did not reveal interactions between HDA1 and AP2-G, ISWI, SNF2L, or LSD2 (**Table 3.3).** This study did reveal however, that HDA1 did in fact co-immunoprecipitate with 7 potential targets of the AP2-G2 transcriptional repressor, a protein that represses genes involved in asexual reproduction, paving the way for gametocytogenesis (section **3.3.10)** (Yuda *et al.*, 2015). Therefore, if HDA1 is interacting with genes that are targeted for repression upon sexual commitment (and more proteins besides), HDA1 may be a crucial facilitator of asexual development. HDA1 involvement in preparing asexually-committed cells for optimum proliferation would in some way explain the phenotype of attenuated asexual growth seen in **Figure 3.8** upon complete KO of the gene (**Figure 3.10).**

3.4.2 The potential role of histone deacetylase 1 (HDA1) in *Plasmodium* spp.

In a recent single-cell transcriptome analysis of wild-type *P. falciparum* NF54 (AP2-G^{+}), an AP2-G-DD line (a *P. falciparum* NF54 parasite line in which removal of the stabilising Shield1 ligand resulted in AP2-G depletion), and an AP2-G-DD line in which Shield1 had been removed (AP2-G^{-}), histone deacetylase 1 (HDA1) was identified as a gene that was upregulated in AP2-G^{+} parasites relative to the AP2-G-depleted line (Poran *et al.*, 2017). In this study, an increase in AP2-G transcript abundance during schizont segmentation resulted in a sharp increase in transcription of HDA1, the histone demethylase LSD2, and the AP2 gene, PF3D7_1139300 (homologous to *P. berghei* ANKA gene, PBANKA_0909600; not yet named (Modrzynska *et al.*, 2017)). The authors of this study propose that sharp upregulation of these three genes upon *pfap2-g* expression at a time-point just prior to schizont segmentation and merozoite egress from the erythrocyte (determined by *P. falciparum* calcium-dependent protein kinase 5 (*Pf*CDPK5) expression (Dvorin *et al.*, 2010)) prepares sexually-committed *Plasmodium* schizonts for segmentation and gametocyte development in the subsequent cell cycle, and that this gene may play a role in mating-type determination (Poran *et al.*, 2017).

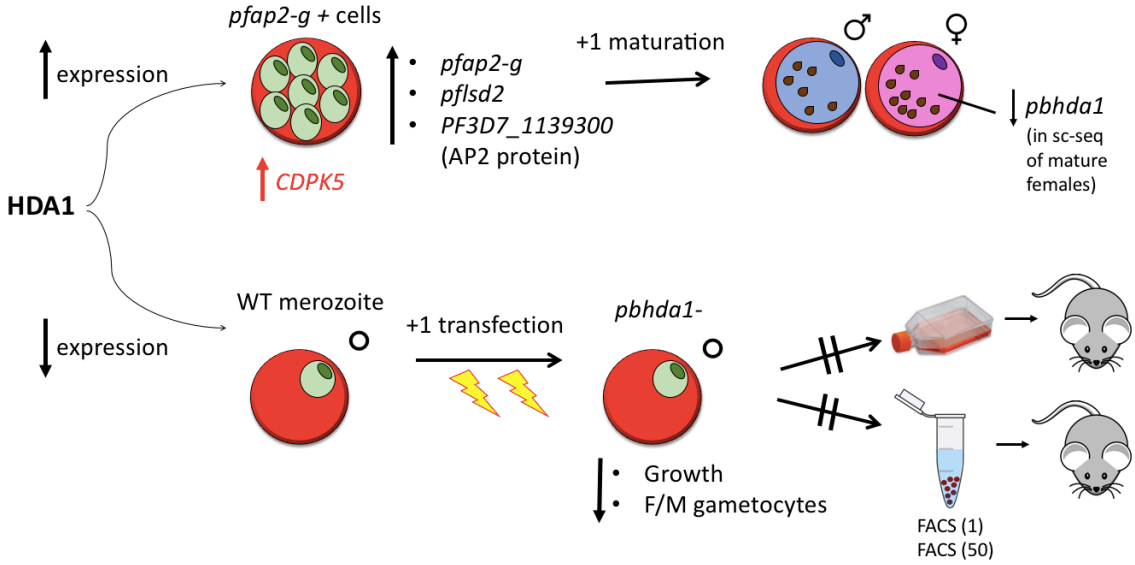
In line with the findings of this recent study (Poran *et al.*, 2017), HDA1 depletion in *P. berghei* ANKA carried out in the present study (section **3.3.6**) did result in a significant decrease in total gametocyte production (P = 0.0109) (**Figure 3.8**), though more striking was the asexual growth deficit seen in all HDA1 KO lines, and the complete inability to clone any of these KO lines by limiting dilution, despite several attempts in three successfully transfected lines (G1579, G1687 and G1688) (**Figure 3.3**). One could propose, based on the present experimental findings, that HDA1 is not only necessary for segmentation of schizonts in sexually-committed *Plasmodium* parasites, but that HDA1 is necessary in both asexual and sexually-committed *Plasmodium* parasites for schizont segmentation or preparation of the schizont-enclosed merozoites for replication upon egress from the erythrocyte. The complete inability to clone HDA1 KO lines could be a product of inability of synchronous schizonts to segment, resulting in loss of asexual replication. However, a defect in *P. berghei hda1*$^{-}$ merozoites to complete invasion of erythrocytes is also possible, as is a defect in the growth of the ring-stage parasite or trophozoite. Although HDA1 transcript abundance steadily increased throughout schizogony alongside *pfap2-g* expression, protein abundance at each point was not examined during the Poran *et al.* study and so it remains to be seen whether HDA1 is in fact translated and exerting its downstream effects (Poran *et al.*, 2017).

In keeping with the suggestion that HDA1 is required during schizont segmentation in *Plasmodium* spp. (Poran *et al.*, 2017), a recent comparative analysis of male and female gametocytes to asexual stage parasites in *P. berghei* ANKA listed HDA1 as being downregulated in female gametocytes relative to asexual parasites (Yeoh *et al.*, 2017). If the hypothesis that HDA1 is involved in schizont segmentation is true, it would be correct that mature female gametocytes would not be expressing HDA1 to the same extent as asexual-stage parasites that may be preparing for schizogony. However, a down-regulation of HDA1 expression in female gametocytes may also indicate that the primary role for HDA1 is to regulate or maintain correct DNA expression throughout asexual replication in *Plasmodium*.

With regard to previously-characterised histone deacetylases in *Plasmodium* spp., depletion of *Pf*SIR2A and *Pf*SIR2B resulted in de-repression of *var* genes and increased telomeric length in *P. falciparum* (Tonkin *et al.*, 2009), while

HDAC2/IPK1 was determined as being essential for asexual replication and repression of commitment to sexual development in *P. falciparum*, with conditional knockdown resulting in both *var* gene and *ap2-g* dysregulation (Coleman *et al.*, 2014). In the present study, complete KO of HDAC2/IPK1 and HDAC1 in *P. berghei* ANKA could not be carried out, with HDA1 KO lines showing severely attenuated asexual growth after transfection. Cloning of HDA1 KO lines by limiting dilution after synchronisation of mature schizonts in culture, or generation of an isogenic line by fluorescence-activated cell sorting (FACS) of single cells, was not possible. Taken together, these findings show that, in general, histone deacetylases are important regulators of both asexual and sexual development in *Plasmodium* spp., and that HDA1 is potentially involved in regulating asexual parasite growth, with transcription of the *hda1* gene increasing steadily until a point just prior to merozoite egress from the schizont, whether the schizont be committed to asexual or sexual replication (Poran *et al.*, 2017) (**Figure 3.10**).

a



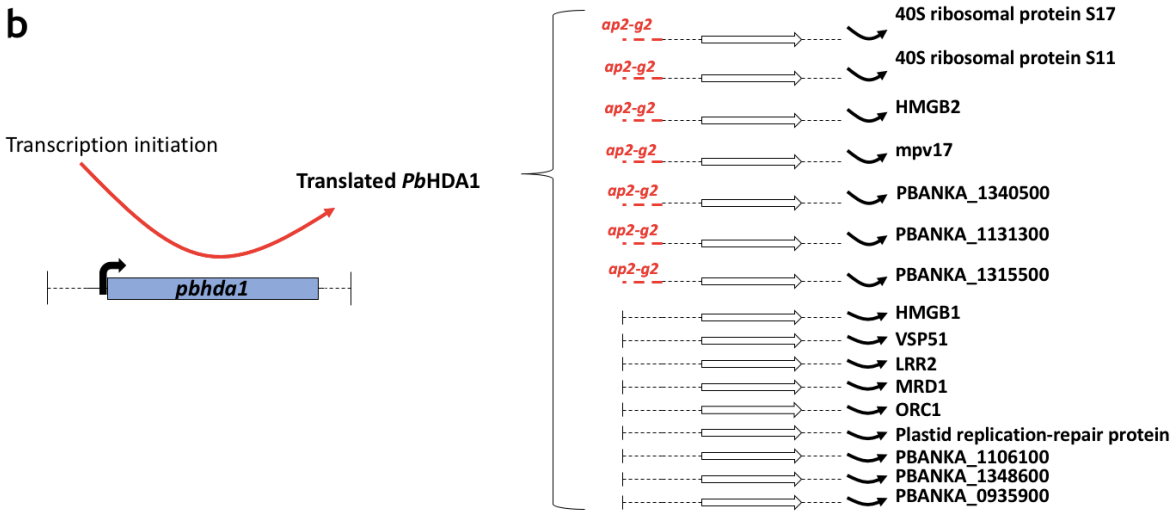


Figure 3.10: What we know about histone deacetylase 1 (HDA1) so far. In this figure, **(a)** depicts our current knowledge of the life-cycle stages of *Plasmodium* spp. at which *hda1* gene transcription occurs in greatest abundance, the genes that appear to be transcribed alongside *hda1* in *P. falciparum*, and the phenotypic traits that occurred upon repeated attempts to knockout *hda1* in *P. berghei*. To begin on the left, and moving right along the diagram, HDA1 transcript expression appears to co-inside with transcription of *ap2-g, lsd2,* and *Pf3D7_1139300* (an AP2 family protein) in *P. falciparum* (cells that are free to express AP2-G (*pfap2-g* + cells)). The abundance of these transcripts increases alongside increasing expression of *Pf*CDPK5, a marker of schizont egress (Poran *et al.*, 2017). After sexually-committed merozoites have been released and have become mature female and male gametocytes, mature female gametocytes have been shown to have a downregulation of *hda1* transcripts, this study being carried out using single-cell sequencing (sc-seq) of *P. berghei* parasites (Yeoh *et al.*, 2017). Below this is shown what occurred in the present study with repeated attempts to KO HDA1 in *P. berghei* ANKA and clone the resulting parasite lines. Following transcription of a linearised *hda1* KO plasmid into a WT asexual (**o**) merozoite by electroporation, growth and gametocytogenesis were both reduced (taking the PbEGAM KO repeat results into consideration (section **3.3.7**)). Attempts to clone these *hda1*$^{-}$ lines by both synchronisation of parasite cultures and limiting dilution, or by fluorescence-activated cell sorting (FACS) of 1 or 50 single asexual parasites, failed to yield any resulting clonal or isogenic lines (section **3.3.7**). **(b)** shows *pbhda1*$^{-}$ transcription, translation to the *Pb*HDA1 protein, and the interacting factors with which tagged-HDA1 co-immunoprecipitated in the present study. The genes are depicted by white arrows, with the 7 AP2-G2 motif-associated proteins indicated.

Taking all of the information contained in **Figure 3.10 (a)** into consideration, it would appear that HDA1 expression occurs just prior to schizont egress. These results also correspond to *P. berghei* ANKA transcriptome data which show greatest expression of *pbhda1* at the 4-hour ring stage (Otto *et al.*, 2014) and greater abundance in asexual parasites when compared with male and female sexual-stage parasites (Yeoh *et al.*, 2017). The defect in asexual growth and inability to produce a clonal line in *P. berghei* upon complete gene KO also

indicates an asexual role for HDA1. In **Figure 3.10 (b),** HDA1 co-immunoprecipitating factors are shown, 7 of which are potential targets of the AP2-G2 transcriptional repressor (Yuda *et al.*, 2015). If the repression of these genes is a requirement for sexual development, so too would their expression indicate asexual growth. Again, these findings hint at a function of HDA1 in asexual development in *P. berghei*. In addition, immunofluorescence results did not show nuclear localisation of HDA1 at mixed asexual or mature schizont stage (**Figure 3.9 (c)** and **(d))**. If HDA1 function is required at an early stage in asexual development, such as in 4-hour ring-stage parasites, examination of this parasite stage by immunofluorescence microscopy may be of greater use in revealing any localisation of the HDA1 protein.

Proteins that co-immunoprecipitated with HDA1 might also provide insight into HDA1 function. Of the 16 HDA1 co-immunoprecipitated proteins, 10 have a known or putative role. Two proteins that are co-immunoprecipitated with both HDA1 (**Table 3.3**) and HDAP (**Table 3.4**) are the putative 40S ribosomal subunit proteins S17 (RPS17; PBANKA_1456100) and S11 (RPS11; PBANKA_1231000), both of which are transcribed in greatest abundance in intraerythrocytic asexual 4-hour ring-stage parasites compared to later asexual stages or gametocytes (Otto *et al.*, 2014; Yeoh *et al.*, 2017). Of note also is the fact that in a *P. berghei* ANKA KO line in which the DDX6 class RNA helicase, DOZI (development of zygote inhibited), has been knocked out (*pbdozi*$^{-}$), RPS11 transcript expression is significantly upregulated compared to a *P. berghei* WT line (Mair *et al.*, 2006). With DOZI's role as a translational repressor of mRNAs prior to zygote formation in the mosquito midgut, these results suggest that RPS11 mRNA is translationally repressed by DOZI prior to sexual development. Remarkably, the RPS11 gene was also noted as having an upstream AP2-G2 binding site (Yuda *et al.*, 2015). Together, an AP2-G2 binding site at the RPS11 gene, and association of DOZI with RPS11 mRNA, suggest that this ribosomal protein must be repressed prior to sexual development in *P. berghei,* and as such, it is a protein heavily associated with asexual replication, i.e. this protein must be repressed to enable sexual development to occur.

Along with RPS17 and RPS11, HDA1 co-immunoprecipitated with HMGB1 (PBANKA_0601900) and HMGB2 (PBANKA_0712900), the first protein of which is

transcribed predominantly in 22 h schizonts (Otto *et al.*, 2014) and mature male gametocytes (Yeoh *et al.*, 2017), and the second of which is transcribed chiefly in mature female gametocytes (Otto *et al.*, 2014; Yeoh *et al.*, 2017) and is both a regulator of oocyst development, and a contributor to the pathogenesis of cerebral malaria in mouse models (Gissot *et al.*, 2008; Briquet *et al.*, 2015). Unlike RPS11 and RPS17, co-immunoprecipitation of HDA1 with these factors suggests a role for this epigenetic regulatory protein in modulation of gametocytogenesis. Perhaps HDA1 transcription at the point of merozoite egress from the schizont and its interaction with proteins that are instrumental to the choice between asexual and sexual development in *P. berghei*, suggests that this protein is involved in sexual commitment or sex determination.

Of the remaining HDA1 co-immunoprecipitating proteins with a putative or known function; mpv17, ORC1, MRD1, LRR2, VSP51, and plastid replication-repair enzyme (PREX), 2 are transcribed predominantly in male gametocytes, ORC2 (PBANKA_0803000) and LRR2 (PBANKA_1108800). In eukaryotes, and in *P. falciparum,* ORC2 is a DNA replication initiation protein (Sharma *et al.*, 2018). With regard to LRR2, relatively little is known, however, a difference appears to be present between the expression of LRR2 in human-infective *P. falciparum* and the *P. berghei* ANKA rodent malaria parasite. Using real-time quantitative PCR throughout the *P. falciparum* asexual intraerythrocytic development, LRR2 was expressed with greatest relative abundance in late trophozoites (Daher, Pierce and Khalife, 2007). In *P. berghei* ANKA, with RNA-sequencing analysis and the addition of transcriptome data from sexual stage parasites (both male and female gametocytes), LRR2 expression was seen predominantly in 22 h schizonts and male gametocytes (Otto *et al.*, 2014; Yeoh *et al.*, 2017). Although late trophozoite and 22 h schizonts are relatively close within the *Plasmodium* spp. asexual cycle, it remains to be seen whether LRR2 is transcribed, translated, or otherwise exerting an effect in *P. falciparum* male gametocytes.

Finally, transcription of mvp17, MRD1, VSP51, and PREX in *P. berghei* ANKA was greatest throughout the asexual cycle; at 4 h ring-stage for both mvp17 and MRD1, and at 16 h trophozoite stage for both VSP51 and PREX (Otto *et al.*, 2014; Yeoh *et al.*, 2017).

Taking all HDA1 data from this study and *Plasmodium* spp. literature at present, it appears that this histone deacetylase is necessary for optimal asexual and sexual parasite development, though the precise epigenetic mechanism by which it exerts its effects are still unknown. From examination of all co-immunoprecipitated proteins and their roles in the *Plasmodium* life cycle, it is likely that the downstream effects of HDA1 are far-reaching, and an alternative conditional protein or gene degradation system is required to pinpoint the exact point at which HDA1 elicits its effects, and what these effects are.

3.4.3 The role of histone acetyltransferase 1 (HAT1) in *Plasmodium* spp.

With regard to HAT1 experiments carried out in this study, all HAT1 KO *P. berghei* ANKA clonal lines showed attenuated asexual growth in both the 820 and PbEGAM background lines, though this decrease was not statistically significant in any experiment despite repeated independent infections (**Figure 3.7).** *In vivo* competition growth assays comparing mutant and WT parasites are required to truly quantify the apparent growth defect of the mutant. With regard to gametocytaemia in HAT1 KO lines, 2 out of 3 KO lines showed a decreased gametocytaemia (in both males and females) ($G1529_{cl6}$ and $G1693_{cl6}$). In a HAT1 KO in an 820 background *P. berghei* ANKA line, the decreased percentage of female gametocytes was significant (P = 0.0048; unpaired, two-tailed t-test) when compared to the WT 820 line (**Figure 3.7**).

When comparing the results of this study to a comparative transcriptomic analysis of male and female *P. berghei* ANKA gametocytes (Yeoh *et al.*, 2017), both mature male and female gametocytes showed decreased HAT1 transcription (log fold changes of -2.06 and -0.63 in males and female gametocytes respectively) when compared to asexual-stage parasites (Yeoh *et al.*, 2017). When comparing data from the AP2-G-DD study (Poran *et al.*, 2017), HAT1 (encoded by PF3D7_0416400 in *P. falciparum*) was not flagged as being differentially expressed between AP2-G^{+} and AP2-G-depleted *P. falciparum* NF54 lines.

With these observations, it may be that the decreased gametocytaemia seen in HAT1 KO lines resulted from the overall attenuated asexual growth phenotype, and not as a direct effect of HAT1 on gametocyte commitment or maturation. In

a study of the *P. falciparum Pf*MYST histone acetyltransferase, disruption of the *pfmyst* gene was not possible, suggesting that this histone acetyltransferase was essential for intraerythrocytic parasite growth (Miao *et al.*, 2010). Similarly, histone H3 acetylation by the *Pf*GCN5 histone acetyltransferase is essential for gene expression and cell cycle progression in *P. falciparum* (Cui *et al.*, 2007).

Though not as detrimental to *Plasmodium* spp. parasite growth as a HDA1 KO (section **3.3.7**), HAT1 KO does appear to result in attenuated asexual growth in *P. berghei* (**Figure 3.7**). With regard to the downregulation of HAT1 in both gametocyte sexes upon comparative transcriptomic analysis (Yeoh *et al.*, 2017), it may also be possible that HAT1 function is disproportionately involved in commitment to female gametocytogenesis at some point during asexual growth in *Plasmodium*. HAT1 expression as an important component of female gametocyte commitment would explain a greater decrease in female gametocytaemia in a HAT1 KO line when compared to asexual parasites and male gametocyte numbers. Involvement in commitment to gametocytogenesis at this early stage would also explain the decreased presence of HAT1 transcripts observed in mature gametocytes (Yeoh *et al.*, 2017) and an absence of changes to HAT1 expression in response to AP2-G depletion (Poran *et al.*, 2017) (**Figure 3.11**).

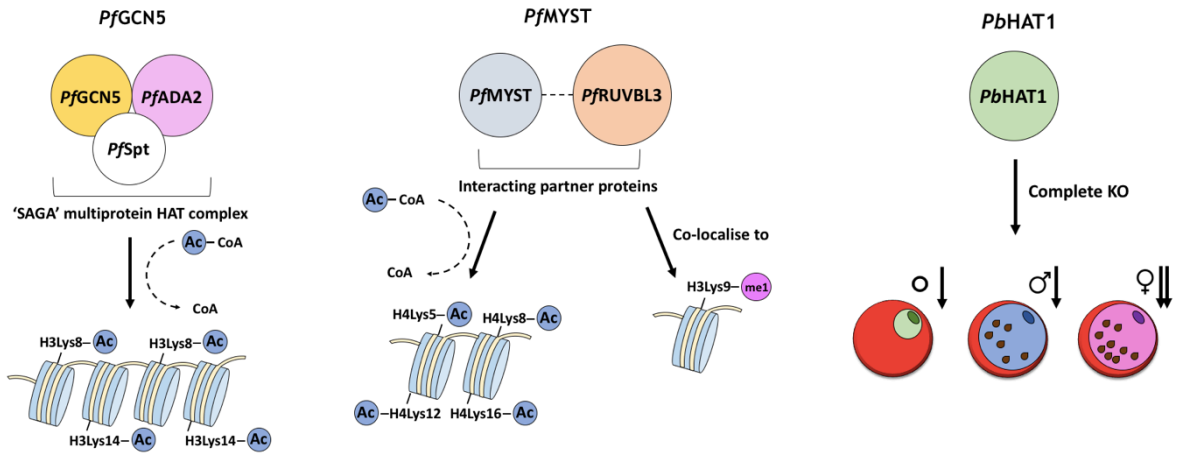


Figure 3.11: Current knowledge of *Plasmodium* spp. histone acetyltransferases. This image summarises known *Plasmodium* spp. histone acetyltransferase molecular mechanisms at the present time. On the left, the first *Plasmodium* spp. histone acetyltransferase to be functionally characterised, the *P. falciparum* GCN5 (general control of amino acid synthesis 5) histone acetyltransferase (*Pf*GCN5), is shown in complex with *P. falciparum* adenosine deaminase 2 (*Pf*ADA2) and a *P. falciparum* Spt (suppressor of Ty) homologue. Together, these proteins form an

'Spt-ADA-GCN5-acetylatransferase' ('SAGA') complex that, in *P. falciparum,* is currently known to acetylate histone 3 (H3), lysines 8 and 14 (Fan, An and Cui, 2004), though the SAGA complex has been shown to mediate further chromatin modification in *Drosophila* (Li *et al.*, 2017). In the centre of the image, the *P. falciparum* MYST (MOZ, Ybf1/Sas3, Sas2, and Tip60) (*Pf*MYST) histone acetyltransferase is shown with its known interacting partner in *P. falciparum,* the RuvB-like protein 3 (*Pf*RUVBL3). Alone, recombinant *Pf*MYST showed a predilection for acetylation of histone 4 (H4), lysines 5, 8, 12, and 16 (Miao *et al.*, 2010). With its interacting partner, *Pf*RUVBL3, *Pf*MYST appears to co-localise with H3K9me1 during the ring stages of intraerythrocytic *P. falciparum* development (Sen *et al.*, 2018). To the right of the image, the findings from the present study into *P. berghei* histone acetyltransferase 1 (HAT1) function are shown. In this study, though no molecular mechanism is known, a complete KO of *Pb*HAT1 resulted in a reduction in asexual growth, male gametocytaemia, and a statistically significant decrease in the production of mature female gametocytes (section **3.3.6**).

As summarised in **Figure 3.11,** three histone acetyltransferase domain-containing proteins in *Plasmodium*, GCN5, MYST, and HAT1, have been characterised in *P. falciparum* or *P. berghei*, out of a total of 11 proteins with putative N-acetyltransferase or histone acetyltransferase domains (listed in **Supplemental Table S1).** Of these three proteins, only HAT1 contains more than one histone acetyltransferase domain (it has 2 HAT domains), and only HAT1 is dispensable during asexual growth (**Figure 3.7**) (Bushell *et al.*, 2017). In *P. berghei* ANKA, the GCN5 acetyltransferase (PBANKA_0707300) is essential for asexual growth of the parasite, and is transcribed throughout asexual stages of the life-cycle (Otto *et al.*, 2014; Bushell *et al.*, 2017; Yeoh *et al.*, 2017). The *P. berghei* MYST protein (encoded for by PBANKA_0929500), is similarly transcribed predominantly throughout the asexual parasite life-cycle, being downregulated in mature male and female gametocytes comparatively (Otto *et al.*, 2014; Yeoh *et al.*, 2017). It is also an essential protein for asexual intraerythrocytic growth (Bushell *et al.*, 2017).

In the case of *P. berghei* HAT1, this enzyme was shown to be transcribed predominantly at 16 h trophozoite stage, decreasing in abundance as the cycle progressed towards schizogony, and downregulated, though still present, in gametocytes (Otto *et al.*, 2014). In keeping with the phenotype seen upon complete KO (section **3.3.6**), comparative transcriptome analysis revealed that HAT1 is expressed to a greater extent in both mature female gametocytes and mixed asexual stages (Yeoh *et al.*, 2017). With these findings, it may be that this

HAT1 acetyltransferase is disproportionately involved in asexual growth and female gametocytogenesis when compared with male gametocytogenesis. In keeping with the findings of the present study, a large screening of essential genes in *P. berghei* ANKA also noted a slow growth rate when *Pbhat1* was knocked out (66% growth rate compared to WT; 95% CI: 0.38-0.93) and a disruption at the gametocyte, oocyst, ookinete, and liver stages (Bushell *et al.*, 2017). Remarkably, of all 11 putative N-acetyltransferases or histone acetyltransferases currently identified in *P. berghei*; the former three proteins plus PBANKA_1201700, PBANKA_0204100, PBANKA_1225200, PBANKA_1127600, PBANKA_0611800, PBANKA_1338500, PBANKA_1442500, and PBANKA_0504900, only one more acetyltransferase domain-containing protein, the GNAT family N-acetyltransferase encoded by PBANKA_1225200, has been shown to be dispensable for asexual growth (Bushell *et al.*, 2017). And similarly to *Pb*HAT1, this GNAT acetyltransferase also showed attenuated growth in asexual stages (96% growth compared to WT; 95% CI: 0.86-1.06), and a disruption of sexual development at gametocyte and mosquito sexual stages.

To summarise, further experimentation is needed to shed light on the exact mechanism of action of HAT1 in *Plasmodium*. ChIP and co-IP experiments in future could elucidate the downstream effects of this acetyltransferase, the identity of the epigenetic modification(s) that HAT1 meditates, and interaction partners of HAT1. Furthermore, with HAT1 KO clonal lines, competitive growth assays and comparative transcriptomic analyses are also possible, as are experiments in transmission of HAT1 KO lines through mosquitoes. Initial experiments in transmission of the $G1529_{cl6}$ *P. berghei* ANKA line have shown a decrease in ookinete conversion compared to a WT 820 line (personal correspondence with Dr Scott Millar, Wellcome Centre for Integrative Parasitology), which again suggests that HAT1 KO may have a detrimental effect on sexual-stage growth in *P. berghei*. A second HAT1 KO in the 820 background line was also successfully generated by transfection (G1689; **Figure 3.2**), which may prove to be a useful resource for future experimentation. When cloned, the G1689 clonal line could confirm the female gametocyte-specific phenotype observed upon $G1529_{cl6}$ examination (**Figure 3.7**).

In conclusion, both HDA1 KO and HAT1 KO in *P. berghei* ANKA background lines proved to be incredibly useful tools in the elucidation of the function of these putative epigenetic modifiers. With the ability to successfully clone HAT1 KO lines, histone extraction and PTM analysis by LC-MS/MS could be used to identify HAT1-mediated acetyl marks on histones in future. These KO studies were also able to confirm the essentiality of previously screened epigenetic modifiers by *Plasmo*GEM (**Table 3.1**), with the HDA1 KO providing compelling evidence for the role of this protein in early development of the parasite, just prior to, or at the point of merozoite egress from the schizont, in both asexually- and sexually-committed cells (Poran *et al.*, 2017). Future studies using these *P. berghei* ANKA lines may contribute significantly to our knowledge of both early intraerythrocytic development and gametocyte commitment in the malaria parasite.

3.4.4 Localisation of putative epigenetic modifiers in *P. berghei* ANKA

Following failure to conditionally degrade the putative epigenetic regulators of interest, fluorescence microscopy was used to determine any organellar localisation of successfully tagged proteins from both asynchronous infections *in vivo* and synchronous mature schizonts *in vitro* (section **3.3.9**). Using anti-HA primary antibodies with two different secondary antibodies demonstrated that HDAC2/IPK1 was indeed concentrated at foci adjacent to the nucleus, as had previously been observed in *P. falciparum* **(Figure 3.9 (a)** and **(b))** (Coleman *et al.*, 2014). However, neither HDA1 or HDAP showed the same nuclear localisation in *P. berghei* asexual-stage cells. HDA1 appeared to faintly localise to the outer trophozoite membrane; whereas HDAP was located diffusely throughout the parasite cytoplasm at both trophozoite and schizont stage (**Figure 3.9**).

It should be mentioned that in this study, live imaging of parasites in a timed manner throughout the entire parasite life-cycle was not carried out. However, taking into account both KO and fluorescence microscopy results upon tagging with HA and mCherry (section **3.9**), live imaging of mCherry fluorescence from HDA1-tagged parasites in the AID background line (G1516) may provide powerful results in determining if HDA1 does in fact function at the point of schizont segmentation, or during early ring-stage development of the asexual parasite (Poran *et al.*, 2017). It may also be of particular benefit to future studies to

generate conditional protein or gene knockdown lines using a different molecular system from auxin degradation, such as the rapamycin-inducible knocksideways (KS) system in *P. berghei* (Hughes and Waters, 2017). Should this conditional displacement system prove successful, the role of HDA1 in *P. berghei* ANKA could be determined in greater detail and the presence of a HDA1::GFP-containing line could complement previous imaging results.

3.4.5 Co-immunoprecipitation findings: putative interacting factors and issues

In co-IP experiments with tagged HDA1 and HDAP, unique proteins identified from LC-MS/MS of independent biological samples provided consistent results, though in all cases, the 'bait' protein used, i.e. HA-tagged HDA1 and HDAP, were not pulled down with their putative interacting factors (full lists of results in **Electronic Supplemental Material 1 (ES1)).**

To combat all of these issues in one fell swoop, the homology regions with unique restriction sites created as shown in **Supplementary Images 1E-1I** could be re-digested and inserted into the pG0079 KS vector by Gibson Assembly (Hughes and Waters, 2017), facilitating further conditional degradation attempts, particularly those that would complement HDA1 and HAT1 KO studies. The additional tagging of putative epigenetic modifying enzymes with GFP would enable further confirmatory studies to be carried out by both fluorescence imaging and co-IP. In recent studies, anti-GFP antibodies have been used successfully to precipitate tagged *Plasmodium* proteins, though issues with proteins of high molecular weight, low abundance, and unusual solubility characteristics have been noted, and may account for the loss of large tagged histone deacetylases during co-IP sample preparation, as was seen in the present study (Zhao *et al.*, 2016; Zhang *et al.*, 2017; Filarsky *et al.*, 2018). Further conditional protein degradation systems that may aid in characterisation of both HDA1 and HAT1 are discussed in **Chapter 6.**

In the present study, co-IP experiments were important providers of data to aid in determining the functions of both HDA1 and HDAP in *P. berghei*, in cases where complete KO lines could not be cloned (HDA1) or where the protein-of-interest was essential or no KO plasmid was available (as was the case with HDAP) **(**section

3.3.10). To confirm the findings of these results, GFP- or HA-tagging of the either of the shared co-immunoprecipitated ribosomal proteins, RPS11, or RPS17, could provide a method by which HDA1 and HDAP could be pulled-up in a co-IP experiment. The interaction of these proteins, and at what stage in the *P. berghei* life-cycle they come together, could provide valuable data in elucidating the downstream effects of these histone deacetylases in the *Plasmodium* life-cycle.

4. Identification of histone modifications during asexual and sexual development in *P. berghei*

4.1 Chapter Aim

The aim of this chapter was to map all histone post-translational modifications (PTMs) for asexual- and sexual-stage *P. berghei* ANKA parasites. These data were then used to identify histone PTMs that were specific to sexual-stage parasites (gametocytes).

4.2 Introduction

4.2.1 Histones and their modifications

In eukaryotes, DNA strands are packaged with histone proteins to form chromatin, the basic unit of which is the histone octamer around which ~146 bp of DNA are wrapped, i.e. the nucleosome (Hammond *et al.*, 2017). In eukaryotes, the DNA-nucleosome complex occurs in a repeated fashion at approximately 200 ± 40 bp intervals, with each nucleosome composed of two copies each of histone 2A (H2A), histone 2B (H2B), histone 3 (H3), and histone 4 (H4) (**Figure 4.1 (a)**) (Luger *et al.*, 1997; Luger, Dechassa and Tremethick, 2012). Nucleosome complexes further assemble into higher order chromatin structures, and are the primary determinant of DNA accessibility in the nucleus (Luger *et al.*, 1997).

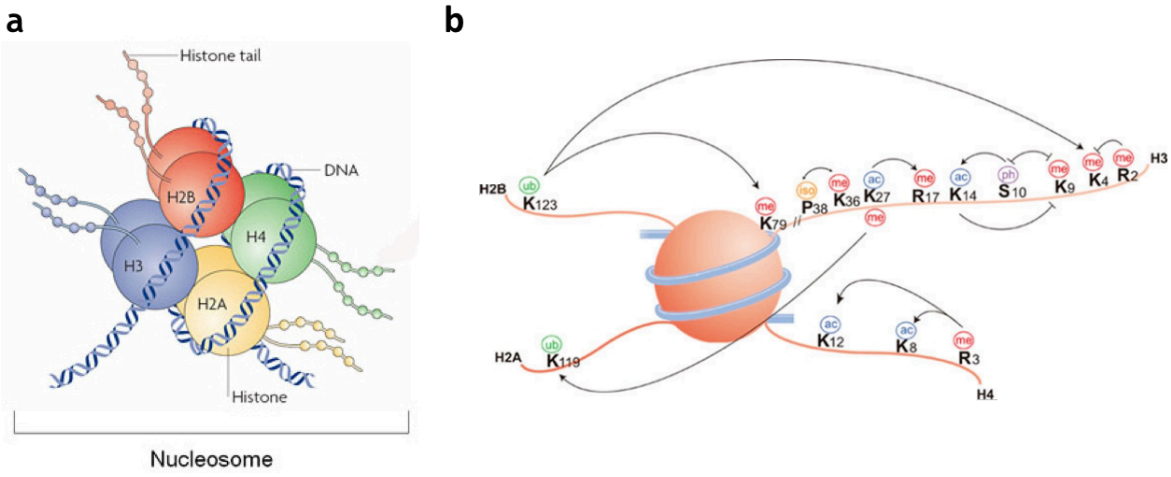


Figure 4.1: An overview of histone conformation within the nucleosome and histone 'crosstalk'. Image **(a)** is a cartoon depiction of the structure of a eukaryotic nucleosome, showing histone pairs (colour-coded). The primary structure of chromatin is the DNA-nucleosome nucleoprotein complex, with a chromatin fibre on average being 10 nm in diameter (Fussner, Ching and Bazett-Jones, 2011). In **(a)**, DNA is shown wrapped around the nucleosome.

Nucleosomes consist of two canonical or variant H3-H4 'tetramers' and two canonical or variant H2A-H2B 'dimers' to create the final histone octamer (Hammond *et al.*, 2017). In this image, N- or C-termini of histones are shown extending from the nucleosome in the same colour as the histone of origin. Histone post-translational modifications (PTMs) often occur at these amino acid termini. In **(b)**, examples of PTMs to N- or C-termini of histones H2B, H2A, H3, and H4 are shown. Termini extend from the nucleosome (in orange) and a number of histone modifications are shown on histone lysines (K), arginines (R), and single proline (P) and serine (S) residues. Black lines are used to indicate interacting histone 'marks', with arrows indicating a positive effect, and flat ends indicating a negative association between 'marks'. Modifications to histones shown include methylation (me), acetylation (ac), phosphorylation (ph), ubiquitination (ub), and isomerisation (iso). Source for **(a):** stock image from irbbarcelona.org. Source for **(b)**: Bannister and Kouzarides, 2011.

The incorporation of variant histones and the organisation of chromatin fibres into secondary and tertiary structures play a large role in determining the accessibility of DNA to transcriptional machinery, as do the vast array of post-translational modifications available to histone amino acid side-chains (Bannister and Kouzarides, 2011; Luger, Dechassa and Tremethick, 2012). In addition, recent studies have also shown that PTMs to the histone amino acid core may determine overall DNA accessibility by affecting histone-histone and histone-DNA interactions (Tessarz and Kouzarides, 2014).

At present, a vast number of post-translational histone modifications have been identified, including, but not limited to, methylation, acetylation, phosphorylation, citrullination (deamination), ubiquitination/ubiquitinylation, SUMOylation, formylation, succinylation, and ADP-ribosylation (Bannister and Kouzarides, 2011; Tessarz and Kouzarides, 2014). The role of many of these histone modifications (also referred to as histone 'marks') and the enzymes that are responsible for their manifestation, are currently under investigation, and may differ between species (Bannister and Kouzarides, 2011; Luger, Dechassa and Tremethick, 2012; Tessarz and Kouzarides, 2014; Prakash and Fournier, 2017). It is also accepted that histone PTMs influence each other, a concept referred to as 'histone cross-talk' **(Figure 4.1 (b))** (Kouzarides, 2007; Bannister and Kouzarides, 2011), for example, the disruption of a histone modification by an adjacent, antagonistic histone 'mark' or competition of different histone modifications for the binding site of the same enzyme.

4.2.2 The histones of *P. berghei*

It is known from studies into the human *P. falciparum* parasite, and from later comparative analyses of encoding genes, that *Plasmodium* parasite nucleosomes are made up of four canonical histones (14-18 kDa) (H2A, H2B, H3, H4), with four possible variant histones (H2A.Z, H2B.Z (sometimes referred to as H2Bv), H3.3, and centromere-specific H3 (CenH3)) (Longhurst and Holder, 1997; Miao *et al.*, 2006). A great number of histone modifications have also been identified on *Plasmodium* histones, and it has been determined that *Plasmodium* parasites do not contain a 'linker' histone 1 (H1), but do possess a heat-unstable (HU) bacterial histone-like protein (denoted *Pf*HU in *P. falciparum*) that is capable of DNA condensation, particularly in the apicoplast (Miao *et al.*, 2006; Ram *et al.*, 2008; Salcedo-Amaya *et al.*, 2009; Trelle *et al.*, 2009; Luah *et al.*, 2010; Jiang *et al.*, 2013; Coetzee *et al.*, 2017; Gupta *et al.*, 2017; Sindikubwabo *et al.*, 2017). The nucleosomes of *P. falciparum* have also been shown to have reduced histone octamer stability in comparison to the nucleosomes of other eukaryotes, and have a nucleosome repeat length of ~155 bp both *in vitro* and *in vivo* (Silberhorn *et al.*, 2016).

A complete comparative analysis of all eight *P. berghei* ANKA core and variant histones with those of all five *Plasmodium* species known to infect humans is located in **Supplemental Section S1.** These comparisons of peptide sequences and structures also include the orthologous host histones from *Homo sapiens sapiens* (modern human), *Rattus norvegicus* (the rat species used in the present study as a *P. berghei* host), and *Mus musculus* (the mouse species commonly used as a *P. berghei* host in the present study). (**Supplemental section S1**).

4.2.3 Experimental methods for the isolation and extraction of histones from whole proteomes

As a result of their importance in the regulation of gene expression among all eukaryotes, a number of methods have been tried and tested for the extraction, isolation, and purification of histones from animal, plant, or other eukaryotic cell types (Shechter *et al.*, 2007). Histones are highly alkaline, acid-insoluble proteins and most histone extraction protocols focus on extraction

using acids, mainly dilute hydrochloric acid (HCl) or dilute sulphuric acid (H_2SO_4) (Shechter *et al.*, 2007). Alternatively, high-salt extraction of histones may prove useful as a result of the neutral pH environment, which may protect acid-labile post-translational modifications, such as phosphorylation marks, from being lost. Procedures have also been developed in which increasing levels of salt (NaCl) concentration have been used to separate H2A/H2B nucleosome components from the even more insoluble H3/H4 nucleosome component, allowing these fractions to be analysed separately thereafter (Shechter *et al.*, 2007; Rodriguez-Collazo, Leuba and Zlatanova, 2009).

Furthermore, both acid and high-salt techniques may differentially extract the histone components from different chromatin environments depending on the heterochromatic or euchromatic conformation of the chromatin. As such, tandem extractions of histones using both methods may provide a more complete picture of the histone PTM profile of a particular cell type (Shechter *et al.*, 2007). After extraction of histones by the chosen method, histones can be analysed by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), acid-urea (AU) gel electrophoresis, or mass spectrometry (MS), with or without purification by any of a number of liquid chromatography procedures (**Figure 4.3**).

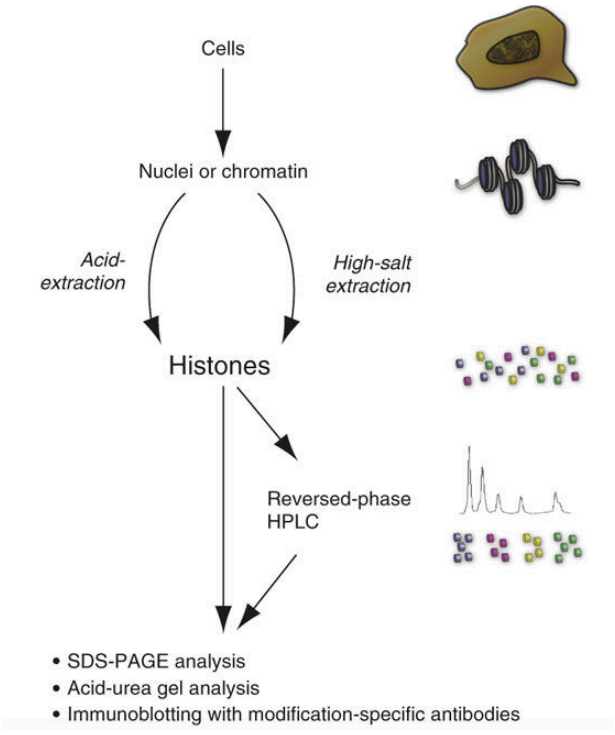


Figure 4.2: Sequence of techniques used for the extraction, purification, and analysis of histones from eukaryotic cells. The image above is a cartoon representation of the processes

involved in extracting and analysing the histones of eukaryotic cells. To begin with, cells are cultured in the appropriate manner (dependent upon cell type) and the nuclear fraction or whole cells taken at the life cycle stage of interest for either acid- or high-salt extraction. Acid-extracted or high-salt-extracted histones can be purified by liquid chromatographic methods (such as reversed-phase high-performance liquid chromatography (HPLC)). Both purified or non-purified histones can be analysed by SDS-PAGE electrophoresis, acid-urea gel electrophoresis, followed by staining or immunoblotting as desired. Examination of histone PTMs is often carried out by combining extraction and purification with mass spectrometry (not shown as an option in this diagram). Source: Shechter *et al.*, 2007.

In the present study, *P. berghei* ANKA histones were extracted from asexual (mature, synchronous schizont) stage or sexual (mature gametocyte) stage, using acid-extraction with dilute HCl as described in Miao *et al.*, 2006 (Miao *et al.*, 2006). Acid-extracted histones were then examined, first by SDS-PAGE to confirm isolation of histones, and then by liquid chromatography-coupled tandem mass spectrometry (LC-MS/MS). The procedure used for the extraction of histones from *P. berghei* ANKA cells is described in section **2.2.4,** with sample preparation, LC-MS/MS, and analysis of histone PTMs carried out as described from section **2.2.8** to **2.2.10** inclusive**.**

4.3 Results

4.3.1 Optimising schizont and gametocyte production in a *P. berghei* ANKA line

In the present study, acid extraction of histones was carried out on mature asexual *P. berghei* ANKA schizonts and mature *P. berghei* ANKA gametocytes (male and female gametocyte mix), followed by examination of histone PTMs by LC-MS/MS data. To ensure enough histone material to investigate PTMs at both asexual and sexual stages of the *P. berghei* ANKA life cycle, infection of Wistar rats was carried out, with asexual parasites synchronised *in vitro* (after Plasmodipur filtration of blood; described in section **2.1.14**) to mature schizont stage, followed by purification from culture by a 55% Nycodenz density gradient (detailed in section **2.1.5.4**) or magnetic separation (detailed in section **2.1.15**). To ensure that a parasite culture was not contaminated by *P. berghei* parasites of other life cycle stages (such as ring stages or trophozoites), synchronisation and maturation of a high-producer of gametocytes (HP) *P.*

berghei ANKA line in culture was carried out in the first instance, with two biological replicates carried out using the G1142 AP2-G overexpressor line, in which transcription of AP2-G was not induced by rapamycin addition. These protocols ensured that minimal (in the case of HP), or absolutely no contaminating gametocytes (in the case of G1142) were present in the samples (the G1142 gametocyte overexpressor line is discussed further in section **1.7.2.3.1**).

Conversely, initial examination of *P. berghei* ANKA gametocytes was carried out using a HP *P. berghei* line in which asexually replicating parasites were killed off by treatment with sulfadiazine for 24 hours at 25mg/l (Beetsma *et al.*, 1998). The preparation of sulfadiazine for administration to infected animals is detailed in section **2.1.8**. With the development of the G1142 AP2-G overexpressor line, 2 biological replicates were carried out in which gametocyte production was ensured by rapamycin induction (procedure detailed in section **2.1.9**). All parasites were purified from blood or culture following removal of contaminating blood products by Plasmodipur filtration (section **2.1.14**) by a 53% Nycodenz density gradient (detailed in section **2.1.5.4**) or magnetic separation (section **2.1.15**).

4.3.2 *P. berghei* ANKA schizont and gametocyte samples

After optimisation of *in vivo* and *in vitro* methods to ensure a maximal amount of the desired parasite stage with no contamination by other life-cycle stages, maturation of asexual and sexual *P. berghei* ANKA parasites was monitored by Giemsa staining (section **2.1.12**). Images of Giemsa stains taken just prior to acid extraction of histones are shown in **Figure 4.3.** A selection of five parasites from each culture were imaged as a representation of the parasites that were seen throughout the entire *in vitro* culture (in the cases of all schizont samples and G1142 gametocytes) and gametocytes seen *in vivo* after Sulfadiazine selection are shown in **Figure 4.3 (B).**

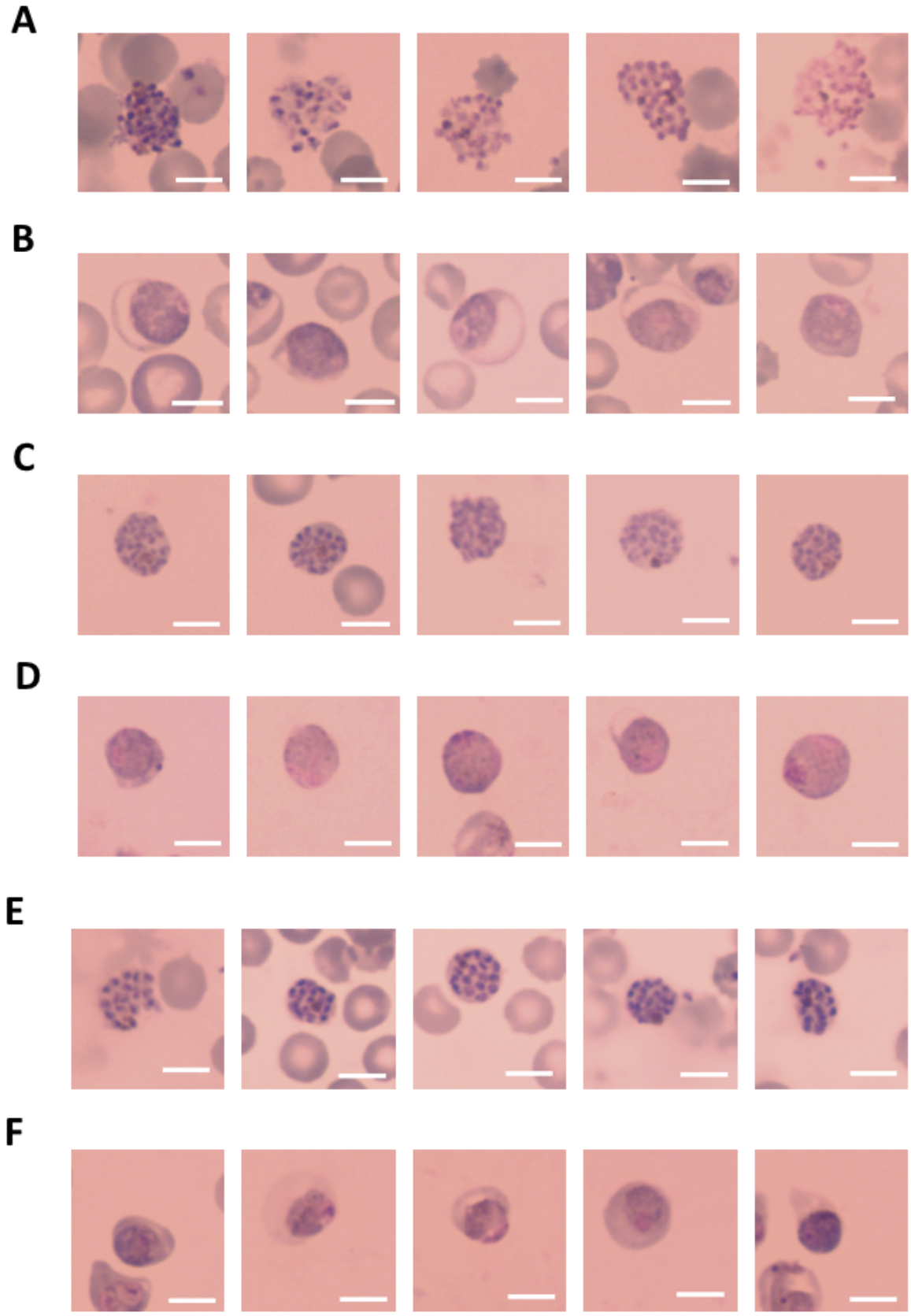


Figure 4.3: Giemsa staining of *P. berghei* ANKA mature asexual-stage parasites (schizonts) and mature sexual-stage parasites (gametocytes). Each row of images from **A-F** shows a selection of the mature asexual parasites (schizonts) and mature sexual-stage gametocytes taken as representative examples of parasites observed throughout an entire *in vitro* culture (**A** and **C-F**) and after *in vivo* Sulfadiazine enrichment of gametocytes (row **B**). All scalebars (in

white) in the right-hand bottom corner of each image represent 10 μm in length. Row **A** shows *P. berghei* ANKA mature schizonts from an *in vitro* culture (~24 h *in vitro*). These schizonts were taken from a *P. berghei* ANKA HP line after *in vitro* synchronisation and growth. *P. berghei* ANKA schizonts typically contain ~12-18 merozoites and, though some schizonts have been partially burst by the Giemsa staining procedure (smearing on a glass slide), the schizonts in row **A** were mature and ready for purification and acid-extraction of histones. The selection of parasites in Row **B** are gametocytes taken from a rat infection after *in vivo* growth to a parasitaemia of ~5%, followed by 24 h treatment with 25 mg/ml Sulfadiazine *in vivo*. After these images were taken, blood was taken from the rat, Plasmodipur-filtered, and the gametocytes purified by 53% Nycodenz density gradient.

Row **C** and row **E** both show images of Giemsa-stained schizonts from *P. berghei* ANKA line G1142 (the rapamycin-inducible AP2-G overexpressor line). These schizonts were taken from *in vivo* infections of rats with no rapamycin induction, followed by Plasmodipur filtration of blood and maturation *in vitro*. In both cases, all parasites in culture followed asexual growth, with no AP2-G expression in the background line, and therefore, no gametocytes to contaminate the culture.

Rows **D** and **F** show a selection of gametocytes after rapamycin-induction *in vivo* (using the *P. berghei* ANKA G1142 overexpressor line), followed by Plasmodipur-filtration and culturing *in vitro* under rapamycin pressure**.** These rapamycin-induced gametocytes resembled native *P. berghei* ANKA HP gametocytes closely (row **B**), with the only difference being that, after rapamycin-induction, the *in vitro* cultures did contain some gametocytes with multiple nuclei, i.e. an erythrocyte that had been infected with multiple merozoites which were then induced to become gametocytes after rapamycin-induction. Such gametocytes with multiple nuclei were not visible in *P. berghei* ANKA HP native samples (as shown in Row **B**).

4.3.3 *P. berghei* ANKA schizont and gametocyte histones

Before either schizont or gametocyte histones were sent for analysis by LC-MS/MS, acid-extracted histones were denatured at 95ºC for 10 min and run on a 15% SDS-PAGE gel that was then stained with Coomassie® Brilliant Blue R-250 (Sigma-Aldrich). Samples were run next to a commercial 10-250 kDa prestained protein ladder (PageRuler™ Plus Prestained Protein Ladder, ThermoFisher Scientific) and a sample of re-suspended calf thymus histones at a known concentration of 1 mg/ml (also denatured at 95ºC for 10 min; lyophilised calf thymus histone from Sigma-Aldrich). An example of acid-extracted G1142 *P. berghei* schizont and gametocyte histones is shown in **Figure 4.4.**

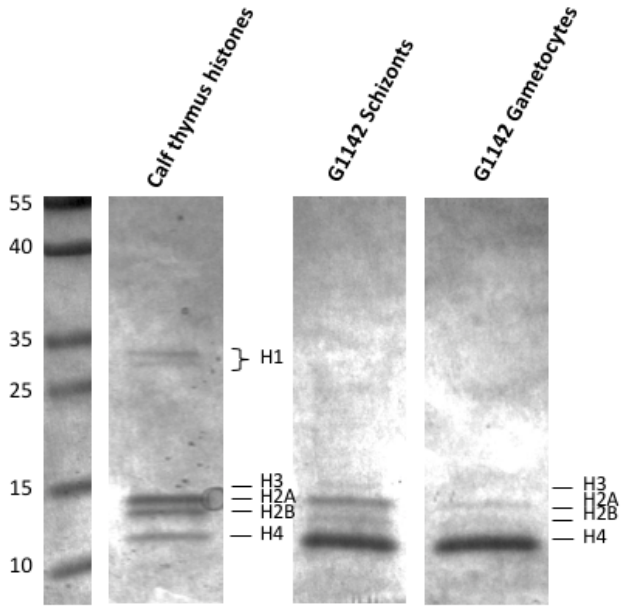


Figure 4.4: Acid-extracted *P. berghei* schizont and gametocyte histones. The image above depicts a Coomassie-stained 15% SDS-PAGE gel (image taken under white light using a Gel Doc® XR+ Gel Documentation System, Bio-Rad Laboratories, Inc.) to which 5 ul of protein ladder, 3 ul of calf histone (at 1 mg/ml) and 5 ul each of a sample of G1142 schizont and G1142 gametocyte histones were added (all proteins denatured at 95ºC for 10 min; histone samples of an unknown concentration). Protein ladder standards are measured in kiloDaltons (kDa) to the left of the image. Histones are labelled H1-H4, with two subtypes of H1 (H1.1 and H1.2) seen in the calf thymus sample. Loss of the H1 histone in both *P. berghei* ANKA schizonts and gametocytes is consistent with previous findings in *P. falciparum* (Miao *et al.*, 2006). The masses of the H3, H2A, H2B, and H4 histones in kDa (roughly 11-15 kDa) are consistent with predicted masses of 15.4, 14.1, 13.3, and 11.5 kDa respectively (PlasmoDB). Variant histones (H2A.Z, H2B.Z, H3.3, and CenH3) could not be distinguished from canonical histones by SDS-PAGE, and H4 appeared to be the most abundant histone present in the acid-extracted samples (again consistent with findings in *P. falciparum* (Miao *et al.*, 2006)). From examination of this gel, it could be determined that both *P. berghei* histone samples were of ~0.75-1 ug/ul (or mg/ml) concentration in solution. These samples were then sent for LC-MS/MS.

As can be seen in **Figure 4.4**, no histone H1 isoforms were isolated from *P. berghei* schizont or gametocyte acid extracts. This SDS-PAGE image was a clear indication that no contaminant histones from rat cells were present in these extracts (clear H1 sub-fractions are visible when acid extraction is carried out on rat cells) (Bucci, Brock and Meistrich, 1982). Visible *P. berghei* histones ranged from 11.5 kDa (H4) to 16.5 kDa (H2A.Z) in mass, with the *P. berghei* centromeric H3 protein (CenH3) not visible in this image above the H3 band (CenH3 has a mass of 19.1 kDa). These results are again consistent with the findings of Miao *et al.* (2006) upon extraction of *P. falciparum* histones. In *P.*

falciparum, CenH3 gene expression remained relatively unchanged across the intraerythrocytic parasite life-cycle apart from a clear reduction at mid to late trophozoite stages (Miao *et al.*, 2006). However, the lack of a CenH3 band in **Figure 4.4** may not be a result of stage-specific decreases in CenH3 abundance at mature schizont or gametocyte stages of the life-cycle. Rather, this protein displays relatively low expression across all eukaryotic cell types. CenH3 is found exclusively at functioning centromeres during cell division, with even pericentromeric chromosome regions being devoid of this protein in lieu of canonical or variant H3 (Maheshwari *et al.*, 2017). As such, failing to observe a CenH3 band upon histone extraction from *P. berghei* cells was not unexpected.

Another clear difference between the calf thymus histones seen in **Figure 4.4** and the histones of *P. berghei* (line G1142) schizonts and gametocytes was the greater abundance of H4 in *P. berghei* samples. This dark staining of the H4 band on an SDS-PAGE gel was also observed when histones were acid extracted from *P. falciparum* intraerythrocytic cells in the same manner as in this experiment (Miao *et al.*, 2006). A number of reasons may be responsible for the relatively dark band of H4 in *P. berghei* samples. To begin with, 0.25 HCl acid extraction of all eight alkaline histones (isoelectric points ranging from pH 10.84 to pH 11.8) should not result in the greater isolation of one histone over another, though H4 is the most alkaline of the histones in *P. berghei*. Histone acid extraction with HCl twice (as described in section **2.2.4**) should have successfully extracted all histones from protein lysates. The SDS-PAGE gel in **Figure 4.4** was also a 15% polyacrylamide gel, a standard percentage used for both the visualisation and extraction of histones from eukaryotic cells, including *Plasmodium* spp. (Miao *et al.*, 2006; Shechter *et al.*, 2007). Despite this, 4-20% or 4-25% gradient SDS-PAGE gels could also be used, and in this case, may be beneficial in isolating any alkaline proteins of a similar mass from H4 out of the extract (Shechter *et al.*, 2007; Coetzee *et al.*, 2017). A Western blot of the gel shown in **Figure 4.4** with an anti-H4 antibody may also be beneficial in discerning the presence of any contaminating alkaline proteins of similar mass to *P. berghei* H4.

Outside of technical reasons for the greater Coomassie® staining of H4 in **Figure 4.4** (the Coomassie® stain itself is also acidic), a greater abundance of

H4 may result from the fact that it is the only histone within the *P. berghei* extract of which there is no variant histone of a slightly different size and isoelectric point (as with H3, H2A, and H2B). Isolation of histones at mature schizont and mature gametocyte stage show that H4 was equally abundant in the merozoite-filled schizont prior to egress and in mature sexual stages prior to transmission, with DNA replication having ceased at these points within the parasite life-cycle. Histone 4 acetylation is required for chromatin decompaction during DNA replication in eukaryotes, though the amount of total H4 protein remains unchanged (Ruan *et al.*, 2015). One other possibility for the relative H4 abundance in *P. berghei* samples (**Figure 4.4**), was that canonical and variant H3, H2A, and H2B histones were somehow being degraded during either the *P. berghei* cell lysis or acid extraction processes, with the same happening upon cell lysis and acid extraction of *P. falciparum* histones (Miao *et al.*, 2006). To rectify this issue, further experiments may require that cell lysis buffers and SDS-PAGE sample buffers contain additional protease inhibitors, such as a mixture of phenylmethylsulfonyl fluoride (PMSF), Aprotinin, Leupeptin, and Pepstatin that has been suggested for rapid purification of recombinant histones (Klinker *et al.*, 2014).

4.3.4 Identification of *P. berghei* ANKA schizont and gametocyte histone PTMs by MS/MS

Post-translational modifications (PTMs) on *P. berghei* ANKA schizont and gametocyte histones are shown in the figures below. The procedures used for preparation of acid-extracted histones for LC-MS/MS, the LC-MS/MS procedure itself, and the software used for identification of histones and their PTMs are described from section **2.2.8** to section **2.2.10.** In all samples examined, no peptides from H3-like centromeric protein CSE4 (CenH3) were identified, most likely due to the low abundance of this histone-like protein in parasites as discussed above (section **4.3.3**), and the fact that these mature parasite stages would not be expressing CenH3 (mitosis has ceased in these mature stages) (Hoeijmakers *et al.*, 2012; Verma and Surolia, 2013).

4.3.4.1 *P. berghei* ANKA schizont histone PTMs

H4		
0	$_{NtermAc}MS_{ph}GR_{ac/me1/me2/me3}GK_{ac/me2/me3}GGK_{ac/me1/me2/me3}GLGK_{ac/me2/me3}GGAK_{ac/me1}.R$	17
8	$K.GLGK_{ac/me2/me3}GGAK_{ac/me1}R.H$	18
23	$R_{ac}.DNIQGITK_{ac/me3}PAIR_{me1}.R$	36
23	$R_{ac}.DNIQGITK_{ac/me3}PAIR_{me1}R_{me1}.L$	37
39	$R.R_{ac/me2}GGVK_{me1/me3}R_{ac}ISGLIYEEIR_{ac/me1}.G$	56
44	$K.R_{ac}ISGLIYEEIR_{ac/me1}.G$	56
45	$R.ISGLIYEEIR_{ac/me1}.G$	56
45	$R.ISGLIYEEIR_{ac/me1}GVLK.V$	60
55	$R.GVLK_{ac}VFLENVIK.D$	68
59	K.VFLENVIK.D	68
59	$K.VFLENVIKDSIMYTEHAK_{ac}.R$	78
67	$K.DSIMYTEHAK_{ac}.R$	78
77	$K.R_{ac/me1}K_{me1}TVTAMDIVYS_{ph}LK.R$	92
78	$R.K_{me1}TVTAMDIVYS_{ph}LK.R$	92
79	$K.TVTAMDIVYS_{ph}LK_{ac}.R$	92
79	$K.TVTAMDIVYS_{ph}LK_{me1}R_{me1}.Q$	93

H3		
0	$MAR_{ac}TK_{ac/me1}QT_{ph}AR_{me1/me2}.K$	9
8	$R.K_{ac/me1}S_{ph}T_{ph}AGK_{ac/me1/me2}APR_{me1/me2/me3}.K$	18
17	$R.K_{ac/me1/me3}QLAS_{ph}K_{ac/me2}AAR_{me1/me2}.K*$	27
26	$R.K_{ac/me1/me2}S_{ph}APIS_{ph}AGIK_{ac}.K$	37
27	$K.S_{ph}APIS_{ph}AGIK_{ac}.K$	37
36	$K.K_{me1}PHR_{me1}YR_{me1}PGT_{ph}VALR_{ac}.E*$	50
40	$R.YRPGTVALR_{ac}.E*$	50
40	$R.YRPGTVALR_{ac}EIR_{ac}.R$	53
52	$R.R_{ac}YQK_{ac}STDLLIR_{ac}K.L$	65
56	$K.STDLLIR_{ac}.K*$	64
56	$K.STDLLIR_{ac}K.L*$	65
72	$R.EIAQDYK_{me1}.T$	80
72	$R.EIAQDYK_{me1}TDLR_{me1}.F$	84
83	R.FQSSAVMALQEAAEAYLVGLFEDTNLCAIHAK.R	116
115	K.RVTIMPK.D*	123
122	K.DIQLARR.I*	130

H3.3		
2	$R.T_{ph}K_{ac/me1}QT_{ph}AR_{ac/me3}K_{ac/me1/me2/me3}S_{ph}T_{ph}GGK_{ac/me1/me3}.A$	15
8	$R.K_{ac/me1/me2/me3}S_{ph}T_{ph}GGK_{ac/me1/me3}APR_{me1/me2/me3}.K$	18
9	$K.S_{ph}TGGK_{ac/me1/me3}APR_{me1/me2/me3}.K$	18
8	$R.KS_{ph}TGGK_{ac/me1/me3}APR_{me1/me2/me3}K_{ac/me1/me3}.Q$	19
17	$R.K_{ac/me1/me3}QLAS_{ph}K_{ac/me2}AAR_{me1/me2/me3}.K*$	27
26	$R.K_{ac/me1/me2/me3}S_{ph}APVS_{ph}T_{ph}GIK_{ac}.K$	37
27	$K.S_{ph}APVS_{ph}T_{ph}GIK_{ac}.K$	37
36	$K.K_{ac/me1/me3}PHR_{me1}YRPGT_{ph}VALR.E*$	50
40	R.YRPGTVALR.E*	50
40	R.YRPGTVALREIR.K	53
52	$R.K_{ac}FQK_{me1}S_{ph}T_{ph}DLLIR_{me1}K_{me1}.L$	65
56	$K.S_{ph}T_{ph}DLLIR_{me1}.K*$	64
56	$K.S_{ph}T_{ph}DLLIR_{me1}K_{me1}.L*$	65
69	R.LVREIAQEYK.T	80
72	R.EIAQEYK.T	80
72	R.EIAQEYKTDLR.F	84
83	R.FQSQAVLALQEAAEAYLVGLFEDTNLCAIHAK.R	116
115	$K.R_{me1}VTIMPK_{ac/me1}.D*$	123
122	K.DIQLARR.I*	130

H2A		
0	$MSAK_{me1/me3}GK_{ac/me3}T_{ph}GR_{me1/me2}K_{me3}.K$	10
8	$R.K_{me3}K_{me1}AVK_{ac/me1}GT_{ph}S_{ph}NS_{ph}AK_{ac/me1}.A$	21
10	$K.AVK_{ac/me1}GT_{ph}S_{ph}NS_{ph}AK_{ac}.A$	21
13	$K.GT_{ph}S_{ph}NS_{ph}AK_{ac/me1}AGLQFPVGR_{me1}.I$	30
20	$K.AGLQFPVGR_{me1}.I$	30
29	$R.IGR_{me1}Y_{ph}LK_{me1}K_{ac}.G$	37
32	$R.YLK_{me1}KGKY_{ph}AKR.V$	43
41	K.RVGAGAPVYLAAVLEYLCAEILELAGNAAR.D	72
42	R.VGAGAPVYLAAVLEYLCAEILELAGNAAR.D	72
81	$R.HIQLAVR_{me1}.N$	89
81	$R.HIQLAVRNDEELNK_{me2}.F$	96
88	R.NDEELNKFLAGVTFASGGVLPNIHNVLLPK.K	119
95	K.FLAGVTFASGGVLPNIHNVLLPK.K	119
118	$K.KS_{ph}QLKS_{ph}GAT_{ph}ANQDY$	132

H2A.Z

0	$_{NtermAc}$MEVPGK$_{ac/me2}$IIGGK$_{ac/me1/me3}$.V	11
5	K.IIGGK$_{ac/me1/me3}$VGGK$_{ac}$.V	15
5	K.IIGGK$_{ac/me1/me3}$VGGK$_{ac}$VGGK$_{ac}$.V	19
5	K.IIGGK$_{ac/me1/me3}$VGGK$_{ac}$VGGK$_{ac}$VLGLGK$_{ac}$.G	25
14	K.VGGK$_{ac}$VLGLGK$_{ac}$.G	25
18	K.VLGLGK$_{ac}$GGK$_{ac/me1}$.G	28
18	K.VLGLGK$_{ac}$GGK$_{ac/me1}$GK$_{ac/me1}$T$_{ph}$GS$_{ph}$GK$_{ac/me1/me2}$.T	35
24	K.GGK$_{ac/me1}$GK$_{ac/me1}$T$_{ph}$GS$_{ph}$GK$_{ac/me1/me2}$T$_{ph}$K$_{ac/me1}$.K	37
29	K.T$_{ph}$GS$_{ph}$GK$_{ac/me1/me2}$T$_{ph}$K$_{ac/me1}$K$_{ac/me3}$.A	38
37	K.APLS$_{ph}$R$_{me1}$AS$_{ph}$R$_{me1}$.A	46
45	R.AGLQFPVGR.V	55
68	R.VGSTAAVYAAAILEYLTAEVLELAGNATK.D	98
100	K.VKR$_{me1}$ITPR$_{me1}$HLQLAIR.G	115
107	R.HLQLAIR.G	115
107	R.HLQLAIRGDEELDTLIK.A	125
114	R.GDEELDTLIK.A	125
124	K.ATIAGGGVIPHIHK.A	139
138	K.ALMNKVPVPPPQTK.K	153
138	K.ALMNKVPVPPPQTKKPK.K	156
143	K.VPVPPPQTK.K	153
143	K.VPVPPPQTKKPK.K	156

H2B

0	MVS$_{ph}$K$_{ac}$K$_{me1/me3}$PAK$_{ac/me1/me2}$.E	8
3	K.K$_{ac/me3}$PAK$_{ac/me2}$EK$_{ac/me1}$K$_{ac/me1}$.A	11
10	K.AT$_{ph}$NGAT$_{ph}$DGK$_{ac/me1/me2/me3}$.K	20
19	K.AT$_{ph}$NGAT$_{ph}$DGK$_{ac/me1/me2/me3}$K$_{ac/me1/me2/me3}$K$_{ac/me1}$.R	22
24	K.SRYDSYGLYIFK.V	37
26	R.YDSYGLYIFK.V	37
26	R.YDSYGLYIFKVLK.Q	40
36	K.VLKQVHPDTGISR.K	50
39	K.QVHPDTGISR.K	50
39	K.QVHPDTGISRK.S	51
49	R.K$_{me1}$SMNIMNSFLVDTFEK.I	66
50	K.SMNIMNSFLVDTFEK.I	66
50	K.SMNIMNSFLVDTFEKIAT$_{ph}$EASR.L	73
72	R.LCK$_{ac/me1}$Y$_{ph}$T$_{ph}$K$_{ac/me1/me3}$R$_{ac/me1/me3}$DTLS$_{ph}$S$_{ph}$R.E	86
75	K.Y$_{ph}$T$_{ph}$K$_{ac/me1/me3}$R$_{ac/me1/me3}$DTLS$_{ph}$S$_{ph}$R.E	86
78	K.R$_{ac/me1/me3}$DTLS$_{ph}$S$_{ph}$R.E	86
79	R.DTLS$_{ph}$S$_{ph}$REIQT$_{ph}$AIR$_{ac/me1}$.L	93
85	R.EIQT$_{ph}$AIR$_{me1}$.L	93
85	R.EIQT$_{ph}$AIR$_{me1}$LVLPGELAK$_{ac}$.H	102
92	R.LVLPGELAK$_{ac}$.H	102
101	K.HAVSEGT$_{ph}$K$_{ac/me1/me3}$.A*	110
101	K.HAVSEGT$_{ph}$K$_{ac/me1/me3}$AVT$_{ph}$K$_{ac}$.F	114

H2B.Z

0	$_{NtermAc}$MS$_{ph}$GK$_{ac/me1/me2}$GPAQK$_{ac/me1/me3}$.S	9
3	K.GPAQK$_{ac/me1/me3}$S$_{ph}$QAAK$_{ac/me1/me3}$.K	14
3	K.GPAQK$_{ac/me1/me3}$S$_{ph}$QAAK$_{ac/me1/me3}$K$_{ac/me1/me2}$T$_{ph}$AGK$_{ac/me1/me2/me3}$.T	19
8	K.SQAAK$_{ac/me1/me3}$K$_{ac/me1/me2}$T$_{ph}$AGK$_{ac/me1/me2/me3}$T$_{ph}$LGPR$_{me3}$.H	24
13	K.K$_{ac/me1/me2}$T$_{ph}$AGK$_{ac/me1/me2/me3}$T$_{ph}$LGPR$_{me3}$.H	24
14	K.T$_{ph}$AGK$_{ac/me1/me2/me3}$T$_{ph}$LGPR$_{me3}$.H	24
14	K.T$_{ph}$AGK$_{ac/me1/me2/me3}$T$_{ph}$LGPR$_{me3}$HK$_{me1/me3}$R$_{me2}$.K	27
28	R.R$_{ac}$T$_{ph}$ES$_{ph}$FS$_{ph}$LYIFK$_{ac}$.V	40
29	R.T$_{ph}$ES$_{ph}$FS$_{ph}$LYIFK$_{ac}$.V	40
39	K.VLK$_{ac.me1}$QVHPETGVTK.K	53
42	K.QVHPETGVTK.K	53
42	K.QVHPETGVTK$_{ac/me1}$K.S	54
52	K.KSMNIMNS$_{ph}$FINDIFDR$_{me1}$.L	69
53	K.SMNIMNS$_{ph}$FINDIFDR$_{me1}$.L	69
53	K.SMNIMNS$_{ph}$FINDIFDR$_{me1}$LVTEATR.L	76
78	R.Y$_{ph}$NK$_{ac}$K$_{ac}$R$_{me1}$T$_{ph}$LSSR.E	89
88	R.EIQTAVR$_{me1}$.L	96
95	R.LLLPGELS$_{ph}$K$_{ac/me1/me2}$.H	105
104	K.HAVSEGT$_{ph}$K$_{ac/me1/me3}$.A*	113
104	K.HAVSEGT$_{ph}$K$_{ac/me1/me3}$AVT$_{ph}$K$_{ac/me1/me3}$.Y	117
112	K.AVT$_{ph}$K$_{ac/me1/me3}$YT$_{ph}$T$_{ph}$S$_{ph}$GA	122

Figure 4.5: Complete profile of PTMs on *P. berghei* ANKA mature schizont histones. Listed above are the peptides identified by LC-MS/MS using Mascot protein identification software version 2.6.2 (Matrix Science) (Perkins *et al.*, 1999) and searching a *P. berghei* ANKA amino acid database (version 35). Numbers at the start and end of each peptide refer to the amino acid position of the first and last residues shown. In each histone, the first methionine (M) (corresponding to the start codon of each sequence) is not counted and sequences beginning with this amino acid have a starting number of zero (0) to the left of the first peptide. Full stops after the first amino acid and before the last amino acid in each peptide sequence indicate where trypsin cut each peptide from the complete protein sequence. Modifications to histone amino acid residues are labelled in subscript to the right of the modified residue. In cases where peptides had more than one modification present at a particular amino acid residue, a forward slash is used (/) to separate modifications. An asterisk to the right of the peptide indicates that this peptide was not unique to the histone in question, but shared with the variant histone; for example, shared peptides between histones H3 and H3.3, or between H2B and H2B.Z. Separate searches were run for mono-methylation (me1), di-methylation (me2), tri-methylation (me3), phosphorylation (ph), acetylation (ac) and protein N-terminal acetylation (NtermAc).

Histone	Acetylation	Methylation			Phosphorylation
		Mono-	Di-	Tri-	
H4	**N-term, K5, K8, K12, K16,** R23, **K31**, R40, R45, R55, **K59**, K77, R78, **K91**	**R3, K5, K8, K16,** R35, **R36,** K44, **R55,** R78, **K79, K91,** R92	**R3, K5, K8, K12, K16** R40	**R3, K5, K8, K12,** K31, K44	**S1, S89**
H3	R2, **K4, K9, K14, K18, K23, K27, K36**, R49, R53, **K56,** R63	**K4, R8, K9, K14, R17, K18, R26, K27,** K37, R40, **R42, K79,** R83	**R8, K14, R17, K23, R26, K27**	**R17, K18**	**T6, S10, T11, S28, S32**
H3.3	**K4,** R8, **K9, K14, K18, K23, K27, K36, K37,** K53, K122	**K4, K9, K14, R17, K18, R26, K27, K37, R40, K56,** R63, **K64,** R116, K122	**K9, R17,** K23, **R26,** K27	**R8, K9, K14, R17, K18,** R26, **K27, K37**	**T3, T6, S10, T11, S22, S28, S32, T33, T45, S57, T58**
H2A	**K13, K20**, K36	**K3, R8, K10, K13, K20, R29,** R32, K35, **R88**	**R8,** K95	**K3, K5, K9**	**T6, T15, S16, S18,** Y39, S120, S124*, T127*
H2A.Z	**N-term, K5, K10, K14, K18, K24, K27, K29, K34, K36, K37**	K10, **K27, K29, K34, K36, R42, R45, R103, R107**	**K5, K34**	**K10, K37**	**T30, S32, S41, S44**
H2B	**K3, K7, K9, K10, K20,** K21*, K75*, K78*, R79*, R92*, K101*, K109*, K113*	**K4, K7, K9, K10, K19, K20,** K21*, K50*, K75*, K78*, R79*, R92*, K109*	**K7, K19, K20**	**K4, K19, K20,** K78*, R79*, K109*	**S2,** T12*, T16*, T68*, Y76*, T77*, **S83**, S84*, T89*, T108*, T112*
H2B.Z	**N-term, K3, K8, K13, K14, K18,** R29, K39, K42, K52, **K81,** K82, **K104, K112**, K116	**K3, K8, K13, K14, K18, K25,** K42, **K52,** R68, **R83, K104,** K112, K116	**K3, K14, K18, R26,** K104	**K8, K13, K18, R23,** K25, K112, K116	**S1, T15, T19, T30, S32, S34,** S60, Y79, **T84,** S103, **T111, T115,** T118, **T119,** S120

Table 4.1: The histone modifications of *P. berghei* ANKA mature schizonts. This table lists all histone PTMs identified in *P. berghei* ANKA schizonts. All histone amino acid modifications that have previously been identified in *P. falciparum* studies are shown in bold text. An asterisk (*) next to a histone modification indicates that this particular amino acid is unique to *P. berghei*, or that the corresponding amino acid in *P. falciparum* is at a different position of the histone amino acid chain.

4.3.4.2 *P. berghei* ANKA gametocyte histone PTMs

H4

0	$MS_{ph}GR_{ac/me1/me3}GK_{ac/me1/me2/me3}GGK_{ac/me1/me2}GLGK_{ac}$.G	13
3	$R.GK_{ac/me1/me3}GGK_{ac/me1/me2}GLGK_{ac/me2}GGAK_{ac/me2}$.R	17
8	$K.GLGK_{ac/me2}GGAK_{ac/me2}$R.H	18
12	$K.GGAK_{ac/me2}RHRK_{me1/me3}ILR_{ac/me1}$.D	24
23	$R.DNIQGITK_{ac/me1}PAIR_{me1}$.R	36
23	$R.DNIQGITK_{ac/me1}PAIR_{me1}R_{me1}$.L	37
44	K.RISGLIYEEIR.G	56
45	R.ISGLIYEEIR.G	56
45	$R.ISGLIYEEIRGVLK_{me1/me2}$.V	60
55	$R.GVLK_{me1/me2}VFLENVIK_{me1/me3}$.D	68
59	$K.VFLENVIK_{me1/me3}$.D	68
59	$K.VFLENVIK_{me1/me3}DSIMYTEHAK_{me1}$.R	78
67	$K.DSIMYTEHAK_{me1}$.R	78
77	$K.R_{me1}K_{ac}TVTAMDIVYSLK$.R	92
78	$R.K_{ac}TVTAMDIVYSLK_{me3}$.R	92
79	$K.TVTAMDIVY_{ph}SLK_{me3}$.R	92
79	$K.TVTAMDIVY_{ph}SLK_{me3}R_{me3}$.Q	93

H3

0	$_{NtermAc}MAR_{me1/me2/me3}T_{ph}K_{ac/me1/me3}QT_{ph}AR_{me1}$.K	9
8	$R.K_{ac/me1/me3}ST_{ph}AGK_{ac/me1/me3}APR_{me1/me3}$.K	18
17	$R.K_{ac/me1/me3}QLAS_{ph}K_{ac/me1}AAR_{ac/me1}$.K*	27
26	$R.K_{ac/me1/me2}S_{ph}APIS_{ph}AGIK_{ac/me1/me3}$.K	37
27	$K.S_{ph}APIS_{ph}AGIK_{ac/me1/me3}$.K	37
27	$K.S_{ph}APIS_{ph}AGIK_{ac/me1/me3}K_{ac/me1/me3}PHR_{me1}$.Y	41
36	$K.K_{ac/me1/me3}PHR_{me1/me2}Y_{ph}R_{me2/me3}PGT_{ph}VALR_{me2/me3}$.E*	50
40	$R.Y_{ph}R_{me2/me3}PGT_{ph}VALR_{me2/me3}$.E*	50
40	$R.Y_{ph}R_{me2/me3}PGT_{ph}VALR_{me2/me3}EIR_{me1/me3}$.R	53
42	$R.PGT_{ph}VALR_{me2/me3}EIR_{me1/me3}R_{me2}YQK_{me1/me3}$.S	57
56	$K.STDLLIR_{me2}$.K*	64
56	$K.STDLLIR_{me2}K_{ac}$.L*	65
72	$R.EIAQDYK_{me1}$.T	80
72	$R.EIAQDYK_{me1}TDLR_{me1}$.F	84
83	R.FQSSAVMALQEAAEAYLVGLFEDTNLCAIHAK.R	116
115	$K.R_{me1}VTIMPK_{ac}$.D*	123
116	$R.VTIMPK_{ac}$.D*	123
116	$R.VTIMPK_{ac}DIQLAR$.R*	129
122	$K.DIQLARR_{me2}$.I*	130
128	$R.RIR_{me2}GER$.S	135

H3.3

0	$_{NtermAc}MAR_{me1/me2/me3}T_{ph}K_{ac/me1/me3}QT_{ph}AR_{me1}$.K	9
8	$R.K_{ac/me1/me2/me3}ST_{ph}GGK_{ac/me1/me2/me3}APR_{me1}$.K	18
17	$R.K_{ac/me1}QLAS_{ph}K_{ac}AAR_{ac/me1/me2}$.K*	27
26	$R.K_{ac/me1/me3}S_{ph}APVS_{ph}T_{ph}GIK_{ac}$.K	37
26	$R.K_{ac/me1/me3}S_{ph}APVS_{ph}T_{ph}GIK_{ac/me1/me3}K_{me1}PHR_{me1}$.Y	41
27	$K.S_{ph}APVS_{ph}T_{ph}GIK_{ac/me1/me3}$.K	37
36	$K.K_{ac/me1/me3}PHR_{me1/me2/me3}Y_{ph}R_{me3}PGT_{ph}VALR.E$*	50
40	$R.Y_{ph}R_{me3}PGT_{ph}VALR.E$*	50
40	$R.Y_{ph}R_{me3}PGT_{ph}VALREIR.K$	53
42	$R.PGT_{ph}VALREIRK_{me2}FQK_{me1/me3}S_{ph}TDLLIR.K$	64
56	K.STDLLIR.K*	64
69	R.LVREIAQEYK.T	80
72	$R.EIAQEYK_{ac}$.T	80
72	$R.EIAQEYK_{ac}TDLR.F$	84
83	R.FQSQAVLALQEAAEAYLVGLFEDTNLCAIHAK.R	116
115	$K.R_{me1}VTIMPK_{ac}$.D*	123
116	$R.VTIMPK_{ac}DIQLAR.R$*	129
122	$K.DIQLARR_{me2}$.I*	130
122	$K.DIQLARR_{me2}IRGER_{me2}$.S	135

H2A

0	$_{NtermAc}MS_{ph}AK_{ac/me1/me3}GK_{ac/me2/me3}T_{ph}GR_{me1/me2}K_{me1/me2}$.K	10
8	$R.K_{me1/me2}K_{ac/me1/me3}AVK_{ac/me1/me2/me3}GTS_{ph}NS_{ph}AK_{ac/me1/me2}$.A	21
13	$K.GTS_{ph}NS_{ph}AK_{ac/me1/me2}AGLQFPVGR_{me1}$.I	30
20	$K.AGLQFPVGR_{me1}$.I	30
20	$K.AGLQFPVGR_{me1}IGR.Y$	33
32	$R.YLK_{ac/me1}K_{me1}GK_{ac/me1}Y_{ph}AK_{ac/me1}R.V$	43
42	R.VGAGAPVYLAAVLEYLCAEILELAGNAAR.D	72
74	$K.K_{me1}S_{ph}R_{me1}IT_{ph}PRHIQLAVR_{me1}$.N	89
81	$R.HIQLAVR_{me1}$.N	89
88	R.NDEELN.K	95
88	R.NDEELNKFLAGVTFASGGVLPNIHNVLLPK.K	119
95	K.FLAGVTFASGGVLPNIHNVLLPK.K	119
118	$K.KS_{ph}QLKS_{ph}GATANQDY_{ph}$	132

H2A.Z

0	$_{NtermAc}$MEVPGK$_{me3}$IIGGK$_{ac}$.V	11
5	K.IIGGK$_{ac}$VGGK$_{ac/me2}$VGGK$_{ac/me2}$.V	19
10	K.VGGK$_{ac}$VGGK$_{ac/me2}$VLGLGK$_{ac/me2}$GGK$_{ac/me3}$.G	28
14	K.VGGK$_{ac/me2}$VLGLGK$_{ac/me2}$GGK$_{ac/me3}$.G	28
18	K.VLGLGK$_{ac/me2}$GGK$_{ac/me3}$.G	28
18	K.VLGLGK$_{ac/me2}$GGK$_{ac/me3}$GK$_{ac/me1}$T$_{ph}$GS$_{ph}$GK$_{ac/me1/me2/me3}$.T	35
29	K.T$_{ph}$GS$_{ph}$GK$_{ac/me1}$T$_{ph}$K$_{ac/me1/me2}$K$_{ac/me1/me2}$APLS$_{ph}$.R	42
36	K.K$_{ac/me1/me2}$APLS$_{ph}$R$_{me1}$AS$_{ph}$R$_{me1/me3}$.A	46
45	R.AGLQFPVGR.V	55
68	R.VGSTAAVYAAAILEYLTAEVLELAGNATK.D	98
102	K.R$_{me3}$ITPR$_{me3}$HLQLAIR$_{me3}$.G	115
107	R.HLQLAIR$_{me3}$.G	115
107	R.HLQLAIR$_{me3}$GDEELDTLIK$_{me2/me3}$.A	125
114	R.GDEELDTLIK$_{me2/me3}$.A	125
124	K.ATIAGGGVIPHIHK.A	139
138	K.ALMNK$_{me1}$VPVPPPQT$_{ph}$K.K	153
143	K.VPVPPPQT$_{ph}$K.K	153
143	K.VPVPPPQT$_{ph}$K$_{ac/me2}$K$_{ac/me3}$PK$_{ac/me2}$.K	156
143	K.VPVPPPQT$_{ph}$K$_{ac/me2}$K$_{ac/me3}$PK$_{ac/me2}$K$_{ac/me2}$N	157

H2B

0	$_{NtermAc}$MVS$_{ph}$K$_{ac/me1/me3}$K$_{ac/me1/me3}$PAK$_{ac/me1}$.E	8
3	K.K$_{ac/me1/me3}$PAK$_{ac/me1}$EK$_{ac/me1/me2}$K$_{ac/me1/me2}$AT$_{ph}$NGAT$_{ph}$DGK$_{ac/me1/me2/me3}$.K	20
10	K.AT$_{ph}$NGAT$_{ph}$DGK$_{ac/me1/me2/me3}$K$_{ac/me1/me2/me3}$.K	21
10	K.AT$_{ph}$NGAT$_{ph}$DGK$_{ac/me1/me2/me3}$K$_{ac/me1/me2/me3}$K$_{ac/me3}$RKK$_{ac}$SR.Y	27
24	K.SRY$_{ph}$DSYGLYIFK.V	37
26	R.YDSYGLYIFK.V	37
36	K.VLK$_{ac}$QVHPDT$_{ph}$GISR$_{me1}$.K	50
39	K.QVHPDT$_{ph}$GIS$_{ph}$R$_{me1}$.K	50
39	K.QVHPDT$_{ph}$GIS$_{ph}$R$_{me1}$K$_{ac/me1}$.S	51
49	R.KSMNIMNSFLVDT$_{ph}$FEK.I	66
50	K.SMNIMNSFLVDT$_{ph}$FEK.I	66
50	K.SMNIMNSFLVDT$_{ph}$FEKIATEASR.L	73
65	K.IATEASRLCK.Y	76
72	R.LCK$_{ac}$Y$_{ph}$T$_{ph}$K$_{ac}$R$_{me1}$DT$_{ph}$LS$_{ph}$SR$_{me1}$.E	86
75	K.YTK$_{me1/me2}$R$_{me1}$DTLSSREIQTAIR.L	93
78	K.R$_{me1}$DT$_{ph}$LS$_{ph}$SR$_{me1}$.E	86
85	R.EIQTAIR$_{me1}$.L	93
92	R.LVLPGELAK$_{ac/me1}$.H	102
101	K.HAVS$_{ph}$EGT$_{ph}$K$_{me3}$.A*	110
101	K.HAVS$_{ph}$EGT$_{ph}$K$_{me3}$AVT$_{ph}$K$_{ac/me1}$.F	114
109	K.AVT$_{ph}$K$_{ac/me1}$FTS.K	117

H2B.Z

0	$_{NtermAc}$MS$_{ph}$GK$_{ac/me1/me2}$GPAQK$_{ac/me1/me3}$.S	9
8	K.S$_{ph}$QAAK$_{ac/me1/me3}$K$_{ac/me1/me2}$T$_{ph}$AGK$_{ac/me1/me2}$T$_{ph}$LGPR$_{me1}$.H	24
13	K.K$_{ac/me1/me2}$T$_{ph}$AGK$_{ac/me1/me2}$T$_{ph}$LGPR$_{me1}$.H	24
14	K.TAGK$_{ac/me1/me2}$T$_{ph}$LGPR$_{me1}$.H	24
14	K.TAGK$_{ac/me1/me2}$T$_{ph}$LGPR$_{me1}$HK$_{ac/me3}$.R	26
28	R.RTESFSLYIFK.V	40
29	R.TESFSLYIF.K	39
29	R.TESFSLYIFK.V	41
39	K.VLK$_{me1}$QVHPETGVT$_{ph}$K$_{me1}$.K	53
42	K.QVHPETGVT$_{ph}$K$_{me1/me2}$.K	53
42	K.QVHPETGVT$_{ph}$K$_{me1/me2}$K.S	54
52	K.KSMNIMNSFINDIFDR.L	69
53	K.SMNIMNSFINDIFDR.L	69
83	R.TLS$_{ph}$S$_{ph}$R$_{me1}$EIQTAVR$_{me1}$.L	96
88	R.EIQTAVR$_{me1}$.L	96
95	R.LLLPGELSK.H	105
95	R.LLLPGELSK$_{ac/me3}$HAVSEGTK.A	113
104	K.HAVSEGTK$_{me1/me2/me3}$.A*	113
104	K.HAVSEGTK$_{me1/me2/me3}$AVTK$_{ac}$.Y	117

Figure 4.6: Complete profile of PTMs on *P. berghei* ANKA gametocyte histones. Listed above are mature gametocyte histone peptides identified by LC-MS/MS using Mascot protein identification software version 2.6.2 (Matrix Science) (Perkins *et al.*, 1999) and searching a *P. berghei* ANKA amino acid database (version 35). Numbers at the start and end of each peptide refer to the amino acid position of the first and last residues shown. In each histone, the first methionine (M) (corresponding to the start codon of each sequence) is not counted and sequences beginning with this amino acid have a starting number of zero (0) to the left of the first peptide. Full stops after the first amino acid and before the last amino acid in each peptide sequence indicate trypsin cleavage. Modifications to histone amino acid residues are labelled in subscript to the right of the modified residue. In cases where peptides had more than one modification present at a particular amino acid residue, a forward slash is used (/) to separate modifications. An asterisk to the right of the peptide indicates that this peptide was not unique to the histone in question, but shared with the variant histone; for example, shared peptides between histones H3 and H3.3, or between H2B and H2B.Z. Separate searches were run for mono-methylation (me1), di-methylation (me2), tri-methylation (me3), phosphorylation (ph), acetylation (ac) and protein N-terminal acetylation (NtermAc).

Histone	Acetylation	Methylation			Phosphorylation
		Mono-	Di-	Tri-	
H4	R3, **K5, K8, K12, K16**, R23, **K31,** K79	**R3, K5, K8, K20,** R23, K31, R35, **R36,** K59, K67, K77, R78	**K5, K8, K12, K16,** K59	**R3, K5, K20,** K67, K91, R92	**S1**, Y88
H3	**N-term, K4, K9, K14, K18, K23, R26, K27, K36, K37, K122**	**R2, K4, R8, K9, K14, R17, K18, K23, R26, K27, K36,** K37, R40, R52, **K56, K79,** R83, R116	**R2, R40,** R42, R49, R53, **R63**, R129, R131	R2, **K4, K9, K14, R17, K18, K36, K37,** R42, R49, R52, **K56**	**T3, T6, T11, S22, S28, S32,** Y41, T45
H3.3	**N-term, K4, K9, K14, K18, K23,** R26, **K27, K36, K37, K79,** K122	**R2, K4, R8, K9, K14, R17, K18, R26, K27,** K36, **K37, R40, K56,** R116	**R2, K9, K14, R26,** R40, K53, R129, **R134**	**R2, K4, K9, K14, K27, K36, K37,** R40, R42, K56	**T3, T6, T11, S22, S28, S32, T33,** Y41, **T45, S57**
H2A	**N-term, K3, K5, K10, K13, K20**, K35, K38, K41	**K3, R8, K9, K10, K13, K20, R29**, K35, K36, K38, K41, **K75,** R77, **R88**	**K5, R8, K9, K13,** K20	**K3, K5, K10, K13**	**S1, T6, S16, S18**, Y39, S76, **T79,** S120, S124*, Y132*
H2A.Z	**N-term, K10, K14, K18, K24, K27, K29, K34, K36, K37,** K155, K156	**K29, K34, K36, K37, R42, R45,** K143	K14, **K18,** K24, **K34, K36, K37, K124,** K152, K155**, K156**	**K5, K27, K34, R45,** R103, R107, R114, K124, K153	**T30, S32, T35**, **S41**, **S44,** T151*
H2B	**N-term, K3, K4, K7, K9, K10, K19, K20,** K21, K24, K39*, K50*, K75*, K78*, K101*, K113*	K3, **K4**, **K7, K9, K10, K19, K20,** R49*, K50*, R79*, R85*, R92*, K101*, K113*	**K9, K10, K19, K20,** K78*	K3, **K4, K19, K20, K21,** K109*	**S2,** T12*, T16*, Y27*, T45*, S48*, T62*, Y76*, T77*, T81*, S83*, S105*, T108*, T112*
H2B.Z	**N-term, K3, K8, K13, K14, K18, K25, K104,** K116	**K3, K8, K13, K14, K18, R23,** K42, **K52,** R88, **R95**, K112	**K3, K14, K18,** K52, K112	**K8, K13,** K25, **K104,** K112	**S1, S9**, **T15, T19,** T51, **S86, S87**

Table 4.2: The histone modifications of *P. berghei* ANKA mature gametocytes. This table lists all histone PTMs identified in *P. berghei* ANKA mature gametocytes. All histone amino acid modifications that have previously been identified in *P. falciparum* studies are shown in bold text. An asterisk (*) next to a histone modification indicates that this particular amino acid is unique to *P. berghei*, or that the corresponding amino acid in *P. falciparum* is at a different position of the histone amino acid chain.

4.3.5 Differences between *P. berghei* schizont and gametocyte histone PTMs

The aim of these acid extractions was to identify histone PTMs in both *P. berghei* ANKA mature asexual-stage parasites (schizonts) and mature sexual-stage parasites (gametocytes) to reveal the identity of shared and differing histone modifications between parasite life stages (**Table 4.3).** Although some histone modifications may simply have not been detected by LC-MS/MS, other changes appeared to be consistent between samples, and differed to histone PTMs seen at the opposing life-cycle stage. Changes to the *P. berghei* ANKA histone modification profile can clearly be seen when comparing figures **4.5** and **4.6**, with all differences listed again in **Table 4.3.**

***P. berghei* ANKA schizont PTMs**		***P. berghei* ANKA gametocyte PTMs**	
H4	**N-termAc**, **R3me2, K8me3**, K12me3, K16me1, K31me3, R40ac/me2, K44me1/me3, R45ac, R55ac/me1, K59ac, K77ac, R78ac, K79me1, S89ph, K91ac/me1, R92me1	**H4**	R3ac, **K20me1**/me3, **R23me1**, K31me1, **K59me1**/me2, K67me1/me3, K77me1, K79ac, Y88ph, K91me3, R92me3
H3	**R2ac**, R8me2, S10ph, K14me2, **R17me2**, K23me2, **R26me2, K27me2**, **R42me1**, **R49ac, R53ac, K56ac, R63ac**	**H3**	N-termAc, **R2me1**/me2/me3, T3ph K4me3, K9me3, K14me3, R17me3, **S22ph**, K23me1, R26ac, **K36me1**/me3, K37ac/me3, R40me2, **Y41ph**, R42me2/me3, **T45ph**, R49me2/me3, R52me1/me3, R53me2, K56me1/me3, R63me2, **R116me1**, K122ac, R129me2, R131me2
H3.3	R8ac/me3, **S10ph**, **R17me2/me3**, **K18me3**, **K23me2**, R26me3, K27me2, K53ac, T58ph, R63me1, R64me1, K122me1	**H3.3**	N-termAc, **R2me1**/me2/me3, K4me3, **R8me1**, K14me2, R26ac, **K36me1**, R40me2/me3, Y41ph, R42me3, K53me2, K56me3, K79ac, R129me2, R134me2
H2A	K9me3, T15ph, **R32me1**, K36ac, **K95me2**, T127ph*	**H2A**	N-termAc, **S1ph**, K3ac, **K5ac**/me2, **K9me1**/me2, **K10ac**/me3, K13me2/me3, K20me2, **K35ac**, K36me1, K38ac/me1, **K41**ac/**me1**, K75me1, S76ph, R77me1, T79ph, Y132ph*
H2A.Z	**K5ac**/me2, K10me1/me3, **K27me1**, K37me3, R103me1, R107me1	**H2A.Z**	K5me3, K14me2, K18me2, K24me2, K27me3, T35ph, K36me2, **K37me1/me2**, R45me3, R103me3, R107me3, R114me3, K124me2/me3, K143me1, T151ph*, K152me2, K153me3, **K155ac**/me2, K156ac/me2
H2B	**K7me2**, **K21me1*,** T68ph*, K75me1*, K78me1/me3*, R79ac/me3*, **S84ph*, T89ph*,** R92ac*, K109ac/me1*	**H2B**	N-termAc, K3me1/me3, K4ac, K9me2, **K10me2**, **K19ac**, K21me3, K24ac, Y27ph*, K39ac, T45ph*, S48ph*, R49me1* K50ac*, T62ph*, K78me2*, **T81ph***, R85me1*, K101me1*, S105ph*, K113me1*
H2B.Z	**K18me3**, **R23me3,** K25me1, R26me2, R29ac, T30ph, S32ph, S34ph, K39ac, K42ac, K52ac, S60ph, R68me1, Y79ph, K81ac, K82ac, R83me1, T84ph, S103ph, **K104me1**/me2, T111ph, K112ac, T115ph, K116me1/me3, T118ph, T119ph, S120ph	**H2B.Z**	**S9ph**, **R23me1**, K25ac, T51ph, K52me2, S86ph, S87ph, R88me1, **R95me1**, K104me3, K112me2

Table 4.3: Differences between *P. berghei* ANKA asexual (schizont) and sexual (gametocyte) histone modifications. The table above lists all histone modification differences that were identified between *P. berghei* ANKA mature asexual-stage parasites (schizonts) and mature sexual-stage parasites (gametocytes). An asterisk (*) next to a histone modification indicates that this particular amino acid is unique to *P. berghei*, or that the corresponding amino acid in *P. falciparum* is at a different position of the histone amino acid chain. Shared histone PTMs are not shown. All histone modifications listed underneath the *P. berghei* ANKA schizont column were seen only in schizonts and not gametocytes. Conversely, all histone modifications listed in the *P. berghei* ANKA gametocyte PTMs column were only found at this parasite stage and not in mature schizonts. Bold black text indicates that the PTM appeared in 2 out of 3 histone extract samples. Red and bold histone PTMs were observed from all three histone extract samples of the same parasite stage.

In total, 430 histone PTMs were identified from *P. berghei* ANKA schizonts and gametocytes together, 267 of which had been observed in *P. falciparum* and were shared with this parasite (Coetzee *et al.*, 2017). This *P. berghei* study identified 163 newly identified histone PTMs, 44 of which were unique to *P. berghei*, with 39 of these unique marks related to the H2B variant histone, H2B.Z. Importantly, H2A.Z/H2B.Z variant nucleosomes have been shown to be involved in the active transcription of *var* genes in *P. falciparum*, and PTMs to these histones could exert important effects in the maintenance of euchromatin (Hoeijmakers *et al.*, 2013; Petter *et al.*, 2013). Identification of histone PTMs on the H2B.Z variant of *P. berghei* ANKA is important to our

understanding of human-infective *Plasmodium* species because this variant histone is 100% identical to the H2B.Z histones of *P. knowlesi*, *P. vivax* Sal-1, and *P. ovale curtisi* GH01 **(Supplemental section S1; Figure S8).**

Of the 430 identified histone PTMs, 281 were observed in mature *P. berghei* schizonts, 197 of which had previously been determined in *P. falciparum*. From mature gametocytes, 323 histone PTMs were identified, 211 of which had previously been determined in *P. falciparum* (Coetzee *et al.*, 2017). Of particular interest in this study were the differential histone marks shown in **Table 4.3** above. These differential histone marks may provide clues as to the epigenetic regulatory proteins involved in commitment to gametocytogenesis. To identify histone PTMs that were heavily involved in sexual development, histone modifications that appeared in two or more samples are highlighted in **Table 4.3.** These histone PTMs are shown again in **Table 4.4**, though in this case, whether or not the histone PTM was unique to *P. berghei* is also indicated.

***P. berghei* ANKA schizont PTMs**		***P. berghei* ANKA gametocyte PTMs**	
H4	N-termAc, R3me2, K8me3	H4	K20me1, R23me1*, K59me1*
H3	R2ac*, R17me2, R26me2, K27me2, R42me1, R49ac*, R53ac*, K56ac, R63ac*	H3	R2me1, K36me1, Y41ph*, T45ph*, R116me1*
H3.3	S10ph, R17me2/me3, K18me3, K23me2*	H3.3	R2me1, R8me1, K36me1*
H2A	R32me1*, K95me2*	H2A	S1ph, K5ac, K9me1, K10ac, K35ac*, K41me1*
H2A.Z	K5ac, K27me1	H2A.Z	K37me1/me2, K155ac*
H2B	K7me2, K21me1**, S84ph**, T89ph**	H2B	K10me2, K19ac, T81ph**
H2B.Z	K18me3, R23me3, K104me1	H2B.Z	S9ph, R23me1, R95me1

Table 4.4: Repeat differential histone PTMs between asexual- and sexual stage *P. berghei* parasites. This table shows the histone PTMs that were unique to either *P. berghei* ANKA mature schizonts or *P. berghei* ANKA mature gametocytes, and appeared in at least 2 biological replicates. PTMs listed under each heading appeared at one life cycle stage and not the other. PTMs listed in black text were observed in two of three biological replicates. PTMs listed in red were observed in all three biological replicates. A red asterisk indicates that this histone PTM was not observed in *P. falciparum* previously. A black asterisk indicates that that particular amino acid position differs in *P. falciparum* and so is unique to *P. berghei*.

Of all differential histone PTMs identified in this study **(Table 4.3)**, one histone PTM consistently appeared in all schizont samples and was not present in mature gametocyte samples: H3.3 K18 trimethylation (H3.3K18me3). As with

the H2A.Z/H2B.Z variant nucleosome in *P. falciparum, Pf*H3.3 also appears to be a variant histone with specific activity, primarily demarcating euchromatic regions and subtelomeric repetitive sequences, including that of the active and poised *var* gene (Fraschka, Henderson and Bártfai, 2016). With regard to the specific trimethylation of lysine 18 on this histone (and not on the canonical H3); little is known besides the fact that H3.3 in eukaryotes appears to accumulate as cells age (Tvardovskiy *et al.*, 2017), and that acetylation of H3.3K18 is also an indicator of active transcription (McKittrick *et al.*, 2004). In *P. falciparum*, this histone modification was identified, though its stage-specificity was not monitored throughout the life cycle of the parasite (Coetzee *et al.*, 2017).

In *P. berghei* mature gametocytes, three differential histone PTMs appeared in all three biological replicate samples: H2BK19ac, H2B.ZS9ph, and H2B.ZR95me1 (**Table 4.4).** All three of these marks were identified in *P. falciparum* previously, though the stage-specificity of these marks was not commented upon, and these PTMs were not evaluated quantitatively or shown to be co-occurring with other histone PTMs (Coetzee *et al.*, 2017). The first gametocyte-specific mark, H2BK19ac, occurs at the *Plasmodium*-specific H2B lysine 19 with no obvious analogous lysine in either *Mus musculus, Rattus norvegicus,* or *Homo sapiens* (**Supplemental Section S1; Figure S7).** Both H2B.ZS9ph and H2B.Z95me1 are equally located at *Plasmodium*-specific amino acid residues, meaning that, at present, these gametocyte-specific marks have no known enzyme linked to their function and appear to be specific to *Plasmodium* spp.

Outside of these marks that appeared in all biological replicates, a further 27 histone PTMs were identified only at *P. berghei* mature schizont stage from two biological replicates, and a further 23 histone PTMs appeared solely in gametocyte samples in two biological replicates **(Table 4.4).** All of these histone PTMs and any known associations to biological function in eukaryotes are listed in **Table 4.5** below.

P. berghei ANKA schizont PTMs		**Function**	**P. berghei ANKA gametocyte PTMs**		**Function**
H4	N-term Ac	Regulator of arginine methylation and chromatin silencing in S. cerevisiae (Schiza et al., 2013).	**H4**	K20me1	A function of SET8 in other eukaryotes (the only mono-methyltransferase of H4K20); implicated in DNA damage response and cell cycle progression (Izzo and Schneider, 2010; van Nuland and Gozani, 2016).
	R3me3	Asymmetric dimethylation of H4R3 is a function of Hmt1 (*Pb*SET1) methyltransferase activity which associated with decreased N-terminal acetylation. This promotes rDNA silencing (Schiza et al., 2013).		R23me1	H4R23 is recognised by SET8 in other eukaryotes to facilitate H4K20 mono-methylation (Weirich et al., 2015). When R23 is mutated, SET8 affinity for H4 increased exponentially (Couture et al., 2005). Mono-methylation of H4R23 may therefore also facilitate mono-methylation of H4K20 by SET8.
	K8me3	Acetylation of H4K5 counteracts H4R3 methylation. As with H4K5me3, H4K8me3 should facilitate chromatin silencing as mediated by H4R3me3 (Feng et al., 2011).		K59me1	H4K59 is located at the solvent-accessible surface of the nucleosome and in *S. cerevisiae,* replacing this lysine with a different amino acid reduced heterochromatin stability and disrupted telomeric silencing at specific loci (Yu *et al.*, 2011). Substitutions to H4K59 also lead to hypersensitivity to DNA-damaging agents and the positive charge on an unmethylated H4K59 is necessary for effective ribosomal DNA silencing (Hyland *et al.*, 2005).
H3	R2ac	Unknown role in transcription; site is conserved across eukaryotes and bound by PHD12 zinc finger protein when unmodified (Qiu et al., 2012).	**H3**	R2me1	Transcriptional repression; antagonises H3K4me3 in other eukaryotes (Hyllus et al., 2007).

	R17me2	Asymmetric di-methylation of H3R17 is essential for mammalian DNA demethylation, i.e. transcriptional activation (Bauer et al., 2002; Di Lorenzo and Bedford, 2011; Hatanaka et al., 2017).		K36me1	The role of H3K36me1 is unclear. Trimethylation and dimethylation of H3K36 have been associated with both transcriptional activity and chromatin silencing mechanisms (Wagner and Carpenter, 2012; Suzuki *et al.*, 2016). In *P. falciparum,* H3K36me3 (but not H3K36me2) was associated with transcriptional silencing of *var* genes, with trimethylation mediated by the SET2 methyltransferase (Jiang et al., 2013). Unlike *P. falciparum,* the *P. berghei* ANKA genome does not encode for a SET2 protein.
	R26me2	Transcriptional activation (Wu et al., 2012).			
	K27me1	Transcriptional repression (Wiles and Selker, 2017).			
	R42me1	Transcriptional activation (Casadio et al., 2013).			
	R49ac	Unknown; conserved arginine in eukaryotes.			
	R53ac	Unknown; H3R53 mutations in model eukaryotic systems result in defects in telomeric and rDNA silencing (Mersfelder and Parthun, 2006).			
	K56ac	In eukaryotes, H3K56 acetylation increases the binding affinity of CAF-1 and Rtt106 chromatin assembly proteins to package DNA into nucleosomes during DNA replication (Li *et al.*, 2008). H3K56 acetylation in yeast acts as a marker for new nucleosome synthesis during DNA replication and repair (Stejskal *et al.*, 2015).			

	R63ac	This conserved eukaryotic arginine lies at the lateral surface of the histone octamer next to H3K64, which, when acetylated, marks transcriptionally active chromatin (Di Cerbo *et al.*, 2014). No studies of R63 acetylation alone have so far been carried out.			
				Y41ph	Transcriptional activation (Dawson et al., 2009).
				T45ph	Associated with DNA replication (Baker, 2010)
				R116me1	This conserved eukaryotic arginine is located within the L1L2 loop at the nucleosome dyad, and a substitution of this residue is lethal to *S. cerevisiae*. H3R116 forms a salt bridge with H3D123, stabilising the L1L2 loop (Hainer and Martens, 2011). The specific effects of mono-methylation of this amino acid are unknown.
H3.3	S10ph	Phosphorylation of H3.3 serine 10 occurs during the prophase of mitosis in eukaryotes and, alongside H3.3S28ph, these marks localise to the outermost peripheral region of condensed DNA (Hake *et al.*, 2005). Phosphorylation of H3.3S10 is a function of 'checkpoint kinase 1' (ChK1) in other eukaryotic cell lines, with H3S10ph playing a role in chromosome condensation and transcriptional activity (Liokatis *et al.*, 2012; Chang *et al.*, 2015).	**H3.3**	R2me1	Transcriptional repression (as in H3R2me1) (Hyllus *et al.*, 2007).
	R17me2/me3	Possible transcriptional activation (as in H3R17me2).		R8me1	Possibly transcriptional repression; may be associated with decreased acetylation of H3K9 (Schwämmle et al., 2016).

	K18me3	Active transcription (McKittrick *et al.*, 2004).		36me1	Unclear; as in H3K36me1.
	K23me2	Unknown; conserved arginine in eukaryotes. Trimethylation of H3K23 in eukaryotes is associated with protection from DNA damage in pericentric heterochromatin during meiosis (Papazyan *et al.*, 2014).			
H2A	R32me1	Unknown; conserved arginine in eukaryotes.	**H2A**	S1ph	In eukaryotic cell lines, H2AS1ph is a function of mitogen- and stress-induced kinase 1 (MSK1), also known as ribosomal protein S6 kinase A5 (RPSKA5) in human cells. Phosphorylation of serine 1 on H2A is negatively associated with transcription, making MSK1 an inhibitor of transcription (Zhang *et al.*, 2004), In *Xenopus,* H2A and H4 S1 phosphorylation events are highly enriched at distinct points in embryo development: H2AS1ph is associated with repression of zygotic gene expression, while H4S1 phosphorylation correlates with the beginning of maternal gene expression (Wang *et al.*, 2014).
				K5ac	Transcriptional activation (associated with gene promoters) (Rajagopal *et al.*, 2014). Associated with early response to DNA damage (Ikura *et al.*, 2007).
				K9me1	Unknown; conserved lysine in eukaryotes.
				K10ac	Unknown; conserved lysine in eukaryotes.

				K35ac	Unknown; *Plasmodium*-specific lysine.
	K95me2	Unknown; K95 conserved across all eukaryotes.			
				K41me1	Unknown; *Plasmodium*-specific lysine.
H2A.Z	K5ac	Transcriptional activation (Draker et al., 2012)	**H2A.Z**	K37me1/me2	Unknown; K37 is conserved across eukaryotes and is located at the nucleosome α1-helix (Liang *et al.*, 2016).
	K27me1	Unknown; *Plasmodium*-specific lysine.		K155ac	Unknown; *Plasmodium*-specific lysine.
H2B	K7me2	Unknown; equivalent K6 in higher eukaryotes is suggested in transcriptional activation when acetylated (Parra et al., 2006).	**H2B**	K10me2	Unknown; *Plasmodium*-specific lysine.
				K19ac	Unknown; *Plasmodium*-specific lysine.
				T81ph*	Unknown; *P. berghei*-specific threonine.
	K21me1*	Unknown; *P. berghei*-specific lysine.			
	S84ph*	Unknown; *P. berghei*-specific serine.			
	T89ph*	Unknown; *P. berghei*-specific threonine.			
H2B.Z	K18me3	Unknown; *Plasmodium*-specific lysine.	**H2B.Z**	S9ph	Unknown; *Plasmodium*-specific serine.
	R23me3	Unknown; *Plasmodium*-specific arginine.		R23me1	Unknown; *Plasmodium*-specific arginine.
	K104me1	Unknown; *Plasmodium*-specific lysine.		R95me1	Unknown; P*lasmodium*-specific arginine.

Table 4.5: Summary of the roles of stage-specific histone modifications between *P. berghei* ANKA schizonts and gametocytes (current knowledge). This table lists all histone modification differences that were identified between *P. berghei* ANKA mature asexual-stage parasites (schizonts) and mature sexual-stage parasites (gametocytes) in at least two biological replicates next to their known functions in eukaryotes. An asterisk (*) next to a histone modification indicates that this particular amino acid is unique to *P. berghei,* or that the corresponding amino acid in *P. falciparum* is at a different position of the histone amino acid chain. Shared histone post-translational modifications (PTMs) and those that did not appear in two or more biological replicates are not shown. All histone modifications listed underneath the *P.*

berghei ANKA schizont column were seen only in schizonts and not gametocytes. Conversely, all histone modifications listed in the *P. berghei* ANKA gametocyte PTMs column were only found at this parasite stage and not in mature schizonts. References are provided next to the reported function of each histone modification.

4.3.6 Confirming a number of PTMs by immunoblotting

In addition to determining the histone modification profiles for mature asexual (schizont) and mature sexual (gametocyte) *P. berghei* ANKA parasites, the present study aimed to identify differences in the epigenetic marks underlying asexual and sexual development, and to take a number of these histone modifications forward into chromatin immunoprecipitation experiments with the aim of identifying genes of DNA loci that are regulated by these histone modifications.

To confirm a number of histone modifications that were identified in this study, in combination with a number of histone modifications identified as having a role in regulating gene expression in *P. falciparum,* commercial chromatin immunoprecipitation (ChIP)-grade antibodies were obtained for H3K122ac (Abcam; Cat. No. ab33309), H4K8ac (Merck Millipore; Cat. #07-328; Lot #2301651), H3K9ac (Merck Millipore; Cat. #07-352; Lot #2658782), H3K9me3 (Merck Millipore; Cat. #07-442; Lot #2664282), H3K27me3 (Merck Millipore; Cat. #07-449; Lot #2653203), H3K36me3 (Abcam; Cat. No. ab9050), H3K4me3 (Merck Millipore; Cat. #17-614; Lot #2618837), and H3K64me3 (Abcam; Cat. No. ab195483). All antibodies were tested against acid-extracted histones from the same G1142 schizont and gametocyte samples. Images of Western blots in which antibody-binding could be discerned are shown in **Figure 4.7** (the distribution of acid-extracted histones on an SDS-PAGE gel from these samples is also provided in this figure).

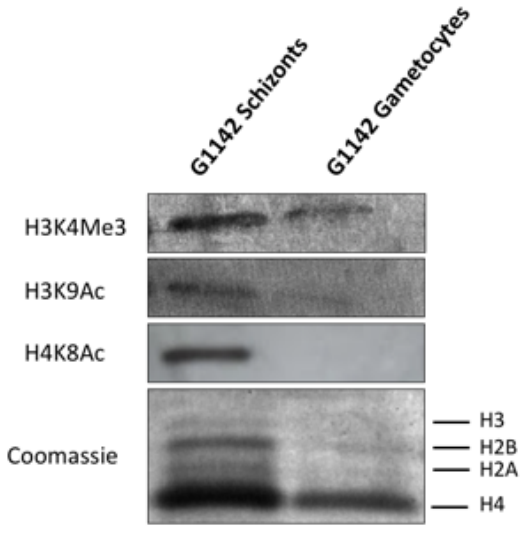


Figure 4.7: Immunoblotting to confirm a number of histone PTMs for chromatin immunoprecipitation experiments. This image shows the combined results of three successful

Western blots and a Coomassie-stained 15% SDS-PAGE gel showing the distribution of histone in the same schizont and gametocyte samples. The top panel shows the results of a H3K4me3 commercial antibody, confirming the presence of this histone modification in both asexual and sexual stage *P. berghei* ANKA parasites (this modification was detected by LC-MS/MS in both asexual- and sexual-stage parasites (see figures **4.5** and **4.6**)). Below that, the results of H3K9ac antibody-staining is shown, showing a stronger presence of this modification in asexual stages when compared to sexual-stage parasites. Below H3K9ac-staining is shown the H4K8ac commercial antibody staining of schizonts and gametocytes, with a strong signal seen in asexual parasites, and no signal observed in gametocytes.

After determining their ability to bind *P. berghei* histones (**Figure 4.7**), commercial antibodies for H4K8ac and H3K9ac were chosen for ChIP experiments. H4K8 acetylation would also later be identified as playing an important regulatory role in the intraerythrocytic developmental cycle of *P. falciparum* (Gupta *et al.*, 2017). Both H4K8ac, H3K9ac were detected in both asexual and sexual-stage *P. berghei* ANKA parasites by LC-MS/MS, with H3K9me3 detected in only one gametocyte replicate by LC-MS/MS (**Figure 4.5** and **Figure 4.6)**. Despite only being observed once, the H3K9me3 antibody was taken forward for ChIP experiments to complement the opposing H3K9ac ChIP results, in line with previous ChIP-on-ChIP and ChIP-sequencing experiments undertaken in *P. falciparum* (Cui *et al.*, 2007; Salcedo-Amaya *et al.*, 2009; Karmodiya *et al.*, 2015).

Though low in abundance, and undetectable by Western blot, the identification of H3K122ac as a potential differential histone modification between asexual and sexual *P. berghei* parasites (only seen in one gametocyte sample, with H3K122me1 in one schizont sample (**Table 4.3)**), and its discovery as a mark of active enhancers and transcriptional activation (Tropberger *et al.*, 2013; Pradeepa *et al.*, 2016) compelled a study of ChIP using H3K122ac to be undertaken. It should also be noted that a study in which this commercial H3K122ac antibody was used in *Drosophila melanogaster* experiments, demonstrated non-specific binding of the antibody *in vivo* (Graves *et al.*, 2016). The authors of this study suggest that acetylation of H3 residues outside of K122 may be sufficient for antibody binding in metazoans, though in later studies, issues with H3K122ac antibody cross-reactivity were not reported (Devaiah *et al.*, 2016; Pradeepa *et al.*, 2016; Assiri *et al.*, 2017).

To conclude, four commercial antibodies corresponding to histone modifications; H4K8ac, H3K9ac, H9K9me3, and H3K122ac, were used to determine the global distribution of these epigenetic marks throughout the mature asexual (schizont) genome in *P. berghei* by both ChIP-sequencing and ChIP-sequencing coupled to library preparation using Tn5 transposase (ChIPmentation) (**Chapter 5**).

4.4 Discussion

The key findings from this chapter were:

- 430 histone PTMs were identified from both *P. berghei* schizonts and gametocytes.
- 182 histone PTMs were shared between both asexual- and sexual- stage *P. berghei* parasites.
- 248 differentially-expressed histone PTMs were identified; 104 PTMs unique to schizonts and 144 PTMs unique to gametocytes.
- 28 schizont-specific histone PTMs were identified from 2 or more of three biological replicates.
- 26 gametocyte-specific histone PTMs were identified from 2 or more of three biological replicates.
- H3.3K18 trimethylation was enriched in all *P. berghei* schizont samples.
- H2BK19ac, H2B.ZS9ph, and R95me1 were all enriched only at mature gametocyte stages.

4.4.1 Differential histone PTMs and potential epigenetic regulatory enzymes in *P. berghei*.

Just as separation of male and female gametocytes and analysis of their respective proteomes can provide an insight into the proteins involved in sexual development and sex-specificity in *Plasmodium* (Khan *et al.*, 2005), determination of sex-specific histone PTMs can provide information about the nucleosome landscape and epigenetic regulators involved in sexual commitment and maturation of gametocytes. In the present study, the focus was to identify histone PTMs that may be involved in gametocytogenesis in *P. berghei*, and, using either enrichment by sulfadiazine treatment, or the use of a parasite line in which the promoter of

the *ap2-g* transcriptional regulator was irreversibly induced by rapamycin administration (section **4.3.2**), gametocytes were selected and subjected to acid extraction of histones and LC-MS/MS to determine histone PTMs. Twenty-six differential histone PTMs were discovered in *P. berghei* gametocytes (shared between 2 or more biological replicates): 5 acetyl marks, 14 methyl marks, 2 dimethyl marks, and 5 phosphoryl marks. A summary of known associations of these particular histone modifications can be found in **Table 4.5.** However, of considerable interest, in addition to merely mapping the histone PTM landscape, was to identify potential epigenetic regulatory enzymes that may be responsible for deposition of these marks, and therefore, may play a role in gametocytogenesis in *P. berghei*. Listed in **Table 4.6,** are the 26 histone modifications, known enzyme responsible for their presence in other eukaryotes, and enzymes that share homology with known epigenetic regulators outside of *Plasmodium* spp. In addition, where a potential homologous enzyme is determined in *P. berghei* ANKA, their shared identity to the known protein, and any stage specificity of the transcript of this gene are shown adjacent to the result.

***P. berghei* ANKA gametocyte-specific PTMs**		**Regulatory enzyme (if known)**	**PlasmoDB IDs of homologous *P. berghei* genes**	**Stage specificity of gene/protein (if known)/Notes**
H4	K20me1	PR-Set7/SET8/KMT5A (Beck *et al.,* 2012)	PBANKA_1001600 (SET8)	Dispensable in asexual stages; greater transcription in asexual stages and female gametocytes.
	R23me1	Unknown	--	--
	K59me1	Unknown	--	--
H3	R2me1	Unknown	--	Symmetric demethylation of H3R2 is a function of PRMT5 and PRMT7 (Migliori *et al.,* 2012).
	K36me1	SET2 (Xu *et al.,* 2007)	No SET2/PBANKA_1128500 (SET1) (53/143 (37%) sequence identity)	No SET2 gene in *P. berghei* or other rodent *Plasmodium* spp. SET1 had greatest sequence homology to *Pf*SET2 but is indispensable in *P. berghei.*
	Y41ph	JAK2 (Dawson *et al.,* 2009)	PBANKA_0616700 (NEK4) (61/204 (30%) sequence identity)	Nima-related kinase 4 (NEK4); This is transcribed largely in female gametocytes, is dispensable for asexual development, but it essential for zygote development (Reininger *et al.,* 2005).
	T45ph	AKT1 (Lee *et al.,* 2015)	PBANKA_1460000 (PKB) (159/336 (47%) sequence identity)	Putative RAC-beta serine/threonine protein kinase (PKB); Indispensable in asexual stages; proteomic evidence from male gametocytes (*Plasmo*DB).
	R116me1	Unknown	--	--
H3.3	R2me1	Unknown	--	Possibility of PRMT7 methylation of H3 and H3.3 H3R2 (Hyllus *et al.,* 2007). Nearest *P. berghei* homologue was PRMT1 (PBANKA_1018400).
	R8me1	Unknown	--	Dimethylation of H3R8 in other eukaryotes is a function of PRMT5. *P. berghei* PRMT5 homologue is PBANKA_1137100.
	K36me1	SET2 (Xu *et al.,* 2007)	No SET2/PBANKA_1128500 (SET1) (53/143 (37%) sequence identity)	No SET2 gene in *P. berghei* or other rodent *Plasmodium* spp. SET1 had greatest sequence homology to *Pf*SET2 but is indispensable in *P. berghei.*

H2A	S1ph	MSK1 (Wang *et al.,* 2014)	PBANKA_1460000 (PKB) (152/318 (48%) sequence identity)	Putative RAC-beta serine/threonine protein kinase (PKB); Indispensable in asexual stages; proteomic evidence from male gametocytes (*Plasmo*DB).
	K5ac	HAT1 (Tafrova and Tafrov, 2014)	PBANKA_0718400 (HAT1)	Dispensable in asexual growth, reduced female gametocyte development (this study).
	K9me1	Unknown	--	--
	K10ac	ESA1 (Keogh *et al.,* 2006)	PBANKA_0929500 (MYST) (157/399 (39%) sequence identity)	Indispensable for asexual growth; recruited to active *var* gene promoter in *P. falciparum* (Miao *et al.,* 2010).
	K35ac	Unknown	--	--
	K41me1	Unknown	--	--
H2A.Z	K37me1	Unknown	--	--
	K37me2	Unknown	--	--
	K155ac	Unknown	--	--
H2B	K10me2	Unknown	--	--
	K19ac	Unknown	--	--
	T81ph*	Unknown	--	--
H2B.Z	S9ph	Unknown	--	--
	R23me1	Unknown	--	--
	R95me1	Unknown	--	--

Table 4.6: Gametocyte-specific histone PTMs and potential regulatory enzymes based on homology. This table listed all gametocyte-specific histone PTMs that were identified in 2 or more biological replicates. The histone type is listed on the left, followed by the histone modification written in standard Brno nomenclature (Turner, 2005). Any enzymes known to deposit the histone modification in question, in any eukaryote, is listed to the right of the modification with references. Any *P. berghei* ANKA protein with homology to the known epigenetic regulator (based upon a blastp search of the enzyme's amino acid sequence) is listed to the right of the enzyme, with the sequence identity for the nearest *P. berghei* homologue in brackets. The sequence identity was the number of amino acids that were exactly identical between the query sequence and the *P. berghei* protein, taking no gaps into account, and with the total number based on the shorter of two sequences. Notes and possible stage-specificity of gene transcripts or proteins from previous data are detailed in the furthest right-hand column.

In total, only 6 enzymes were identified that may be responsible for these differential gametocyte-specific PTMs (based on current knowledge). In the absence of SET2 in rodent lineages of the *Plasmodium* genus (including *P. berghei*), SET1 is the nearest homologous methyltransferase enzyme that may be responsible for lysine 36 mono-methylation on both canonical histone H3, and variant histone H3.3 **(Table 4.6)**, though the *P. berghei* ANKA genome encodes for eight other SET domain-containing methyltransferases **(Supplemental Table S1),** and none of these enzymes can be ruled out as possible regulators of the H3/H3.3 K36me1 modification.

With the H4K20me1 histone modification, much more information is available from studies in other eukaryotic cell lines and model organisms. H4K20 is a conserved eukaryotic lysine that can be mono-, di-, or tri-methylated, with each methylation state being representative of a different point in mitosis. H4K20 mono-methylation occurs at the initial mitotic (M) phase of mitosis in dividing

cells, and only one mono-methyltransferase of H4K20 is known: SET8 (Evertts *et al.*, 2013; Jørgensen, Schotta and Sørensen, 2013). SET8 and H4K20 methylation are crucial for maintaining chromatin compaction during the transition from M to Gap 1(G1) phases of mitosis (Evertts *et al.*, 2013; Shoaib *et al.*, 2018). Mono- and di-methylated H4K20 are involved in DNA replication and DNA damage repair respectively, with tri-methylation of H4K20 marking silenced heterochromatic regions (Jørgensen, Schotta and Sørensen, 2013).

In *P. falciparum*, a ChIP-on-chip study identified H4K20me1 as being positively associated with transcription and occupancy at open reading frames (ORFs) during the intra-erythrocytic asexual cycle, though gametocytes were not examined during this study (Gupta *et al.*, 2013). In all other studies in *Plasmodium* however, methylation of H4K20 is associated with heterochromatin, though SET8 loss does not disrupt *var* gene silencing (Duffy *et al.*, 2013). SET8 has mono-, di-, and tri-methylation capabilities (Sautel *et al.*, 2007) and is dispensable during the *P. berghei* life cycle, with disruption effecting sexual stages, and transcription comparatively upregulated in asexual stages parasites and female gametocytes (Bushell *et al.*, 2017; Yeoh *et al.*, 2017). These results warrant further elaboration, and an assay of *P. berghei* SET8 methyltransferase activity, taking the full protein or the methyltransferase catalytic core, with unmethylated H4 as a substrate, could easily determine the specific activity of SET8 (Sautel *et al.*, 2007). In addition, immunoblotting with a H4K20me1 antibody or fluorescent tagging of SET8 would add to our knowledge of this enzymes stage-specificity, and the association of H4K20me1 with gametocytogenesis.

Certain phosphorylation modifications also differ between asexual and sexual stage *P. berghei* histones. Dual phosphorylation of tyrosine 41 (H3Y41ph) and threonine 45 (H3T45ph) only occurred at the gametocyte stage, as did phosphorylation of H2A serine 1 (H2AS1ph), H2B threonine 81 (H2BT81ph), and H2B.Z serine 9 (H2B.ZS9ph). Though both H2BT81 and H2B.ZS9 were *P. berghei* and *Plasmodium* spp.-specific amino acids respectively, and therefore their regulatory enzymes remain unknown, that was certainly not the case for H3Y41ph, H3T45ph, and H2AS1ph. In a study of human peripheral blood stem cells, Janus Kinase 2 (JAK2) was shown to directly phosphorylate H3Y41, a conserved eukaryotic tyrosine (Dawson *et al.*, 2009). The direct result of this histone PTM

was to block heterochromatin protein 1 α (HP1α), but not human HP1β, from binding to this region of H3. In *Plasmodium* spp. parasites, only one HP1 protein is encoded for, and its regulation is essential for parasite asexual and sexual development (Fraschka *et al.*, 2018). Remarkably, in *P. falciparum,* regulation of sexual commitment by the *ap2-g* transcription factor is dependent upon removal of HP1 from the *ap2-g* locus, a function of gametocyte development 1 (GDV1) protein (Filarsky *et al.*, 2018). However, as with SET2, the GDV1 protein is not encoded for in rodent *Plasmodium* spp. genomes. In addition, the trigger for GDV1 activation remains unknown.

However, removal of HP1α from H3 upon phosphorylation of H3Y41 by JAK2 may be a conserved process in eukaryotes. Taking the JAK2 amino acid sequence and comparing it to the protein sequences of *P. berghei* (full blastp results not shown), the closest homologous protein sequence belonged to Nima-related kinase 4 (NEK4), a kinase that is indispensable for sexual development in *P. berghei* (Reininger *et al.*, 2005). In a complete KO of NEK4 from *P. berghei,* gametocyte and gamete development occurred as normal, though measurements of nuclear DNA content indicated that in the zygote, the point of DNA replication from two fused haploid genomes (microgamete and macrogamete haploid genomes) to a tetraploid ookinete nucleus was inhibited (Reininger *et al.*, 2005). This phenotype also appeared to originate from a developmental issue in the female macrogamete, with male *nek4*$^{-}$ microgametes capable of fertilising WT macrogametes, with subsequent zygote and ookinete development. Despite pinpointing the point at which sexual development fails, this study could not determine the exact function of NEK4. It should therefore be investigated whether NEK4 functions in the phosphorylation of H3Y41, and that this process effects HP1 binding downstream.

In addition to H3Y41 phosphorylation, H3T45 phosphorylation was also identified in mature gametocytes but not in mature schizonts of *P. berghei*. Like H3Y41, H3T45 is located at a point of DNA entry/exit into the nucleosome (Tropberger and Schneider, 2010), and phosphorylation of this amino acid in human cells antagonises HP1-mediated transcriptional repression (Jang *et al.*, 2014). In other eukaryotic cells, H3T45 phosphorylation was mediated by a dual specificity tyrosine phosphorylation regulated kinase 1A (no homologue in *P. berghei*), AKT

serine threonine kinase 1 (AKT1) (homologous to PKB (PBANKA_1460000) in *P. berghei*), and human protein kinase C δ (PKCδ) (PKB/PBANKA_1460000 was closest homologous gene in *P. berghei*) (Baek, 2011; Jang *et al.*, 2014; Lee *et al.*, 2015). In *P. berghei,* the homologous protein kinase B (PKB) was not amenable to KO, was shown to be essential to parasite growth in *P. falciparum* (Kumar *et al.*, 2004), and was picked up in a proteomic study in male gametocytes (in comparison with female gametocytes) (Khan *et al.*, 2005). Together, phosphorylation of both H3Y41 and H3T45 likely alter DNA accessibility at the αN loop of H3 within the nucleosome, leading to DNA unwrapping (Bowman and Poirier, 2015). If the effect of these PTMs on nucleosome dynamics hold true in *P. berghei,* the enzymes that regulate these modifications may play a crucial role in DNA transcription in gametocytogenesis.

One final gametocyte-specific phosphorylation mark, H2AS1ph, is again a function of a protein in which the nearest homologous gene was PKB (PBANKA_1460000) in *P. berghei* **(Table 4.6).** In multiple eukaryotes, dual H2A/H4 serine 1 phosphorylation was shown to be indicative of chromosome condensation during mitosis (Barber *et al.*, 2004). The roles of these histone PTMs were further elucidated using *X. laevis* eggs and embryos. In *X. laevis,* H2AS1ph was highly enriched in blastula chromatin during repression of zygotic gene expression, with the beginning of maternal gene expression marked by H4S1 phosphorylation (Wang *et al.*, 2014). With H2S1ph being identified differentially in *P. berghei* gametocytes, it is possible that the role of this mark is conserved. In *Plasmodium* spp., mRNAs responsible for zygote development have been shown to be translationally repressed by DOZI-CITH ribonucleoprotein complexes (Vembar, Droll and Scherf, 2016), and so the presence, in gametocytes, of a histone PTM associated with repression of zygotic gene expression is not unexpected. In *P. berghei,* LC-MS/MS identified H4S1ph at both schizont and gametocyte stages **(Table 4.1** and **Table 4.2),** and so combinatorial H2AS1/H4S1 phosphorylation is present in *P. berghei* gametocytes. In yeast, in response to DNA double-strand breaks (DSBs), H4S1 is phosphorylated by casein kinase II (CK2) (Cheung *et al.*, 2005), sporulation-specific protein 1 (SPS1) (Krishnamoorthy *et al.*, 2006), or both.

The *P. falciparum* genome also contains one protein kinase CK2 alpha subunit orthologue (PF3D7_1108400; PBANKA_0938600 in *P. berghei* ANKA) and two CK2

beta subunit orthologues, *Pf*CK2β1 and, *Pf*CK2β2 (PF3D7_1103700 and PF3D7_1342400; PBANKA_0943200 and PBANKA_1355500 in *P. berghei* ANKA) with roles in both asexual parasite development and chromatin assembly (Holland *et al.*, 2009; Dastidar *et al.*, 2012). The essentiality of CK2 has resulted in its catalytic alpha subunit becoming a promising target for novel antimalarial therapies, though a possible role for this protein in *Plasmodium* gametocytogenesis remains undetermined (Ruiz-Carrillo *et al.*, 2018).

Remarkably, taking the SPS1 gene of *S. cerevisiae* and comparing it to the kinases of *Plasmodium spp.* resulted in two potential SPS1 orthologues: Nima-related kinase 2 (NEK2) (PBANKA_1240700) and Nima-related kinase 4 (NEK4) (PBANKA_0616700) (blastp results not shown). Studies in both *P. falciparum* and *P. berghei* have already determined that NEK2 is not essential for gametocytogenesis, but is required for ookinete development (Reininger *et al.*, 2009). As previously discussed regarding phosphorylation of H3Y41, NEK4 in *P. berghei* ANKA has been shown to be essential for the correct replication of zygotes into the tetraploid form that precedes meiosis, with knockout parasites incapable of developing to ookinete stage (Reininger *et al.*, 2005). Considering the role of yeast SPS1 in the maturation of four haploid nuclei within a spore, it is possible that NEK4 in *Plasmodium spp.* shares the orthologous role of this enzyme (Govin *et al.*, 2010). As such, an assay of NEK4 phosphorylation activity is warranted, and a determination of this enzymes preference for phosphorylation of either H4S1 or H3Y41, or both, could be examined (Peck, 2006).

In addition to the possible roles of *P. berghei* PKB (PBANKA_1460000), NEK4 (PBANKA_0616700), SET1 (PBANKA_1128500), and SET8 (PBANKA_1001600) in regulating gametocyte-specific PTMs that are essential for progression of sexual development, two more gametocyte-specific PTMs at the H2A N-terminal tail, K5ac and K10ac, are likely the product of *Pb*HAT1 and *Pb*MYST (**Figure 4.6**). In human cell lines, H2AK5 acetylation was shown *in vivo* to be a result of HAT1 activity (Tafrova and Tafrov, 2014). In *P. berghei*, the *Pb*HAT1 amino acid sequence had greatest sequence homology to the human HAT1, and, as demonstrated in this study (section **3.3.6**), a complete KO of this gene resulted in a significant decrease in female gametocyte conversion, though gametocytogenesis was not inhibited entirely (suggesting that HAT1 function is

redundant to some extent). With regard to acetylation of K10 on H2A, studies in *S. cerevisiae* have identified the histone acetyltransferase ESA1 as catalysing H2AK5 acetylation as part of the NuA4 HAT complex in this species (Keogh *et al.*, 2006; Mehta *et al.*, 2010). When searching for amino acid homology, *Pb*MYST was the protein with nearest sequence identify (**Figure 4.6;** blastp results not shown). To test these hypotheses, *in vitro* histone methyltransferase assays (Fingerman, Du and Briggs, 2008) could be carried out, or in the case of *Pb*HAT1, histone extraction from a *pbhat1*$^{-}$ line generated in this study (**Chapter 3)** may be used to identify missing or reduced H2AK5Ac presence from mature gametocyte histones.

In summary, a number of leads have been determined in the identification of potential epigenetic regulatory enzymes that play a role in gametocytogenesis, based upon gametocyte-specific histone PTMs and known eukaryotic mediators of these histone PTMs. Though these findings are potentially important in discovering an enzyme that modulates commitment to gametocytogenesis, the histone PTMs identified in this study alone provide an insight into the conformation of the nucleosome in mature *P. berghei* schizonts and mature *P. berghei* gametocytes. These marks themselves can influence DNA accessibility and monoclonal antibodies to these histone PTMs can be used to immunoprecipitate genes or loci that are associated with these histone modifications (Collas, 2010).

4.4.2 Sex-specific histone PTMs and their effect on the nucleosome

Genome-wide nucleosomal maps demonstrate that gene ORFs are nucleosome-dense, and so changes to nucleosome architecture, such as changes resulting from histone PTMs, can greatly affect gene transcription (Campos and Reinberg, 2009). Each histone alone has a histone fold domain that consists of three alpha helices (α1-α3) that are linked together by short loops (L1 and L2) that facilitate heterodimerisation of histone dimers (H2A with H2B, and H3 with H4) (Hammond *et al.*, 2017). Following oligomerisation of H2A-H2B and H3-H4 dimers to form a complete nucleosome, the H3-H4 tetramer occupies the central 'dyad' of the nucleosome, with two H2A-H2B dimers forming entry and exit points for nucleosome-bound DNA. In addition to α1-α3 helices, H2A contains two short helices, H2B contains one C-terminal helix, and H3 harbours an N-terminal helix (H3αN) (Hammond *et al.*, 2017) (**Figure 4.8 C** and **D)**. A number of schizont- and

gametocyte- specific histone PTMs identified in this study were located at positions of DNA-histone and histone-histone contact that could influence nucleosome dynamics (Bowman and Poirier, 2015). These histone PTMs are shown in **Figure 4.8 A** and **B.**

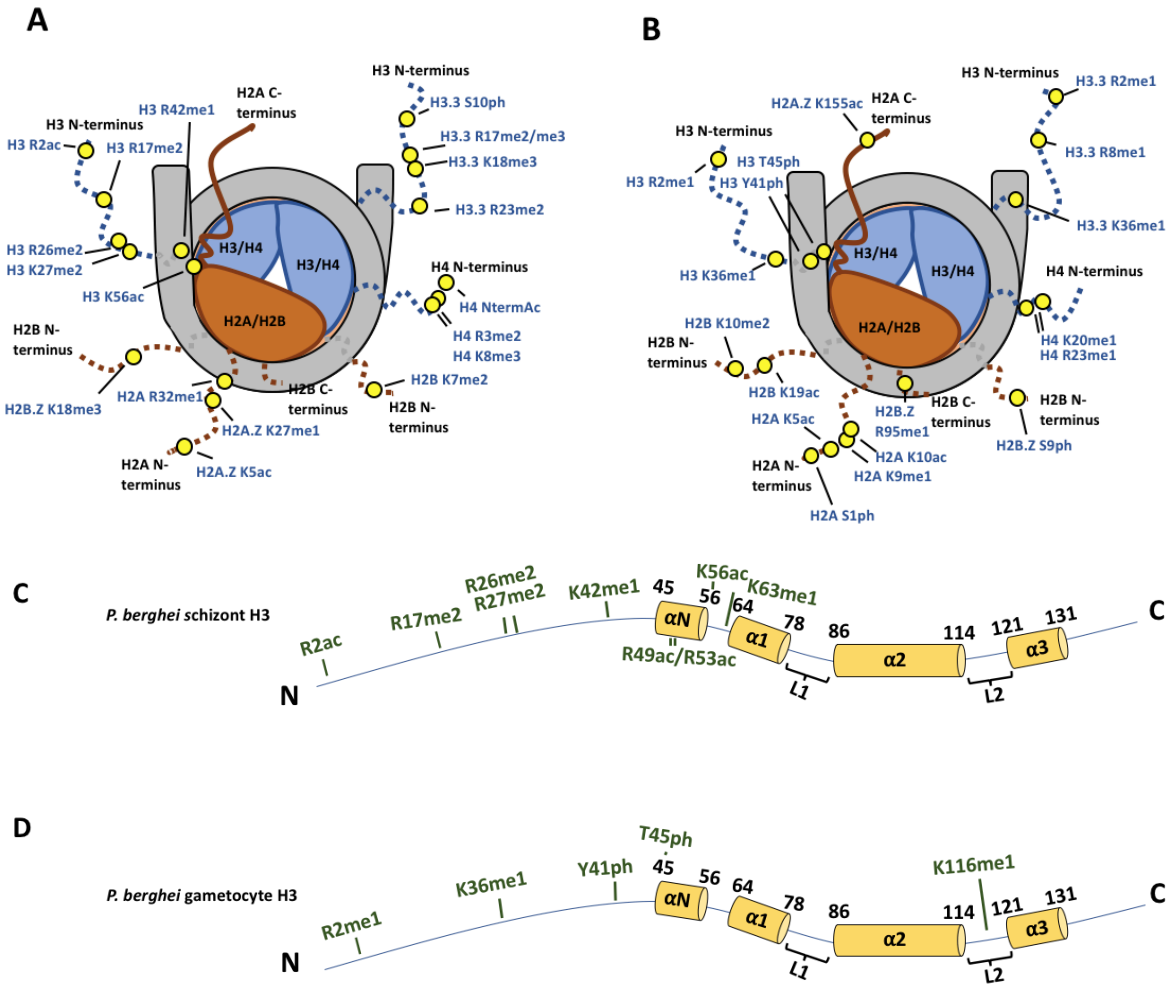


Figure 4.8: Overview of *P. berghei* ANKA schizont and gametocyte nucleosomes with stage-specific histone modifications. Images **A** and **B** are cartoon representations of *P. berghei* ANKA schizont and gametocyte nucleosomes respectively. In both **A** and **B**, stage-specific modifications identified using LC-MS/MS in the present study are indicated in dark blue text. All histone modifications on N- and C-terminal tails, and those with a putative role in regulating nucleosome dynamics are shown. In this figure, the H3/H4 heterodimer represents all canonical and variant H3 and H4 combinations. The H2A/H2B heterodimer cartoon also represents all combinations of canonical and variant H2A and H2B histones. Histone modifications however are labelled as to whether they were present in the canonical or variant histone type. This cartoon representation of nucleosomes and their histones was drawn in the style of Figure 2 (B) from Bowman and Poirier, 2015. Images **C** and **D** are cartoon representations of the histone modification changes observed in *P. berghei* schizont and gametocyte H3. In each histone, the N-terminus is indicated with an 'N' and the C-terminal end of the histone is indicated with a 'C'. The positions of all histone modifications that differed between schizont and gametocyte H3s are indicated in dark green text.

The structurally important loop 1 and 2 domains of histones are labelled L1 and L2 respectively. The α-helical domains of H3 histones are indicated in yellow and labelled N-terminal α-helix (αN), α1, α2, and α3, based on both recombinant and human histone structural data (Luger *et al.*, 1997; Iwasaki *et al.*, 2013). Amino acid positions that mark the beginning and end of α-helical domains are numbered next to the α-helices in bold text.

Figure 4.8 A and **B** illustrate the locations of schizont- and gametocyte-specific *P. berghei* histone PTMs that are known to influence nucleosome dynamics in eukaryotes (Bowman and Poirier, 2015). Though only 28 histone PTMs were identified as schizont-specific in 2 or more biological replicates, with 26 differential histone PTMs in gametocyte histones, many of these modifications appear at nucleosome DNA entry/exit points, particularly, the histone 3 N-terminal tail and the unique αN helix. The positions are again visualised in **Figure 4.8 C** and **D,** with **Figure 4.8 C** showing the location of schizont-specific histone PTMs and **Figure 4.8 D** showing the location of gametocyte-specific H3 PTMs.

Using the *X. laevis* model, mutation of any residues from H3K36 to H3S57 resulted in nucleosome sliding; a process in which nucloeosomal DNA accessibility is increased and the histone octamer is relocated to adjacent DNA segments (Becker, 2002; Ferreira *et al.*, 2007). In mature schizonts (**Figure 4.8 A** and **C**), the H3 αN helix is flanked and marked by 4 histone PTMs: H3K42me1, H3K49ac, H3K53ac, and H3K56ac. Of these marks, H3K56 acetylation has been extensively studied, and acetylation of this amino acid considerably increases nucleosome unwrapping and alters H3/H4 binding to histone chaperones (Bowman and Poirier, 2015). H3K56 acetylation leads to unwrapping of DNA from the histone octamer and is a marker of DNA replication (Stejskal *et al.*, 2015).

Conversely, in the gametocyte nucleosome (depicted in **Figure 4.8 B**), the histone 3 residues associated with chromatin accessibility (H3K36 to H3S57) in other eukaryotes are marked by three different histone PTMs: H3K36me1, H3Y41ph, and H3T45ph (also depicted in **Figure 4.8 D**). H3K36 is located in the H3 N-terminal tail just ahead of the DNA entry/exit site of the nucleosome, though H3K36 itself does not interact directly with DNA. Instead, modification of H3K36, particularly trimethylation of this mark, is associated with binding of the Tudor domain-containing PHD (plant homeodomain) finger protein 1 (PHF1), with results in downstream gene silencing (Musselman *et al.*, 2013). The effect of H3K36 mono-

methylation remains unclear. However, the adjacent H3T45 residue makes direct hydrogen bonds with the DNA backbone, with phosphorylation of both this mark and H3Y41 expected to result in DNA unwrapping from the histone octamer and exclusion of HP1 from cells (Dawson *et al.*, 2009; Bowman and Poirier, 2015). In human cell lines, phosphorylation of H3Y41 appears tissue-specific (Dawson *et al.*, 2012).

To summarise, mature asexual (schizont) stage and mature sexual (gametocyte) stage *P. berghei* ANKA parasites bear distinct histone PTMs **(Figure 4.8).** The extensively studied histone H3, and PTMs at the N-terminal domain and αN helix of this histone, suggests differential nucleosome-DNA binding at this point between these *P. berghei* life cycle stages. Though in both cases, histone marks on H3 are largely indicative of nucleosome unwrapping, the histone-DNA interactions at these stages may result in divergent gene expression profiles, with the H3K56ac mark, for example, marking asexual-stage genes for transcription (Bowman and Poirier, 2015). In gametocytes, the phosphorylation of H3Y41 and H3T45 in particular, may be associated with the expulsion of HP1 from repressed genes that are necessary for gametocyte conversion and later sexual maturation (Dawson *et al.*, 2009; Brehove *et al.*, 2015). Though shown to similarly promote nucleosome unwrapping and DNA accessibility in other eukaryotes (Brehove *et al.*, 2015), the results of this study suggest that these marks are stage-specific in *P. berghei*. ChIP-sequencing using H3K56ac and H3Y41 antibodies in *P. berghei* asexual stages and gametocytes would determine whether these histone PTMs associate with asexual-stage genes or genes upregulated during gametocytogenesis.

In addition to further ChIP-sequencing studies, assays to determine the potential enzyme activity of *P. berghei* NEK4 and HAT1, as complete KO lines are already available for further examination, are warranted.

4.4.3 Additional experimental methods to improve extraction and identification of histone modifications from *Plasmodium* spp.

As alluded to in section **4.2.3**, acid-extracted or high-salt-extracted histones can be analysed by immunoblotting or MS for identification (and possible

quantification) of histone PTMs. Immunoblotting is a limited method of PTM detection because detection is limited by the availability of antibodies, and that, even when an antibody is available, cross-reactivity between similar PTMs (such as di- or tri-methylation of a specific lysine) can occur (Egelhofer *et al.*, 2011).

At present, MS is the most useful tool for the identification of known or novel PTMs on histones. When preparing acid-extracted or high-salt-extracted histones for MS, the intact proteins must first be digested by proteolysis into peptides, ideally in the 5-20 aa size range, that can be more easily separated by LC, and more easily ionised for MS detection (Sidoli *et al.*, 2016). The use of trypsin as the digesting enzyme (cleaving proteins at lysine and arginine residues, except where either is followed by a proline) is a standard approach in proteomics but can be an issue in the analysis of histones, which are lysine-rich and can yield tryptic fragments that are too small for LC retention.

To overcome this problem, lysine derivatisation techniques using propionic anhydride have been developed in which the ε-amino groups of unmodified and mono-methylated lysine residues are derivatised to propionyl amides, protecting them from trypsin digest (Garcia *et al.*, 2007; Sidoli *et al.*, 2015). Trypsin digest then takes place only at the C-termini of arginine residues (Sidoli *et al.*, 2016). A second round of derivatisation is then used to modify the amino groups of newly-generated N-termini, increasing peptide hydrophobicity and facilitating better interaction with LC reverse-phase columns (Karch, Sidoli and Garcia, 2016; Sidoli *et al.*, 2016). Though not used in the present study, the addition of a lysine derivatisation step, and the addition of a greater number of protease inhibitors to cell lysis buffers and SDS-PAGE buffer (as described in section **4.3.3**) may result in a greater number of histone PTMs detected by LC-MS/MS, particularly at the lysine-rich N-terminal tails of *Plasmodium* spp. histones.

Using tandem MS (MS/MS), PTMs on histone amino acids are identified based on changes to the peptide mass-to-charge ratio (m/z), with isobaric peptides (peptides of identical m/z; for example, peptides containing K18ac or K23ac are isobaric) being separated only at MS/MS level, i.e. according to their unique fragment ions at a second (or later) round of MS detection (Sidoli *et al.*, 2016). Several software options are available for the examination and interpretation of

MS/MS spectra from highly modified proteins such as histones (Sidoli, Cheng and Jensen, 2012). Quantification of high-resolution MS/MS data obtained from non-labelled histones (for example, to quantify and compare the extent of post-translational modifications between asexual and sexual *P. berghei* ANKA parasites) is also possible with the development of the 'EpiProfile' program (Yuan *et al.*, 2015). Such quantification software could also be applied in future to the dataset produced in this study, enabling quantification of a number of histone PTMs at asexual and sexual stages.

4.4.4 Biological limitations to *P. berghei* histone extraction

In the present study, a number of issues were taken into consideration when interpreting the results of LC-MS/MS spectra from *P. berghei* ANKA asexual and sexual stage parasites. In this study, the AP2-G over-expressor *P. berghei* ANKA line (the *P. berghei* G1142 line) was used to decrease contamination of either mature asexual stage parasites (schizonts) or mature sexual stage parasites (gametocytes) with parasites of an undesired life cycle stage (discussed in section **4.3.1**). One could argue that, though asexual or sexual development was assured by the absence or administration of rapamycin respectively, the forced manner in which gametocytogenesis was either prevented or induced produced histone modifications that were not natural to the *P. berghei* parasite. In a wild-type *P. berghei* line, commitment to gametocytogenesis may be the result of histone modifications (and subsequent downstream effects) that are bypassed when AP2-G is conditionally induced in the G1142 line.

In addition, assuring gametocyte maturity in both WT and G1142 *P. berghei* ANKA lines may have resulted in the loss of early gametocyte histone modifications that are only present at the moment of sexual commitment. To ensure that this is not the case, future experiments would need to be carried out in a non-inducible *P. berghei* ANKA line, and with a timed infection of a rodent model by synchronous merozoites. In the present study, an attempt was made to infect a rat with synchronous merozoites for examination of synchronous trophozoite histone modifications, with intravenous infection of a Wistar rat carried out in the same method as intravenous infection of synchronous, purified *P. berghei* parasites into the lateral tail veins of mice (section **2.1.3**). Unfortunately, owing to the

thickened skin overlying the vasculature of the adult rat tail, attempts to infect rats intravenously with *P. berghei* failed. Infection of rats with synchronous, purified *P. berghei* parasites for the production of schizonts and gametocytes was carried out by intraperitoneal (IP) injection (as described in section **2.1.2**).

4.4.5 Limitations to identification and quantification of histone PTMs from LC-MS/MS data using Mascot and EpiProfile software

In the present study, identification of histone PTMs was carried out using a so-called 'bottom-up' proteomics analysis approach, i.e. histones were extracted from *P. berghei* ANKA cells, digested with a protease (trypsin), and analysed by LC-MS/MS (Yuan *et al.*, 2014). In a 'top-down' proteomics approach, intact proteins, i.e. proteins that have not been lysed by trypsin or any other protease, are introduced to the mass spectrometer. The latter approach, though requiring more starting material and longer running times, may be more useful in the identification of proteoforms compared to a 'bottom-up' approach (Catherman, Skinner and Kelleher, 2014). In the present study, a 'bottom-up' proteomics approach was sufficient.

As was recommended for discovery of peptide PTMs using this approach, Mascot software version 2.6.2 (Matrix Science) was used to identify stage-specific histone modifications in *P. berghei* parasites (Sidoli *et al.*, 2016). Mascot is a probability-based protein identification software, written in ANSI C, that calculates the probability that the observed match between the experimental dataset and a sequence database entry is a chance event, and the match with the lowest probability is given the highest score and presented to the user (Perkins *et al.*, 1999). In the present study, Mascot software was used in combination with a *P. berghei* ANKA amino acid FASTA database (version 35), with a significance threshold of $p<0.05$, and a false discovery rate (FDR) less than 1%. An allowance was made for up to five missed cleavages by trypsin digest which enabled detection of the lysine-rich N-termini of *P. berghei* ANKA histones. Without this allowance being made, PTMs such as H2AK5ac and H3K9ac were not identified.

Though developed following the commencement of this study, quantification of chemically-unlabelled histone PTMs in *P. berghei* ANKA may be possible by using

the EpiProfile software tool, a feat not possible using Mascot software alone (Yuan *et al.*, 2015). This software is freely available upon request and was used to quantify a selection of PTMs in *P. falciparum* that were identified by LC-MS/MS in a similar manner to the present study (Coetzee *et al.*, 2017). EpiProfile quantifies histone PTMs based upon chromatographic elution retention times of known histone peptides. Raw MS1 and MS2 files can be imported into EpiProfile and quantification of histone peptides carried out. Quantification of *P. berghei* ANKA histone PTMs identified in the present study would be possible using this software (Yuan *et al.*, 2015).

5. Identification of genes regulated by histone modifications using chromatin immunoprecipitation and high throughput sequencing

5.1 Chapter Aim

The aim of this chapter was to determine *P. berghei* ANKA genes that were targeted by four epigenetic modifications (H3K9ac, H3K9me3, H4K8ac, and H3K122ac) in both schizonts and gametocytes to examine their stage-specificity and possible involvement in the regulation of gametocytogenesis.

5.2 Introduction: Chromatin immunoprecipitation (ChIP)

The identification of differences between mature asexual stage (schizont) and sexual stage (gametocyte) *P. berghei* histone PTMs **(Chapter 4)** warranted an investigation into the possible target genes of these differential marks. To facilitate such an examination, chromatin immunoprecipitation (ChIP) was employed; a procedure that enables investigators to probe the relationship of a protein-of-interest (POI), or a modified POI, to a genomic DNA region (Carey, Peterson and Smale, 2009). A ChIP procedure can be carried out on samples in which the DNA and protein have not been chemically crosslinked (called native ChIP or 'N-ChIP') or in samples in which a protein and its target DNA have been crosslinked chemically (sometimes referred to as 'XChIP'). Formaldehyde is the most commonly-used protein-DNA crosslinker used in ChIP analyses (Hoffman *et al.*, 2015).

In fact, a ChIP procedure is quite straightforward, and it is only with multiple options available at each step of the procedure that the complete protocol can become complicated. For example, the investigator must choose between native- or crosslinked- ChIP; must choose a crosslinking chemical to work with; must choose between Tris or glycine for quenching of the crosslinking reaction; must choose to fragment cell pellets by douncing, micrococcal nuclease (MNase) digestion, sonication, or a combination of these methods (Haring *et al.*, 2007). Even more choices are required after these preparation steps; a control sample to normalise results against must be chosen, and library preparation before high-throughput sequencing can be undertaken in multiple different ways, generating the separate techniques: standard ChIP-seq, ChIP-exo, or ChIPmentation (to name but a few) (Carey, Peterson and Smale, 2009; Schmidl *et al.*, 2015; Perreault and

Venters, 2016). Both standard ChIP-seq and ChIPmentation procedures were attempted during the present study (experimental procedures detailed in sections **2.2.11** and **2.2.12** respectively).

Though standard ChIP-seq and ChIPmentation protocols in the present study are described for sequencing using Illumina sequencing platforms, a standard ChIP procedure, followed by immunoprecipitation and de-crosslinking, would produce DNA fragments amenable to other forms of DNA sequencing, with only slight changes to library preparation. Sequencing of ChIP DNA by sequencing platforms from Pacific Biosciences or Oxford Nanopore Technologies, though not common, is possible (Rhoads and Au, 2015; Jain *et al.*, 2018). In the present study, all ChIP-sequencing and ChIPmentation experiments were carried out as described in sections **2.2.11** and **2.2.12** respectively, with bioinformatic analysis carried out as described in section **2.2.13.**

5.3 Results

5.3.1 Optimising ChIPmentation in *P. berghei:* Cell number, sonication, and de-crosslinking

Although novel protocols for ChIP-seq experiments are currently being devised for fewer and fewer cell numbers; for example, a protocol for ultra-low-input native ChIP-seq from as few as 1000 cells (Brind'Amour *et al.*, 2015), or the amalgamation of drop-based microfluidics and chromatin fragment barcoding to enable the first single-cell ChIP-seq analysis from as few as 50 embryonic stem cells (Rotem *et al.*, 2015); standard ChIP-seq protocols still require abundant starting material, often in the range of 1-20 million single cells per immunoprecipitation (Gilfillan *et al.*, 2012).

To overcome this limitation, and to study cells from a single stage of the *P. berghei* ANKA life-cycle, the present study was designed to test the reliability of a ChIPmentation protocol using the Tn5 transposase from a Nextera DNA library preparation kit (Illumina, Inc.) (Schmidl *et al.*, 2015). Though the initial aim of this chapter was to examine both asexual- and sexual- stage *P. berghei* parasites, time restrictions allowed only for optimisation of ChIPmentation using *P. berghei*

schizonts. Synchronous *P. berghei* ANKA schizonts provided the greatest abundance of starting genetic material, and multiple antibodies could be tested from the same schizont sample at 250,000, 2.5 million, and 25 million cells, with later analyses revealing the optimum cell number and antibody concentration for a successful ChIPmentation experiment. Standard ChIP-seq analysis was also attempted using the same cell numbers as described in section **2.2.11**, though there was never enough genetic material left after the library preparation procedure for high-throughput sequencing to be carried out.

Other alterations to a standard ChIP protocol used when studying *P. berghei* ANKA in the present study included a change to sonication conditions (detailed in sections **2.2.11.2** and **2.2.12.2**) which were adapted according to consultation with Dr Agnieszka Religa at the University of Glasgow. When undertaking ChIP in *P. berghei* schizonts, this sequence of sonication pulses was often repeated to obtain chromatin-DNA fragments in the range of 100-1000 bp, with regions of 100-300 bp excised by gel excision for further immunoprecipitation.

One final note must be made regarding the optimisation of de-crosslinking and enzyme denaturation temperatures when undertaking ChIP experiments and ChIP DNA library preparation in *Plasmodium* parasites. As alluded to in Fraschka *et al.* (Fraschka *et al.*, 2018), the extreme AT-richness of *Plasmodium* genomes requires that a lowered temperature be used when de-crosslinking chromatin from DNA in a ChIP experiment (45ºC instead of 65ºC). A standard 65ºC de-crosslinking temperature (usually carried out from 6 h to overnight), results in denaturation of ChIP DNA, particularly in standard ChIP-seq samples or inputs (personal experience). It is possible that ligation of Illumina adapters prior to de-crosslinking in ChIPmentation experiments somewhat ameliorated the loss of genetic material by increasing the GC content of DNA fragments during the 'tagmentation' process. This may explain why ChIPmentation experiments in this study proved successful, while a standard ChIP-seq protocol failed.

At the outset of the present study, all experiments were carried out with overnight de-crosslinking at 65ºC, resulting in the loss of many input samples. Upon publication of Fraschka *et al*, 2018 (Fraschka *et al.*, 2018), in which the de-crosslinking of ChIP DNA from *Plasmodium spp*. at 45ºC was noted, the cause for

loss of ChIP material in the present study had also been narrowed down to this point in the experimental procedure. When undertaking a standard ChIP-seq experiment, altering enzyme denaturation temperatures when using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (Catalogue No. #E7645S; New England BioLabs Inc.) was also attempted upon consultation with Dr Kathryn Crouch and Craig Lapsley at the University of Glasgow. In this case, the replacement of 65ºC incubation steps with SPRIselect bead (Beckman Coulter, Inc.) clean-ups had considerably reduced the loss of genetic material when studying subtelomeric regions in the genome of trypanosomatid parasites. Despite this, using a standard ChIP-seq protocol in the present study failed to produce amplifiable DNA libraries.

5.3.2 Comparative analysis of 250K, 2.5M, and 25M schizonts

Following an initial pilot study of H4K8 acetylation with 22.7 million schizont-stage *P. berghei* parasites per sample (results not shown), further optimisation of the ChIPmentation procedure was carried out to identify an optimum cell number for further experiments. Although cultivation of 22.7 million synchronous schizonts was possible, the collection of such cell numbers from other *P. berghei* life-cycle stages was not possible without a great number of animals being infected and resulting samples pooled. As such, it was decided to attempt the same procedure with 250,000, 2.5 million, and 25 million synchronous schizont-stage parasites. For each number of cells, an input was also taken, and each cell amount was investigated with 4 separate antibodies: H4K8ac, H3K9me3, H3K9ac, and H3K122ac.

Owing to issues with de-crosslinking temperature and loss of ChIP DNA (discussed in section **5.3.1**), only one input sample produced a usable library for sequencing following tagmentation, de-crosslinking and amplification steps; an input sample for 250,000 schizont cells. Consequently, this sample was used as a control for all initial UpSetR visualisation of ChIPmentation samples from 250,000, 2.5 million, and 25 million cells. Later ChIPmentation repeats (with 2.5 million cells) would also be normalised to input data downloaded from NCBI (https://www.ncbi.nlm.nih.gov/sra/SRR5935720) (Fraschka *et al.*, 2018), as this input provided greater genome coverage than the input obtained in the present study. Peak-calling software options were then tested as described in section

5.3.4. The number of peaks called and the number of peaks shared between different software packages are shown for all samples as an UpSet plot (separate UpSet plots are shown for examination of each histone PTM).

5.3.3 Correlating ChIPmentation results to RNA-seq data

It is important for the present study that ChIPmentation results using each of the four antibodies tested (H4K8ac, H3K9ac, H3K9me3, and H3K122ac) were aligned alongside transcriptomic data (i.e. RNA-sequencing (RNA-seq) data) to evaluate whether the presence of each histone modification correlated to gene expression, gene repression, or a mixture of both (for example, if the modification is dependent upon co-localisation of a second histone mark; such as the relationship between H3K9ac and H3K4me3 (Gates *et al.*, 2017)).

In the present study, because optimisation of the ChIPmentation technique took place only with mature asexual (schizont-stage) *P. berghei* parasites owing to time limitations, comparative analysis between the previous developmental stage (16 h trophozoites) and 22 h schizonts was undertaken to identify transcripts that were specific to mature schizonts. Down-regulation of transcription in 22 h schizonts indicated that greater expression of a particular gene took place during earlier asexual development. Upregulation of transcription in schizonts indicated that a gene was expressed to a greater extent in the mature 22 h asexual parasites (Otto, Böhme, *et al.*, 2014). Methods used to analyse RNA-seq data are described in section **2.2.14.**

In *Plasmodium* spp. parasites so far, H4K8ac has been linked to transcriptional activation in euchromatic regions of the genome, and at heterochromatic regions, where it marks transcription of the active *var* gene at the promoter (Gupta *et al.*, 2017). Trimethylation of H3K9 is associated with silent multicopy variant gene families at the nuclear periphery and transcriptional repression (Lopez-Rubio, Mancio-Silva and Scherf, 2009; Karmodiya *et al.*, 2015), while acetylation of H3K9 in *Plasmodium* is associated with transcriptional activation (Salcedo-Amaya *et al.*, 2009; Srivastava *et al.*, 2014). The recently characterised H3K122ac modification has never before been studied in a *Plasmodium* parasite, but has been associated with transcriptional activation and marking a subset of active enhancers in human

cell lines (K562 and MCF-7 lines) and mouse embryonic stem cells (Tropberger *et al.*, 2013; Pradeepa *et al.*, 2016).

5.3.4 Optimisation of peak-calling software

In the present study, identification of a 'peak-caller' that could adequately identify the broad peaks associated with some histone PTMs was carried out using *P. berghei* schizont ChIPmentation data. It is well-noted that a number of PTMs, for example H4K20me1 and H3K27me3, associate with large genomic regions (Rintisch *et al.*, 2014), and so a number of algorithms have been developed to assist in the identification of broad peaks from ChIP-sequencing experiments, with or without *a priori* knowledge of the genomic distribution of a histone PTM to be analysed (Wang, Lunyak and King Jordan, 2013; Heinig *et al.*, 2015; Starmer and Magnuson, 2016; Khomtchouk, Van Booven and Wahlestedt, 2017).

In this study, three ChIP peak-calling software packages were tested for their ability to identify broad peaks associated with histone modifications in *P. berghei*. The peak-calling software packages, MACS2 (Zhang *et al.*, 2008), HOMER (Heinz *et al.*, 2010), and Epic (Xu *et al.*, 2014) were tested and the 'UpSetR' R package was used to display unique and shared peak data as described by Conway *et al.* (Conway, Lex and Gehlenborg, 2017). HOMER ChIP peak annotation software was later used to label ChIP peaks relative to their location adjacent to, or within, an ORF (http://homer.ucsd.edu/homer/ngs/annotation.html). In the case of H3K9me3, ChIP data from this study was also mapped alongside HP1 ChIP data which was downloaded from NCBI (https://www.ncbi.nlm.nih.gov/sra/SRX3095675[accn]) (Fraschka *et al.*, 2018). Results from these analyses are shown in the sections below.

5.3.5 Identifying outlying ChIPmentation samples by multidimensional scaling

As described in section **2.2.15,** and often used to examine heterogeneity between RNA-seq datasets (bulk or single-cell samples) (Dey, Hsiao and Stephens, 2017), multidimensional scaling (MDS) was carried out in this study to map similarity between ChIPmentation sample datasets and inputs. Using edgeR, outlying ChIPmentation datasets for each of the four histone marks analysed, at 2.5 million

schizonts per experiment, were identified and excluded from the study (Robinson, Mccarthy and Smyth, 2010). MDS results are shown in **Figure 5.1.**

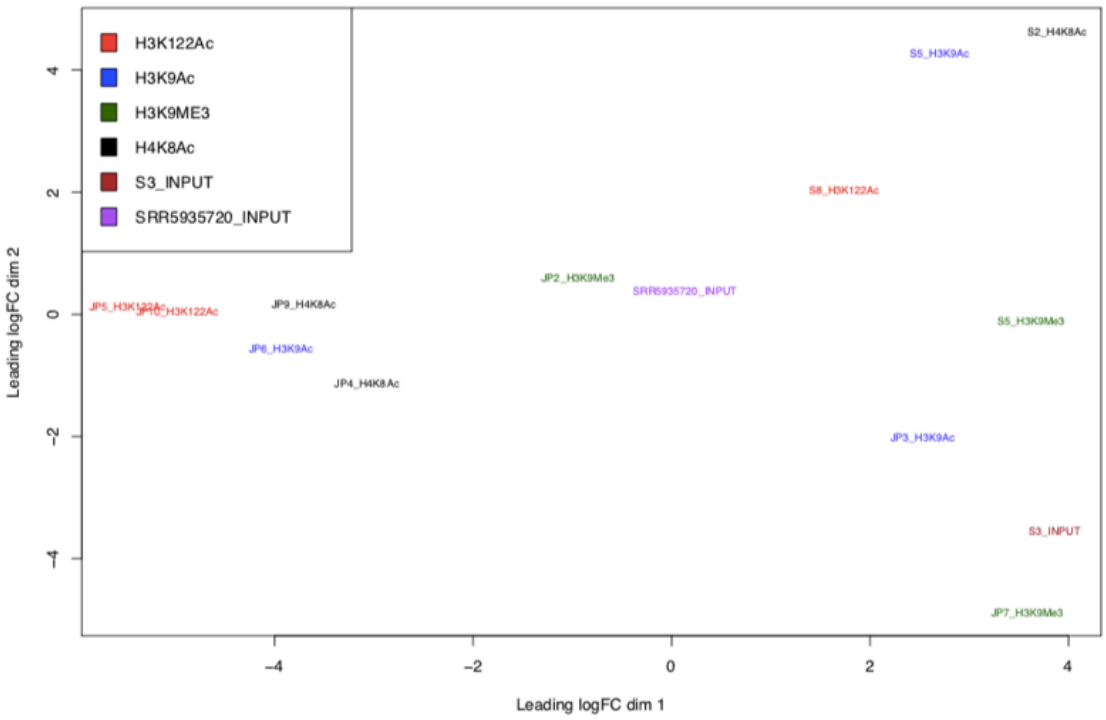


Figure 5.1: Multidimensional scaling (MDS) of triplicate ChIPmentation samples to determine outliers. The MDS plot shown was generated using edgeR (Robinson, McCarthy and Smyth, 2010). The 'plotMDS' function was applied to all three replicate datasets from ChIPmentation with four separate histone modification antibodies: H4K8ac, H3K9me3, H3K9ac, and H3K122ac. All 12 of these samples were derived from 2.5 million *P. berghei* ANKA schizonts. In the case of inputs; the first input (S3_INPUT) was derived from 250,000 mature *P. berghei* schizonts during the present study. As no further inputs were successfully collected during the present study (owing to the 65°C de-crosslinking temperature used), a second input was also downloaded for *P. berghei* schizonts from (https://www.ncbi.nlm.nih.gov/sra/SRR5935720) (Fraschka *et al.*, 2018). This input was used to normalise triplicate ChIPmentation studies using 2.5 million schizont cells. X- and y- axes represent the leading log fold change (logFC) dimension 1 (dim 1) and leading logFC dimension 2 (dim 2). Samples are colour-coded in addition to labelling.

As can be seen from **Figure 5.1**, inputs cluster separately within the MDS plot and in the case of each histone modification antibody, two samples in each case appear to cluster together, with a third sample appearing to be an outlier. In the case of H3K122ac, two samples (samples JP5_H3K122Ac and JP10_H3K122Ac) are clustered closely. A third replicate, S8_H3K122Ac, appears to be an outlier among the triplicate samples and so final H3K122ac results are based on the former two clustered samples. In the case of H4K8 acetylation, two samples clearly cluster together (samples JP4_H4K8Ac and JP9_H4K8Ac), while the third replicate

(S2_H4K8Ac) was a clear outlier and so was again discarded during in-depth analysis of H4K8 acetylation in later sections.

Regarding histone 3, lysine 9 modifications; in each case, two samples cluster vertically at the x-axis: S5_H3K9Ac and JP3_H3K9Ac, and S5_H3K9me3 and JP7_H3K9me3. In both of these cases also, outlying samples (JP6_H3K9Ac and JP2_H3K9me3) were discarded in final results sections.

Though outlying samples were removed in the text, a complete list of merged peaks for all four antibodies, between all three biological replicates, are provided in **Supplemental Material ES3** alongside peaks called for each sample separately, and peaks called using Epic software for clustered duplicate samples.

5.4 H4K8 acetylation in *P. berghei* schizonts

5.4.1 H4K8ac: Cell number and genomic distribution

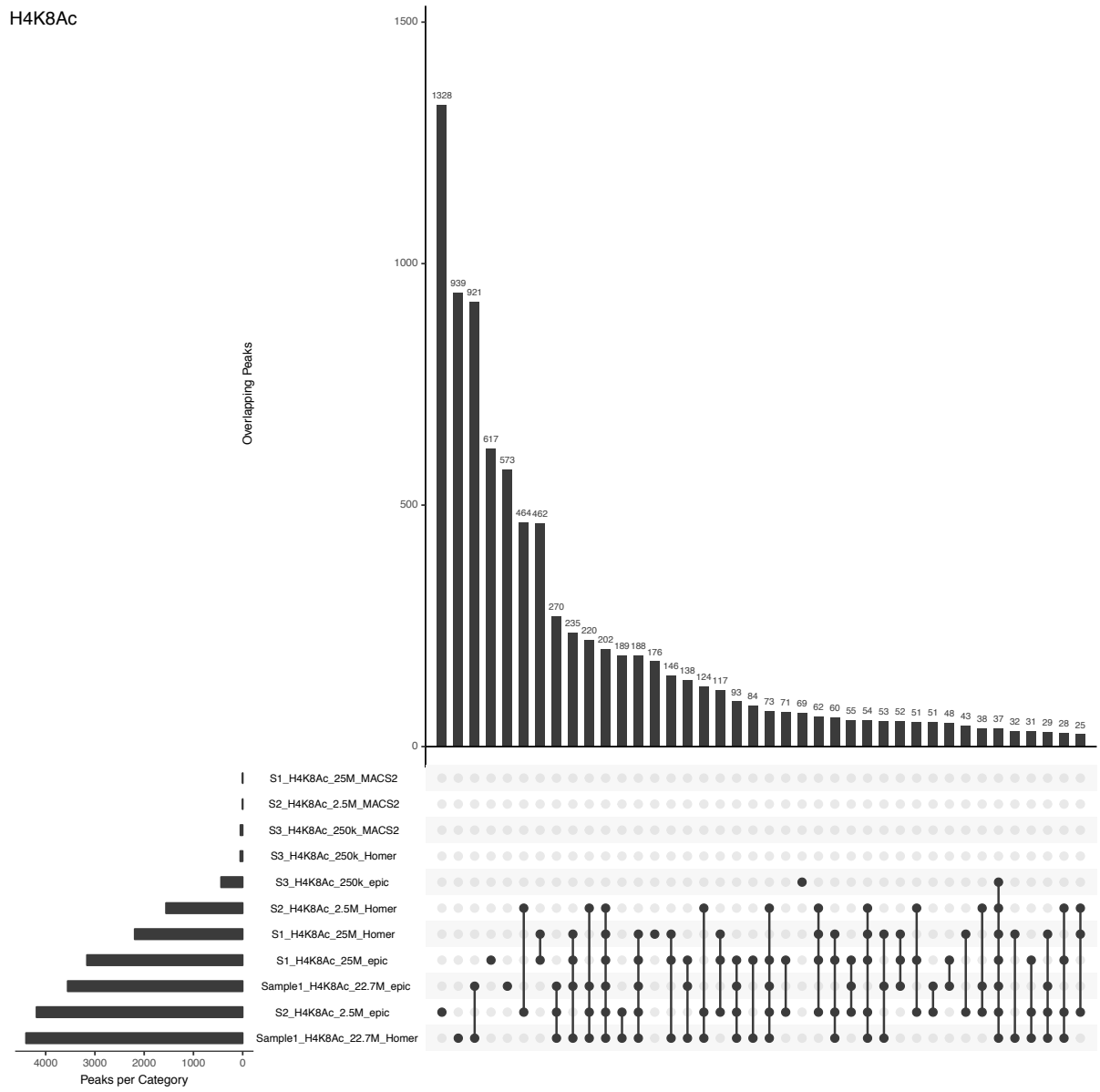


Figure 5.2: A comparison of shared and unique peaks from H4K8ac ChIP samples using MACS2, HOMER, and Epic software. This UpSet plot shows the number of unique peaks called using MACS2, HOMER, and Epic software on the left-hand side. Each sample name is shown next to the number of unique peaks per category, with cell numbers and software name clearly labelled. The main histogram depicts the number of overlapping peaks in each H4K8ac sample. At the base of the histogram, overlapping peaks shared between samples are indicated using circles, or circles joined with a straight line. From the histogram above, it is very clear that MACS2 software (Zhang *et al.*, 2008) was identifying a small number of unique peaks, but no overlapping peaks, in the H4K8ac replicate samples (no peaks were detected for the 22.7 million cell sample; not shown). In contrast, using HOMER (Heinz *et al.*, 2010) and Epic (Xu *et al.*, 2014), a much greater number of unique and overlapping peaks were identified. Using both HOMER and Epic, peaks were called for each of the normalised H4K8ac ChIP samples.

As shown in **Figure 5.4,** the greatest number of overlapping peaks were identified using Epic, the most recently developed peak-calling software for disperse ChIP-enriched regions, based on the SICER peak-calling algorithm (Xu *et al.*, 2014). Epic called 1328 overlapping peaks across the *P. berghei* ANKA genome from a sample of 2.5 million schizonts. Using Epic, a score is generated for each overlapping peak and these scores are listed alongside the gene identifiers and peak position in **Electronic Supplemental Material 2 (ES2).**

To further assess the accuracy of the Epic peak-calling software and to compare the results from all three different cell numbers (and the 22.7 M cell pilot study), alignment of all bigwig (H4K8ac/control) and Epic peak files was undertaken and these results are shown in **Figure 5.3.**

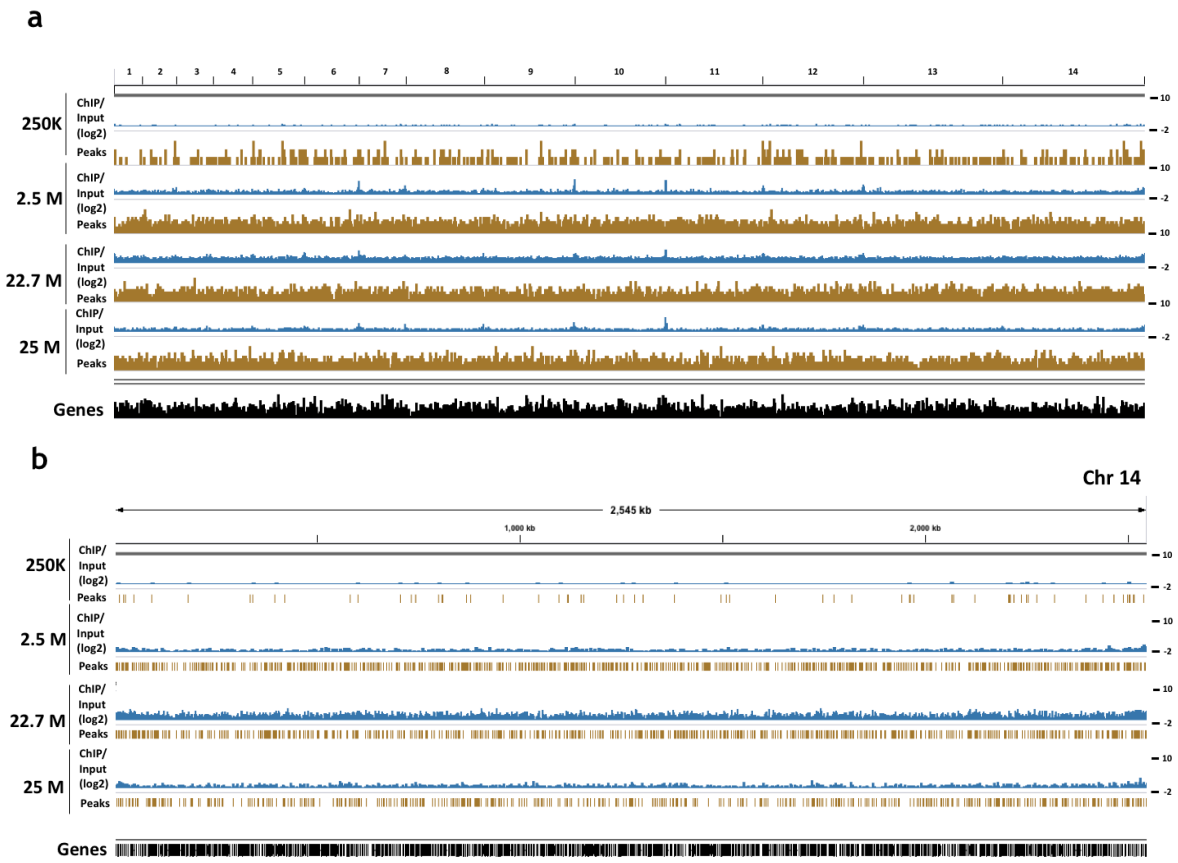


Figure 5.3: H4K8ac peaks across the *P. berghei* ANKA genome using Epic peak-calling software. Using the Integrative Genome Viewer (IGV) (Robinson *et al.*, 2011), H4K8ac-bound (sample/input) reads (for 250K, 2.5 million, 22.7 million, and 25 million schizonts) are shown across the entire *P. berghei* ANKA genome (version 34) alongside peaks called using the 'Epic' peak-calling software (in .bed format; results are normalised to the 250K schizont ChIP input sample) **(a).** At the base of these tracks, the general feature format (.gff) file shows the location of genes as black bars. For each sample, aligned ChIP data is displayed in blue. Epic peaks are shown in a separate track

underneath ChIP data and are displayed in a dark yellow colour. In **(b)**, aligned reads and peak files are again shown as in **(a)**, though in this case, only the developmentally important (during gametocytogenesis) chromosome 14 is shown in closer detail (Kooij *et al.*, 2005; Lopez-Rubio, Mancio-Silva and Scherf, 2009; Fraschka *et al.*, 2018). Again, colour-coding is carried out as in **(a).**

As seen in **Figure 5.3,** both the whole genome view of aligned ChIP data **(Figure 5.3 (a))** and a zoomed-in visualisation of *P. berghei* ANKA chromosome 14 (syntenic to *P. falciparum* chromosome 12; a chromosome that has previously been heavily analysed due to its role in gametocytogenesis) (**Figure 5.3 (b);** (Kooij *et al.*, 2005; Lopez-Rubio, Mancio-Silva and Scherf, 2009; Salcedo-Amaya *et al.*, 2009; Fraschka *et al.*, 2018)**)** showed quite an even distribution of raw data and Epic peaks between all samples except for 250,000 schizont cells. In **Figure 5.3 (a)**, it is evident that quite a few of the peaks in raw data seen with >2.5 million cells were not enriched in this sample. Taking a closer look at chromosome 14 (**Figure 5.3 (b))** also showed that fewer peaks were observed when ChIPmentation was carried out with 250,000 cells. With these data, it was decided to repeat the ChIPmentation protocol using two more independent *P. berghei* ANKA samples of 2.5 million synchronous schizonts. With these data, a comparative analysis of H4K8ac-binding across the *P. berghei* genome could be undertaken with a greater degree of certainty. These data were also aligned to an input with greater genome coverage (downloaded from NCBI (Fraschka *et al.*, 2018)).

5.4.2 H4K8ac analysis in triplicate (2.5 M schizonts)

Following a comparison of three peak-calling algorithms across samples of either 250,000, 2.5 million, 22.7 million (a pilot study), or 25 million cells, a final analysis of H4K8ac-binding during mature asexual schizont stage in *P. berghei* was carried out using 2.5 M cells in triplicate. All peaks called using Epic peak-calling software for all independent replicate experiments are listed in detail in **Electronic Supplemental Material 3 (ES3).** All samples shown were normalised to input data obtained as part of a previous *P. berghei* ANKA schizont ChIP experiment (Fraschka *et al.*, 2018). Final results show shared peaks between duplicate biological replicates JP4_H4K8Ac and JP9_H4K8Ac, following removal of outlying sample S2_H4K8Ac based on MDS analysis (**Figure 5.3.5).** An overview of read and peak distribution for these H4K8ac samples of 2.5 million schizont cells is shown in **Figure 5.4.**

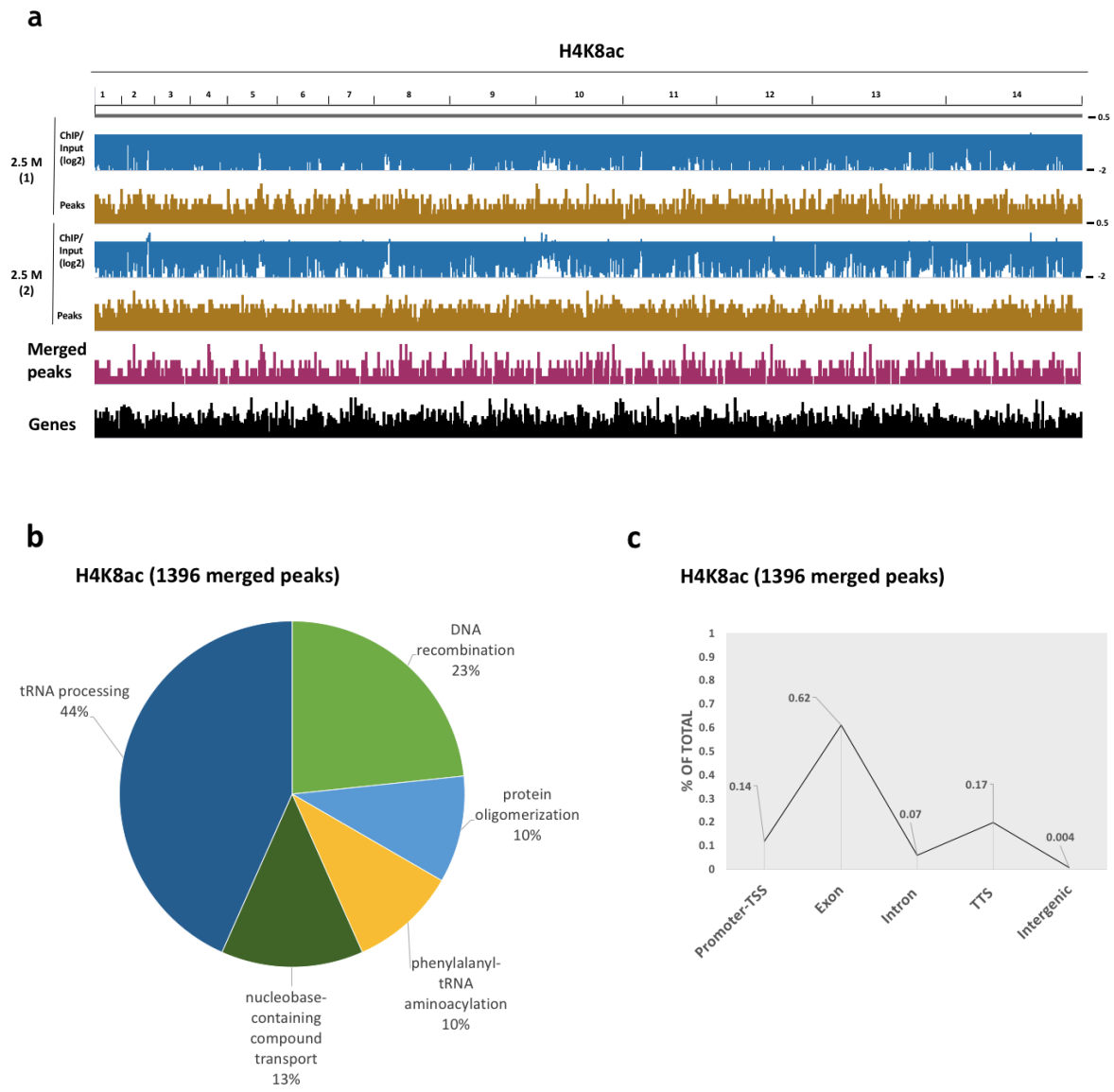


Figure 5.4: H4K8ac peak distribution across the *P. berghei* ANKA genome. Using the Integrative Genome Viewer (IGV) (Robinson *et al.*, 2011), H4K8ac-bound reads (relative to an input sample from Fraschka *et al.*, 2018) from two independent samples of 2.5 million schizonts are shown in **(a)**. Raw data are displayed in blue and Epic peaks are displayed in yellow underneath the raw data track. Merged peaks between both samples are shown in pink underneath individual results. Genes are displayed in general feature format (.gff) in the final track as black bars. **(b)** is a pie-chart showing gene ontology (GO) terms for biological processes that were associated with the genes to which H4K8ac was bound (with a P-value < 0.025). Image **(c)** shows the same data as **(b)**, though this time, the location of the area of H4K8ac enrichment is shown relative to the nearest gene ORF.

Using the Epic algorithm, ChIP-enriched 'islands' are delineated and assessed for statistical significance by i) partitioning the genome into non-overlapping windows, ii) identifying an 'eligible' region of enrichment according to a read count threshold (which takes into consideration the average number of reads

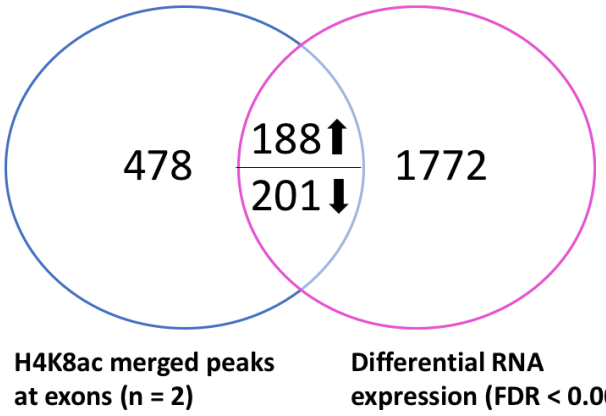
within a window, within the library, and the genome length), iii) determining clusters of windows as contiguous ChIP-enriched islands by searching for gaps of a pre-determined basepair value, iv) identifying 'candidate' islands that are unlikely to appear by chance (assigned an initial 's(l)' score), and v) if a control is available, candidate islands of ChIP-enrichment are further filtered based on the control signal, with values assigned a *P*-value based on Poisson distribution and a false discovery rate (FDR) reported using the *P*-value adjusted for multiple-testing (Xu *et al.*, 2014). In the **ES3** material, the 's(l)' score, log2 fold change (log2FC), *P*-value, and FDR are shown in columns O, P, Q, and R respectively for each independent experiment.

To annotate peaks, the Epic output (.bed) file was mapped to the reference genome using the HOMER (Hypergeometric Optimization of Motif Enrichment) script downloaded from http://homer.ucsd.edu/homer/ngs/annotation.html. With regard to the genome locations shown in **Figure 5.4 (c),** the promoter-transcription start site (promoter-TSS) is defined by default as -1 kb to +100 bp around the TSS using the *P. berghei* ANKA genome (version 34) annotation. The transcription termination site (TTS) is defined as -100 bp to +1 kb either side of the TTS. The exon and intron locations are defined according to the *P. berghei* ANKA genome annotation (version 34). An 'intergenic' region is defined as anything up- or down- stream of exons and introns that is not covered by the 'promoter-TSS' or 'TTS' parameters.

Figure 5.4 (b) depicts the GO terms for biological processes associated with H4K8ac-bound genes. When analysing GO terms using *Plasmo*DB for merged H4K8ac peaks, a *P*-value cut-off was set at 0.025. Taking the 1396 merged peaks between two independent experiments, H4K8ac-bound genes were associated with tRNA processing (44%), DNA replication (23%), nucleobase-containing compound transport (13%), phenylalanyl-tRNA aminoacylation (10%) and protein oligomerisation (10%) (**Figure 5.4 (b)),** all processes suggesting active DNA replication in mature schizonts**.** When an analysis of H4K8 acetylation was carried out in *P. falciparum*, this mark was associated in schizonts with transcription of genes involved in DNA replication, excision-repair, and apicoplast/mitochondrial import (Gupta *et al.*, 2017). However, it should also be noted that a recent transcriptomic analysis between multiple different *P. falciparum* strains at

schizont stage determined that there were significantly differentially-expressed genes between schizonts from laboratory-adapted *P. falciparum* strains and clinical isolates, particularly at genes involved in gametocytogenesis; meaning that caution should be taken when comparing *P. falciparum* H4K8ac results to *P. berghei* ANKA H4K8ac results seen in this experiment (Tarr *et al.*, 2018). Differences in H4K8ac distribution may not only occur between human and rodent *Plasmodium* lineages, but also between strains among the same *Plasmodium* species.

In addition, transcriptional activity in *P. falciparum* was noted as occurring only when H4K8ac-occupancy shifted from the 5'-end of the ORF to within the protein-coding region (Gupta *et al.*, 2017). In **Figure 5.4 (c),** it can be seen that of the merged H4K8ac peaks, H4K8ac-occupancy was mainly located at exons (62% of peaks), TTS regions (17%), and promoter-TSS regions (14% of peaks), almost exclusively. Taking into account the association of H4K8ac occupancy in these regions and altered transcription in *P. falciparum,* it may be the case that H4K8ac merged peaks associated with exons correlate with differential RNA expression in *P. berghei*. To examine this possible relationship, all H4K8ac merged peaks located at exons from duplicate schizont samples (1396 merged peaks) were searched among differentially-expressed genes (as discussed in section **5.3.3**). The results of this comparative analysis are shown in **Figure 5.5** below.



GO term enrichment (P < 0.001)

GO ID	GO term	Fold enrichment
GO: 0006259	DNA metabolic process	4.17
GO: 0051704	Multi-organism process	3.65
GO: 0044419	Interspecies interaction between organisms	3.96
GO: 0044403	Symbiosis, encompassing mutualism through parasitism	3.96

GO ID	GO term	Fold enrichment
GO: 0006457	Protein folding	4.2
GO: 0010467	Gene expression	1.83
GO: 0043170	Macromolecule metabolic process	1.46
GO: 0034470	ncRNA processing	3.4
GO: 0034660	ncRNA metabolic process	2.65

Figure 5.5: H4K8ac-binding at *P. berghei* schizont exons and association with differential RNA expression. The Venn diagram on the left of this image depicts the overlap between H4K8ac-associated exons (thought to be associated with greater transcriptional regulation in *P. falciparum*

(Gupta *et al.*, 2017)) and differentially-expressed genes between schizont- and trophozoite- stage *P. berghei* ANKA parasites (with a false discovery rate (FDR) less than 0.005). 1396 merged peaks across duplicate H4K8ac ChIPmentation experiments were compared to 2161 differentially-expressed genes with FDR < 0.005. Within these two groups, 389 H4K8ac-associated exons were associated with differentially-expressed genes; 188 genes that were up-regulated in schizonts compared to trophozoites, and 201 genes that were down-regulated in schizonts compared to trophozoites. Up-regulation and down-regulation is indicated by upward and downward facing arrows next to the figure respectively. To the right of this Venn diagram, GO terms for up-regulated and down-regulated genes associated with H4K8ac-binding are shown (P < 0.001). Again, arrows indicate up-regulation or down-regulation of RNA transcription, with the tables shown also depicting the GO ID for these particular datasets alongside the GO term and fold enrichment for each term.

As shown in **Figure 5.5,** H4K8ac-binding at *P. berghei* ANKA exons is not solely associated with up-regulation of RNA at a particular gene. These data corroborate the results of Gupta *et al.* (Gupta *et al.*, 2017) who, in their study, showed a similar pattern of overlap between H4K8ac occupancy and differential expression, with H4K8 acetylation being a potential regulator of chromatin-linked transcriptional changes associated with both euchromatin and heterochromatin. Of the H4K8ac-associated genes in *P. falciparum*, up-regulated genes were associated with chromatin modification, merozoite surface proteins, and genes encoding for transport proteins. Of the down-regulated H4K8ac-associated genes, genes involved in excision repair and DNA replication were effected to the greatest extent (Gupta *et al.*, 2017).

In the present study, H4K8ac-bound exons that correlated with greatest gene up-regulation in *P. berghei* ANKA were associated with DNA metabolic processes (which includes DNA repair and excision), multi-organism processes (which include adhesion of symbionts to a host), interspecies interactions, and symbiosis (**Figure 5.5)**. Of the down-regulated H4K8ac-associated exons, greatest down-regulation was associated with GO terms encompassing gene expression, protein folding, and non-coding RNA processes, all of which are consistent with previous results in *P. falciparum* (Gupta *et al.*, 2017; The Gene Ontology Consortium, 2017).

With regard to all H4K8ac-bound regions identified in the present study, the highest-scoring shared peaks for H4K8ac-bound genes from duplicate independent samples (according to the Epic algorithm) are shown in **Table 5.1.**

	Gene ID	Function	H4K8ac location	Highest Epic peak score (of 3)	Differential RNA expression at schizont stage	Log_2FC (if present)	P-value (if present)	FDR (if present)
1	PBANKA_0006900.1	PIR protein, pseudogene	Exon/TTS	1712.52	No	-	-	-
2	PBANKA_1100051.1	PIR protein, pseudogene	Promoter-TSS/Exon	1572.35	No	-	-	-
3	PBANKA_1442200.1	Protein SOC2	Exon	618.84	Yes	0.56	0.13	0.19
4	PBANKA_API00055.1	Small subunit ribosomal RNA	TTS	612.70	Yes	2.05	0.06	0.10
5	PBANKA_0100041.1	Conserved rodent malaria protein, unknown function	Exon	582.02	Yes	-0.08	0.94	0.96
6	PBANKA_0600071.1	Conserved rodent malaria protein, unknown function	Exon	536.22	Yes	0.89	0.2	0.28
7	PBANKA_0201051.1	Conserved rodent malaria protein, unknown function	Exon	514.42	Yes	0.12	0.91	0.94
8	PBANKA_0501000.1	Reticulocyte binding protein, putative (Pb235)	Exon	510.00	No	-	-	-
9	PBANKA_1365721.1	Conserved rodent malaria protein, unknown function	Exon	489.36	Yes	0.29	0.61	0.68
10	PBANKA_0214100.1	Conserved protein, unknown function	Promoter-TSS/Exon	420.30	Yes	-2.29	0.0007	0.002

Table 5.1: Highest-scoring H4K8ac peaks in *P. berghei* schizonts. This table displays the top 10 H4K8ac-bound genes according to their Epic peak-calling score. All genes are shared among duplicate H4K8ac ChIPmentation experiments and the highest score between replicates is shown in the fifth column. The list of genes is numbered from highest to lowest score and the first and second columns of the table list gene IDs and current known function (if any) of the gene in *P. berghei* ANKA [as of 09-07-18]. The fourth column shows the location of H4K8ac peaks relative to the gene in question. The sixth column indicates differentially-expressed gene regulation in schizonts at 22 h compared to 16 h trophozoite expression (Otto, Böhme, *et al.*, 2014). The following columns list the log_2 fold change (log_2FC), P-value, and false discovery rate (FDR) (where applicable) of the differential RNA expression.

One striking feature of **Table 5.1** is the fact that H4K8 acetylation is associated with differential RNA expression in all genes other than PIR protein pseudogenes and the reticulocyte binding protein, Pb235 (encoded by PBANKA_0501000.1. All of these H4K8ac-bound genes were found in both duplicate independent experiments and of the only two gene that have so far been characterised (SOC2 and a small rRNA), the SOC2 protein is associated with mitotic spindle assembly, and therefore cell replication, alongside the small rRNA (Fang *et al.*, 2017). However, of the highest scoring peaks according to Epic, the five H4K8ac-bound regions that also coincide with differential expression of genes do not meet the threshold for statistically significant differential expression, a Log_2FC false discovery rate < 0.05.

With regard to the PIR genes/pseudogenes, these genes comprise the largest multigene family in *Plasmodium* spp., with a recent study in *P. chabaudi* AS implicating a number of *'cir'* proteins in rosetting of infected RBCs and merozoite

invasion (Yam *et al.*, 2016). The H4K8ac-bound '*bir*' genes identified in **Table 5.1** did not correlate with differential RNA expression. However, like other multigene families implicated in antigenic variation, such as *P. falciparum var* genes, H4K8ac-binding may only indicate that a gene is poised for expression, with a unique set of both activating and repressive histone modifications necessary for mutually exclusive transcription (Karmodiya *et al.*, 2015). Mapping of repressive histone modifications, such as H3K9me3 shown below in section **5.5,** may provide more evidence as to which '*bir*' family members are in fact transcribed, and which are silenced.

5.5 H3K9 trimethylation in *P. berghei* schizonts

Using identical sequencing methods and bioinformatic analysis as those used to delineate the role of H4K8ac-binding in mature schizonts (section **5.4**), the location of H3K9me3-bound regions in the *P. berghei* ANKA schizont genome were examined**.** Unlike H4K8 acetylation, H3K9me3-binding in *Plasmodium spp*. was previously shown to demarcate large heterochromatic domains in all subtelomeric regions and a few internal regions in some chromosomes (Salcedo-Amaya *et al.*, 2009; Fraschka *et al.*, 2018).

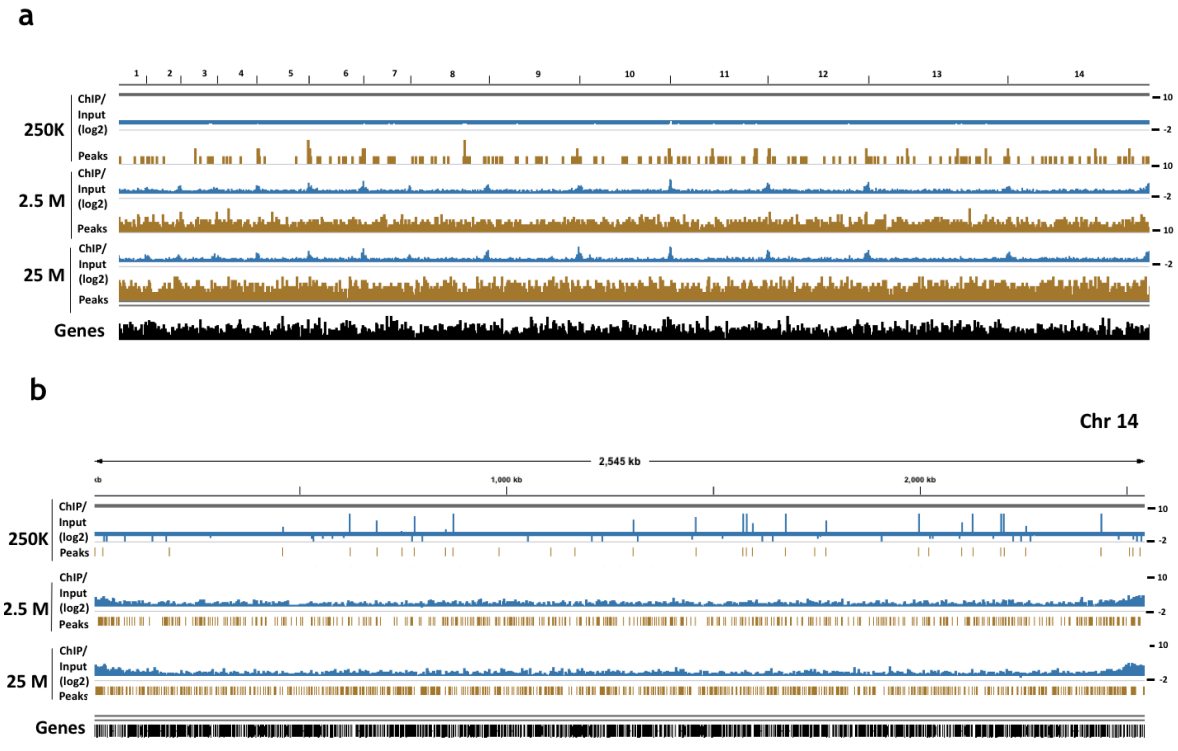


Figure 5.6: Distribution of H3K9 trimethylation across the *P. berghei* ANKA genome at schizont stage. Using the IGV browser, H3K9me3-bound (sample/input) reads (for 250K, 2.5 million, and 25

million schizonts) are shown across the entire *P. berghei* ANKA genome (version 34) alongside peaks called using the 'Epic' peak-calling software (in .bed format; results are normalised to the 250K schizont ChIP input sample) **(a).** At the base of these tracks, the general feature format (.gff) file shows the location of genes as black bars. For each sample, aligned ChIP data is displayed in blue. Epic peaks are shown in a separate track underneath ChIP data and are displayed in a dark yellow colour. In **(b)**, aligned reads and peak files are again shown as in **(a)**, though in this case (as for **Figure 5.3**, only the developmentally important chromosome 14 is shown in closer detail (homologous to *P. falciparum* chromosome 12) (J.-J. Lopez-Rubio, Mancio-Silva and Scherf, 2009; Fraschka *et al.*, 2018). Colour-coding is carried out as in **(a).**

From an initial comparison of *P. berghei* H3K9me3 ChIPmentation results in **Figure 5.6 (a),** it appears that, as in *P. falciparum*, the H3K9me3 modification demarcates heterochromatic regions at chromosome telomeres (Lopez-Rubio, Mancio-Silva and Scherf, 2009). The repressive H3K9me3 mark appears enriched at the telomeres of all 14 chromosomes, though when focussing in on chromosome 14 (**Figure 5.6 (b),** three H3K9me3-enriched loci at the centre of the chromosome are not present as in the homologous *P. falciparum* chromosome 12 (Lopez-Rubio, Mancio-Silva and Scherf, 2009). In the *P. falciparum* 3D7 strain used by Lopez-Rubio *et al.*, H3K9 trimethylation is present at the AP2-G transcription factor locus (PF3D7_1222600), at a subtelomeric locus of *rif, stevor,* and *var* group B genes (Chr12: 2,184,000 bp – 2,250,000 bp), and at a second upstream locus of *var* group C genes (Chr12: 1,686,000 bp – 1,746,000 bp). In *P. berghei* ANKA and other non-Laveranian *Plasmodium* spp., *var* genes are absent. However, in *P. berghei*, the AP2-G transcription factor is located on chromosome 14 (PBANKA_1437500) and this gene will be analysed in greater detail in repeat experiments of 2.5 million cells below**.**

As with H4K8 acetylation, the three separate ChIP peak-callers, MACS2, HOMER, and Epic, were used to determine H3K9me3-enriched regions of the genome. In addition to these results, and the alignment profile shown in **Figure 5.6,** an optimum cell number and peak-calling software were chosen to analyse repeat experiments. Shared H3K9me3 peaks for 250,000, 2.5 million, and 25 million schizonts are shown in **Figure 5.7.**

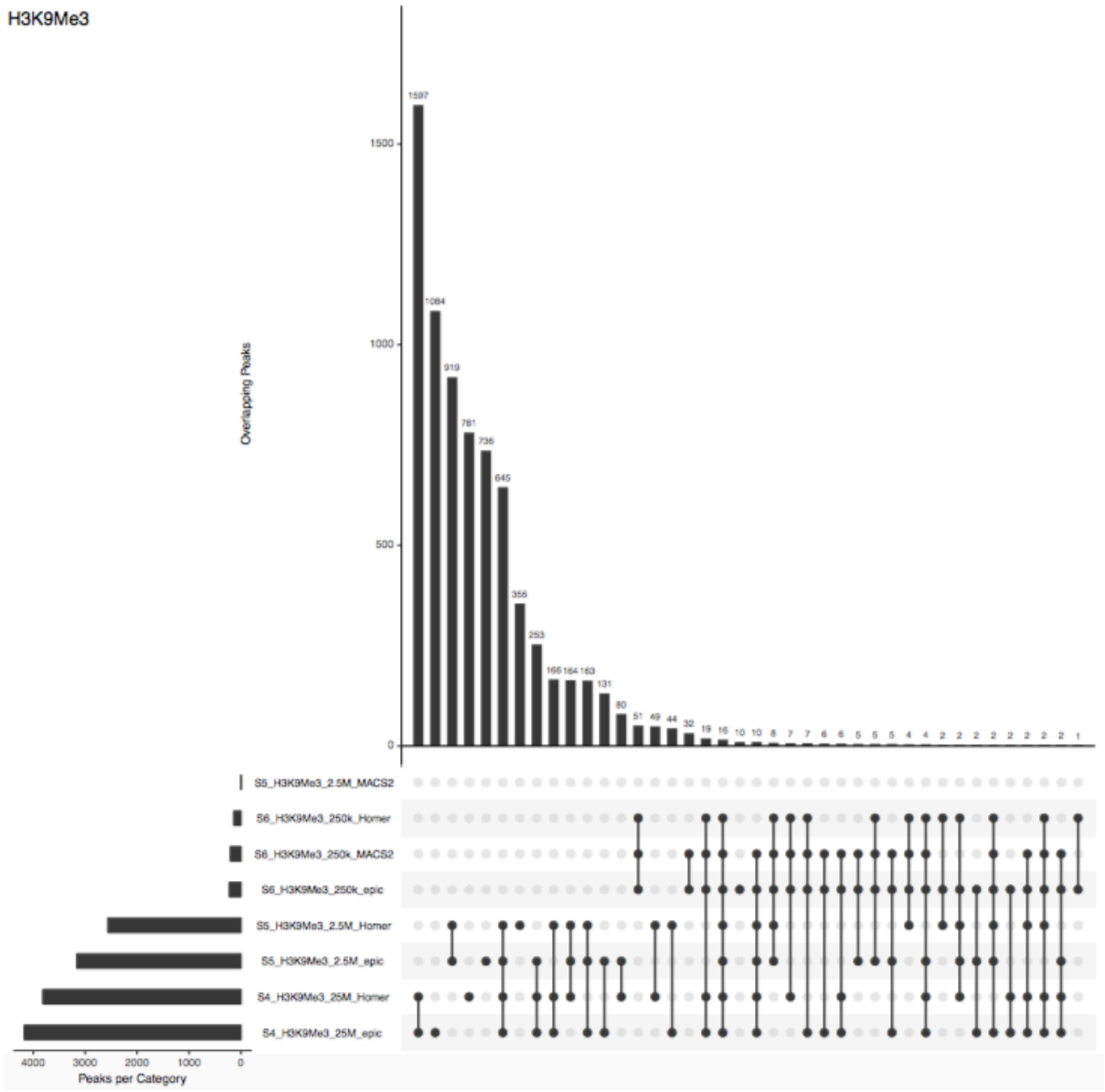


Figure 5.7: A comparison of shared and unique peaks from H3K9me3 ChIP samples using MACS2, HOMER, and Epic software. The present UpSet plot shows the number of unique peaks called using MACS2, HOMER, and Epic software on the left-hand side. Each sample name is shown next to the number of unique peaks per category, with cell numbers and software name clearly labelled. The main histogram depicts the number of overlapping peaks in each H3K9me3 sample (from 25 million, 2.5 million, and 250K *P. berghei* schizonts). At the base of the histogram, overlapping peaks shared between samples are indicated using circles, or circles joined with a straight line. Similar to the H4K8ac results (**Figure 5.2**), MACS2 software (Zhang *et al.*, 2008) identified only a small number of unique peaks for both 250K and 2.5 million cells, but with overlapping peaks only detected using 250K schizonts. For 25 million schizonts, MACS2 identified no unique or overlapping peaks. In contrast, both HOMER (Heinz *et al.*, 2010) and Epic (Xu *et al.*, 2014) identified peaks in all H3K9me3 ChIPmentation samples. 16 peaks were shared between all samples using both peak-callers.

Using the 'Epic' peak-caller (Xu *et al.*, 2014), 1084 overlapping peaks were identified using 25 million schizonts; 736 overlapping peaks were identified from 2.5 million schizonts; and 10 overlapping peaks were identified from 250,000 schizont-stage parasites. As can be observed in **Figure 5.7 (a),** the distribution of reads across the entire genome for both 25 million and 2.5 million schizonts were also similar, with only the 250K schizont sample falling behind with regards to coverage across all 14 *P. berghei* chromosomes. Because of this, H3K9me3 ChIPmentation was again repeated three times with 2.5 million schizonts from independent *P. berghei* infections. Sample JP2_H3K9me3 was then discarded from final results according to MDS plot results shown in **Figure 5.1.** Remaining results from duplicate H3K9me3 experiments of equal cell number are shown in **Figure 5.8.**

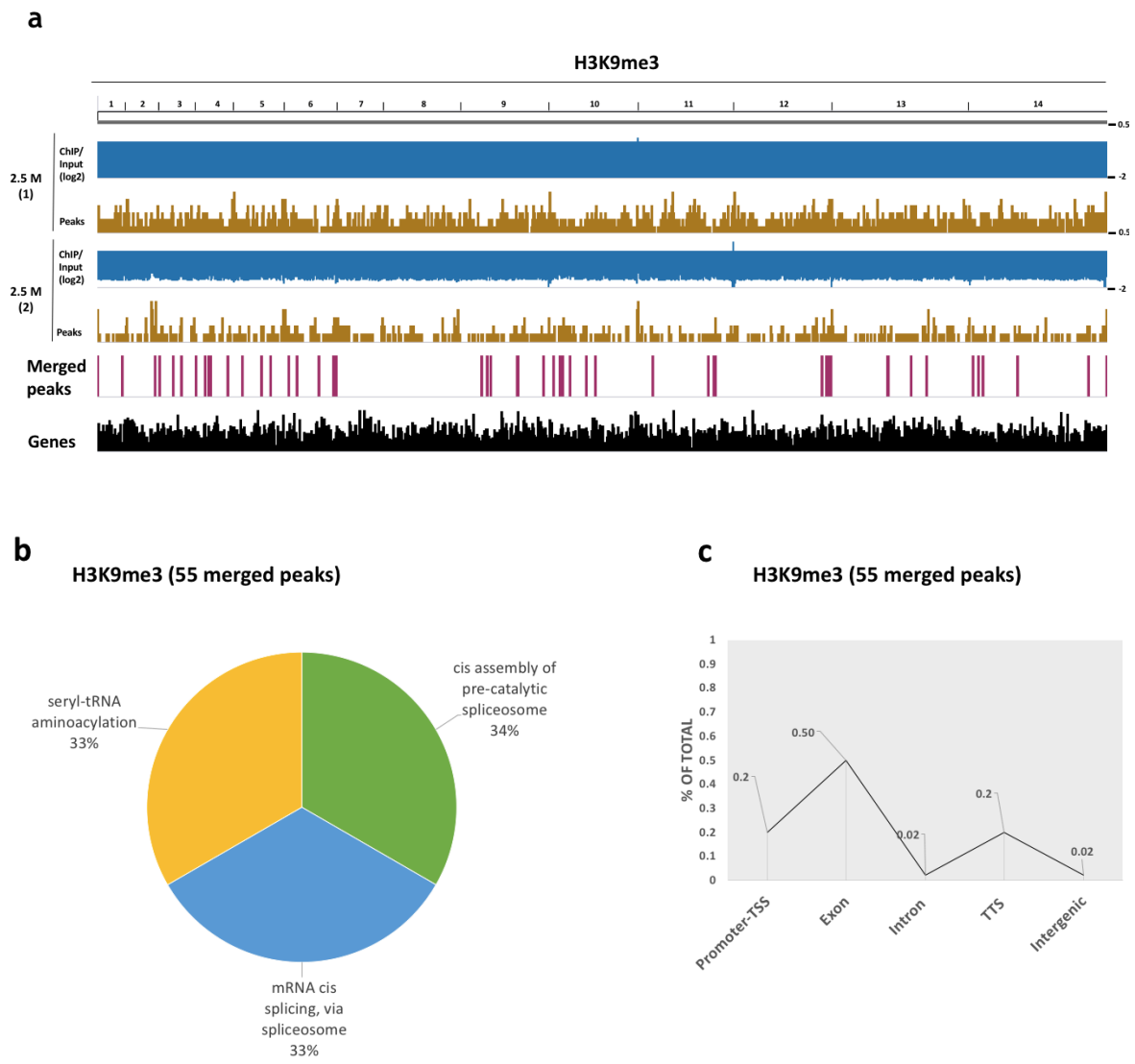
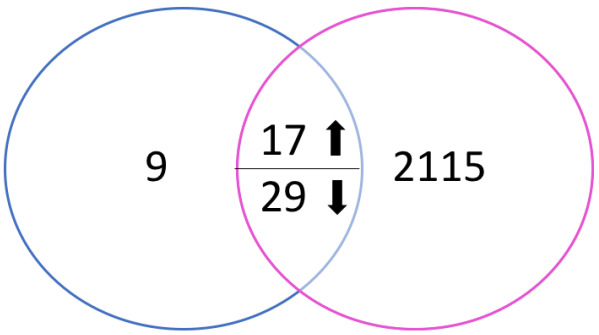


Figure 5.8: H3K9me3 merged peak distribution across the *P. berghei* ANKA genome at schizont stage. In this series of images, **(a)** shows H3K9me3 peak distribution across the entire genome of *P. berghei* using IGV. H3K9me3-bound reads (relative to an input sample from Fraschka *et al.*,

2018) from duplicate samples of 2.5 million schizonts are shown. Raw data are displayed in blue and Epic peaks are displayed in yellow underneath the ChIP track. Merged peaks for both replicates are shown in pink. Genes are displayed in general feature format (.gff) as black bars. **(b)** is a pie-chart representation of gene ontology (GO) terms for biological processes that were associated with the genes to which H3K9me3 was bound across both independent experiments (with a P-value < 0.025). Image **(c)** shows the same data as **(b)**, though this time, the location of the area of H3K9me3 enrichment is shown relative to the nearest gene ORF.

Unlike H4K8 acetylation, H3K9me3-enrichment occurs often at the promoter of a gene, or at the exon (Salcedo-Amaya *et al.*, 2009; Biga *et al.*, 2017). In *P. falciparum,* it has been suggested that clonally variant multicopy gene families (such as *var* genes) remain un-transcribed if the H3K9me3-binding site is located to the 3' end of the ORF (Karmodiya *et al.*, 2015). Of the duplicate H3K9me3 experiments undertaken during this study, 20% of merged peaks between both experiments were located at promoter-TSS regions (-1kb to +100 bp either side of the TSS), while 50% of merged peaks were located within exons (**Figure 5.8 (c)).**

When examining GO terms of biological processes associated with H3K9me3-bound loci, this repressive mark is enriched equally for seryl-tRNA aminoacylation, mRNA cis splicing via the spliceosome, and cis-assembly of the pre-catalytic spliceosome (**Figure 5.8 (b)**) (The Gene Ontology Consortium, 2017). To examine the relationship between H3K9 trimethylation and differential RNA expression in schizonts compared to 16 h trophozoites, a comparison of the overlap of merged peaks with differentially-expressed genes was undertaken and the results are shown in **Figure 5.9.**



GO term enrichment (P < 0.05)

GO ID	GO term	Fold enrichment
GO: 0006928	Movement of cell or subcellular component	11.98
GO: 0022904	Respiratory electron transport chain	46.73
GO: 0022900	Electron transport chain	29.21
GO: 0071976	Cell gliding	23.37

Gene ID	Name (location)	log_2FC
GO: 0000354	Cis assembly of pre-catalytic spliceosome *	147.58

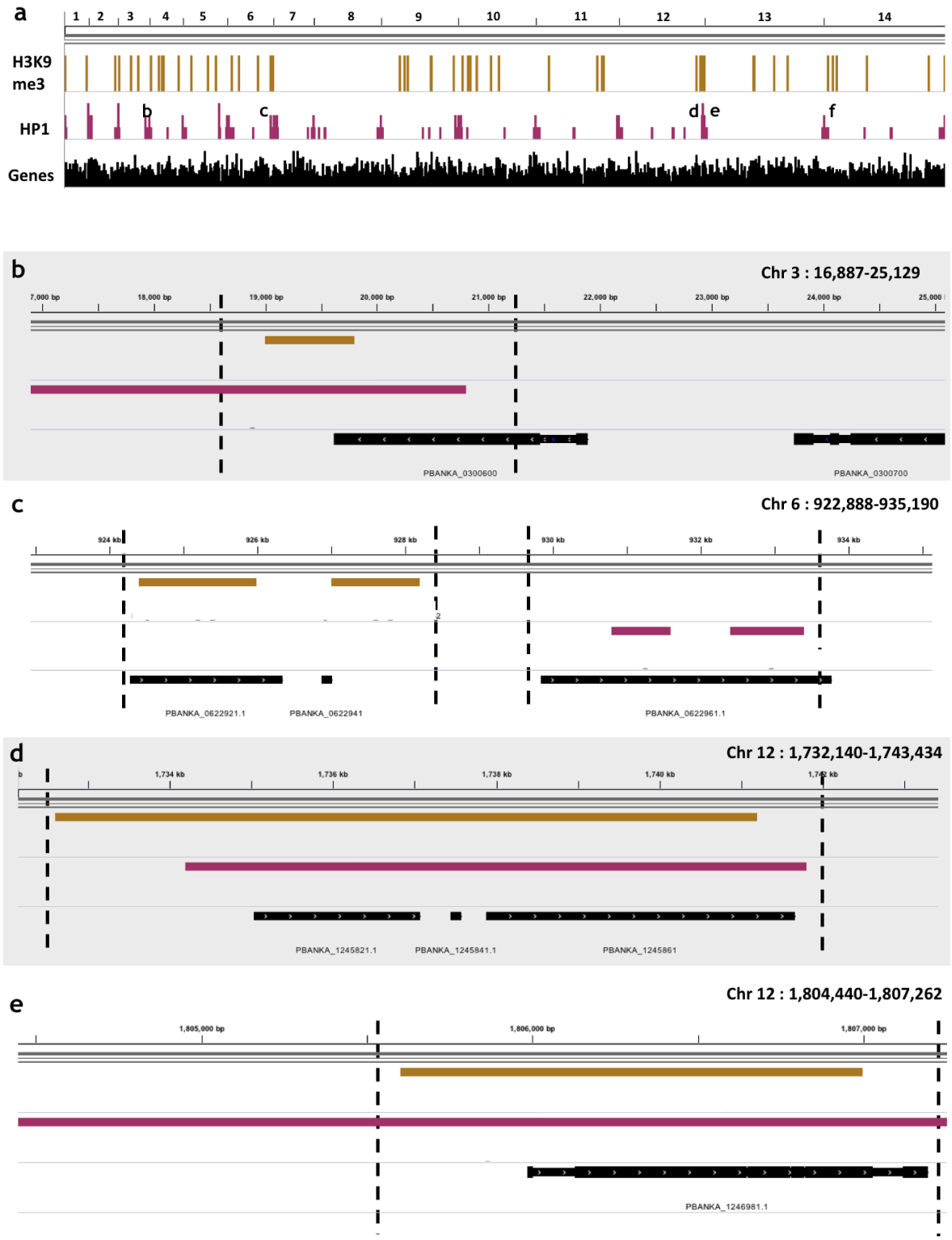
* < 0.01

Figure 5.9: H3K9me3-binding in the *P. berghei* ANKA schizont genome and association with differential RNA expression. The Venn diagram on the left of this image depicts the overlap between all H3K9me3 merged peaks and differentially-expressed genes between schizont- and trophozoite- stage *P. berghei* ANKA parasites (with a false discovery rate (FDR) less than 0.005). 55 merged peaks across two independent H3K9me3 ChIPmentation experiments were compared to 2161 differentially-expressed genes with FDR < 0.005. Of the 46 overlapping differentially-expressed genes, 17 were associated with up-regulation of RNA in schizonts and 29 were associated with down-regulation of RNA in schizonts. Only 9 H3K9me3 merged peaks did not correlate to any change in RNA expression between trophozoites and schizonts. Black arrows indicate increased or decreased RNA expression according to the direction of the arrow. Tables to the right of the Venn diagram show GO terms associated with up- and down- regulated genes that are bound at any location by H3K9 trimethylation. GO identifiers (GO IDs), description of GO term, and fold enrichment are listed for these genes.

As shown in **Figure 5.9,** 55 merged peaks were identified between both replicates of 2.5 million schizonts when ChIPmentation was carried out for H3K9 trimethylation. Though a correlation has been shown in *P. falciparum* between the presence of H3K9me3 at virulence genes and their epigenetic silencing via perinuclear repressive centres at the nuclear periphery (Lopez-Rubio, Mancio-Silva and Scherf, 2009), **Figure 5.9** shows that in *P. berghei* ANKA schizonts, H3K9 trimethylation was associated with both increased and decreased RNA expression in schizonts compared to trophozoites, as was the case with H4K8 acetylation (**Figure 5.5**). However, the subset of up-regulated genes associated with H4K8ac-binding at the exon, and H3K9me3-binding, were different. Whereas H4K8 acetylation was associated with up-regulation at genes responsible for DNA replication and host-parasite symbiosis (**Figure 5.5**), H3K9me3-marked genes associated with RNA up-regulation (where function was known) were associated with movement of a cell or subcellular component, cell gliding, and the electron transport chain ($P < 0.05$). As for H3K9me3 merged peaks associated with down-regulation of RNA expression in schizonts, only one GO term was enriched with a $P < 0.01$, and that was cis-assembly of the pre-catalytic spliceosome (GO: 0000354) (**Figure 5.9)** (The Gene Ontology Consortium, 2017)**.**

As previously mentioned, both H3K9me3- and HP1- binding is associated with monoallelic expression of clonally variant multicopy gene families, including those that are associated with antigenic variation (Lopez-Rubio, Mancio-Silva and Scherf, 2009; Fraschka *et al.*, 2018). To identify H3K9me3-enriched regions, and

to determine if any identifiable peaks are associated with HP1 and reduced transcription in schizonts, duplicate H3K9me3 experiments with 2.5 million schizonts were aligned to both HP1-bound peaks and RNA-seq data obtained from previous studies (Otto, Böhme, *et al.*, 2014; Fraschka *et al.*, 2018). All ChIP experiments were normalised using the input from Fraschka *et al.*, and Epic was used to determine ChIP peaks. Results are shown in **Figure 5.10.**



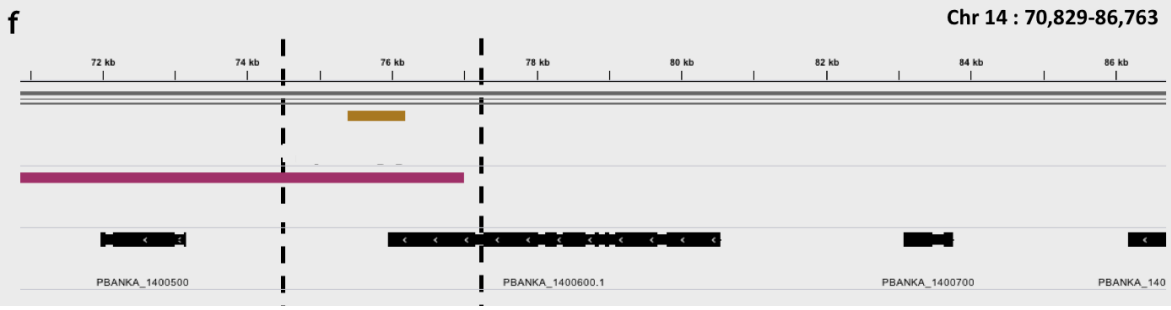


Figure 5.10: Comparative analysis of H3K9me3 binding and HP1 binding in *P. berghei* ANKA schizonts. This series of images depicts merged H3K9me3-enriched loci (Epic peaks) from duplicate independent experiments of 2.5 million schizonts. Merged peaks are aligned to the *P. berghei* ANKA genome (version 34) alongside Epic peaks called for HP1-enriched regions (Fraschka *et al.*, 2018). In each image, H3K9me3 merged peaks are shown in dark yellow, HP1 peaks are shown in pink, and the locations of genes are shown in black. In **(a),** the complete coverage across all 14 chromosomes is shown for H3K9me3 merged peaks and HP1 merged peaks in dark yellow and pink respectively. Five regions are labelled **'b'** to **'f'** in this image and these areas correspond to genes at which a H3K9me3 merged peak and HP1 peak are present at the same gene. **(b)** to **(e)** show the resulting genes and peaks corresponding to the lettering provided in **(a).** From **(b)** to **(f)**, dashed vertical lines are used to highlight loci at which peaks are located.

From results shown in **Figure 5.10 (b)** to **(f),** it can be seen that, though there is a clear enrichment of both H3K9me3 merged peaks and HP1 peaks at chromosome telomeres, only 5 genes in fact shared H3K9me3 and HP1 enrichment according to ChIPmentation and broad peak analysis carried out in the present study (**Figure 5.10 (a)).** These genes that shared both merged H3K9me3 peaks and HP1 peaks were PBANKA_0300600 **(Figure 5.10 (b)**, the locus containing PBANKA_1245821.1, PBANKA_1245841.1, and PBANKA_1245861 (**Figure 5.10 (d)**, PBANKA_1246981.1 **(Figure 5.10 (e),** the promoter-TSS of PBANKA_1400600.1 **(Figure 5.10 (f)),** and a cluster of genes on chromosome 6 at which duplicate peaks for located immediately adjacent to one another, PBANKA_0622921.1 and PBANKA_0622941 (both marked by a merged H3K9me3 peak), and PBANKA_0622961.1 which is immediately down-stream of these genes and is marked by two HP1 peaks **(Figure 5.10 (c)).**

Marked by H3K9me3/HP1-binding at its 3' end (TTS) (and with HP1-binding encompassing the exon also), PBANKA_0300600 is a *Plasmodium* exported protein of unknown function [as of 13-09-18] **(Figure 5.10 (b)**. Interestingly, this protein

appears to be completely specific to rodent malaria parasites, with orthologues so far being identified in *P. yoelii* (strains YM, 17X, and 17XNL) only (www.PlasmoDB.org, 2018). A protein with a highly-repetitive internal amino acid sequence, PBANKA_0300600 was shown to be non-essential in *P. berghei* sexual commitment and ookinete conversion (Kehrer, Frischknecht and Mair, 2016). However, PBANKA_0300600 does appear to be a factor of the *Plasmodium* gamete "egressome", i.e. it is released when the gamete egresses from the host erythrocyte within the mosquito vector (Kehrer, Frischknecht and Mair, 2016).

In **Figure 5.10 (c),** a particular genetic locus is shown at which two peaks each for H3K9me3- and HP1- binding are located immediately adjacent to one another, suggesting that these genes are located within a heterochromatic region of chromosome 6. Marked by H3K9me3 merged peaks, genes PBANKA_0622921.1 and PBANKA_0622941 encode for an 18S ribosomal RNA and a 5.8S ribosomal RNA, the former of which has been disrupted in *P. berghei* knockout experiments, resulting in a strongly delayed maturation of oocysts in the mosquito midgut, and a subsequently reduced number of salivary gland sporozoites [correct as of 13-09-18] (www.PlasmoDB.org, 2018) (van Spaendonk *et al.*, 2001). The adjacent HP1-bound PBANKA_0622961.1 encodes for a 28S ribosomal RNA, which, when disrupted, results in the same phenotype as the upstream H3K9me3-bound 18S rRNA: a strong delay of oocyst maturation, and reduced salivary gland sporozoites (though these sporozoites did remain infective in NIH Swiss mice) [correct as of 13-09-18] (www.PlasmoDB.org, 2018).

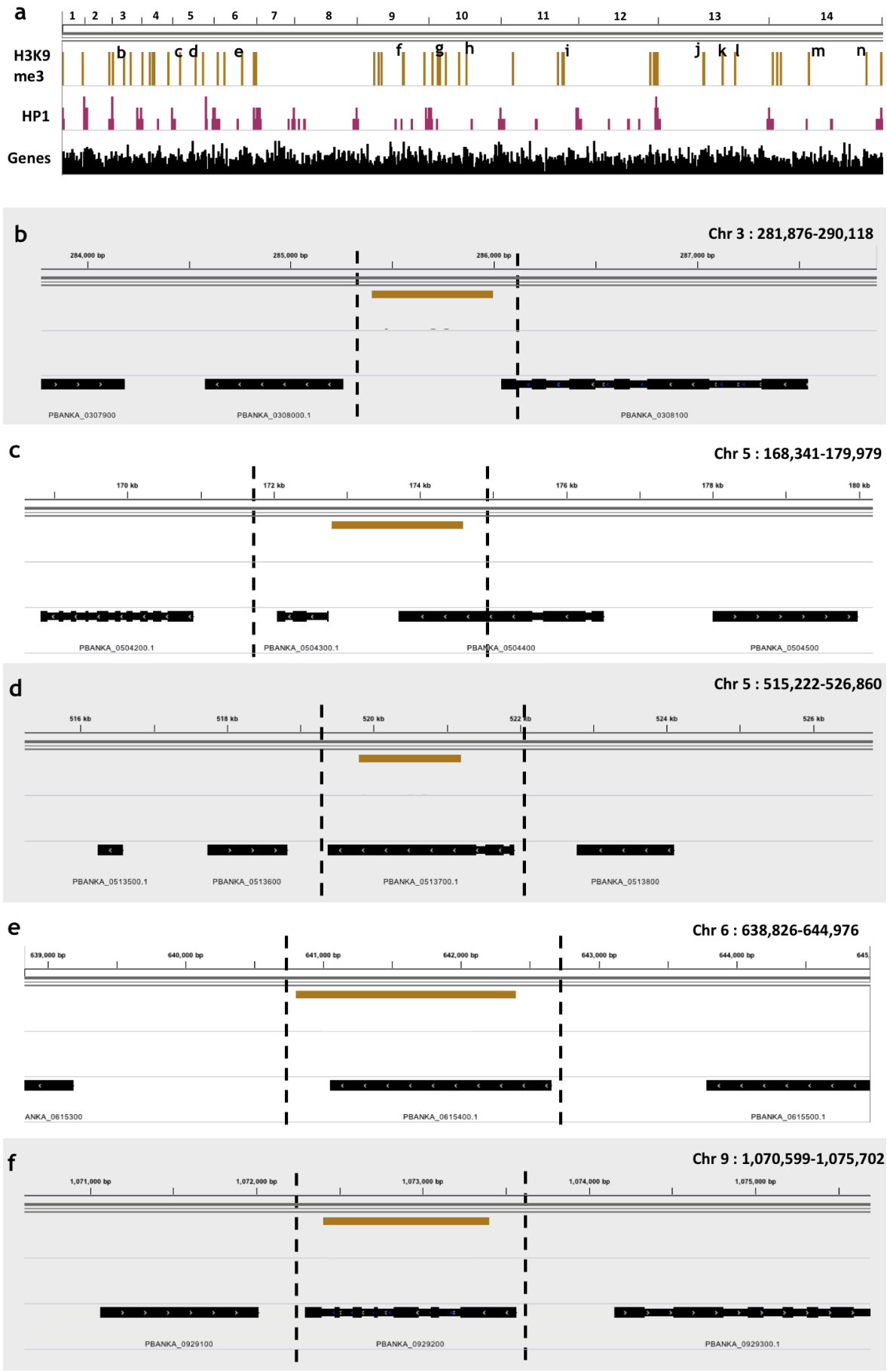
As for **Figure 5.10 (d),** a subtelomeric cluster of three genes on chromosome 12 are marked by both H3K9 trimethylation and HP1-binding. These genes are PBANKA_1245821.1, PBANKA_1245841.1, and PBANKA_1245861, encoding for an 18S ribosomal RNA, a 5.8S ribosomal RNA, and a 28S ribosomal RNA once more [correct as of 13-09-18] (www.PlasmoDB.org, 2018). Unlike with the previous cluster of rRNA-encoding genes on chromosome 6 **(Figure 5.10 (c))**, this cluster of rRNA genes have yet to be disrupted, or their proteins characterised [correct as of 13-09-18] (www.PlasmoDB.org, 2018).

In **Figure 5.10 (e)** and **Figure 5.10 (f),** two single genes are marked at the exon and exon/TTS respectively by both H3K9me3- and HP1- binding:

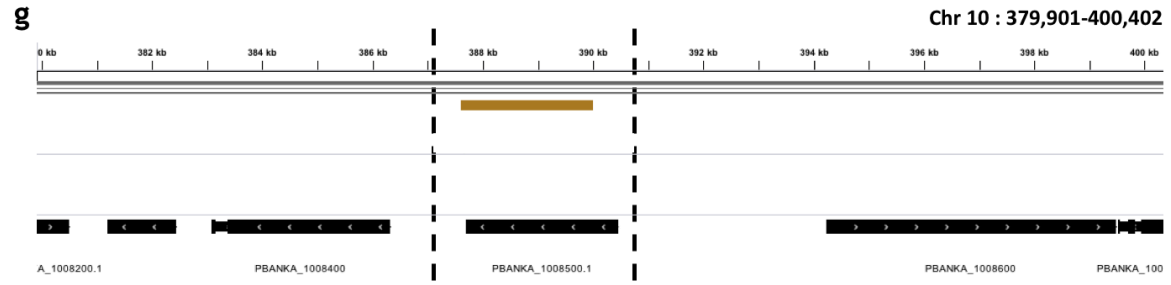
PBANKA_1246981.1 and PBANKA_1400600.1, encoding for a PIR protein pseudogene and a cytoadherence-linked asexual protein respectively [correct as of 13-09-18] (www.PlasmoDB.org, 2018). The former PIR protein pseudogene on chromosome 12 remains uncharacterised at present, but the latter cytoadherence-linked asexual gene (a member of the '*clag*' multigene family or sometimes '*rhoph1*' gene family) on chromosome 14 has been deemed as essential in erythrocytic stages of *P. berghei* development, and may be homologous in its function to the PF3D7_0302200 *clag 3.2* which remains invariably HP1-bound between multiple *Plasmodium* species (Gupta, Thiruvengadam and Desai, 2015; Fraschka *et al.*, 2018).

What is evident from **Figure 5.10 (a)** is that HP1-enrichment is not exclusively associated with the presence of H3K9me3-enrichment and *vice versa*. Also evident is the fact that HP1-binding, alone or in combination with H3K9me3, is not solely an indicator of transcriptional repression, with up-regulation of RNA expression in schizonts compared to trophozoites seen in the cases of PBANKA_1245861 (encoding for a 28S rRNA) (**Figure 5.10 (c)**) and PBANKA_1400600.1 (encoding for a *clag* protein) **(Figure 5.10 (f))**, though only the latter is up-regulated in schizonts to a significant extent (log_2FC of 2.16; P-value = 4.97×10^{-7}) **(Supplemental Material E4)**.

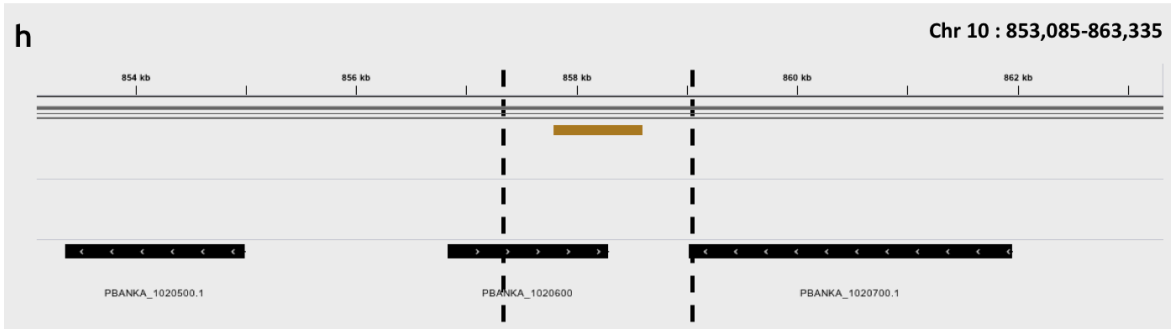
Similarly, a number of H3K9me3-enriched loci can be seen in **Figure 5.10 (a)** that are not associated with HP1-binding. Taking into consideration previous findings in *P. falciparum* showing that H3K9me3-enrichment occurs typically at clusters of clonally variant gene families involved in parasite antigenic variation (Lopez-Rubio, Mancio-Silva and Scherf, 2009; Salcedo-Amaya *et al.*, 2009), a study of H3K9me3-enriched loci in *P. berghei* ANKA schizonts was undertaken. Results are shown in **Figure 5.11.**


a
1 2 3 4 5 6 7 8 9 10 11 12 13 14
H3K9 me3
b c d e f g h i j k l m n
HP1
Genes
b
Chr 3 : 281,876-290,118
284,000 bp 285,000 bp 286,000 bp 287,000 bp
PBANKA_0307900 PBANKA_0308000.1 PBANKA_0308100
c
Chr 5 : 168,341-179,979
170 kb 172 kb 174 kb 176 kb 178 kb 180 kb
PBANKA_0504200.1 PBANKA_0504300.1 PBANKA_0504400 PBANKA_0504500
d
Chr 5 : 515,222-526,860
516 kb 518 kb 520 kb 522 kb 524 kb 526 kb
PBANKA_0513500.1 PBANKA_0513600 PBANKA_0513700.1 PBANKA_0513800
e
Chr 6 : 638,826-644,976
639,000 bp 640,000 bp 641,000 bp 642,000 bp 643,000 bp 644,000 bp
ANKA_0615300 PBANKA_0615400.1 PBANKA_0615500.1
f
Chr 9 : 1,070,599-1,075,702
1,071,000 bp 1,072,000 bp 1,073,000 bp 1,074,000 bp 1,075,000 bp
PBANKA_0929100 PBANKA_0929200 PBANKA_0929300.1

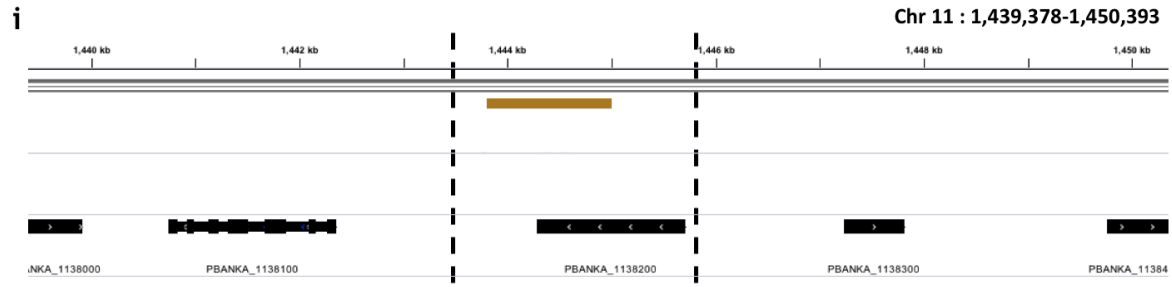
g


Chr 10 : 379,901-400,402
0 kb
382 kb
384 kb
386 kb
388 kb
390 kb
392 kb
394 kb
396 kb
398 kb
400 kb
A_1008200.1
PBANKA_1008400
PBANKA_1008500.1
PBANKA_1008600
PBANKA_100

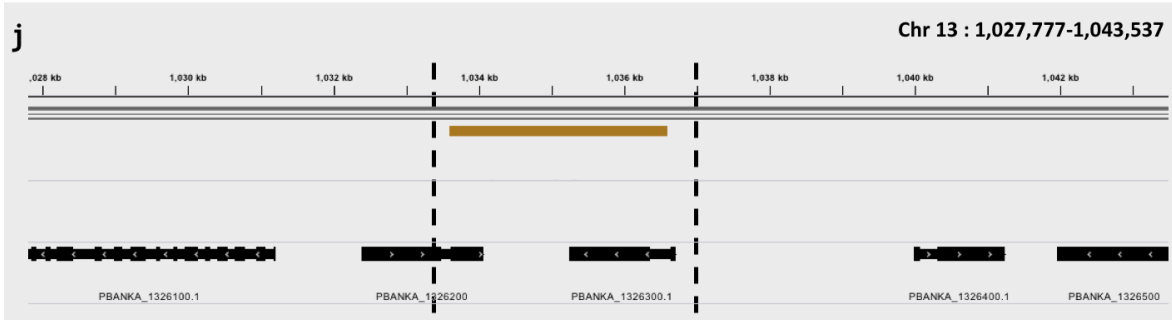
h


Chr 10 : 853,085-863,335
854 kb
856 kb
858 kb
860 kb
862 kb
PBANKA_1020500.1
PBANKA_1020600
PBANKA_1020700.1

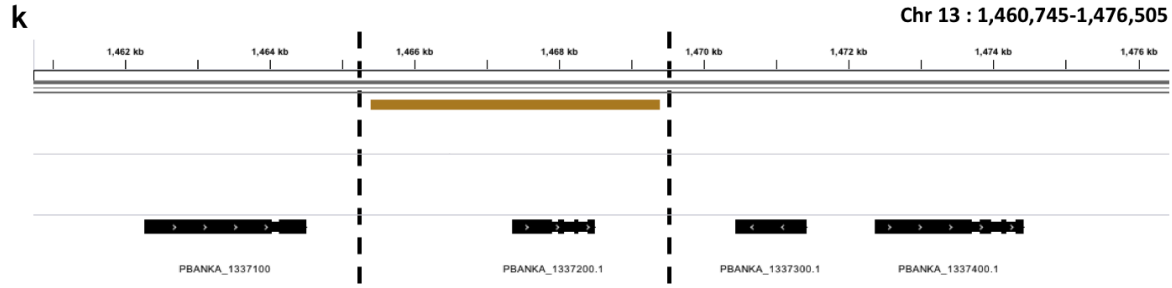
i


Chr 11 : 1,439,378-1,450,393
1,440 kb
1,442 kb
1,444 kb
1,446 kb
1,448 kb
1,450 kb
NKA_1138000
PBANKA_1138100
PBANKA_1138200
PBANKA_1138300
PBANKA_11384

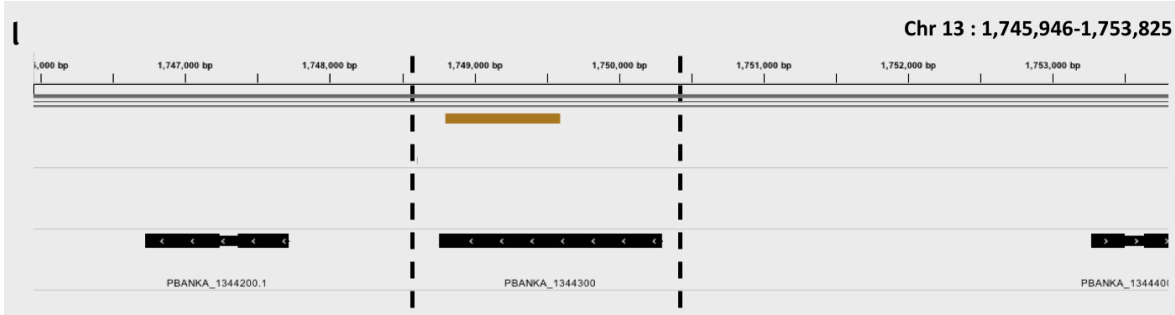
j


Chr 13 : 1,027,777-1,043,537
,028 kb
1,030 kb
1,032 kb
1,034 kb
1,036 kb
1,038 kb
1,040 kb
1,042 kb
PBANKA_1326100.1
PBANKA_1326200
PBANKA_1326300.1
PBANKA_1326400.1
PBANKA_1326500

k


Chr 13 : 1,460,745-1,476,505
1,462 kb
1,464 kb
1,466 kb
1,468 kb
1,470 kb
1,472 kb
1,474 kb
1,476 kb
PBANKA_1337100
PBANKA_1337200.1
PBANKA_1337300.1
PBANKA_1337400.1

l


Chr 13 : 1,745,946-1,753,825
,000 bp
1,747,000 bp
1,748,000 bp
1,749,000 bp
1,750,000 bp
1,751,000 bp
1,752,000 bp
1,753,000 bp
PBANKA_1344200.1
PBANKA_1344300

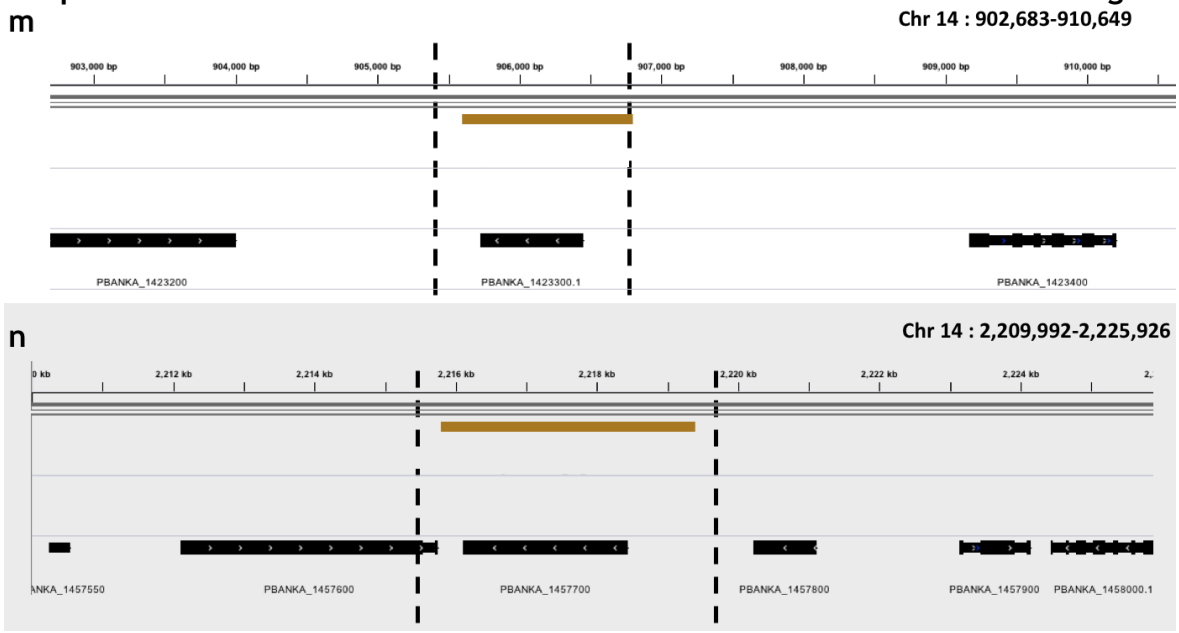


Figure 5.11: *P. berghei* schizont loci enriched for H3K9 trimethylation without HP1-enrichment. This series of images depicts merged H3K9me3-enriched loci (Epic peaks) from duplicate experiments of 2.5 million schizonts. Merged peaks are aligned to the *P. berghei* ANKA genome (version 34) alongside a track in which Epic peaks are called for HP1-enriched regions (Fraschka *et al.*, 2018). In each image, H3K9me3 merged peaks are shown in dark yellow, HP1 peaks are shown in pink, and the locations of genes are shown in black. From images **(b)** to **(n)**, no HP1 peaks are present at these H3K9me3-enriched loci and so no pink peaks are portrayed. Image **(a)** shows the merged H3K9me3 peak results for all 14 chromosomes above a track showing HP1 peaks across the *P. berghei* ANKA genome also. In **(a),** H3K9me3 merged peaks for which no corresponding HP1 peaks are seen are labelled **(b)** to **(n).** Zoomed-in views of these peaks are then depicted from **(b)** to **(n)** below the overview image. Genome coordinates are displayed in the upper right-hand corner of each image. All images were taken using the IGV genome browser. As in **Figure 5.10,** dashed vertical lines are used to highlight loci at which peaks are located.

As can be seen in **Figure 5.11 (a),** there are a number of H3K9me3-enriched loci in *P. berghei* ANKA schizonts that are not also enriched for HP1-binding. In **Figure 5.11 (a),** 13 of these solely H3K9me3-enriched loci are labelled **(b)** to **(n)**. In **Figure 5.11 (b),** a clear H3K9me3 peak is located at the intergenic region between PBANKA_0308100 and PBANKA_0308000.1, encompassing the promoter-TSS region of PBANKA_0308000.1 (encoding for a putative Ras-related protein Rab-5A). Expression of this gene is greatest at 22 h schizonts (Otto *et al.*, 2014), and its function is haemoglobin uptake at the parasite plasma membrane and transport to the food vacuole (Ebine *et al.*, 2016).

In **Figure 5.11 (c)** and **Figure 5.11 (d),** H3K9 trimethylation is associated with the 5' region of PBANKA_0504400, the promoter-TSS region of PBANKA_0504300.1, and the exon of PBANKA_0513700.1, encoding for a conserved *Plasmodium* protein of unknown function, a putative flagellar outer arm dynein-associated protein, and a second conserved *Plasmodium* protein of unknown function [correct as of 13-09-18] (www.PlasmoDB.org, 2018). Regarding the characterised PBANKA_0504300.1 **(Figure 5.11 (c)),** this protein is heavily associated with the male gametocyte (Otto *et al.*, 2014; Yeoh *et al.*, 2017) and is also subject to AP2-G2 expression (Yuda *et al.*, 2015). Regarding both conserved *Plasmodium* proteins of unknown function, the PBANKA_0504400 gene, marked at the exon and TTS by H3K9 trimethylation, appears to be involved in female gametocytogenesis, in direct contrast to its neighbour, PBANKA_0504300.1 (Otto *et al.*, 2014; Yeoh *et al.*, 2017). This gene is also under control of the AP2-G2 transcriptional regulator (Yuda *et al.*, 2015), suggesting that the locus depicted in **Figure 5.11 (c)** is regulated by AP2-G2, and H3K9 trimethylation (both suggestive of transcriptional repression), but not with HP1.

Depicted in **Figure 5.11 (e) and Figure 5.11 (f)** are two genes that are completely associated at the ORF with merged H3K9me3 peaks from repeat experiments: PBANKA_0615400.1 and PBANKA_0929200. These genes encode for a putative serine-tRNA ligase, and a conserved *Plasmodium* protein of unknown function, both of which remain uncharacterised, though the former is known to be expressed at intra-erythrocytic parasite stages and is amenable to disruption [correct as of 13-09-18] (www.PlasmoDB.org, 2018). From **Figure 5.11 (b)** to **Figure 5.11 (f),** all but one gene (PBANKA_0615400.1) are associated with up-regulation in 22 h schizonts compared to 16 h trophozoites, though all are marked with H3K9 trimethylation **(Supplemental Material ES4)** (Otto *et al.*, 2014)**.**

Though H3K9me3-enrichment is associated with repression of transcription (Lopez-Rubio, Mancio-Silva and Scherf, 2009; Salcedo-Amaya *et al.*, 2009), H3K9 di-methylation and trimethylation fail to bind HP1 in the presence of H3S10 phosphorylation (Hirota *et al.*, 2005). In the present study, it was observed that both H3S10 and H3.3S10 were phosphorylated at schizont stage in *P. berghei* **(Table 4.3**). Therefore, it may not be the level of H3K9 trimethylation that is

important for heterochromatin formation alone. It may be that differential transcription between *P. berghei* ANKA stage transitions is facilitated by a H3K9me3-H3S10ph ratio, with H3S10ph resulting in heterochromatin dissociation (Hirota *et al.*, 2005; Bannister and Kouzarides, 2011). This histone crosstalk may also explain why, in *P. berghei* schizonts, the differential expression occurring at genes with H3K9me3 peaks is not uniquely repressive in nature **(Figure 5.9)**.

In **Figure 5.11 (g)** and **Figure 5.11 (h),** two more genes are associated with H3K9 trimethylation at the exon: PBANKA_1008500.1 and PBANKA_1020600, encoding for translocon component PTEX150 and putative actin-related protein ARP4a [correct as of 13-09-18] (www.PlasmoDB.org, 2018). The translocon component PTEX150 is refractory to gene deletion and essential for intra-erythrocytic growth in its role as a core component of the PTEX protein export complex (Garten *et al.*, 2018). The actin-related ARP4a protein is also refractory to deletion in *P. berghei,* and both ARP4 and ARP6 are suggested to be subunits of a chromatin-remodelling complex in Apicomplexa (Gordon and Sibley, 2005). Both genes are down-regulated in 22 h schizonts in comparison with 16 h trophozoites, though only ARP4a to a statistically significant degree (log_2FC: -1.73; P-value = 3.73 x 10^{-6}) **(Supplemental Material ES4).**

Of the final 6 H3K9me3-bound genes (**Figure 5.11 (i)** to **Figure 5.11 (n)),** all merged peaks encompass almost the entire ORF of each gene. These genes are: PBANKA_1138200 (putative RuvB-like helicase 3), PBANKA_1326300.1 (putative cytochrome c1 precursor), PBANKA_1337200.1 (putative protein transport protein SFT2), PBANKA_1344300 (a putative chromatin assembly factor 1 subunit), PBANKA_1423300.1 (DNA/RNA-binding protein Alba1), and PBANKA_1457700 (ookinete-expressed protein PSOP26), 3 of which are down-regulated, and 3 of which are up-regulated in 22 h schizonts compared to 16 h trophozoites (www.PlasmoDB.org, 2018) **(Supplemental Material ES4)** (Otto *et al.*, 2014)**.**

With regard to their expression and functions, RuvB-like helicase 3 is expressed greatly in female gametocytes and appears to be AP2-G2-regulated, in addition to being marked by H3K9me3 (Otto *et al.*, 2014; Yuda *et al.*, 2015; Yeoh *et al.*, 2017). The putative cytochrome c1 precursor (PBANKA_1326300.1) is refractory to disruption and essential for intra-erythrocytic growth, while protein transport

protein SFT2 in *P. berghei* ANKA is non-essential, and expressed in all parasite stages except female gametocytes (reduced RNA transcription at this life-cycle stage) (Otto *et al.*, 2014; Yuda *et al.*, 2015; Yeoh *et al.*, 2017) (www.PlasmoDB.org, 2018).

Of the final three H3K9me3-bound proteins, the chromatin assembly factor 1 subunit (PBANKA_1344300) is essential for asexual growth, and is expressed predominantly in 22 h schizonts and male gametocytes, with a suggested interaction with AP2-G2 (Otto *et al.*, 2014; Yuda *et al.*, 2015; Yeoh *et al.*, 2017). Alba1 (encoded by PBANKA_1423300.1) is predominantly expressed during asexual *P. berghei* life-cycle stages, while PBANKA_1457700 (the ookinete-expressed PSOP26) is dispensable, unregulated by AP2-G2, and associated with gametocytes, particularly male gametocytes, along with being characterised as ookinete-specific and under control by the AP2-O transcriptional regulator (Otto *et al.*, 2014; Yuda *et al.*, 2015; Zheng *et al.*, 2016; Yeoh *et al.*, 2017).

In summary, of the merged H3K9me3 peaks from duplicate experiments of *P. berghei* mature schizonts without concomitant HP1-binding, 5 were positively associated with expression in gametocytes (PBANKA_0504300.1, PBANKA_0504400, PBANKA_1138200, PBANKA_1344300, and PBANKA_1457700) while one was negatively associated with RNA expression in female gametocytes (PBANKA_1337200.1). All 5 H3K9me3-bound genes that are associated with positive transcription in gametocyte stages were also associated with regulation by either AP2-G2, or AP2-O (in the case of PBANKA_1457700) (Otto *et al.*, 2014; Yuda *et al.*, 2015; Zheng *et al.*, 2016; Yeoh *et al.*, 2017). One of the more interesting results can be seen in **Figure 5.11 (c)**, in which two adjacent H3K9me3-bound genes on chromosome 13 are both AP2-G2-associated, one being expressed in male gametocytes (PBANKA_0504300.1), and the other being expressed almost solely in female gametocytes (PBANKA_0504400) (Otto *et al.*, 2014; Yuda *et al.*, 2015). Though both H3K9me3- and AP2-G2- associated, it remains to be seen how both adjacent genes can be associated with differential transcription between male and female gametocytes.

To complete a study of H3K9me3 merged peaks, a list of the genes to which the highest-scoring H3K9me3 Epic peaks are associated are depicted in **Table 5.2.** All

further peaks discovered, and a list of merged peaks common to both duplicate H3K9me3 experiments are listed in **Electronic Supplemental Material ES3.**

	Gene ID	Function	H3K9me3 location	Highest Epic peak score (of 3)	Differential RNA expression at schizont stage	Log_2FC (if present)	P-value (if present)	FDR (if present)
1	PBANKA_0007100.1	PIR protein pseudogene	Promoter-TSS/exon	2163.49	No	-	-	-
2	PBANKA_1100051.1	PIR protein pseudogene	Promoter-TSS/exon	1602.50	No	-	-	-
3	PBANKA_0001501.1	PIR protein	Intergenic/TTS	1079.45	No	-	-	-
4	PBANKA_0008001.1	PIR protein	Intergenic/intron	1054.49	No	-	-	-
5	PBANKA_0700021.1	PIR protein pseudogene	Promoter-TSS	499.2	No	-	-	-
6	PBANKA_1246921.1	PIR protein	Exon	438.82	No	-	-	-
7	PBANKA_1300061.1	PIR protein	Promoter-TSS/exon/TTS	433.6	Yes	1.15	0.4	0.49
8	PBANKA_1300021.1	PIR protein	TTS	429.43	No	-	-	-
9	PBANKA_0600021.1	Fam-a protein	Promoter-TSS/TTS	428.68	No	-	-	-
10	PBANKA_MIT01900.1	Cytochrome B (CYTB)	Promoter-TSS	414.5	Yes	1.33	0.009	0.018

Table 5.2: Highest-scoring H3K9me3 peaks in *P. berghei* schizonts. This table displays the top 10 H3K9me3-bound genes according to their Epic peak-calling score. Genes are shared among duplicate independent H3K9me3 ChIPmentation experiments. The highest score between replicates is shown in the fifth column. The list of genes is numbered from highest to lowest score and the first and second columns of the table list gene IDs and current known function (if any) of the gene in *P. berghei* ANKA [as of 26-08-18]. The fourth column shows the locations of the H3K9me3 peak relative to the gene in question across both independent experiments. The sixth column indicates differentially-expressed gene regulation in schizonts at 22 h compared to 16 h trophozoite expression (Otto *et al.*, 2014). The following three columns list the log_2 fold change (log_2FC), P-value, and false discovery rate (FDR) (where applicable) of the differential RNA expression.

Remarkably, all but one of the H3K9me3-enriched genes shown in **Table 5.2** are members of the PIR protein multicopy gene family, or the fam-a multicopy gene family in *P. berghei* (Fougère *et al.*, 2016; Yam *et al.*, 2016; Fraschka *et al.*, 2018). Regarding PIR and fam-a proteins, genes from both of these subtelomeric multigene families have a wide array of functions throughout the parasite developmental cycle, nearly all of which remain uncharacterised (Fougère *et al.*, 2016; Yam *et al.*, 2016). Of the genes shown in **Table 5.2**, the PIR protein pseudogene PBANKA_1100051.1, appears in the top 10 H4K8ac-enriched genes also (**Table 5.1**), though without differential RNA expression between trophozoite and schizont stages.

Of the ten H3K9me3-enriched genes shown in **Table 5.2**, only one gene (PBANKA_MIT01900.1; mitochondrial cytochrome b (CYTB) protein), showed

significant differential RNA expression in 22 h schizonts compared to 16 h trophozoites; a significant up-regulation of transcription (log_2FC of 1.33; P-value: 0.009; FDR: 0.02). This gene is suggested as being responsible for susceptibility and resistance of *P. falciparum* to 8-aminoquinolines and hydroxynapthoquinones (Vaidya *et al.*, 1993), with new research determining that *P. falciparum* resistance to the napthoquinone antimalarial, atovaquone, is in fact a result of mitochondrial heteroplasmy, and the failure of atovaquone to inhibit the cytochrome *bc1* complex of the electron transport chain (https://doi.org/10.1101/232033). Though *var* genes are indeed regulated epigenetically, it is unknown as to whether epigenetic marks such as H3K9 trimethylation play an additional role in mitochondrial heteroplasmy, or regulation of mitochondrial genome replication in general (Deshmukh, Srivastava and Dhar, 2013).

The results shown in **Table 5.2** suggest that, though H3K9 trimethylation has been implicated in transcriptional repression in *Plasmodium* spp. (Lopez-Rubio, Mancio-Silva and Scherf, 2009; Salcedo-Amaya *et al.*, 2009; Fraschka *et al.*, 2018), it is most likely that clonally variant multicopy gene families are regulated by a conglomeration of histone modifications and other regulatory processes, such as long non-coding RNAs (lncRNAs), antisense RNAs, or exoribonucleases (Liao,Q., Shen,J., Liu,J., Sun,X., Zhao,G., Chang,Y., Xu,L., Li,X., Zhao,Y., Zheng,H. Zhao, Y. and Wu, 2014; Zhang *et al.*, 2014; Filarsky *et al.*, 2018). However, with a number of HP1 and H3K9me3-enriched loci located at, or adjacent to, genes involved in sexual development **(Figures 5.10 and 5.11),** it may be that some of the PIR and fam-a genes shown in **Table 5.2** could also be involved in the transmission stage of *P. berghei* ANKA. Therefore, these genes are good candidates for further functional studies.

5.6 H3K9 acetylation in *P. berghei* schizonts

To complement the findings of ChIPmentation studies using an anti-H3K9me3 antibody, identical examinations of 250,000, 2.5 million, and 25 million mature synchronous schizonts were carried out to map the binding of an anti-H3K9ac antibody. In *Plasmodium spp.*, H3K9ac-binding correlates with transcriptional activation at gene promoters and H2A.Z binding (Bártfai *et al.*, 2010). Using the MACS2, Epic, and HOMER peak-callers once more, H3K9ac-bound regions were

identified and the resulting peak numbers are depicted as an UpSet plot in **Figure 5.12.** An overview of H3K9ac peak distribution is also depicted in **Figure 5.13.**

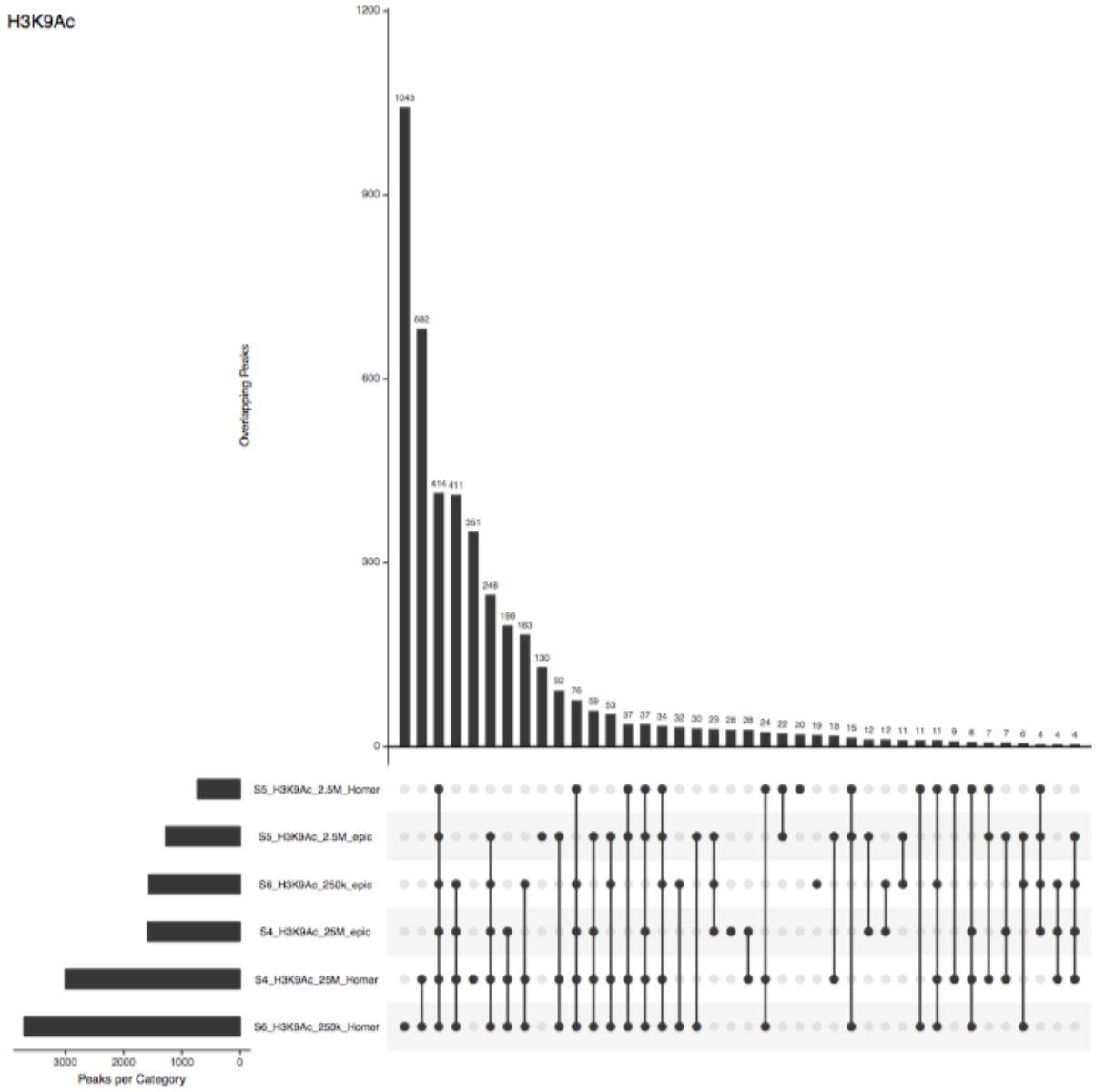


Figure 5.12: A comparison of shared and unique peaks from H3K9ac ChIP samples using MACS2, HOMER, and Epic software. The present UpSet plot shows the number of unique peaks called using MACS2, HOMER, and Epic software on the left-hand side. Each sample name is shown next to the number of unique peaks per category, with cell numbers and software name clearly labelled. The main histogram depicts the number of overlapping peaks in each H3K9as sample (from 25 million, 2.5 million, and 250K *P. berghei* schizonts). At the base of the histogram, overlapping peaks shared between samples are indicated using circles, or circles joined with a straight line. Using the anti-H3K9ac antibody, MACS2 software (Zhang *et al.*, 2008) identified no unique peaks for any of the samples analysed. However, when using Epic and HOMER peak-calling software, 414 overlapping peaks were identified across all samples, with HOMER identifying 1043 overlapping peaks from the 250K schizont sample alone.

From a comparison of H3K9ac ChIPmentation results using both HOMER and Epic peak-calling software, it is clear that HOMER identified more samples from 250K schizonts than any other sample or software used (**Figure 5.12**). However, for the purposes of comparative analysis following visualisation of both H4K8ac and H3K9me3 results, Epic was again used for peak-calling of H3K9ac-enriched regions of the genome. Epic peaks for 250,000, 2.5 million, and 25 million schizonts are shown in **Figure 5.12**, with results from 2.5 million schizonts in triplicate (normalised to the Fraschka *et al.*, 2018 input once more) carried out subsequently. Taking into consideration MDS analysis (**Figure 5.1**), final results shown in **Figure 5.14** include duplicate samples S5_H3K9Ac and JP3_H3K9Ac, with third samples, JP6_H3K9Ac being discarded.

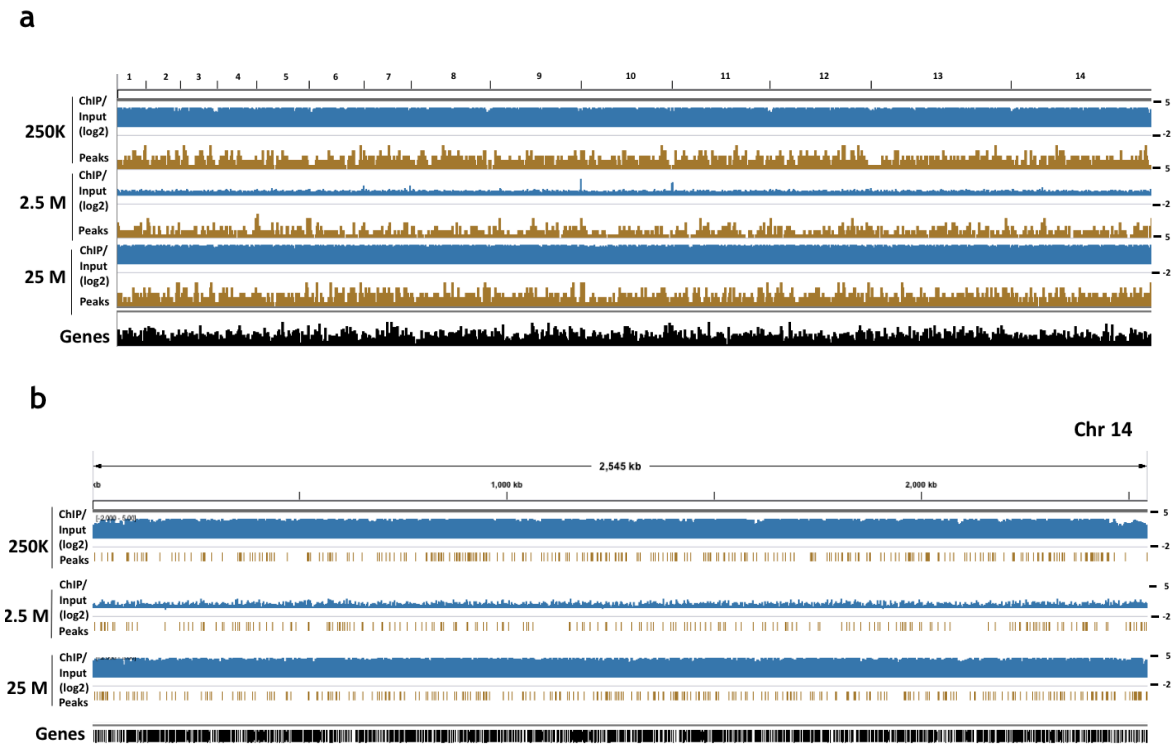
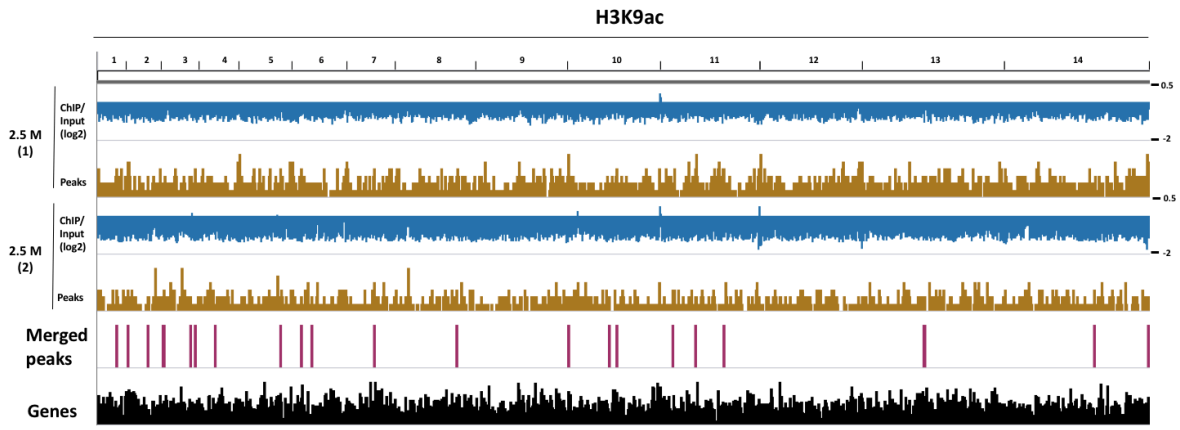


Figure 5.13: Distribution of H3K9 acetylation across the *P. berghei* ANKA genome at schizont stage. Using the IGV browser, H3K9ac-bound (sample/input) reads (for 250K, 2.5 million, and 25 million schizonts) are shown across the entire *P. berghei* ANKA genome (version 34) alongside peaks called using the 'Epic' peak-calling software (in .bed format; results are normalised to the 250K schizont ChIP input sample) **(a).** At the base of these tracks, the general feature format (.gff) file shows the location of genes as black bars. For each sample, aligned ChIP data is displayed in blue. Epic peaks are shown in a separate track underneath ChIP data and are displayed in a dark yellow colour. In **(b)**, aligned reads and peak files are again shown as in **(a)**, though in this case (as for **Figure 5.3** and **Figure 5.6)**, only the developmentally important chromosome 14 is shown in closer

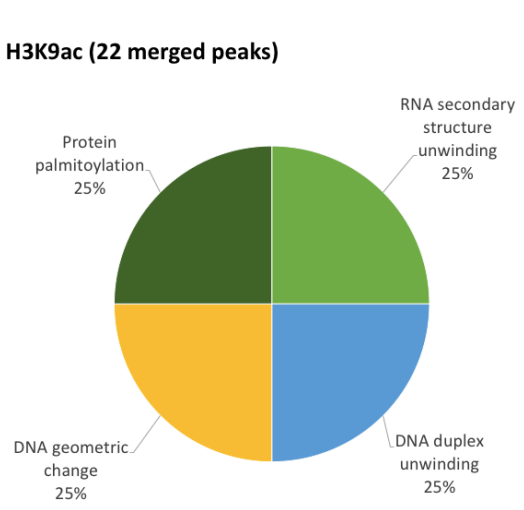
detail (J.-J. Lopez-Rubio, Mancio-Silva and Scherf, 2009; Fraschka *et al.*, 2018). Colour-coding is carried out as in **(a).**

Unlike both H4K8ac and H3K9me3 ChIPmentation experiments, visualisation of raw ChIP data and Epic peaks for H3K9 acetylation across the genome **(Figure 5.12)** indicated that ChIPmentation of 250,000 schizonts worked quite well for examination of antibody enrichment. However, in the case of H3K9ac-enrichment with 250,000 and 25 million schizonts, both the genome overview **(Figure 5.13 (a))** and overview of chromosome 14 **(Figure 5.13 (b))** show IGV tracks that appears to have a considerable amount of background reads that may skew Epic peak results. As a result, and to remain consistent with previous experiments, H3K9ac-enriched regions were once again examined from 2.5 million schizont-stage parasites in duplicate. These results were analysed using Epic, with annotation carried out using a HOMER annotation script (http://homer.ucsd.edu/homer/). Results are shown in **Figure 5.14.**

a



b



c

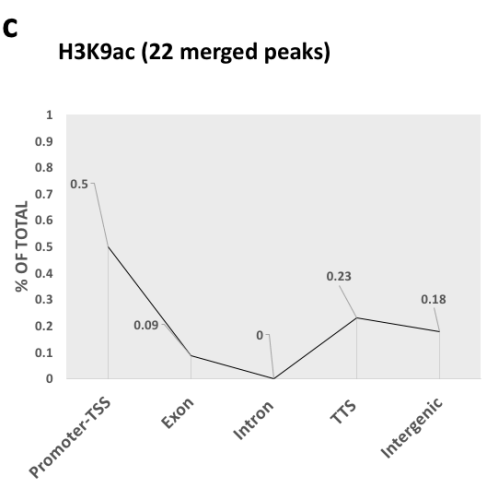


Figure 5.14: H3K9ac peak distribution across the *P. berghei* ANKA genome. In this series of images, **(a)** shows H3K9ac peak distribution across the entire genome of *P. berghei* using IGV. H3K9ac-bound reads (relative to an input sample from Fraschka *et al.*, 2018) from duplicate independent samples of 2.5 million schizonts are shown. Raw data are displayed in blue and Epic peaks are displayed in yellow underneath the ChIP track. Genes are displayed in general feature format (.gff) as black bars. **(b)** is a pie-chart showing gene ontology (GO) terms for biological processes that were associated with the genes to which H3K9ac was bound (showing results for 22 merged peaks across both independent experiments). Image **(c)** shows the same data, but the locations of H3K9ac-enrichment relative to the nearest gene are depicted in graph format (as a percentage of 100%).

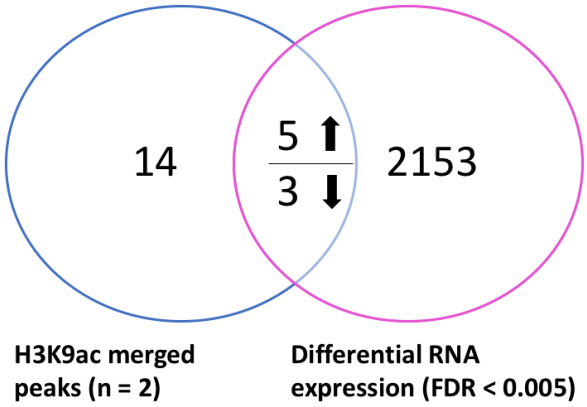
As can be seen in **Figure 5.14 (a),** merged Epic peak distribution from duplicate H3K9ac ChIPmentation experiments is very sparse, suggesting that either few loci are marked by H3K9 acetylation, or that this particular histone modification is very dynamic in nature, with few conserved peaks between samples, even samples of the same *P. berghei* ANKA life-cycle stage. Taking a look at GO terms for biological processes enriched in these samples ($P < 0.05$) using the 22 merged peaks from duplicate experiments (**Figure 5.14 (b)),** it appears that H3K9 acetylation peaks are equally associated with protein palmitoylation, DNA duplex unwinding, DNA geometric change, and RNA secondary structure unwinding (The Gene Ontology Consortium, 2017).

In a previous ChIP-on-chip profiling study, H3K9 acetylation was shown to be homogenous across both active and inactive genes, with the mark moving to the 5'end of active genes in schizonts (Salcedo-Amaya *et al.*, 2009). In the present study, and as seen in **Figure 5.14 (c),** merged peaks are indeed associated with 5' promoter-TSS regions predominantly (50%). Of genes that were associated with up-regulated RNA expression (shown in **Figure 5.15** below), all H3K9ac peaks were located upstream of the gene in question, either at the upstream intergenic/promoter-TSS region, or at the 5' end of the ORF (**Supplemental Material ES3).**

In a separate study, acetylation of H3K9 at promoter regions was also found to be indicative of gene activation in *P. falciparum* (Bártfai *et al.*, 2010). In *P. berghei,* if H3K9ac at promoter regions is to be associated with active gene transcription, the processes identified in **Figure 5.14 (b)** should predominantly be activated in

schizonts. Previous studies suggest that genes encoding for the processes shown in **Figure 5.14 (b),** namely translation (DNA unwinding; DNA geometric change) and other protein-related metabolic processes (protein palmitoylation), are in fact associated with peak transcription and mRNA stabilisation at late trophozoite stage *in vivo* in *P. falciparum* (Bozdech *et al.*, 2003; Painter *et al.*, 2018).

To examine whether H3K9ac enrichment was associated with up-regulation of RNA expression between *P. berghei* ANKA trophozoites and schizonts, merged peaks between duplicate H3K9ac experiments were compared with differentially expressed genes and the results are shown in **Figure 5.15.**



Differentially expressed genes

Gene ID	Name (location)	log_2FC
PBANKA_0108300.1	Palmitoyltransferase DHHC2	6.50
PBANKA_0828300.1	Dynein light chain, putative	5.14
PBANKA_1442100.1	Conserved *Plasmodium* protein, unknown function	2.36
PBANKA_0607400.1	Protein phosphatase PPM11, putative	1.85
PBANKA_1116700.1	AP-2 complex subunit alpha, putative	1.15

Gene ID	Name (location)	log_2FC
PBANKA_1021700.1	Conserved *Plasmodium* protein, unknown function	-1.43
PBANKA_0301500.1	Protein MAK16, putative	-1.67
PBANKA_0316941.1	PIR protein	-2.73

Figure 5.15: Comparison of H3K9ac merged peaks and differential gene expression in *P. berghei* ANKA schizonts. The Venn diagram on the left of this image depicts the overlap of H3K9ac merged peaks and differentially-expressed genes between schizont- and trophozoite- stage *P. berghei* ANKA parasites (with a false discovery rate (FDR) less than 0.005). 22 merged peaks across two independent H3K9ac ChIPmentation experiments were compared to 2161 differentially-expressed genes with FDR < 0.005. Of the 8 overlapping differentially-expressed genes, 5 were associated with up-regulation of RNA in schizonts and 3 were associated with down-regulation of RNA in schizonts. 14 H3K9ac merged peaks did not correlate to any change in RNA expression between trophozoites and schizonts. Black arrows indicate increased or decreased RNA expression. Tables to the right of the Venn diagram show the 5 upregulated genes and 3 down-regulated genes that are bound at any location by the H3K9 acetylation modification. Gene identifiers (Gene IDs), description of gene function, and log_2 fold change (log_2FC) are all listed for these genes.

In **Figure 5.15**, increased RNA expression in schizonts when compared to trophozoites was associated with H3K9 acetylation at only 5 genes, and down-regulation of transcription in only 3 genes. The most up-regulated gene (PBANKA_0108300.1) was associated with H3K9ac at the promoter-TSS region. This

gene encodes for palmitoyltransferase DHHC2, an AP2-G2-associated gene (Yuda *et al.*, 2015), expressed predominantly in female gametocytes (Yeoh *et al.*, 2017), that is essential for post-fertilisation development of *P. berghei* ANKA in the mosquito midgut and ookinete formation (Santos *et al.*, 2015).

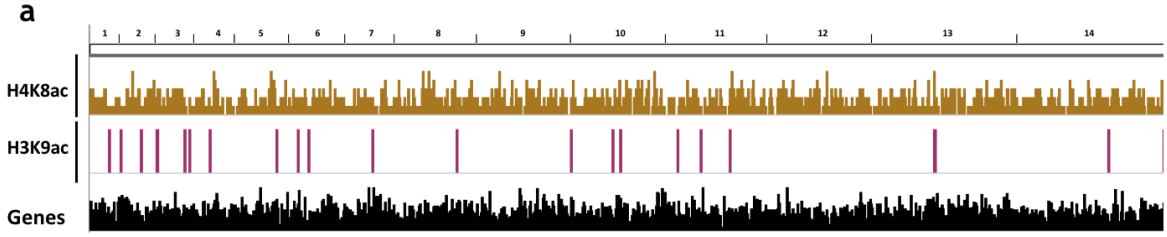
The second H3K9ac-associated, up-regulated gene, PBANKA_0828300, encoding for a putative dynein light chain protein, is once again linked to AP2-G2 expression (Yuda *et al.*, 2015) and transcribed predominantly in male gametocytes (Otto *et al.*, 2014; Yeoh *et al.*, 2017). Of the three remaining up-regulated, H3K9ac-associated genes, PBANKA_1442100.1 (a conserved *Plasmodium* protein of unknown function), PBANKA_0607400.1 (protein phosphatase PPM11), and PBANKA_1116700.1, none are associated with AP2-G2 expression (Yuda *et al.*, 2015), though each shows a negative correlation to transcription in the male gametocyte, and greatest expression in both intraerythrocytic parasite stages and female gametocytes in *P. berghei* ANKA (Otto *et al.*, 2014; Yeoh *et al.*, 2017). With these results, and those of H3K9 trimethylation shown in section **5.5,** could it be possible that changes to the histone H3K9 are a determining factor in AP2-G2 binding and the regulation of genes involved in parasite sexual development and transmission?

Of the H3K9ac-associated, down-regulated genes, one (PBANKA_0316941.1) is an uncharacterised PIR protein that is expressed predominantly during ring-stage asexual development (Otto *et al.*, 2014), one (PBANKA_1021700.1) is a conserved *Plasmodium* protein of unknown function that is again transcribed predominantly during ring-stage development in *P. berghei* ANKA (Otto *et al.*, 2014), and the only characterised gene, PBANKA_0301500.1, encodes for a putative MAK16 protein, an essential RNA-binding protein that is almost exclusively expressed in ring-stage *P. berghei* ANKA parasites (Otto *et al.*, 2014).

It should also be noted that all up-regulated and down-regulated genes shown in **Figure 5.15** are not the same as those differentially expressed and marked by H4K8 acetylation (**Figure 5.5)**. This suggests that, though both histone modifications are posited as having a role in transcriptional activation, these post-translational marks are involved in the regulation of different gene families in *P. berghei* (Gates *et al.*, 2017; Gupta *et al.*, 2017). To demonstrate the differential

global distribution of these two activating marks across the *P. berghei* ANKA genome, and to further analyse genes enriched for H3K9 acetylation, a comparison of H4K8ac and H3K9ac merged peak distribution and the top 10 peaks for H3K9ac merged peaks are shown in **Figure 5.16** below.

a



b

	Gene ID	Function	H3K9ac location	Highest Epic peak score (of 3)	Differential RNA expression at schizont stage	Log_2FC (if present)	P-value (if present)	FDR (if present)
1	PBANKA_0006800.1	PIR protein, pseudogene	Promoter-TSS	1562.21	**No**	-	-	-
2	PBANKA_1100051.1	PIR protein, pseudogene	Exon	1477.66	No	-	-	-
3	PBANKA_API00170.1	DNA-directed RNA polymerase subunit beta'', putative (rpoC2)	Promoter-TSS/exon/TTS	1164.33	No	-	-	-
4	PBANKA_0008901.1	PIR protein, pseudogene	TTS	408.93	No	-	-	-
5	PBANKA_0700021.1	PIR protein, pseudogene	TTS	294.30	No	-	-	-
6	PBANKA_1200600.1	*Plasmodium* exported protein, unknown function	Intergenic/exon	203.06	Yes	2.8	2.82×10^{-10}	2.33×10^{-9}
7	PBANKA_0944121.1	PIR protein, pseudogene	TTS	187.88	No	-	-	-
8	PBANKA_0519000.1	S-antigen, putative	Promoter-TSS/exon	157.12	Yes	7.99	2.84×10^{-24}	8.53×10^{-23}
9	PBANKA_0622921.1	18S ribosomal RNA	Promoter-TSS/exon	111.90	Yes	-0.93	0.24	0.32
10	PBANKA_0006300.1	PIR protein	TTS	111.67	No	-	-	-

Figure 5.16: Differential distribution of H4K8 and H3K9 acetylation marks and highest-scoring peaks for H3K9ac-enriched genes. Image **(a)** is a screenshot of merged H4K8ac peaks (1396 merged peaks) called using Epic software (in dark yellow) adjacent to H3K9ac merged peaks (in pink) called using Epic peak-calling software also. Tracks are labelled to the left and chromosomes are numbered 1-14 at the top of the image. Genes (in .gff) format are displayed in black underneath merged ChIP peak tracks. **(b)** is a table of the top 10 highest-scoring merged peaks identified for H3K9 acetylation across two independent biological replicates. All merged peaks are listed in **Supplemental Material ES3**. Gene IDs, known function [up-to-date as of 27-08-18], highest peak score across all peak-calling analyses, peak location, and corresponding RNA expression data are provided in **(b).** Further differential RNA expression data can be found in **Supplemental Material ES4.**

It is clear from **Figure 5.16 (a),** that many conserved peaks were present across duplicate biological samples when studying H4K8 acetylation in 2.5 million *P. berghei* schizonts, compared to H3K9 acetylation in which only 22 merged peaks were identified. A small number of peak locations do appear to be shared among

these activating histone modifications, i.e. at the distal end of chromosomes 3 and 9, and the centre of chromosome 13. Despite this, the table shown as **Figure 5.16 (b)** reveals that only one of the highest-scoring H3K9ac-enriched genes was shared with H4K8ac: PBANKA_1100051.1, a PIR pseudogene that is marked by all four histone modifications examined during the present study, and that is associated with intraerythrocytic, asexual parasite stages only (Yeoh *et al.*, 2017).

Of all of the highest scoring H3K9ac-enriched genes **(Figure 5.16 (a))**, 6 are PIR pseudogenes/protein of unknown function, 1 is an exported protein of unknown function that is significantly up-regulated in 22 h schizonts (PBANKA_1200600.1), 1 is an 18S ribosomal RNA (PBANKA_0622921.1), one is an apicoplast-located gene, a DNA-directed RNA polymerase subunit (rpoC2), and one is an S-antigen (PBANKA_0519000.1) that is heavily up-regulated in 22 h schizonts, and appears to be dispensable to asexual growth of the parasite [up-to-date as of 14-09-18] (www.PlasmoDB.org, 2018).

Though a number of H3K9ac-enriched genes are discussed in **Figure 5.15** and **Figure 5.16,** analysis of duplicate independent experiments using Epic software discovered 22 common peaks between samples. A complete list of these results are provided in the "H3K9ac_S5JP3mergedpeaks" section of **Supplemental Material ES3.**

5.7 H3K122 acetylation in *P. berghei* schizonts

The final histone modification examined in the present study was H3K122 acetylation, owing to the fact that an analysis of post-translational histone modifications (PTMs) using mass spectrometry showed acetylation of this mark in *P. berghei* gametocytes and methylation of H3K122 only in *P. berghei* schizonts (**Chapter 4**). The H3K122 acetylation mark was only recently discovered in *Plasmodium* by Coetzee *et al.* (Coetzee *et al.*, 2017), and as of the time of writing [27-08-18], no antibody to H3K122 methylation in any species is available commercially. Therefore, to examine the role of H3K122 acetylation in a *Plasmodium* parasite, ChIPmentation with an anti-H3K122ac antibody was attempted in schizonts first, to provide a background level of acetylation at H3K122 at asexual-stage to which future studies in gametocytes could be

compared. Initial alignment of H3K122ac peaks from different numbers of *P. berghei* schizonts are shown below in **Figure 5.17,** with repeat experiments of 2.5 million schizonts shown in **Figure 5.19.** This study represents the first time that this histone modification has been examined in any stage, in any malaria parasite.

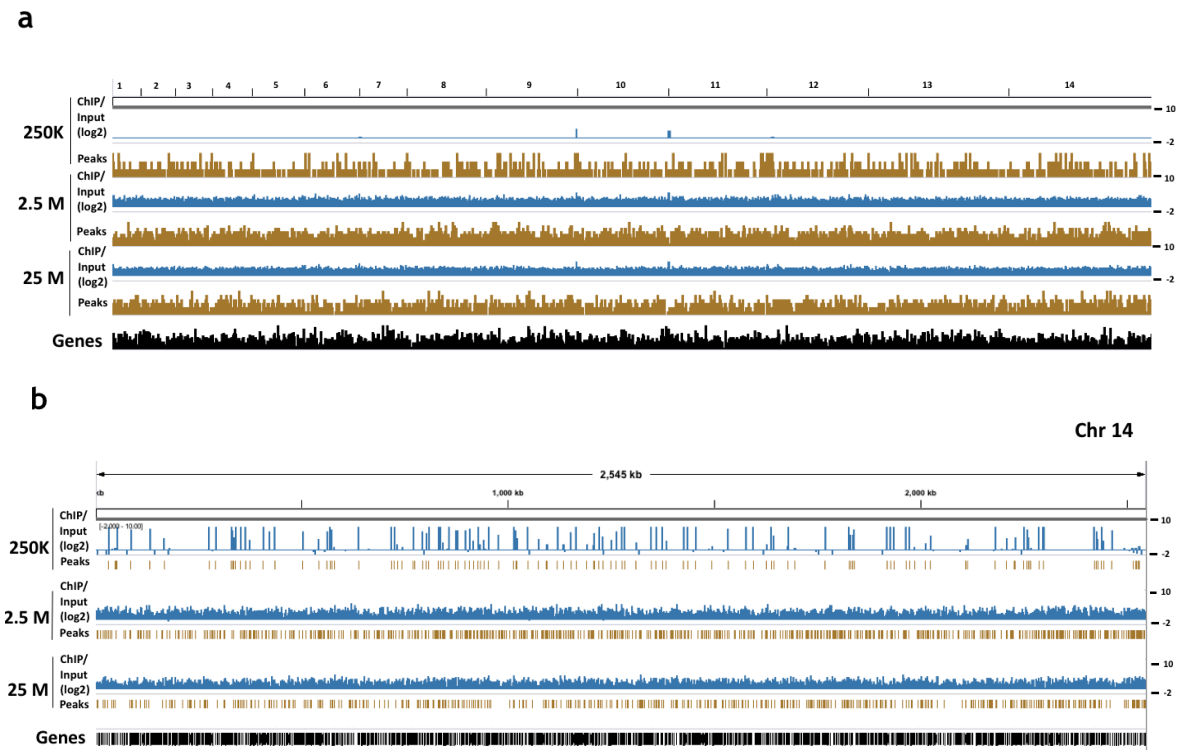


Figure 5.17: Distribution of H3K122 acetylation across the *P. berghei* ANKA genome at schizont stage. Using the IGV browser, H3K122ac-bound reads and peaks (for 250K, 2.5 million, and 25 million schizonts) are shown across the entire *P. berghei* ANKA genome (version 34) (in .bed format; results are normalised to the 250K schizont ChIP input sample) **(a).** At the base of these tracks, the general feature format (.gff) file shows the location of genes as black bars. For each sample, aligned ChIP data is displayed in blue. Epic peaks are shown in a separate track underneath ChIP data and are displayed in a dark yellow colour. In **(b)**, aligned reads and peak files are again shown as in **(a)**, though in this case (as for **Figure 5.3, Figure 5.6** and **Figure 5.13)**, only chromosome 14 is shown in closer detail (J.-J. Lopez-Rubio, Mancio-Silva and Scherf, 2009; Fraschka *et al.*, 2018). Colour-coding is carried out as in **(a).**

As in previous ChIPmentation experiments, similar distribution of H3K122ac-enriched loci were observed for 2.5 million and 25 million schizont cells (**Figure 5.17),** at both the genome level, and when focussing on chromosome 14 alone. Following the precedent set in previous studies of H4K8ac, H3K9me3, and H3K9ac modifications, 2.5 million *P. berghei* schizonts were examined twice more to determine if the ChIPmentation protocol and peak-calling analysis yielded reproducible results. MDS analysis was then carried out (**Figure 5.1)** and sample

'S8_H3K122Ac' was discarded in the final analysis. Genome distribution for H3K122ac-enriched regions in these experiments are shown in **Figure 5.19.**

In studies using yeast and human cell lines, the H3K122 acetylation mark has become associated with nucleosome disassembly and transcriptional activation (Tropberger *et al.*, 2013; Chatterjee *et al.*, 2015). A recent study in human embryonic stem cells also suggested that acetylation of H3K122 was not only present at active gene promoters, but that H3K122ac marks a subset of active enhancers distinct from those marked by H3K27ac (Pradeepa *et al.*, 2016). In the present study, as was the case with H4K8ac, H3K9me3, and H3K9ac, an examination of H3K122 acetylation was carried out using 250,000, 2.5 million, and 25 million *P. berghei* schizonts. All peaks identified using either MACS2, HOMER, or 'Epic' software are shown in **Figure 5.18.**

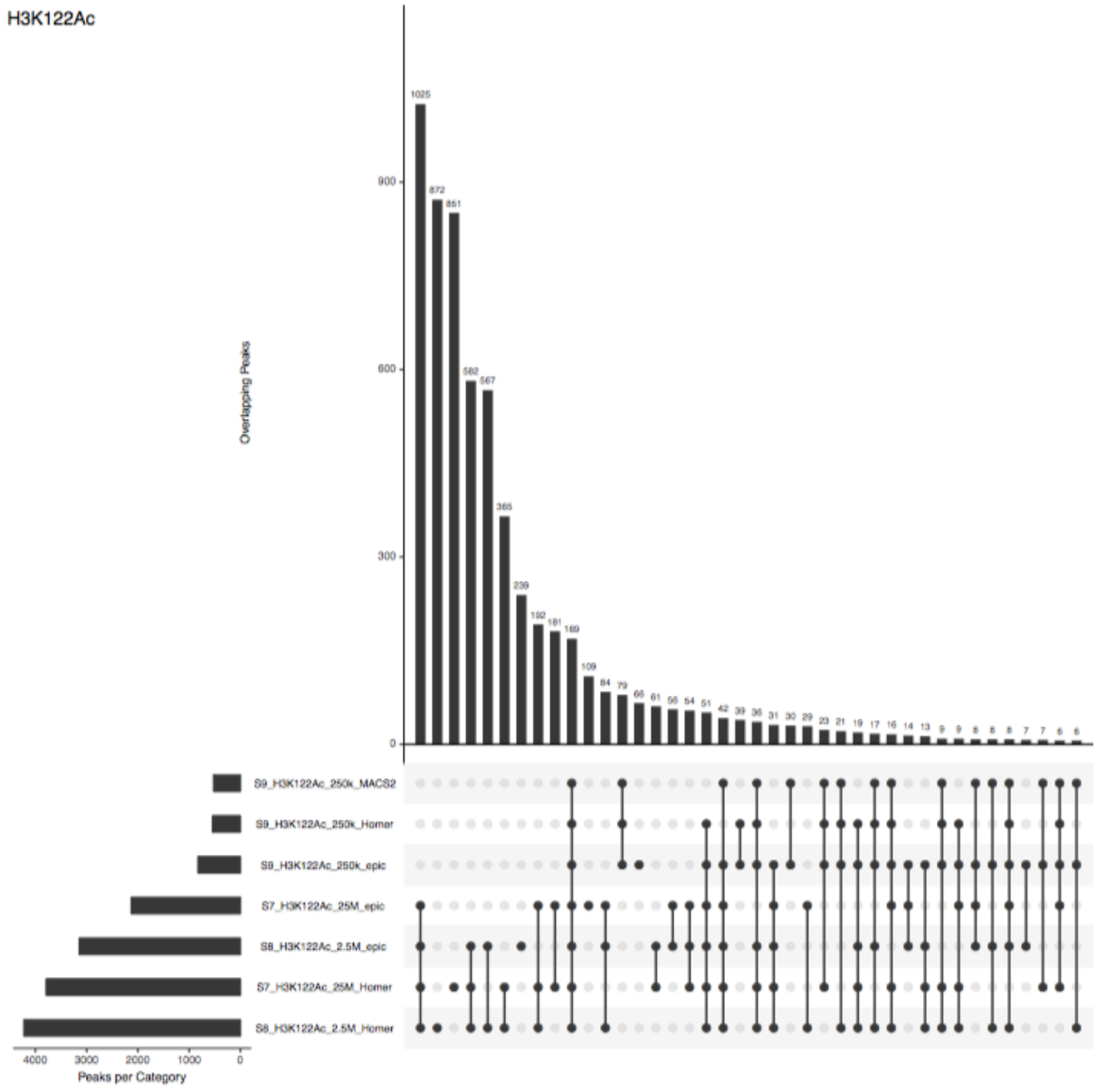
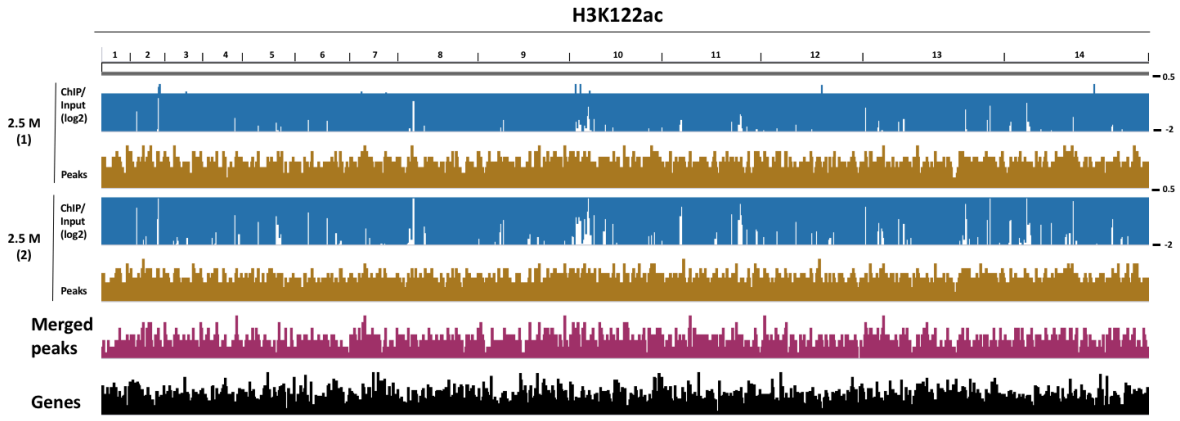


Figure 5.18: A comparison of shared and unique peaks from H3K122ac ChIP samples using MACS2, HOMER, and Epic software. The present UpSet plot shows the number of unique peaks called using MACS2, HOMER, and Epic software on the left-hand side. Each sample name is shown next to the number of unique peaks per category, with cell numbers and software name clearly labelled. The main histogram depicts the number of overlapping peaks in each H3K122ac sample (from 25 million, 2.5 million, and 250K *P. berghei* schizonts). At the base of the histogram, overlapping peaks shared between samples are indicated using circles, or circles joined with a straight line. Using the anti-H3K122ac antibody, MACS2 software (Zhang *et al.*, 2008) identified overlapping peaks only from a sample of 250K schizonts. Using Epic and HOMER, 1025 shared peaks were identified between samples of 2.5 million schizonts and 25 million schizonts.

As with analyses of all ChIPmentation data using three different antibodies (H4K8ac, H3K8me3, and H3K9ac), MACS2 software failed to call broad peaks from nearly all samples; in this case, peaks were only called from a sample of 250K schizonts, with no peaks called from samples of 2.5 million or 25 million *P. berghei* schizonts (**Figure 5.18**). With the Epic peak-caller showing overlapping peaks from all samples, and to align with previous analyses, Epic was once more chosen to examine H3K122ac-enriched areas between duplicate 2.5 million schizont samples (samples 'JP5_H3K122Ac' and 'JP10_H3K122Ac' shown in **Figure 5.1**). Unlike with previous findings, results from a ChIPmentation experiment using an anti-H3K122ac antibody in *Plasmodium* has never been published and so no comparisons can be made between the findings of these experiments to any studies carried out in a human malaria parasite or other *Plasmodium* species. Analysis of H3K122ac peak distribution, GO term associations, and location of peaks relative to the nearest gene are shown in **Figure 5.19.**

a



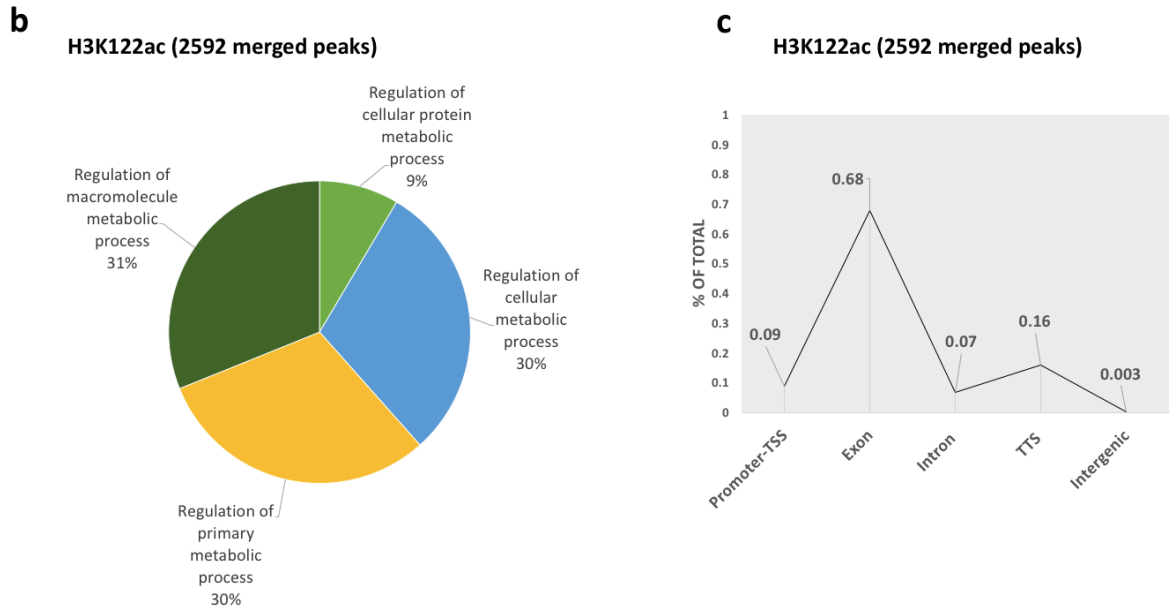
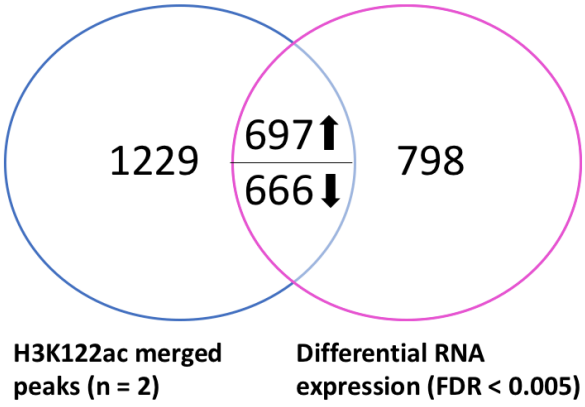


Figure 5.19: H3K122ac peak distribution across the *P. berghei* ANKA genome. In this series of images, **(a)** shows H3K122ac peak distribution across the entire genome of *P. berghei* using IGV. H3K122ac-bound reads (relative to an input sample from Fraschka *et al.*, 2018) from duplicate samples of 2.5 million schizonts are shown. Raw data are displayed in blue and Epic peaks are displayed in yellow underneath the ChIP track. Merged H3K122ac peaks are shown in pink. Genes are displayed in general feature format (.gff) as black bars. **(b)** is a pie-chart representation of gene ontology (GO) terms for biological processes that were associated with the genes to which H3K122ac was bound across both independent experiments. Image **(c)** shows the same data, but the locations of H3K122ac-enrichment relative to the nearest gene are depicted in graph format.

Results in **Figure 5.19 (a)** show an abundance of H3K122 acetylation across the *P. berghei* ANKA genome at schizont stage. Unlike previous histone modifications (particularly H3K9me3), no particular areas of H3K122ac-enrichment can be observed at subtelomeric chromosome regions or otherwise. When establishing enriched GO terms for biological processes (**Figure 5.19 (b),** the 2592 H3K122ac merged peaks covered an array of biological processes (with a $P < 0.005$), with H3K122 acetylation associated with genes involved in the regulation of macromolecule metabolic processes (31%), regulation of primary metabolic processes (30%), regulation of cellular metabolic processes (30%), and regulation of cellular protein metabolic processes (9%), all biological terms correlating to metabolism. These GO terms have previously been associated with transcribed genes during late trophozoite growth and schizogony (Bozdech *et al.*, 2003; Painter *et al.*, 2018), indicating that H3K122 acetylation may indeed be associated with active promoters and enhancers in eukaryotes (Tropberger *et al.*, 2013;

Pradeepa *et al.*, 2016). As seen in **Figure 5.19 (c)**), H3K122ac-enrichment often appeared at the gene exon, with 68% of merged peaks located at exons.

To examine a possible relationship between H3K122 acetylation and RNA expression, all merged H3K122ac peaks were compared to a list of differentially-expressed genes in schizonts compared to trophozoites, as was carried out for previous histone modifications (Otto *et al.*, 2014). The results of this comparison are shown in **Figure 5.20.**



GO term enrichment (upregulated genes) (P < 0.001)

GO ID	GO term	Fold enrichment
GO: 0044419	Interspecies interaction between organisms	2.45
GO: 0044403	Symbiosis, encompassing mutualism through parasitism	2.45
GO: 0016310	Phosphorylation	2.01

GO term enrichment (down-regulated genes) (P < 0.0001)

GO ID	GO term	Fold enrichment
GO: 1901566	Organonitrogen compound biosynthetic process	1.93
GO: 1901564	Organonitrogen compound metabolic process	1.84
GO:0006457	Protein folding	3.07
GO: 0044281	Small molecule metabolic process	1.98
GO:0044271	Cellular nitrogen compound biosynthetic process	1.64
GO: 0043604	Amide biosynthetic process	1.83

Figure 5.20: A comparison of H3K122 acetylation and RNA expression in *P. berghei* ANKA schizonts. The Venn diagram in this image depicts the overlap of H3K122ac merged peaks and differentially-expressed genes between schizont- and trophozoite- stage *P. berghei* ANKA parasites (with a false discovery rate (FDR) less than 0.005). 2592 merged peaks across two independent H3K122ac ChIPmentation experiments were compared to 2161 differentially-expressed genes (FDR < 0.005). Of 1363 overlapping differentially-expressed genes, 697 were associated with up-regulation of RNA in schizonts and 666 were associated with down-regulation of RNA in schizonts. 1229 H3K122ac merged peaks did not correlate to any change in RNA expression between trophozoites and schizonts. Black arrows indicate increased or decreased RNA expression. Tables to the right of the Venn diagram show the GO terms associated with up- and down- regulated genes that are bound at any location by the H3K122 acetylation modification. GO identifiers (GO IDs), description of GO term, and fold enrichment of these terms are all listed for these genes.

As demonstrated in **Figure 5.20,** an association of a gene with H3K122 acetylation is, as with previous histone marks assessed during the present study, not uniquely associated with increased or decreased RNA expression in *P. berghei* 22 h schizonts

when compared to 16 h trophozoites. In all cases, histone marks are associated with genes that are both up- and down- regulated. Of the GO terms shown in **Figure 5.20** that are associated with upregulation of RNA expression and H3K122 acetylation, two GO terms ($P < 0.001$) are shared with up-regulated H4K8ac-marked exons: GO: 0044419 (interspecies interaction between organisms, and GO: 0044403 (symbiosis, encompassing mutualism through parasitism) **(Figure 5.4 and Figure 5.20).** These results would suggest that both H4K8 acetylation and H3K122 acetylation at a gene exon is indicative of transcriptional activation, as was previously suggested for these particular histone modifications (Tropberger *et al.*, 2013; Pradeepa *et al.*, 2016; Gupta *et al.*, 2017). In addition, H3K122 acetylation is associated with upregulation at genes responsible for phosphorylation **(Figure 5.20).**

With regard to H3K122ac-associated genes that are also down-regulated in 22 h schizonts compared to 16 h trophozoites, only one GO term is shared among H4K8ac- and H3K122ac- associated genes, GO: 006457, protein folding **(Figure 5.4** and **Figure 5.20).** Unlike with H4K8ac-associated exons, down-regulated genes associated with H3K122 acetylation are (with $P < 0.0001$), largely associated with nitrogen-containing biosynthetic and metabolic processes **(Figure 5.20),** which ties in with previous results that showed that all H3K122ac-bound genes were enriched for those involved in metabolic processes **(Figure 5.20).** If H3K122 acetylation is associated transcription of enhancers at a separate subset of genes than H4K8ac, or even H3K27ac (Pradeepa *et al.*, 2016), the results shown in **Figure 5.19** and **Figure 5.20** would suggest that, in *P. berghei* ANKA, H3K122 acetylation is associated with the regulation of genes involved in parasite metabolism, whether they are transcriptionally activated, or repressed.

To examine whether H3K122 acetylation displays a similar pattern of binding to putative activating acetylation marks H4K8ac and H3K9ac, the top 10 highest-scoring H3K122ac-enriched genes (according to Epic software) were identified and RNA expression at these loci determined. These results are shown in **Table 5.3.**

	Gene ID	Function	H3K122ac location	Highest Epic peak score (of 3)	Differential RNA expression at schizont stage	Log_2FC (if present)	P-value (if present)	FDR (if present)
1	PBANKA_0006900.1	PIR protein pseudogene	Exon	1543.76	No	-	-	-
2	PBANKA_0600061.1	fam-b protein	Intergenic/exon	1501.50	Yes	-0.69	0.26	0.34
3	PBANKA_1100051.1	PIR protein pseudogene	Exon	1385.24	No	-	-	-
4	PBANKA_0214100.1	Conserved *Plasmodium* protein, unknown function	Exon	941.68	Yes	-2.29	0.0007	0.0018
5	PBANKA_1341000.1	Conserved *Plasmodium* protein, unknown function	Exon	671.04	Yes	3.24	3.58×10^{-15}	4.95×10^{-14}
6	PBANKA_1442200.1	Protein SOC2	Exon/TTS	668.14	Yes	0.56	0.13	0.19
7	PBANKA_0803700.1	Conserved *Plasmodium* protein, unknown function	Exon/intron	606.07	Yes	0.7	0.05	0.08
8	PBANKA_1003100.1	Pre-mRNA-processing-splicing factor 8 (PRPF8), putative	Exon	601.61	Yes	-0.67	0.07	0.10
9	PBANKA_0901000.1	fam-a protein	Exon	567.15	Yes	-0.95	0.007	0.01
10	PBANKA_0600071.1	Conserved rodent malaria protein, unknown function	Exon	564.72	Yes	1.39	0.20	0.28

Table 5.3: Highest-scoring H3K122ac peaks in *P. berghei* ANKA schizonts. This table displays the top 10 H3K122ac-bound genes according to their Epic peak-calling score. Genes are shared among duplicate independent H3K122ac ChIPmentation experiments. The highest score between all replicates is shown in the fifth column. The list of genes is numbered from highest to lowest score and the first and second columns of the table list gene IDs and current known function (if any) of the gene in *P. berghei* ANKA [as of 11-07-18]. The fourth column shows the location of the H3K122ac peak relative to the gene in question. The sixth column indicates differential gene expression in schizonts compared to 16 h trophozoites (Otto *et al.*, 2014). The final three columns show the log_2 fold change (log_2FC), and associated P-value and false discovery rate (FDR) (if applicable).

As was the case with all other antibodies used in the present ChIPmentation studies, the PBANKA_1100051.1 gene (encoding for a PIR protein pseudogene) was found among the top 10 highest Epic peak scores for H3K122ac enrichment **(Table 5.3).** Four more genes; another PIR protein pseudogene (PBANKA_0006900.1), the conserved rodent malaria protein of unknown function encoded by PBANKA_0600071.1, the conserved *Plasmodium* protein of unknown function, PBANKA_0214100.1, and the SOC2 protein, encoded for by PBANKA_1442200.1, also appeared to be marked by both a merged H3K122ac peak as well as a merged H4K8ac peak. However, these genes were not associated with H3K9 acetylation or trimethylation (**Table 5.2** and **Figure 5.16 (b)).** As previously mentioned, the SOC2 protein encoded for by PBANKA_1442200.1 is responsible for mitotic spindle formation during microgametogenesis and was upregulated in 22 h schizonts, but not to a significant degree (Otto *et al.*, 2014; Invergo *et al.*, 2017).

Of the remaining H3K122ac-enriched genes, two were members of multigene families (one fam-a protein and one fam-b protein), two were conserved proteins of unknown function (PBANKA_1341000.1 and PBANKA_0803700.1), and one was the pre-mRNA-processing-splicing factor 8 (PRPF8) (PBANKA_1003100.1). Of all 10 H3K122ac-associated genes shown, only the three conserved *Plasmodium* proteins of unknown function (encoded for by PBANKA_0214100.1 PBANKA_1341000.1, and PBANKA_0803700.1) were significantly up-regulated in 22 h schizonts when compared to 16 h trophozoites, and one gene, PBANKA_0901000.1 (a fam-a gene) was significantly down-regulated **(Figure 5.20)** (Otto *et al.*, 2014; Deng *et al.*, 2016; Fougère *et al.*, 2016).

Of the uniquely H3K122ac-enriched genes mentioned, both the Fam-a and Fam-b family genes have no associated phenotypic data as of yet (PhenoPlasm; up-to-date as of 11-07-18) (Sanderson and Rayner, 2017). As for conserved *Plasmodium* protein PBANKA_1341000.1 (a conserved *Plasmodium* protein of unknown function), this gene was refractory to disruption in *P. berghei* ANKA but amenable to *piggyBac* transposon insertion in *P. falciparum* (Zhang *et al.*, 2018). The second and third H3K122ac-enriched conserved *Plasmodium* proteins, encoded for by PBANKA_0803700.1 and PBANKA_0214100.1, were either refractory to disruption in *P. berghei* ANKA (PBANKA_0803700.1) or no data was available at the present time (*P. falciparum* status also unreported at present) (PhenoPlasm; up-to-date as of 11-07-18). The uniquely H3K122ac-enriched gene, encoding for putative PRPF8 (PBANKA_1003100.1), was refractory to disruption in both *P. berghei* ANKA and *P. falciparum* 3D7 (PhenoPlasm; up-to-date as of 11-07-18).

Taking a look at merged peaks for both H4K8 acetylation and H3K122 acetylation in the present study (found in **ES3)**; of the 1396 H4K8ac-enriched peaks and the 2592 H3K122ac-enriched peaks identified, 1013 were shared between these samples. Of the 22 H3K9ac merged peaks, 6 are shared with H3K122ac merged peaks. Taking all three acetylation marks together, only 2 merged peaks are shared between all three samples. These comparisons are displayed in **Figure 5.21.** As with all previous experiments, all peaks for independent ChIPmentation analyses are listed in **Electronic Supplemental Material ES3** alongside complete merged peak data.

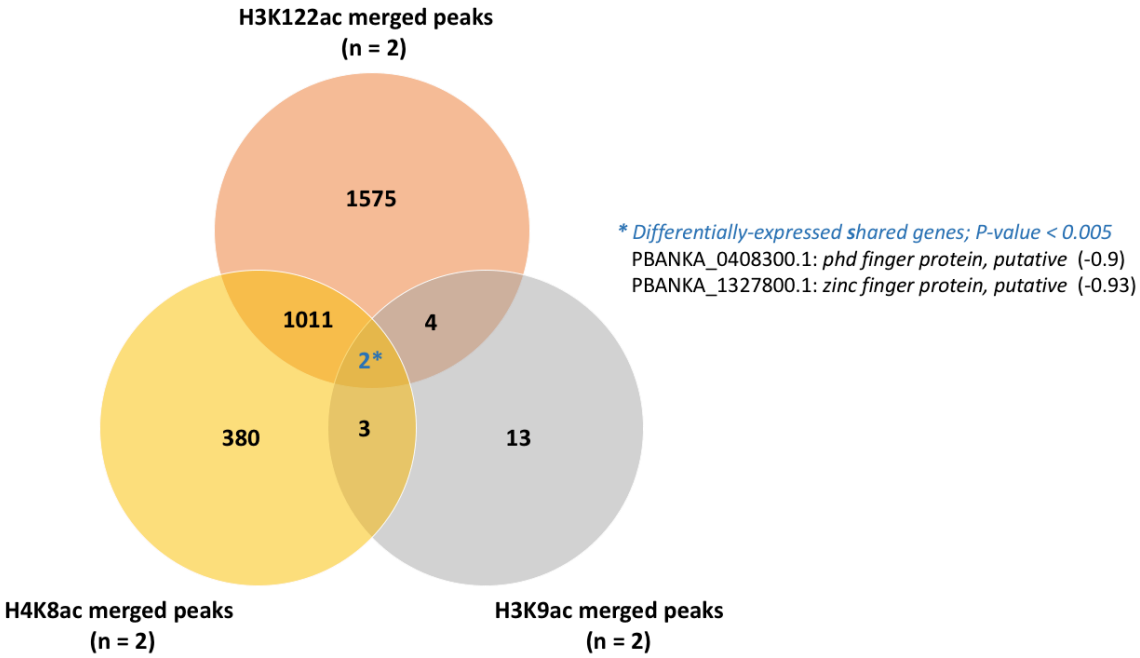


Figure 5.21: A comparison of peaks associated with H4K8ac, H3K9ac, and H3K122ac in *P. berghei* ANKA schizonts. The Venn diagram shown above depicts the overlap between H3K9ac merged peaks, H4K8ac merged peaks, and H3K122ac merged peaks (n = 2 in each case). To the right of the Venn diagram, all genes that were associated with merged peaks from all three antibodies and statistically significant differential RNA expression (P-value < 0.005) in 22 h schizonts (compared to 16 h trophozoites) are listed (Otto *et al.*, 2014). In each case, gene IDs, putative function according to *Plasmo*DB [up-to-date as of 28-08-18], and $\log_2$ fold change (in brackets) are shown.

Despite all three acetylation marks (H4K8ac, H3K9ac, and H3K122ac) shown in **Figure 5.21** being previously associated with active transcription in *Plasmodium* spp. and other eukaryotes ((Lopez-Rubio, Mancio-Silva and Scherf, 2009; Salcedo-Amaya *et al.*, 2009; Tropberger *et al.*, 2013; Karmodiya *et al.*, 2015; Pradeepa *et al.*, 2016; Gupta *et al.*, 2017), a comparison of all genes and differential RNA expression associated with all three marks at schizont stage in *P. berghei* ANKA shows that in fact both shared genes were down-regulated in schizonts compared to trophozoites to a statistically significant degree (Otto *et al.*, 2014) **(Supplemental Material ES4)**. Notably, both genes contained structural DNA/RNA-binding PHD or zinc finger motifs **(Figure 5.21),** though only the former, PBANKA_0408300.1, has been characterised to any extent in *P. berghei*, being dispensible for asexual growth (though a slow growth phenotype was observed), and being expressed predominantly in male gametocytes and ring-stage asexual

parasites (Otto *et al.*, 2014; Yeoh *et al.*, 2017). The zinc finger protein encoded by PBANKA_1327800.1, is expressed predominantly during the 4 h ring stage of asexual parasite growth (Otto *et al.*, 2014; Yeoh *et al.*, 2017).

In addition to identifying shared gene associations between three different histone modifications in *P. berghei*, **Figure 5.21** also demonstrates that each antibody in question (H4K8ac, H3K9ac, and H3K122ac) was enriched at different genes across two independent biological replicates, indicating that in each case, antibodies chosen for ChIPmentation in *P. berghei* were most likely not attributing randomly to GC-enriched areas or particular genome sequences. Merged peak data for all experiments are present in **Supplemental Material ES3**, with all differential RNA expression data provided in **ES4.**

5.8 Discussion

The key findings from this chapter were:

- Considerable optimisation of ChIPmentation in *P. berghei* ANKA limited this chapter to a study of mature asexual parasites (schizonts) only.
- A ChIP procedure of the AT-rich *P. berghei* ANKA genome requires a decrease of de-crosslinking temperature from 65ºC to 45ºC.
- The minimum cell number for a ChIPmentation procedure of *P. berghei* schizonts was 2.5 M cells.
- Peak-calling software packages, 'Epic' and 'HOMER' consistently detected a greater number of shared and unique histone modification peaks when compared with MACS2.
- 1396 shared peaks were identified between two H4K8ac samples in 2.5 M *P. berghei* ANKA schizonts; 62% of which were identified at gene exons, with 389 peaks associated with differential RNA expression between trophozoite and schizont stages of parasite development.
- 55 shared peaks were identified between two H3K9me3 samples from 2.5 M *P. berghei* ANKA schizonts; 50% of which were associated with gene exons, and with only 46 peaks associated with differential gene expression between trophozoite and schizont stages of parasite development.

- 22 shared peaks were identified between two H3K9ac samples from 2.5 M *P. berghei* ANKA schizonts; 50% of which were associated with gene promoter-TSS regions, and with only 8 peaks associated with differential gene expression between trophozoite and schizont stages of parasite development.
- 2592 shared peaks were identified between two H3K122ac samples from 2.5 M *P. berghei* ANKA schizonts; 68% of which were associated with gene exons, and with 1363 peaks associated with differential gene expression between trophozoite and schizont stages of parasite development.
- Two genes with upregulated transcription at schizont stage were marked by all three potentially activating histone modifications: zinc finger domain-containing proteins, PBANKA_0408300.1 and PBANKA_1327800.1.

5.8.1 The AT-rich *Plasmodium* spp. genomes: issues in ChIP sequencing

During the present study, ChIP-seq of *P. berghei* ANKA schizonts using standard next-generation sequencing library preparation was attempted using 10 million and 5 million cells on multiple occasions. Despite completion of the protocol and attempts using two separate library preparation kits, no amplified libraries contained enough DNA for a successful sequencing run (results not shown). Upon much discussion and analysis of published methods, it became clear that the de-crosslinking temperature and library preparation kit temperatures (65ºC for human and other eukaryotic genomes) were proving detrimental to the ChIP DNA obtained from *P. berghei* samples.

With the publication of Fraschka *et al.* (Fraschka *et al.*, 2018), the first publication to acknowledge a decrease in de-crosslinking temperature to 45ºC during a ChIP protocol, attempts to de-crosslink *P. berghei* ChIP samples at 65ºC were halted. In *Plasmodium* spp., it is necessary to decrease the chromatin-DNA de-crosslinking overnight temperature to 45ºC (Fraschka *et al.*, 2018), as opposed to previously published protocols which suggested 65ºC for *Plasmodium spp.* ChIP-seq (Lopez-Rubio, Siegel and Scherf, 2012; Gómez-Díaz *et al.*, 2014). In some cases, the de-crosslinking temperature was simply brushed over when describing experimental

methods in *Plasmodium spp.* ChIP-seq (Kaneko *et al.*, 2015; Karmodiya *et al.*, 2015).

In a similar vein, when undertaking ChIP-seq library preparation using the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (Catalogue No. #E7645S; New England BioLabs Inc.), the 30 min 65ºC incubation between end-repair, dA-tailing, and adapter ligation should be substituted with clean-up using SPRIselect beads (Beckman Coulter, Inc.). The initial denaturation step in PCR amplification of libraries can also be lowered from 98ºC to 94ºC to minimise damage to AT-rich DNA. In addition to ChIP and DNA sequencing work on *Plasmodium spp.*, these temperature changes can also be adopted in experiments in which subtelomeric or telomeric DNA regions (particularly AT-rich regions) are under investigation (consultation with Dr Kathryn Crouch and Craig Lapsley at the University of Glasgow, and technicians from New England BioLabs, Inc.).

5.8.2 Advantages and disadvantages of ChIPmentation

As depicted in **Figure 2.8,** the ChIPmentation approach combines a ChIP procedure with library preparation using a hyperactive Tn5 transposase that enables simultaneous ChIP DNA fragmentation and adapter tagging (Adey *et al.*, 2010; Schmidl *et al.*, 2015). In this procedure, adapter tagging is undertaken while the ChIP DNA is still bound to the attached chromatin-antibody complex. Because de-crosslinking and library amplification can be undertaken directly after this step, ChIPmentation bypasses the multiple purification steps required in a standard ChIP-seq protocol. Therefore, using ChIPmentation, sequencing of ChIP DNA from smaller cell numbers is possible because DNA loss is minimised (Schmidl *et al.*, 2015). This procedure was then particularly suitable for studies of *P. berghei* ANKA parasites which, unlike *P. falciparum*, require the use of experimental animals. ChIPmentation may also be suitable for ChIP studies of mosquito-stage parasites such as ookinetes, oocysts, and sporozoites; parasite stages that are present in relatively low numbers.

Although ChIPmentation in a *Plasmodium spp.* parasite has not yet been published, a protocol for an Assay of Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq) in *Plasmodium* parasites, which takes advantage of the Tn5

transposase from the same Nextera DNA Library Preparation Kit (Illumina, Inc.) has been used to determine genome accessibility throughout asexual blood stages (Toenhake *et al.*, 2018) and sporozoite-stage parasites in *P. falciparum* (Ruiz *et al.*, 2018). As such, it is only a matter of time before this Tn5 transposase-based approach is applied to ChIP studies in the *Plasmodium* parasite and beyond.

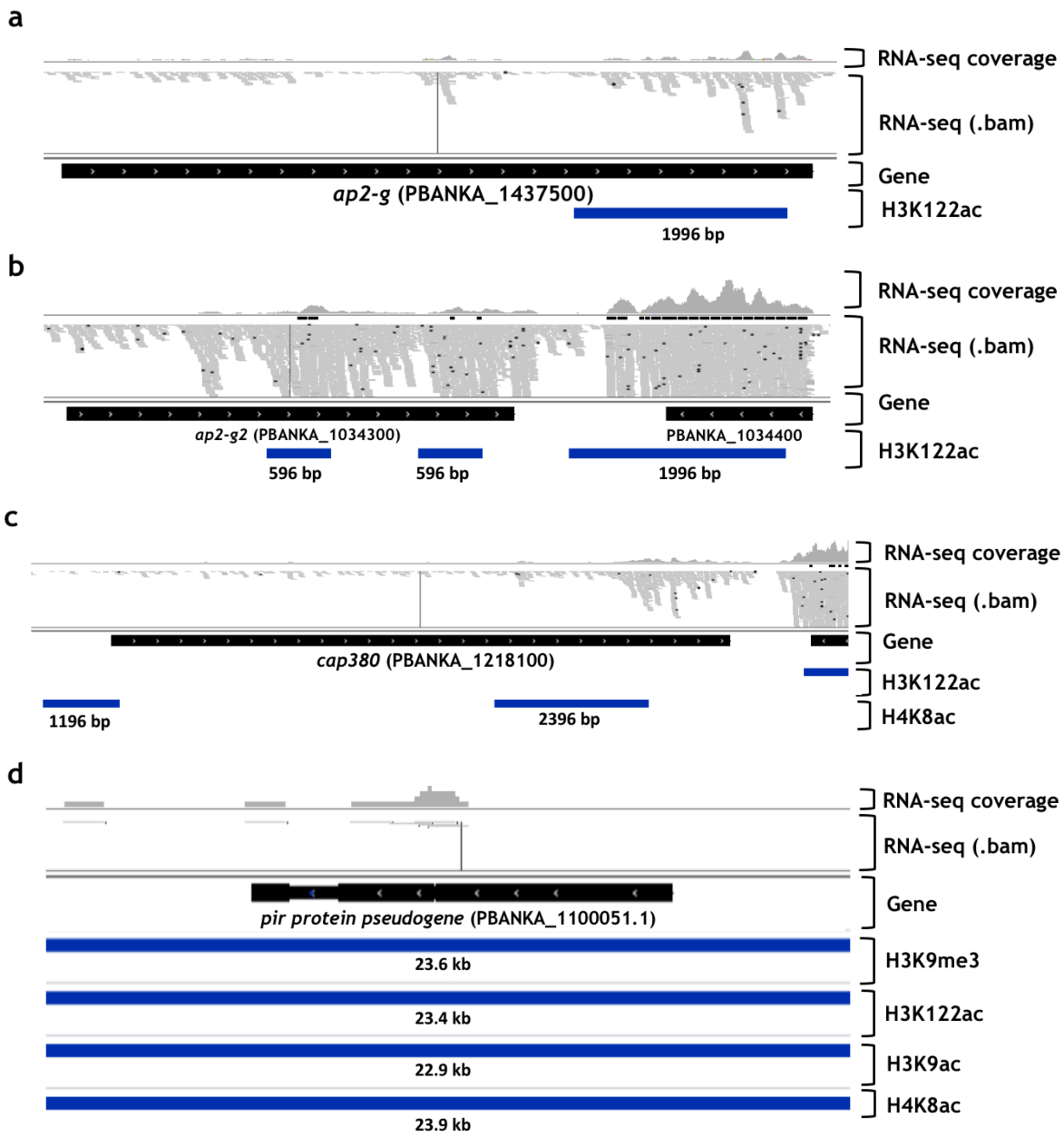
Despite the apparent advantages of ChIPmentation over standard ChIP-seq, the use of a single Tn5 transposase for DNA fragmentation and adapter ligation raises the possibility that a greater number of sequence-dependent biases may be introduced relative to a standard library protocol (Adey *et al.*, 2010). However, extensive investigation of Tn5-transposase-mediated fragmentation and adapter ligation across several organisms, and using different sequencing platforms, revealed only a slightly greater adapter insertion bias in Tn5-transposase-catalysed reactions that had little impact at the level of genomic coverage (Adey *et al.*, 2010; Picelli *et al.*, 2014). The reduction in ChIP DNA loss that usually resulted from the multiple PCR amplification and purification steps in standard ChIP-seq also offset any biases during the ChIPmentation procedure, as PCR-free transposase-based library construction was possible with this method (Adey *et al.*, 2010).

5.8.3 Epigenetic regulation of stage-specific genes from *P. berghei* mature schizonts

With the focus of the present study being commitment to gametocytogenesis, regions around the AP2-G (PBANKA_1437500) and AP2-G2 genes (PBANKA_1034300) were examined closely (Sinha *et al.*, 2014; Yuda *et al.*, 2015). In **Figure 5.22,** the complete repertoire of epigenetic modifications identified in the present study are shown aligned to these two transcriptional regulators, along with other stage-specific developmental genes. These include the oocyst capsule protein Cap380 (PBANKA_1218100) (determined to be HP1/H3K9me3-enriched) (**Figure 5.10)** (Fraschka *et al.*, 2018); the PBANKA_1100051.1 PIR protein pseudogene, enriched for multiple epigenetic marks in the present study,; the *P. berghei* apical membrane antigen 1 (AMA1) gene (PBANKA_0915000), a protein implicated in the parasites invasion of erythrocytes and hepatocytes in *P. falciparum* (Yang *et al.*, 2017); the ring-stage-expressed *P. berghei* ATP-

dependent rRNA helicase SPB4 (PBANKA_1024100); and the DNA repair endonuclease XPF (PBANKA_1144700; also referred to as ERCC4) which is predominantly expressed in 22 h schizonts (Otto, Rayner, *et al.*, 2014).

To complement previous analyses in *P. falciparum*, the array of epigenetic modifications associated with conserved *Plasmodium* protein PBANKA_0902300 (Lopez-Rubio, Mancio-Silva and Scherf, 2009), conserved *Plasmodium* protein UIS27 (PBANKA_1435900) (Fraschka *et al.*, 2018), and the HP1 gene itself (PBANKA_1436100) (Brancucci *et al.*, 2014; Fraschka *et al.*, 2018) are also shown in **Figure 5.22.** In each case, RNA expression of the gene at 22 h post-invasion (schizont stage) compared to 16 h trophozoites is shown (Otto *et al.*, 2014) .



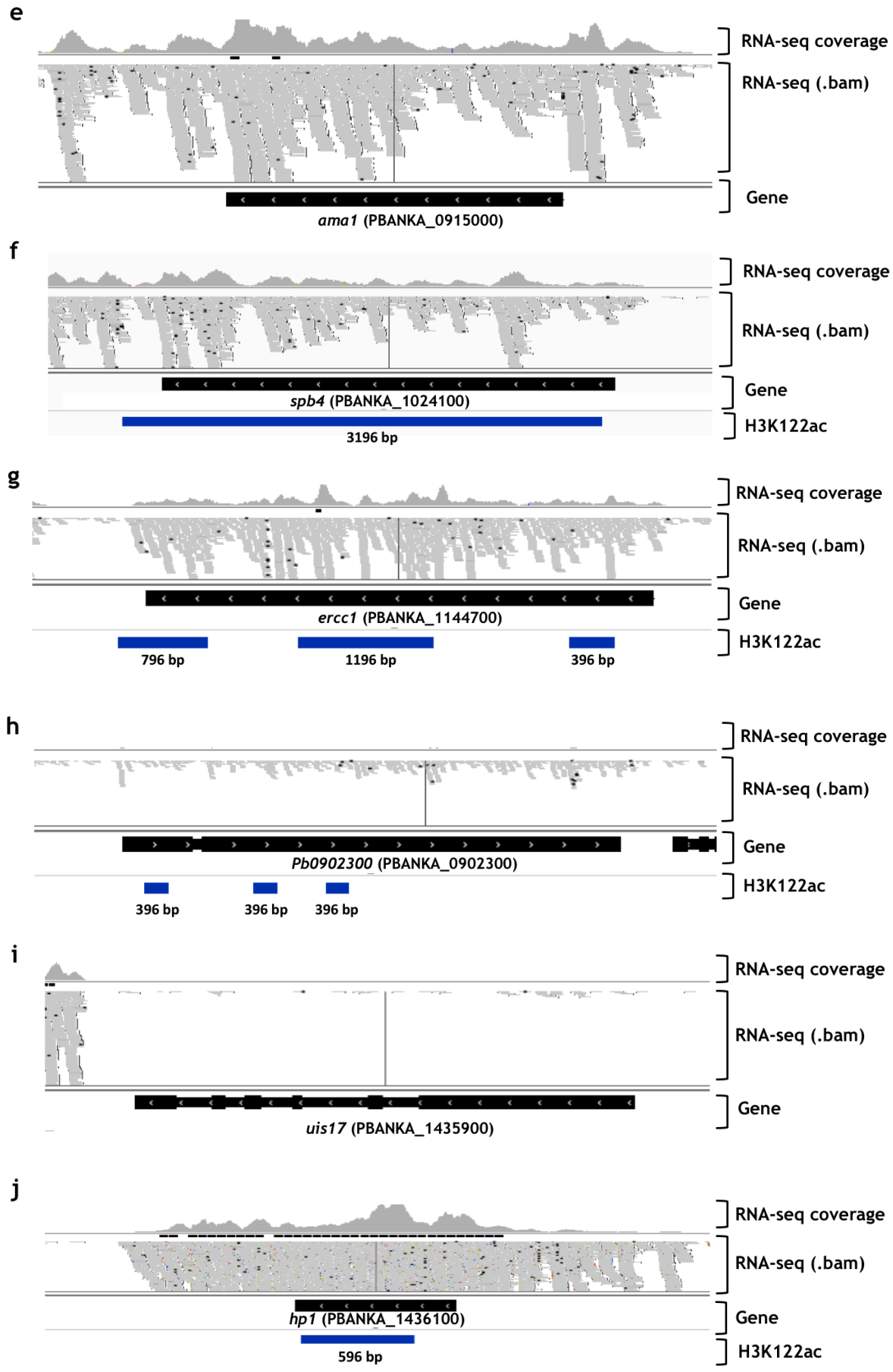

e
RNA-seq coverage
RNA-seq (.bam)
Gene
ama1 (PBANKA_0915000)
f
RNA-seq coverage
RNA-seq (.bam)
Gene
spb4 (PBANKA_1024100)
H3K122ac
3196 bp
g
RNA-seq coverage
RNA-seq (.bam)
Gene
ercc1 (PBANKA_1144700)
H3K122ac
796 bp
1196 bp
396 bp
h
RNA-seq coverage
RNA-seq (.bam)
Gene
Pb0902300 (PBANKA_0902300)
H3K122ac
396 bp
396 bp
396 bp
i
RNA-seq coverage
RNA-seq (.bam)
Gene
uis17 (PBANKA_1435900)
j
RNA-seq coverage
RNA-seq (.bam)
Gene
hp1 (PBANKA_1436100)
H3K122ac
596 bp

Figure 5.22: An overview of four epigenetic modifications at stage-specific genes. Image **(a)** to **(j)** above are screenshots from the IGV genome browser of RNA-seq BAM file and coverage data mapped to *P. berghei* ANKA genes in GFF format. Any peaks present from duplicate ChIPmentation experiments of H3K9me3, H3K9ac, H4K8ac, and H3K122ac, are shown beneath the GFF track in IGV. All tracks are clearly labelled and the size of peaks called using Epic software (in dark blue) in base-pairs (bp) is indicated under the peak (if present). All merged peaks for all three 2.5 M schizont samples, and merged peaks obtained from the two samples based upon MDS clustering are listed in **Supplemental Material ES3**. Further differential RNA expression data can be found in **Supplemental Material ES4.**

As depicted in **Figure 5.22 (a),** the *ap2-g* locus in *P. berghei* ANKA high-producer (HP) mature schizonts is marked at the TTS by H3K122 acetylation, and was down-regulated in 22 h schizonts when compared to 16 h trophozoites (Otto *et al.*, 2014). If the current theory regarding the function of the H3K122ac mark is to be believed, the *ap2-g* locus would be poised for transcriptional activation in this state (**Figure 5.22 (a)**). At this point in the asexual cycle, the gene encoding for the AP2-G2 transcriptional repressor (**Figure 5.22 (b))** is also marked by H3K122 acetylation, with greatest overlap of peaks at the exon and TTS.

In **Figure 5.22 (c),** the cap380 oocyst capsule protein locus is depicted, showing that in mature schizonts, this gene is marked at the promoter-TSS region and the exon by the H4K8ac histone modification. Down-regulation of RNA expression in this case indicates that this gene (associated with the transmission stage of *P. berghei* ANKA) is subject to decreased transcriptional activation (Srinivasan, Fujioka and Jacobs-Lorena, 2008). From analysis of the 10 genes shown in **Figure 5.22,** it becomes clear that H3K122 acetylation alone (**Figure 5.22 (a), (b), (f), (g), (h)** and **(j))** or in tandem with H4K8 acetylation acetylation (**Figure 5.22 (c),** is not indicative of increased RNA expression in *P. berghei* schizonts (Lopez-Rubio, Mancio-Silva and Scherf, 2009; Salcedo-Amaya *et al.*, 2009; Tropberger *et al.*, 2013; Karmodiya *et al.*, 2015; Pradeepa *et al.*, 2016).

With regard to the PIR pseudogene associated with multiple histone modifications (**Figure 5.22 (d)),** it is unclear whether the repetitive nature of the PIR family gene sequences was attractive to immunoglobulin-binding, or whether these genes are in fact heavily associated with epigenetic marks (Yam *et al.*, 2016). Both hypotheses are possible. What is clearest from **Chapter 5** in its entirety is

that single epigenetic modifications alone, without consideration for their location, lncRNAs or histone crosstalk, cannot be used alone as indicators of transcriptional activation or repression.

5.8.4 Issues with peak calling software

In addition to the considerable difficulty in optimisation of a ChIP protocol for the detection of chromatin-bound regions in the AT-rich *Plasmodium* genome, further issues arose in the present study with attempts to optimise bioinformatic analysis of the data produced upon sequencing of histone-bound DNA fragments. It is well-documented that software tools for differential ChIP-seq analysis can considerably influence the outcome of ChIP-sequencing results (Steinhauser *et al.*, 2016; Tu and Shao, 2017). In the present study, three peak-callers were compared: MACS2 (Zhang *et al.*, 2008), HOMER (Heinz *et al.*, 2010), and Epic (based on SICER) (Xu *et al.*, 2014).

With MACS2 widely used to perform peak-calling of transcription factor-bound regions that often have sharp or isolated peaks (Tu and Shao, 2017), HOMER and Epic were tested and compared as described in section **5.3.4.** HOMER and Epic tools are considered useful tools when working with broad peak ChIP data using default settings and when no pre-defined peak parameters have previously been determined (as was the case with ChIPmentation in *P. berghei*) (Steinhauser *et al.*, 2016). Despite this, a number of issues arose when using the HOMER annotation script (http://homer.ucsd.edu/homer/ngs/annotation.html) to define regions of the *P. berghei* ANKA genome. Though HOMER annotates regions based upon information from *Plasmo*DB (version 34 of *P. berghei* ANKA in this case), prediction of promoters within the AT-rich *Plasmodium* genome is a difficult task, requiring specialised software, and with the *Plasmodium* genome often encoding for bi-directional promoters, large upstream promoters, and the ability to form antisense RNAs (Brick *et al.*, 2008). With this in mind, HOMER prediction of peaks located at 'intergenic' or 'intron' regions may in fact encompass regions of *Plasmodium* gene promoters, though relatively few histone marks were determined to be located at these regions in the present study (sections **5.4** to **5.8**).

In addition to *Plasmodium*-specific genetic parameters that cannot be determined by HOMER annotation, a general issue arises in the method by which the 'mergePeaks' option determines overlapping peaks between two or more peak files (for examples, two H4K8ac peak samples). When using the 'mergePeaks' command, HOMER requires a 'primary peak file' to which a second file is compared (http://homer.ucsd.edu/homer/ngs/mergePeaks.html). When choosing a primary peak file, HOMER software then determines 'overlapping' peaks as those with peak centres at an equal or lesser difference than 100 bp by default. This method of comparison may result in the loss of 'overlapping peaks' because the 'primary peak file' has fewer defined peaks than the secondary peak file. This iterative process of discovering overlapping peaks will also effect downstream comparisons when a 'mergedPeaks' file is again used as a primary file to which further peak files are compared. To determine that no peaks are missed when examining a series of samples, each file must be compared to one another in a pairwise fashion, with the resulting 'overlapping' files compared to one another or visualised on a genome browser. Visualisation of normalised BAM files of ChIP data or visual comparison of peak files may also aid in choosing primary peak files.

5.8.5 Future experiments

With optimisation of a 'wet lab' ChIPmentation procedure for *P. berghei* carried out in the present study, experiments to determine the changes to these epigenetic modifications in different life-cycle stages of the *P. berghei* ANKA parasite will be carried out (such as an analysis of changes after gametocytogenesis). The experiments shown above also represent the first time a comparative analysis of multiple peak-calling software packages has been undertaken in relation to *Plasmodium* ChIP experiments. In previous publications, only MACS2 software has been used to call peaks, though the use of this software may limit the scope of epigenetic findings (broader peaks may not be discovered, though they may be present) (Salcedo-Amaya *et al.*, 2009; Karmodiya *et al.*, 2015; Filarsky *et al.*, 2018; Fraschka *et al.*, 2018). Further discussion of ChIPmentation findings in the present study and the implications of these findings to *Plasmodium* biology can be found in **Chapter 6.**

6. Discussion

6.1 Epigenetics and control of *Plasmodium* gametocytogenesis

It has been shown that gametocytogenesis in *Plasmodium spp.* is controlled by the AP2-G and AP2-G2 transcriptional regulators (Kafsack *et al.*, 2014; Sinha *et al.*, 2014; Yuda *et al.*, 2015), with AP2-G under the control of epigenetic factors HP1 and HDAC2/IPK1 (Brancucci *et al.*, 2014; Coleman *et al.*, 2014). In recent studies, the role of gametocyte development protein 1 (GDV1) as an antagonist of HP1-binding to *pfap2-g* has been shown, as well as the effect of lysophosphatidylcholine (LysoPC) levels on the initiation of gametocytogenesis in *P. falciparum* (Brancucci *et al.*, 2017; Filarsky *et al.*, 2018). Interestingly, GDV1 and the effect of LysoPC on gametocytogenesis induction are both lost in *P. berghei* ANKA and other rodent malaria parasites (Brancucci *et al.*, 2017; Filarsky *et al.*, 2018). Loss of both the GDV1 protein and the effect of LysoPC together perhaps suggests that GDV1 activation is triggered by the absence of LysoPC, or that, in abundance, LysoPC is detrimental to GDV1 antisense RNA production in human malaria parasites. This remains to be investigated.

However, upon completion of the present thesis, and during its correction, a further development occurred in the field of *Plasmodium* spp. sexual development and transmission that is important to elaborate upon. The initial discovery of AP2-G and its role in gametocytogenesis (Kafsack *et al.*, 2014; Sinha *et al.*, 2014) was followed by development of a method to irreversibly express AP2-G, therefore irreversibly inducing sexual conversion of all parasites in a *P. berghei* experiment (Kent *et al.*, 2018). This same method of rapamycin induction to induce sexual development was used in the present study to ensure populations of either 100% asexual parasites or 100% gametocytes prior to histone extraction (section **1.7.2.3.1**). However, in a recent study by Bancells *et al.* (Bancells *et al.*, 2019), re-analysis of single-cell *P. falciparum* data determined that there are two distinct routes by which the *Plasmodium* parasite underwent sexual development. In some parasites, early *ap2-g* expression at ring-stage resulted in the growth of a sexual-stage gametocyte within the same developmental cycle. In a second group of cells, *ap2-g* expression at a later stage (trophozoite stage) resulted in sexually-committed parasites that remained morphologically asexual until after schizont development, merozoite egress, and re-invasion into a new erythrocyte. Only

upon re-invasion did the following developmental cycle produce gametocytes (Bancells *et al.*, 2019) (**Figure 6.1**).

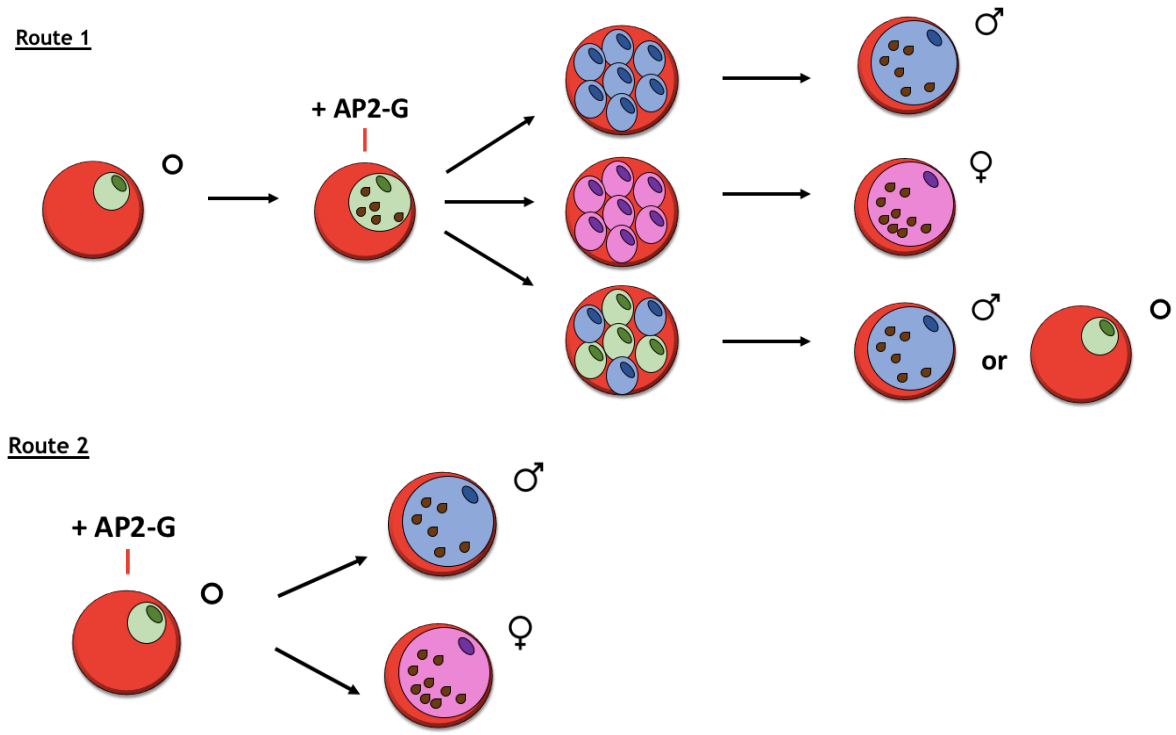


Figure 6.1: Two routes of gametocytogenesis in a *P. berghei* model. Using *P. berghei* as a model in this diagram, the two routes by which *Plasmodium* spp. parasites are postulated to undergo gametocytogenesis are shown. In 'Route 1', the asexual ring-stage parasite (asexual denoted by 'o') undergoes normal asexual growth until the induction of AP2-G by an as yet unknown mechanism. At this stage of asexual development, the parasite becomes committed to sexual development and expressed AP2-G, though the trophozoite stage parasite develops into a schizont containing either sexually-committed merozoites or a mix of asexual and sexual stage merozoites. Upon re-invasion into a new erythrocyte, these parasites develop in the second cycle into male or female gametocytes, or, in the case of the mixed schizont, further asexual parasites are produced. In 'Route 2' of gametocytogenesis, the expression of AP2-G during early ring-stage development results in conversion of the asexual parasite to a gametocyte within the same developmental cycle, and in the same erythrocyte. The mechanism for sex determination in these cases is unknown as of yet (Bancells *et al.*, 2019).

As shown in **Figure 6.1,** a further finding from the Bancells *et al.* study was that *Plasmodium* schizonts were not limited to asexual or sexually-committed merozoites as was previously thought (Smith *et al.*, 2000), but could contain both AP2-G-expressing and AP2-G-negative merozoites. In this study however, the sex-specificity of these mixed schizont stages was not explored, i.e. whether each route towards sexual development produced male and female gametocytes, or were sex-specific in nature (Bancells *et al.*, 2019). In fact, the mixed schizont depicted in **Figure 6.1** could also contain a mixture of asexual merozoites and

those committed to producing female gametocytes. All three possibilities (asexual, male sexually-committed, and female sexually-committed merozoites) may also be found within a single schizont cell, though this remains to be determined (Bancells *et al.*, 2019).

These recent findings also raise implications for the present study. In **Chapter 4,** the rapamycin-inducible *P. berghei* ANKA line was used to ensure a complete conversion of asexual parasites to gametocytes, with no induction producing 100% asexual parasites (**Figure 4.3**). The resulting asexual or sexually-committed nucleosomes demonstrated distinct histone modification profiles as depicted in **Figure 4.8**. However, our method of *ap2-g* induction using rapamycin (detailed in section **2.1.9)** could not distinguish whether the isolated gametocytes were produced as a result of 'Route 1' or 'Route 2' gametocytogenesis, as depicted in **Figure 6.1.** The Bancells *et al.* study also suggests that *ap2-g* transcription by these developmental pathways resulted from epigenetic alterations at the *ap2-g* or *gdv1* loci. To identify histone PTMs that may contribute to these pathways, histone extraction of *P. berghei* ANKA PTMs upon rapamycin induction in timed ring-stage parasites versus trophozoite-stage parasites could be undertaken. In addition, using rapamycin induction of *ap2-g in vivo,* followed by immediate *in vitro* culture may determine whether both mature asexual and mature sexual parasites develop. The availability of the '820' line in *P. berghei* ANKA would also be useful in that the sex of gametocytes produced by induction at either ring-stage or trophozoite-stage could be determined.

With the discovery of two routes of gametocytogenesis also comes the possibility that enzymes that may result in either asexual or sexual-stage histone PTMs, such as those suggested in **Table 4.6,** may play a greater role than previously thought in the induction of sexual development. This is because *ap2-g* has essentially been shown to be an important factor in commitment to gametocytogenesis, but the method by which *ap2-g* is regulated remains unknown and is the subject of much research (Kafsack *et al.*, 2014; Sinha *et al.*, 2014; Bancells *et al.*, 2019). The epigenetic landscape of *P. berghei* asexual and sexual-stage parasites as carried out in the present study provides new leads as to which enzymes may be involved in the process of gametocytogenesis.

With this study, insights have also been gleaned into the roles of both *P. berghei* HAT1 and HDA1 in asexual development and gametocytogenesis (**Chapter 3**). Further studies of the clonal and tagged parasite lines produced in this study will result in the expansion of our current knowledge of the relatively limited repertoire of epigenetic regulatory enzymes in *Plasmodium* spp. and related Apicomplexa.

Finally, with continued optimisation of the ChIPmentation procedure, beginning with the lessons learned from **Chapter 5** of this study, investigation of specific histone PTMs at various stages of the *Plasmodium* life-cycle can be carried out. For instance, proteomic analysis of histone PTMs before and after induction of gametocytogenesis using the G1142 *P. berghei* line (section **1.7.2.3.1**), followed by ChIPmentation of differential PTMs (if present) between gametocytes induced at either ring-stage or trophozoite-stage, could shed light on the processes that produce the two different routes depicted in **Figure 6.1.**

Finally, from **Chapter 5,** it is clear that, using Epic peak-calling software, H4K8ac and H3K122ac peaks are abundantly present in the *P. berghei* mature schizont genome, with H4K8 acetylation found at exons predominantly associated with tRNA processing and DNA recombination (**Figure 5.4)**, and H3K122 acetylation associated with the exons of genes involved in macromolecule and cellular metabolic processes (**Figure 5.19**). Far fewer peaks were called for H3K9 trimethylation and acetylation marks, with the H3K9me3 modification predominantly found at genes involved in spliceosome functions and seryl-tRNA aminoacylation (**Figure 5.8).** H3K9 acetylation marks were then associated with DNA geometric changes, RNA unwinding, and protein palmitoylation. However, these results form the basis for future ChIPmentation studies in *P. berghei* gametocytes, with results then comparable to the findings of **Chapter 5.** This chapter also highlights the difficulty associated with choosing an appropriate peak-calling software for analysis of ChIP-sequencing results. Further optimisation of peak-calling software for analysis of broad peaks in the AT-rich *Plasmodium* genome will aid studies of epigenetic marks in future.

6.2 H3K79 methylation in the absence of DOT1: the potential for unidentified *Plasmodium* histone methyltransferases

One particularly intriguing finding from an analysis of histone post-translational modifications (PTMs) in both asexual-stage (schizont) and sexual-stage (gametocyte) *P. berghei* ANKA parasites, was the presence of mono-methylation of histone 3, lysine 79 (H3K79) in both mature schizonts and mature gametocytes (section **4.3.4**). This histone modification was also found during a quantitative study of histone PTMs in *P. falciparum* (Coetzee *et al.*, 2017).

The reason for the unusual nature of this finding is that *Plasmodium* parasites of both *Plasmodium* and *Laverania* subgenera lack genes homologous to the human disruptor of telomeric-silencing 1 (DOT1), the enzyme responsible for H3K79 methylation (Farooq *et al.*, 2016). DOT1 was previously described as the 'sole' enzyme responsible for methylation of H3K79 in humans, *Drosophila melanogaster,* and *Saccharomyces cerevisiae* (Van Leeuwen, Gafken and Gottschling, 2002; Farooq *et al.*, 2016), though with the findings of the present study, alongside those of Coetzee *et al.* (Coetzee *et al.*, 2017), this statement simply cannot be accurate.

In *Trypanosoma brucei brucei,* two DOT1 homologues, DOT1A and DOT1B, are responsible for dimethylation and trimethylation of the parasite H3K76 (corresponding to H3K79 in other eukaryotes) respectively (Janzen *et al.*, 2006). DOT1B in particular, has been suggested as having an essential role in chromatin remodelling during developmental differentiation of *T. brucei brucei* (Dejung *et al.*, 2016). Taking both *T. brucei brucei* (TREU927 strain) DOT1A (Tb927.8.1920) and DOT1B (Tb927.1.570) protein and BLASTing these sequences against all currently identified *Plasmodium spp.* protein sequences (EuPathDB version 36 for both), DOT1B generated no hits (result not shown). However, three hits were generated when comparing DOT1A to known *Plasmodium spp.* proteins: one uncharacterised protein in *P. coatneyi* Hackeri (a non-human primate malaria) and the heat shock 90 protein (HSP90) from both strains of the non-human *Laverania* species, *P. reichenowi*. Further comparison of these proteins was then carried out using Clustal (Sievers and Higgins, 2014) (**Figure 6.2**).

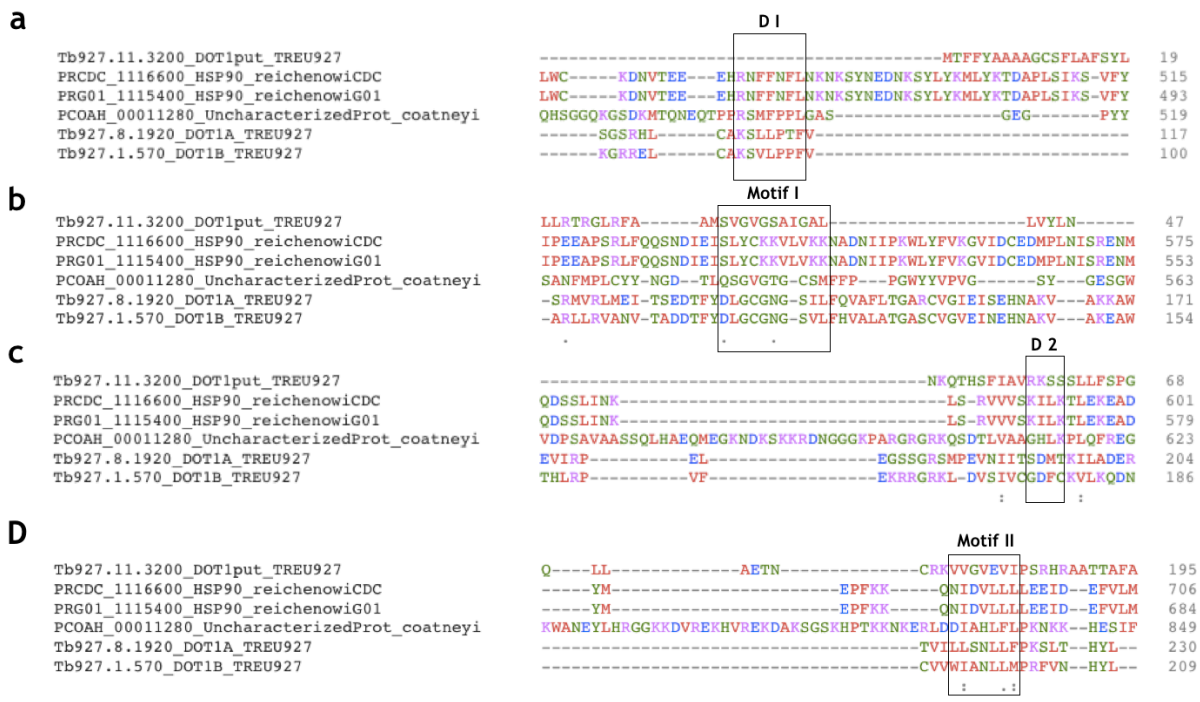


Figure 6.2: Clustal alignment of *T. brucei brucei* DOT1 methyltransferases with *Plasmodium spp*. proteins of greatest homology. In this image, two regions of a Clustal alignment are shown. Three *T. brucei brucei* DOT1 methyltransferases (Tb927.8.1920 (DOT1A); Tb927.1.570 (DOT1B), and a putative DOT1 protein (Tb927.11.3200)) are aligned to three protein hits generated upon BLAST-searching for *T. brucei brucei* DOT1A. These *Plasmodium spp.* proteins were a *P. coatneyi* Hackeri uncharacterised protein (encoded by PCOAH_00011280) and the heat shock protein 90 (HSP90) of *P. reichenowi* CDC and GO1 strains (encoded by PRCDC_1116600 and PRG01_1115400 respectively). In both **(b)** and **(d),** the protein regions corresponding to S-adenosyl-L-methionine (SAM) protein binding motifs are outlined within the Clustal alignment. In **(a)** and **(c),** regions of the peptides corresponding to DOT1-specific motifs, labelled D1 and D2, are outlined. Greatest homology appears between SAM-binding motifs in these particular proteins.

As can be seen in **Figure 6.2,** the regions within *Plasmodium spp.* proteins with greatest homology to three *T. bruci brucei* DOT1 methyltransferases are S-adenosyl-L-methionine (SAM)-binding motifs, and not the DOT1 methyltransferase-specific motifs (Min *et al.*, 2003; Janzen *et al.*, 2006). These results suggest that perhaps these *Plasmodium spp.* proteins maintain SAM-dependent properties, but are not capable of histone methyltransferase activity, as is the case with many other SAM-binding proteins (Gana *et al.*, 2013).

To conclude, there are no proteins homologous to DOT1 found in any *Plasmodium spp.* so far identified. Therefore, it remains to be seen whether methylation of H3K79 in *Plasmodium* spp. is a result of SET methyltransferase activity, or

previously unknown methyltransferase activity of a new class of eukaryotic or *Plasmodium*-specific proteins (SAM-dependent or otherwise).

6.3 Translating epigenetic findings into anti-malarial targets

In the present study, the epigenetic landscape of the rodent malaria parasite, *P. berghei* ANKA, was investigated with the aim of identifying differences between the regulation of gene expression in asexual-stage parasites and sexual-stage parasites (gametocytes). The discovery of the epigenetic regulatory mechanisms underlying commitment to gametocytogenesis in *P. berghei* may provide insight into the biology of malaria transmission.

At present, epigenetic regulation of the clonally variant *var* genes in the most lethal human malaria parasite, *P. falciparum*, has generated interest in the proteins that drive these processes as potential anti-malarial drug targets (Duffy *et al.*, 2014; Ay *et al.*, 2015; Duraisingh and Horn, 2016). In recent years, a number of studies have been published that have demonstrated the effects of approved and novel inhibitors of epigenetic regulators on *Plasmodium* asexual- and sexual-stage parasites (Darkin-Rattray *et al.*, 1996; Agbor-Enoh *et al.*, 2009; Wheatley *et al.*, 2010; Andrews *et al.*, 2012; Sumanadasa *et al.*, 2012; Hansen *et al.*, 2014; Malmquist *et al.*, 2015; Alves Avelar *et al.*, 2017; Chua *et al.*, 2017). These potential antimalarials are listed in **Table 6.1.**

Drug name	**Target name**	**IC_{50} (against *Plasmodium spp*.)**	**Novel/Approved**	***Plasmodium* species tested**	**Publication**
BIX-01294	Histone methyltransferases (non-specific)	280 ± 90 nM (*P. falciparum; In vitro)* 390 ± 90 nM (*P. vivax; In vitro*) 50 mg/kg/day (*P. berghei; In vivo*)	Novel	*P. falciparum* (multiple strains), *P. vivax* (clinical isolates), *P. berghei*	Malmquist et al., 2015
TM2-115 (BIX-01294 analogue)	Histone methyltransferases (non-specific)	340 ± 160 nM (*P. falciparum; In vitro*) 240 ± 70 nM (*P. vivax; In vitro*) 50 mg/kg/day (*P. berghei; In vivo*)	Novel	*P. falciparum* (multiple strains), *P. vivax* (clinical isolates), *P. berghei*	Malmquist et al., 2015
Vorinostat (SAHA)	Plasmodium histone deacetylase 1 (HDAC1)	250 ± 140 nM (P. *falciparum; In vitro)* 370 ± 20 nM (*P. knowlesi; In vitro*) 50 mg/kg/day (*P. berghei; In vivo*)	Approved (FDA/EMA)	*P. falciparum* 3D7, *P. knowlesi* A1H.1, *P. berghei*	Chua et al., 2017
Belinostat (PXD101)	Plasmodium histone deacetylase 1 (HDAC1)	170 ± 4 nM (*P. falciparum; In vitro*) 190 ± 20 nM (*P. knowlesi; In vitro*)	Approved (FDA/EMA)	*P. falciparum* 3D7, *P. knowlesi* A1H.1, *P. berghei*	Chua et al., 2017
Panobinostat (LBH-589)	Plasmodium histone deacetylase 1 (HDAC1)	4 ± 0.2 nM (*P. falciparum; In vitro*) 9 ± 1 nM (*P. knowlesi; In vitro*) 50 mg/kg/day (*P. berghei; In vivo*)	Approved (FDA/EMA)	*P. falciparum* 3D7, *P. knowlesi* A1H.1, *P. berghei*	Chua et al., 2017
Romidepsin (FK228)	Plasmodium histone deacetylase 1 (HDAC1)	140 ± 4 nM (P. *falciparum; In vitro*) 210 ± 20 nM (*P. knowlesi; In vitro*)	Approved (FDA)	*P. falciparum* 3D7, *P. knowlesi* A1H.1, *P. berghei*	Chua et al., 2017
YC-II-88 (WR301801)	Histone deacetylases (non-specific)	1.25 nM (*P. falciparum* 3D7; *In vitro)* 50 mg/kg/day (*P. berghei; In vivo*)	Novel	*P. falciparum* (multiple strains), *P. berghei*	Agbor-Enoh et al., 2009

Pracinostat (SB939)	Histone deacetylases (non-specific; likely Plasmodium HDAC1)	80 ± 30 nM (*P. falciparum* 3D7; *In vitro*) 150 ± 30 nM (*P. falciparum* Dd2; *In vitro*) 50 mg/kg/day (*P. berghei; In vivo*)	Approved (FDA/EMA)	*P. falciparum* (3D7 + Dd2), *P. berghei* ANKA	Sumanadasa et al., 2012
"6h"	Histone deacetylases (non-specific; likely Plasmodium HDAC1)	70 nM (*P. falciparum* 3D7; *In vitro*) 70 nM (*P. falciparum* Dd2; *In vitro*)	Novel	*P. falciparum* (3D7 + Dd2)	Alves Avelar et al., 2017
Trichostatin A (TSA)	Histone deacetylases (non-specific; likely Plasmodium HDAC1)	8-11 nM (*P. falciparum* 3D7; *In vitro*)	Approved (FDA)	*P. falciparum* 3D7	Andrews et al., 2012
2-ASA-9	Histone deacetylases (non-specific; likely Plasmodium HDAC1)	15-39 nM (*P. falciparum* 3D7; *In vitro*)	Novel	*P. falciparum* 3D7	Andrews et al., 2012
Apicidin	Histone deacetylases (non-specific)	91 nM (*P. falciparum* Dd2; *In vitro*) 50 mg/kg/day (*P. berghei; In vivo*)	Experimental only	*P. falciparum* Dd2, *P. berghei* KBG 173	Darkin-Rattray et al., 1996
LMK235	Plasmodium histone deacetylase 1 (HDAC1)	90-112 (*P. falciparum* 3D7; *In vitro*)	Novel	*P. falciparum* 3D7	Hansen et al., 2014
Asu-9	Plasmodium histone deacetylase 1 (HDAC1)	15 nM (*P. falciparum* 3D7; *In vitro*)	Novel	*P. falciparum* 3D7	Wheatley et al., 2010
Asu-13a	Plasmodium histone deacetylase 1 (HDAC1)	14 nM (*P. falciparum* 3D7; *In vitro*)	Novel	*P. falciparum* 3D7	Wheatley et al., 2010

Table 6.1: Chemical inhibitors of *Plasmodium spp*. epigenetic regulatory proteins. This table lists all chemical inhibitors of *Plasmodium spp.* histone/lysine methyltransferases (HKMTs) and histone/lysine deacetylases (HDACs) that have been identified to date. From left to right, each row lists the

name of the compound (both commercial and chemical names if present); the epigenetic regulator to which they are targeted; the half maximal inhibitory concentration (IC_{50}) of each compound (and the *Plasmodium* species in which this result was achieved); the approval of each drug for medicinal use (showing the drug as novel (untested in humans), experimental (established as a laboratory drug but not approved in humans), or approved); the *Plasmodium* species tested; and the publication from which the data were obtained.

As can be seen from **Table 6.1**, quite a few novel and approved chemical inhibitors of epigenetic regulators have been identified as having potential anti-malarial properties, some in multiple *Plasmodium spp.*, both *in vitro* and *in vivo*. What is abundantly clear is that the majority of the compounds listed are histone deacetylase (HDAC) inhibitors, owing to the fact that these drugs have already proven effective in treating human cancers, with drugs listed as 'approved' in **Table 6.1** referring to those HDAC inhibitors that have already been approved for the treatment of cancers only (Li and Seto, 2016).

There are two main issues that arise when studying HDACs as potential targets for anti-malarial therapy: i) the cytotoxicity of novel and approved HDAC inhibitors to host cells, with the potential for intolerable side-effects (Eckschlager *et al.*, 2017); and ii) the paucity of knowledge regarding the role of HDACs in *Plasmodium spp.* (Kanyal *et al.*, 2018). However, both issues can be overcome with the application of basic science; experiments such as those carried out in the present study to characterise potential epigenetic regulators such as histone acetyltransferase 1 (HAT1) and histone deacetylase 1 (HDA1) will inform future studies into the development of *Plasmodium*-specific inhibitors of epigenetic regulators.

6.4 Thesis summary: findings from the present study

Key findings from the present thesis report are listed below:

1. This study represents the first time a *Plasmodium* histone acetyltransferase (HAT1) was shown to play a role in gametocytogenesis. Complete HAT1 KO in the P. berghei 820 line resulted in a statistically significant decrease in mature female gametocytes ($P = 0.048$; $n = 5$).

2. Complete HDA1 knockout in a PbEGAM line resulted in a significant decrease in overall gametocytaemia ($P = 0.0109$; $n = 3$); a decrease in female gametocytes from $2.93\% \pm 1.09\%$ to $1.03\% \pm 0.22\%$ of total parasitaemia ($P = 0.0415$), and a decrease in male gametocytes from $2.07\% \pm 0.38\%$ to $1.21\% \pm 0.32\%$ of total parasitaemia ($P = 0.0402$).

3. *P. berghei* lines were created in which 4 histone deacetylases were tagged with haemagglutinin and a fluorescent marker, enabling future studies to be undertaken regarding these epigenetic factors.

4. The differential landscapes of histone modifications in both asexual-stage (mature schizont) and sexual-stage (gametocyte) *P. berghei* parasites were determined.

5. A protocol for the use of chromatin immunoprecipitation with high-throughput sequencing using a Tn5 transposase (ChIPmentation) was optimised for use in *Plasmodium spp*. parasites.

6. Locations within the entire *P. berghei* ANKA genome at which four histone modifications (H4K8ac, H3K9ac, H3K9me3, and H3K122ac) bind were determined in mature asexual *P. berghei* parasites using Epic peak-calling software.

7. It was determined that in *Plasmodium* subgenera, H3K79 methylation, and H3K36 di- and tri-methylation, are functions of as-yet unidentified histone/lysine methyltransferases, or that a different *Plasmodium* SET protein has evolved to take over the traditional DOT1 and SET2 roles.

6.5 Future work stemming from the present study

With the findings listed above (section **6.4**), and the generation of multiple *P. berghei* ANKA parasite lines throughout the course of this study, many more experiments are possible that could potentially provide important insights into the epigenetic regulation of the *Plasmodium* life-cycle. A number of these experiments are listed below:

1. Transmission of HAT1 knockout *P. berghei* parasites through mosquitoes to determine the role of HAT1 in the generation of gametes, in fertilisation, in oocyst development, and in sporozoite generation.

2. Histone extraction and analysis of histones from HAT1-depleted *P. berghei* schizonts and gametocytes using LC-MS/MS to identify HAT1 acetylation targets.

3. Fluorescence microscopy of HDA1 knockout and tagged parasites throughout the entire asexual life-cycle in a timed manner to identify potential enzyme stage-specificity.

4. Conversion of the AID knockdown plasmids to vectors that could be used for conditional degradation of proteins via the knocksideways system, preferably using the faster Gibson Assembly system.

5. ChIPmentation of *P. berghei* gametocytes to identify differential binding of the same four epigenetic marks (H4K8ac, H3K9ac, H3K9me3, and H3K122ac) to the parasite genome during sexual development.

6. Application of the rapamycin-inducible DiCre system to any of the epigenetic regulatory proteins from this study to complement knockout/knockdown findings.

7. A bioinformatic analysis of potential methyltransferase motifs in rodent *Plasmodium* genomes to identify novel histone/lysine methyltransferases that could take over the methyltransferase function of DOT1 and SET2.

8. Phosphorylation assays with recombinant *P. berghei* ANKA NEK4 and protein kinase B (PKB) with putative histone substrates to identify their function.

9. Histone extraction from parasites of the rapamycin-inducible *ap2-g P. berghei* line before induction of *ap2-g* expression, after induction of *ap2-g* in ring-stage parasites, and after induction of *ap2-g* during trophozoite growth to identify epigenetic differences between parasites that undergo gametocytogenesis in the same developmental cycle versus those that have re-invaded previously uninfected erythrocytes.

7. Appendix

Table 7.1: Primers used in the present study. The table below lists all PCR primer sequences designed and used in the present study alongside their name/number, recommended melting temperature (T_m), restriction site(s) (if present), the gene identifier (ID) of the gene from which the sequence was taken (excluding an introduced restriction enzyme recognition site), and a description of how the primer was used. All names given to primers are a number followed by a 'GU' prefix, standing for 'Glasgow University' (from where the primers were designed and ordered). In primer sequences that contain a restriction enzyme recognition site; the nucleobases of the restriction site are shown in red and in lower case, and the T_m shows the temperature of the sequence not including the restriction site. In primer sequences with nucleobases shown in green text, the adjacent T_m is the temperature of this particular primer sequence (and is also shown in green text). Lower case nucleobase letters in black text (reserved for primers designed for Gibson Assembly) indicates that this portion of the sequence overlaps with that of a second primer sequence. These overlapping fragments would be required for accurate Gibson assembly of a DNA construct into the knocksideways (KS) pL0078 background DNA construct.

Primer name	Sequence (5'→3')	T_m	Restriction site	Gene ID	Description
GU3199	ATTaagcttGTGCCACCACATTTGAATTTTAG	52°C	HindIII	PBANKA_1436100	HP1 KO upstream region forward primer
GU3200	TATccgcggTTATGTTATGCTATATTTCACTATTTTGTT	51°C	SacII	PBANKA_1436100	HP1 KO upstream region reverse primer
GU3201	ATAggtaccATGTATAGAGAATCCTCTTAGGC	52°C	KpnI	PBANKA_1436100	HP1 KO downstream region forward primer
GU3202	AATctcgagATGCATACATATAGCCACTAATTCA	51°C	XhoI	PBANKA_1436100	HP1 KO downstream region reverse primer
GU3219	ATAgggcccGTAGTAGCTATTTGTGAATTTCCATT	52°C	ApaI	PBANKA_0415300	RNAseII KO upstream region forward primer
GU3220	ATAccgcggCACTTTCGTATACAACAACTTACTT	51°C	SacII	PBANKA_0415300	RNAseII KO upstream region reverse primer
GU3221	TTAggtaccTGATTTGCGTCTTTCAGTTCATTC	52°C	KpnI	PBANKA_0415300	RNAseII KO downstream region forward primer
GU3222	ATTctcgagCTTGTATTGGACTTCGATGTGATT	52°C	XhoI	PBANKA_0415300	RNAseII KO downstream region reverse primer

GU3223	ATAgggcccGAAATTTTGGTGTTTCTCCACTTC	52°C	ApaI	PBANKA_1414600	IPK2 KO upstream region forward primer
GU3224	ATAccgcggATGGATATACATATATTCCCTTTGAC	52°C	SacII	PBANKA_1414600	IPK2 KO upstream region reverse primer
GU3225	TATggtaccCACAGTTTATTATTTCCAGTTGATGT	52°C	KpnI	PBANKA_1414600	IPK2 KO downstream region forward primer
GU3226	ATTctcgagTTTTGTACCCCACATGTACATTAC	52°C	XhoI	PBANKA_1414600	IPK2 KO downstream region reverse primer
GU3227	ATAgggcccATTGTATGCATGTACGCTTAGCA	52°C	ApaI	PBANKA_0826500	HDAC1 KO upstream region forward primer
GU3228	ATTccgcggAGTTGCTTATATACTTGGACCATC	52°C	SacII	PBANKA_0826500	HDAC1 KO upstream region reverse primer
GU3229	TATggtaccGGAGTGTAATTTGTGTGTATGAAG	52°C	KpnI	PBANKA_0826500	HDAC1 KO downstream region forward primer
GU3230	ATActcgagTTTCCCCATTTCATTTATTCCCTC	52°C	XhoI	PBANKA_0826500	HDAC1 KO downstream region reverse primer
GU3262	ATTaagcttCATCCCCATATATGCCTGAGT	52°C	HindIII	PBANKA_1206200	HDAC2/IPK1 KO upstream region forward primer
GU3263	ATAccgcggTGCAAATGTATATACTTTATACATACACA	52°C	SacII	PBANKA_1206200	HDAC2/IPK1 KO upstream region reverse primer
GU3264	TATggtaccGAAGCTATCCCTCTATGCTAC	52°C	KpnI	PBANKA_1206200	HDAC2/IPK1 KO downstream region forward primer
GU3265	ATActcgagCGACAGCATATACATAGAATGAC	52°C	XhoI	PBANKA_1206200	HDAC2/IPK1 KO downstream region reverse primer
GU3280	AATggcgcgccCAAGATGATATGGGGAATACTCT	52°C	AscI	PBANKA_1436100	HP1 KD forward primer
GU3281	ATTctcgagAACCGTTCTATATCTAAGTCTTGATA	52°C	XhoI	PBANKA_1436100	HP1 KD reverse primer
GU3282	ATTACATCTAAGAATAAAGAAAATGATTgtcgac CCGTTCCATATGATGTGG	48°C	SalI	PBANKA_0415300	RNAseII KD forward primer
GU3283	ATTctcgagGACAAAACGTAGGTTCCACATAT	52°C	XhoI	PBANKA_0415300	RNAseII KD reverse primer for SalI site insertion

GU3284	AATggcgcgccCATCTCCAATTCGGAGATATATAG	52°C	AscI	PBANKA_0415300	RNAseII KD forward primer for SalI site insertion
GU3285	CCACATCATATGGAACGGgtcgac AATCATTTTCTTTATTCTTAGATGTAAT	48°C	SalI	PBANKA_0415300	RNAseII KD reverse primer
GU3286	CTTATCATCACTTGCAAAAATACgtcgac CAAAAATCTGTAATACAAAAGTAGC	48°C	SalI	PBANKA_1414600	IPK2 KD forward primer
GU3287	ATTctcgagTTCACTTATAGAGCAATTCGAAAAATAAT	52°C	XhoI	PBANKA_1414600	IPK2 KD reverse primer for SalI site insertion
GU3288	AATggcgcgccACTGAAAAGTCTGAAGTTGATGATAA	52°C	AscI	PBANKA_1414600	IPK2 KD forward primer for SalI site insertion
GU3289	GCTACTTTTGTATTACAGATTTTTGgtcgac GTATTTTTGCAAGTGATGATAAG	48°C	SalI	PBANKA_1414600	IPK2 KD reverse primer
GU3290	AATggcgcgccCATGATTATGAATATGTTGATTTTTTATCAT	51°C	AscI	PBANKA_0826500	HDAC1 KD forward primer
GU3291	ATTctcgagAATAATTCCTTGATCTCTGTCTGAC	53°C	XhoI	PBANKA_0826500	HDAC1 KD reverse primer
GU3292	AATggcgcgccTATGCATGGTTATATCTATGCTGTGA	52°C	AscI	PBANKA_1206200	HDAC2/IPK1 KD forward primer
GU3293	ATTctcgagAATCGAGGATTTTACTTTTTCAATAGTTT	52°C	XhoI	PBANKA_1206200	HDAC2/IPK1 KD reverse primer
GU3294	GTAGTGTTTTCGAAGCCTTgtcgac ATGTTGTACATGCAATAAAATAGTA	48°C	SalI	PBANKA_1343800	SIR2A KD forward primer
GU3295	ATTctcgagTTTTCCTCTTTTAATATTTCGATCAAATTAG	52°C	XhoI	PBANKA_1343800	SIR2A KD reverse primer for SalI site insertion
GU3296	AATggcgcgccGCCTTGACAGGGTCTGG	52°C	AscI	PBANKA_1343800	SIR2A KD forward primer for SalI site insertion
GU3297	TACTATTTTATTGCATGTACAACATgtcgac AAGGCTTCGAAAACACTAC	47°C	SalI	PBANKA_1343800	SIR2A KD reverse primer
GU3298	GTATCAAATTATAATCAAAATTTACTCAgtcgac AAAGGGAAAATGTTCAACAAACT	48°C	SalI	PBANKA_1315100	SIR2B KD forward primer

GU3299	ATTctcgagTACACATTCAAATAAGTCGTTGGTG	53°C	XhoI	PBANKA_1315100	SIR2B KD reverse primer for SalI site insertion
GU3300	AATggcgcgccGATGATATTGACGATTCTTATTTTAGC	52°C	AscI	PBANKA_1315100	SIR2B KD forward primer for SalI site insertion
GU3301	AGTTTGTTGAACATTTTCCCTTTgtcgac TGAGTAAATTTTGATTATAATTTGATAC	48°C	SalI	PBANKA_1315100	SIR2B KD reverse primer
GU3302	CGATAACTAATACTATCTCAAAAAAgtcgac ATGCGCACGAAAAGAGAG	48°C	SalI	PBANKA_1335400	HDA1 KD forward primer
GU3303	ATTctcgagATTTTTTAGGTGATTCCATAAATTATTTAAC	51°C	XhoI	PBANKA_1335400	HDA1 KD reverse primer for SalI site insertion
GU3304	AATggcgcgccAGCGGGGTCGTTTTAAATGCA	52°C	AscI	PBANKA_1335400	HDA1 KD forward primer for SalI site insertion
GU3305	CTCTCTTTTCGTGCGCATgtcgac TTTTTTGAGATAGTATTAGTTATCG	48°C	SalI	PBANKA_1335400	HDA1 KD reverse primer
GU3306	TTATATTAGACGTTGATGTGCATgtcgac CAAGGTGATGGAACAGCA	48°C	SalI	PBANKA_1106200	HDAP KD forward primer
GU3307	ATTctcgagTCCATGTTTTCCCTTTGTAGCC	52°C	XhoI	PBANKA_1106200	HDAP KD reverse primer for SalI site insertion
GU3308	AATggcgcgccCAAAAAATCCACCTTATGTTTTTCATC	52°C	AscI	PBANKA_1106200	HDAP KD forward primer for SalI site insertion
GU3309	TGCTGTTCCATCACCTTGgtcgac ATGCACATCAACGTCTAATATAA	48°C	SalI	PBANKA_1106200	HDAP KD reverse primer
GU3310	TTGTCAATTCTAACATAAAAGAAATTAgtcgac CTGTAGAAGATCCAGCAG	48°C	SalI	PBANKA_0718400	HAT1 KD forward primer

GU3311	ATTctcgagAAGTGGGTATAATGACCTCAACTT	52°C	XhoI	PBANKA_0718400	HAT1 KD reverse primer for SalI site insertion
GU3312	AATggcgcgccCAATTGCTGTGATAATGATGAAGAAA	52°C	AscI	PBANKA_0718400	HAT1 KD forward primer for SalI site insertion
GU3313	CTGCTGGATCTTCTACAGgtcgac TAATTTCTTTTATGTTAGAATTGACAA	48°C	SalI	PBANKA_0718400	HAT1 KD reverse primer
GU3399	GACAACCAAGAAATTACATTTCAAAAG	52°C	-	PBANKA_1436100	HP1 KO 5' integration forward primer
GU3400	AGAATTAAGCTGGGCTGCAG	52°C	-	-	*Pbeef1α* promoter reverse primer (to confirm KO construct integration)
GU3401	TTAGCGAAGATTAACACGTTTGC	52°C	-	PBANKA_1436100	HP1 KO 3' integration reverse primer
GU3402	GTTACTGGTGCCCTCGAC	53°C	-	-	*yfcu* sequence forward primer (to confirm KO construct integration)
GU3403	GGTAATACTGGAATGAGATAGTAG	52°C	-	PBANKA_0415300	RNAseII KO 5' integration forward primer
GU3404	AATTATGACACTACAGAAGATTACCAT	52°C	-	PBANKA_0415300	RNAseII KO 3' integration reverse primer
GU3405	CCCATACATAGGACTGGCTTA	52°C	-	PBANKA_1414600	IPK2 KO 5' integration forward primer
GU3406	GAAGTTTTCCCCGAAGAAGG	52°C	-	PBANKA_1414600	IPK2 KO 3' integration reverse primer
GU3407	CTGGCATGTATACACACATATAAG	52°C	-	PBANKA_0826500	HDAC1 KO 5' integration forward primer
GU3408	AGTTTATCCCATGTTTTACCTTTTC	51°C	-	PBANKA_0826500	HDAC1 KO 3' integration reverse primer
GU3409	CCACCTATCTTTCCCTCCATT	52°C	-	PBANKA_1206200	HDAC2/IPK1 KO 5' integration forward primer
GU3410	CATGAAAATATATTATTGTTCAATGGAGG	53°C	-	PBANKA_1206200	HDAC2/IPK1 KO 3' integration reverse primer

GU3433	TCGAATTCAGTATTGCACGACTAGATAT	56°C	-	-	3' *Pb48/45* reverse primer (to confirm KD construct integration)
GU3434	GCGACACGGAAATGTTGAATACTC	56°C	-	-	*Amp* sequence forward primer (to confirm KD construct integration)
GU3871	GTGTATGTATAAAGTATATACATTTGCATATAATTTATTG	56.3°C	-	PBANKA_1206200	HDAC2/IPK1 KD integration forward primer (extra internal primer)
GU3872	CGGACCCTTCAAAAAGTAAATCTGAAG	56.7°C	-	PBANKA_1206200	HDAC2/IPK1 KD integration forward primer (extra internal primer)
GU3435	CAAATACATGGTGAGTACCTTCCCT	56°C	-	PBANKA_1343800	SIR2A KD 5' integration forward primer
GU3436	GACGAAGTAAATGTATACTTTACAAACCC	56°C	-	PBANKA_1343800	SIR2A KD 3' integration reverse primer
GU3437	CAAAAATGTGCATGATGTACTAGACACAG	57°C	-	PBANKA_1315100	SIR2B KD 5' integration forward primer
GU3438	CCAGCCATTCCATGTGGGCA	56°C	-	PBANKA_1315100	SIR2B KD 3' integration reverse primer
GU3439	GAGTGCATGATGTGTCATATATAAAGATG	56°C	-	PBANKA_1335400	HDA1 KD 5' integration forward primer
GU3440	CGTTCATGAGAATATCAAGTTTATTGTGG	56°C	-	PBANKA_1335400	HDA1 KD 3' integration reverse primer
GU3441	GTTATATGGGGATTGGGGTTATTTGG	56°C	-	PBANKA_1106200	HDAP KD 5' integration forward primer
GU3442	TATCATATGTGAACGATTTGGTCGGG	56°C	-	PBANKA_1106200	HDAP KD 3' integration reverse primer
GU3443	CTTTAAAGGATGTGAAGGAAAAATCGAAC	56°C	-	PBANKA_0718400	HAT1 KD 5' integration forward primer
GU3444	CGAACCTTCTATTAACCCTCAAAATG	55°C	-	PBANKA_0718400	HAT1 KD 3' integration reverse primer
GU3445	GCATTTTACACTATTTTGCCATAAGCAC	56°C	-	-	OsTIR1/9Myc construct sequencing forward primer (within HSP70 promoter 3'UTR sequence)
GU3446	CTCCTCAACATCTTGCATTTCTCC	56°C	-	-	OsTIR1/9Myc construct sequencing reverse primer (within OsTIR1 sequence)

GU3447	GGATTTGGGGACAGGGAATTTGAC	57°C	-	-	OsTIR1/9Myc construct sequencing forward primer (within OsTIR1 sequence)
GU3448	CCAAAGGGATCGCATTGTCTCG	57°C	-	-	OsTIR1/9Myc construct sequencing reverse primer (within OsTIR1 sequence)
GU3449	GTATTCGAATTCAGTATTGCACGACTAG	57°C	-	-	OsTIR1/9Myc construct sequencing reverse primer (within OsTIR1 sequence)
GU3450	GCAATATATAGCTATATGTAATGGTGAGTG	56°C	-	-	Additional AID::HA construct reverse primer (within AID::HA sequence)
GU4150	CAGGATCAGATGAAGAATTTGAAATAGG	55.5°C	-	PBANKA_1436100	Extra HP1 KD 5' integration forward primer
GU4151	CCTTGTTTGCAGGCTCATTATTCAG	56°C	-	-	Additional AID::HA construct reverse primer (within AID::HA sequence)
GU4152	CCGATTTAGAGCTTGACGGGG	56.3°C	-	-	Additional AID::HA construct forward primer (within plasmid background sequence)
GU4153	GGATGTTATTAGGTGCCTAAGAGG	55.7°C	-	PBANKA_1436100	Extra HP1 KD 3' integration reverse primer
GU4154	GGTCCAAGTATATAAGCAACTCCC	55.7°C	-	PBANKA_0826500	Extra HDAC1 KD 5' integration forward primer
GU4155	CGACGTCATGAGTCGTAGATCC	56.7°C	-	PBANKA_0826500	Extra HDAC1 KD 5' integration reverse primer
GU4156	GCTCGTATTTTATAAAATGGCGTGTCG	56.7°C	-	PBANKA_0826500	Extra HDAC1 KD 3' integration reverse primer
GU4157	CGAAATGTGCTAAGCTCTATGAGG	55.7°C	-	PBANKA_1206200	Extra HDAC2/IPK1 KD 5' integration forward primer
GU4158	GGTGCTCCGTCCATTGATACC	56.3°C	-	-	Additional AID::HA construct reverse primer (within AID::HA sequence)

GU4159	CGACTCACTATAGGGCGAATTGG	57.1°C	-	-	Additional AID::HA construct forward primer (within plasmid background sequence)
GU4160	GGTATATGATGGTGTAGATACGGG	55.7°C	-	PBANKA_1206200	Extra HDAC2/IPK1 KD 3' integration reverse primer
GU4161	GGAATGTTTGACGAAAACGAGATGG	56°C	-	PBANKA_1343800	Extra SIR2A KD 5' integration forward primer
GU4162	GGAAACCATCACGTTCTTCCGG	56.7°C	-	-	Additional AID::HA construct reverse primer (within AID::HA sequence)
GU4163	GCTAACCAAAGATCGGACATAGC	55.3°C	-	PBANKA_1343800	Extra SIR2A KD 3' integration reverse primer
GU4164	GCATTAAATGCAAGCGGGGTCG	56.7°C	-	PBANKA_1335400	Extra HDA1 KD 5' integration forward primer
GU4165	GGAGCCACTGTATCTCCACC	55.9°C	-		Additional AID::HA construct reverse primer (within AID::HA sequence)
GU4175	GCCACTACCAGGTGAATTTTCCC	57.1°C	-	PBANKA_1335400	Extra HDA1 KD 3' integration reverse primer
GU4167	GGATGAACAAATATGGGACTATCC	54°C	-	PBANKA_1106200	Extra HDAP KD 5' integration forward primer
GU4168	CCCTCTCTTGTTTCCGGTTACC	56.7°C	-	-	Additional AID::HA construct reverse primer (within AID::HA sequence)
GU4169	CATTCGCCATTCAGGCTGCG	55.9°C	-	-	Additional AID::HA construct forward primer (within plasmid background sequence)
GU4170	CTTCGCTTTTTAAATATGGCATAGCC	54.8°C	-	PBANKA_1106200	Extra HDAP KD 3' integration reverse primer
GU4171	CGAAAAGTAGAATGGTTTTATCACTGG	55.2°C	-	PBANKA_0718400	Extra HAT1 KD 5' integration forward primer
GU4172	CCCAGTCACGACGTTGTAAAACG	57.1°C	-		Additional AID::HA construct forward primer (within plasmid background sequence)
GU4173	CTGTTTACACACAATATATTTCCATTCC	54.1°C	-	PBANKA_0718400	Extra HAT1 KD 3' integration reverse primer
GU4062	CATACTAGCCATTTTATGTGTG	57°C	-	-	*Plasmo*GEM GW1 forward primer

GU4063	GACTTTGGTGACAGATACTAC	58°C	-	-	*Plasmo*GEM GW2 reverse primer
GU4260	CTAGGCCCAGGCGCAGAGGA	60°C	-	-	*Plasmo*GEM GT reverse primer
GU4166	TGGGGGCTCTTCTTCATCATCCTCA	59.3°C	-	-	*Plasmo*GEM GT forward primer
GU4315	GCAGCTACCCCTCCTGCATTG	58.3°C	-	PBANKA_1335400	HDA1 5' reverse primer for *Plasmo*GEM KO construct integration
GU4316	CTGTAGAAGATCCAGCAGCATCTTTTAC	58.5°C	-	PBANKA_0718400	HAT1 3' forward primer for *Plasmo*GEM KO construct integration
GU4261	ggggagatggtttccacctgcactcccat CCATGGTGCCACCACATTTGAATTTTAGTTTCC	74.6°C	-	PBANKA_1436100	HP1 KS forward primer (for Gibson Assembly)
GU4262	GGCGGCCGCTATCTAGACACTAGTGAGAACC GTTCTATATCTAAGTCTTGATAAAAGAAAG	71.4°C	-	PBANKA_1436100	HP1 KS reverse primer (for Gibson Assembly)
GU4263	CTTTCTTTTATCAAGACTTAGATATAGAAC GGTTCTCACTAGTGTCTAGATAGCGGCCGCC	71.4°C	-	PBANKA_1436100	HP1 KS forward primer (for Gibson Assembly)
GU4264	GGAAACTAAAATTCAAATGTGGTGGCACCATGG atgggagtgcaggtggaaaccatctcccc	74.6°C	-	PBANKA_1436100	HP1 KS reverse primer (for Gibson Assembly)

8. References

Adey, A., Morrison, H. G., Asan, Xun, X., Kitzman, J. O., Turner, E. H., Stackhouse, B., MacKenzie, A. P., Caruccio, N. C., Zhang, X. and Shendure, J. (2010) 'Rapid, low- input, low-bias construction of shotgun fragment libraries by high-density in vitro transposition', *Genome Biology*, 11(12): R119. doi: 10.1186/gb-2010-11-12-r119.

Adjalley, S. H., Chabbert, C. D., Klaus, B., Pelechano, V. and Steinmetz, L. M. (2016) 'Landscape and Dynamics of Transcription Initiation in the Malaria Parasite *Plasmodium falciparum*.', *Cell reports*, 14(10): 2463–75.

Adl, S. M., Simpson, A. G. B., Lane, C. E., Lukeš, J., Bass, D., Bowser, S. S., Brown, M. W., Burki, F., Dunthorn, M., Hampl, V., Heiss, A., Hoppenrath, M., Lara, E., le Gall, L., Lynn, D. H., McManus, H., Mitchell, E. A. D., Mozley-Stanridge, S. E., Parfrey, L. W., Pawlowski, J., Rueckert, S., Shadwick, L., Schoch, C. L., Smirnov, A. and Spiegel, F. W. (2012) 'The Revised Classification of Eukaryotes', *Journal of Eukaryotic Microbiology*, 59(5): 429–514.

Agbor-Enoh, S., Seudieu, C., Davidson, E., Dritschilo, A. and Jung, M. (2009) 'Novel inhibitor of Plasmodium histone deacetylase that cures *P. berghei*-infected mice', *Antimicrobial Agents and Chemotherapy*, 53(5): 1727–1734.

Alonso, J. M. and Ecker, J. R. (2006) 'Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*', *Nature Reviews Genetics*, 7(7): 524–536.

Alves Avelar, L. A., Held, J., Engel, J. A., Sureechatchaiyan, P., Hansen, F. K., Hamacher, A., Kassack, M. U., Mordmüller, B., Andrews, K. T. and Kurz, T. (2017) 'Design and Synthesis of Novel Anti-Plasmodial Histone Deacetylase Inhibitors Containing an Alkoxyamide Connecting Unit', *Archiv der Pharmazie*, 350(3–4). doi: 10.1002/ardp.201600347.

Aly, A. S. I., Lindner, S. E., MacKellar, D. C., Peng, X. and Kappe, S. H. I. (2011) 'SAP1 is a critical post-transcriptional regulator of infectivity in malaria parasite sporozoite stages', *Molecular Microbiology*, 79(4): 929–939.

Andenmatten, N., Egarter, S., Jackson, A. J., Jullien, N., Herman, J. P. and Meissner, M. (2013) 'Conditional genome engineering in *Toxoplasma gondii* uncovers alternative invasion mechanisms', *Nature Methods*, 10(2): 125–127.

Andreína Pacheco, M., Cranfield, M., Cameron, K. and Escalante, A. A. (2013) 'Malarial parasite diversity in chimpanzees: the value of comparative approaches to ascertain the evolution of *Plasmodium falciparum* antigens', *Malaria Journal*, 12. doi: 10.1186/1475-2875-12-328.

Andrews, K. T., Gupta, A. P., Tran, T. N., Fairlie, D. P., Gobert, G. N. and Bozdech, Z. (2012) 'Comparative gene expression profiling of *P. falciparum* malaria parasites exposed to three different histone deacetylase inhibitors', *PLoS ONE*, 7(2). doi: 10.1371/journal.pone.0031847.

Arumugam, T. U., Takeo, S., Yamasaki, T., Thonkukiatkul, A., Miura, K., Otsuki, H., Zhou, H., Long, C. A., Sattabongkot, J., Thompson, J., Wilson, D. W., Beeson, J. G., Healer, J., Crabb, B. S., Cowman, A. F., Torii, M. and Tsuboi, T. (2011) 'Discovery of GAMA, a *Plasmodium falciparum* merozoite micronemal protein, as a novel blood-stage vaccine candidate antigen.', *Infection and immunity*, 79(11): 4523–32.

Assiri, M. A., Roy, S. R., Harris, P. S., Ali, H., Liang, Y., Shearn, C. T., Orlicky, D. J., Roede, J. R., Hirschey, M. D., Backos, D. S. and Fritz, K. S. (2017) 'Chronic Ethanol Metabolism Inhibits Hepatic Mitochondrial Superoxide Dismutase via Lysine Acetylation', *Alcoholism: Clinical and Experimental Research*, 41(10):1705–1714.

Auburn, S., Campino, S., Clark, T. G., Djimde, A. A., Zongo, I., Pinches, R., Manske, M., Mangano, V., Alcock, D., Anastasi, E., Maslen, G., MacInnis, B., Rockett, K., Modiano, D., Newbold, C. I., Doumbo, O. K., Ouédraogo, J. B. and Kwiatkowski, D. P. (2011) 'An Effective Method to Purify *Plasmodium falciparum* DNA Directly from Clinical Blood Samples for Whole Genome High-Throughput Sequencing', *PLoS ONE*, 6(7): e22213. doi: 10.1371/journal.pone.0022213.

Ay, F., Bunnik, E. M., Varoquaux, N., Vert, J.-P., Noble, W. S. and Le Roch, K. G. (2015) 'Multiple dimensions of epigenetic gene regulation in the malaria parasite *Plasmodium falciparum*: gene regulation via histone modifications, nucleosome positioning and nuclear architecture in *P. falciparum*.', *BioEssays : news and reviews in molecular, cellular and developmental biology*, 37(2): 182–94.

Brancucci, N. M., Goldowitz, I., Buchholz, K., Werling, K. and Marti, M. (2015) 'An assay to probe *Plasmodium falciparum* growth, transmission stage formation and early gametocyte development'. *Nature Protocols,* 10(8): 1131-1142.

Baek, S. H. (2011) 'When Signaling Kinases Meet Histones and Histone Modifiers in the Nucleus', *Molecular Cell*, 42(3): 274-84.

Bahl, A., Brunk, B., Crabtree, J., Fraunholz, M. J., Gajria, B., Grant, G. R., Ginsburg, H., Gupta, D., Kissinger, J. C., Labo, P., Li, L., Mailman, M. D., Milgram, A. J., Pearson, D. S., Roos, D. S., Schug, J., Stoeckert, C. J. and Whetzel, P. (2003) 'PlasmoDB: The *Plasmodium* genome resource. A database integrating experimental and computational data', *Nucleic Acids Research*, 31(1): 212–215.

Bajar, B. T., Wang, E. S., Zhang, S., Lin, M. Z. and Chu, J. (2016) 'A guide to fluorescent protein FRET pairs', *Sensors (Switzerland)*, 16(9): 1488. doi: 10.3390/s16091488.

Bakar, N. A., Klonis, N., Hanssen, E., Chan, C. and Tilley, L. (2010) 'Digestive-vacuole genesis and endocytic processes in the early intraerythrocytic stages of *Plasmodium falciparum*', *Journal of Cell Science*, 123(3): 441-50.

Baker, D. A. (2010) 'Malaria gametocytogenesis.', *Molecular and biochemical parasitology*, 172(2): 57–65.

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

Balu, B., Blair, P. L. and Adams, J. H. (2009) 'Identification of the transcription initiation site reveals a novel transcript structure for *Plasmodium falciparum maebl*', *Experimental Parasitology*, 121(1): 110–114.

Bancells, C., Llorà-Batlle, O., Poran, A., Nötzel, C., Rovira-Graells, N., Elemento, O., Kafsack, B. F. C. and Cortés, A. (2019) 'Revisiting the initial steps of sexual development in the malaria parasite *Plasmodium falciparum*', *Nature Microbiology*. doi: 10.1038/s41564-018-0291-7.

Bannister, A. J. and Kouzarides, T. (2011) 'Regulation of chromatin by histone modifications.', *Cell research*, 21(3): 381–95.

Baragaña, B., Hallyburton, I., Lee, M. C. S., Norcross, N. R., Grimaldi, R., Otto, T. D., Proto, W. R., Blagborough, A. M., Meister, S., Wirjanata, G., Ruecker, A., Upton, L. M., Abraham, T. S., Almeida, M. J., Pradhan, A., Porzelle, A., Martínez, M. S., Bolscher, J. M., Woodland, A., Norval, S., Zuccotto, F., Thomas, J., Simeons, F., Stojanovski, L., Osuna-Cabello, M., Brock, P. M., Churcher, T. S., Sala, K. A., Zakutansky, S. E., Jiménez-Díaz, M. B., Sanz, L. M., Riley, J., Basak, R., Campbell, M., Avery, V. M., Sauerwein, R. W., Dechering, K. J., Noviyanti, R., Campo, B., Frearson, J. A., Angulo-Barturen, I., Ferrer-Bazaga, S., Gamo, F. J., Wyatt, P. G., Leroy, D., Siegl, P., Delves, M. J., Kyle, D. E., Wittlin, S., Marfurt, J., Price, R. N., Sinden, R. E., Winzeler, E. A., Charman, S. A., Bebrevska, L., Gray, D. W., Campbell, S., Fairlamb, A. H., Willis, P. A., Rayner, J. C., Fidock, D. A., Read, K. D. and Gilbert, I. H. (2015) 'A novel multiple-stage antimalarial agent that inhibits protein synthesis', *Nature*, 522(7556): 315–320.

Barber, C. M., Turner, F. B., Wang, Y., Hagstrom, K., Taverna, S. D., Mollah, S., Ueberheide, B., Meyer, B. J., Hunt, D. F., Cheung, P. and Allis, C. D. (2004) 'The enhancement of histone H4 and H2A serine 1 phosphorylation during mitosis and S-phase is evolutionarily conserved', *Chromosoma*, 112(7): 360-371.

Bártfai, R., Hoeijmakers, W. A. M., Salcedo-Amaya, A. M., Smits, A. H., Janssen-Megens, E., Kaan, A., Treeck, M., Gilberger, T. W., Franc ̧oijs, K. J. and Stunnenberg, H. G. (2010) 'H2A.Z demarcates intergenic regions of the *Plasmodium falciparum* epigenome that are dynamically marked by H3K9ac and H3K4me3', *PLoS Pathogens*, 6(12): e1001223. doi: 10.1371/journal.ppat.1001223.

Basco, L. K., Eldin De Pccoulas, P., Wilson ', C. M., Le Bras, J. and Mazabraud, A. (1995) 'Point mutations synthase gene and in the dihydrofolate reductase-thymidylate pyrimethamine and cycloguanil resistance in *Plasmodium falciparum*', *Molecular and Biochemical Parasitology*, 69: 135–138.

Baton, L. A. & Ranford-Cartwright, L. C. (2005) 'Spreading the seeds of million-murdering death: metamorphoses of malaria in the mosquito', *Trends in Parasitology*, 21(12): 573–580.

Baton, L. & Ranford-Cartwright, L. (2005) 'How do malaria ookinetes cross the mosquito midgut wall?', *Trends in Parasitology*, 21(1): 22–28.

Batugedara, G., Lu, X. M., Bunnik, E. M. and Le Roch, K. G. (2017) 'The Role of Chromatin Structure in Gene Regulation of the Human Malaria Parasite'. *Trends In Parasitology,* 33(5): 364-377.

Bauer, U. M., Daujat, S., Nielsen, S. J., Nightingale, K. and Kouzarides, T. (2002) 'Methylation at arginine 17 of histone H3 is linked to gene activation', *EMBO Reports*, 3(1): 39–44.

Baum, J., Papenfuss, A. T., Baum, B., Speed, T. P. and Cowman, A. F. (2006) 'Regulation of apicomplexan actin-based motility', *Nature Reviews Microbiology*, 4(8): 621– 628.

Baum, J., Richard, D., Healer, J., Rug, M., Krnajski, Z., Gilberger, T.-W., Green, J. L., Holder, A. A. and Cowman, A. F. (2006) 'A conserved molecular motor drives cell invasion and gliding motility across malaria life cycle stages and other apicomplexan parasites.', *The Journal of biological chemistry*, 281(8): 5197–208.

Becker, P. B. (2002) 'Nucleosome sliding: Facts and fiction', *EMBO Journal*, 21(18): 4749- 53.

Beeson, J. G., Drew, D. R., Boyle, M. J., Feng, G., Fowkes, F. J. I. and Richards, J. S. (2016) 'Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria.', *FEMS microbiology reviews*, 40(3): 343–72.

Beetsma, A. L., Van De Wiel, T. J. J. M., Sauerwein, R. W. and Eling, W. M. C. (1998) '*Plasmodium berghei* ANKA: Purification of Large Numbers of Infectious Gametocytes', *Experimental Parasitology*, 88(1): 69–72.

Beier, J. C. (1998) 'Malaria Parasite Development In Mosquitoes', *Annual Review of Entomology*, 43(1): 519–543.

Bennink, S., Kiesow, M. J. and Pradel, G. (2016) 'The development of malaria parasites in the mosquito midgut', *Cellular Microbiology*, 18(7): 905–918.

Bermúdez, M., Moreno-Pérez, D. A., Arévalo-Pinzón, G., Curtidor, H. and Patarroyo, M. A. (2018) '*Plasmodium vivax in vitro* continuous culture: the spoke in the wheel', *Malaria Journal*. doi: 10.1186/s12936-018-2456-5.

Biga, P. R., Latimer, M. N., Froehlich, J. M., Gabillard, J.-C. and Seiliez, I. (2017) 'Distribution of H3K27me3, H3K9me3, and H3K4me3 along autophagy-related genes highly expressed in starved zebrafish myotubes', *Biology Open*, 6(11): 1720– 1725.

Billker, O., Dechamps, S., Tewari, R., Wenig, G., Franke-Fayard, B. and Brinkmann, V. (2004) 'Calcium and a Calcium-Dependent Protein Kinase Regulate Gamete Formation and Mosquito Transmission in a Malaria Parasite', *Cell*, 117(4): 503–514.

Billker, O., Lindo, V., Panico, M., Etienne, A. E., Paxton, T., Dell, A., Rogers, M., Sinden, R. E. and Morris, H. R. (1998) 'Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito', *Nature*, 392(6673): 289–292.

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

Bischoff, E. and Vaquero, C. (2010) '*In silico* and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of *Plasmodium falciparum*', *BMC Genomics*, 11(1). doi: 10.1186/1471-2164-11-34.

Blanquart, S. and Gascuel, O. (2011) 'Mitochondrial genes support a common origin of rodent malaria parasites and *Plasmodium falciparum*'s relatives infecting great apes.', *BMC evolutionary biology*, 11(1), p. 70. doi: 10.1186/1471-2148-11-70.

Blomqvist, K., DiPetrillo, C., Streva, V. A., Pine, S. and Dvorin, J. D. (2017) 'Receptor for Activated C-Kinase 1 (PfRACK1) is required for *Plasmodium falciparum* intra- erythrocytic proliferation', *Molecular and Biochemical Parasitology*, 211: 62–66.

Bousema, T. and Drakeley, C. (2011) 'Epidemiology and infectivity of *Plasmodium falciparum* and *Plasmodium vivax* gametocytes in relation to malaria control and elimination.', *Clinical microbiology reviews*, 24(2): 377–410.

Bowman, G. D. and Poirier, M. G. (2015) 'Post-translational modifications of histones that influence nucleosome dynamics', *Chemical Reviews*, pp. 2274–2295. doi: 10.1021/cr500350x.

Bozdech, Z., Llinás, M., Pulliam, B. L., Wong, E. D., Zhu, J. and DeRisi, J. L. (2003) 'The Transcriptome of the Intraerythrocytic Developmental Cycle of *Plasmodium falciparum*', *PLoS Biology*, 1(1), p. e5. doi: 10.1371/journal.pbio.0000005.

Bozdech, Z., Mok, S., Hu, G., Imwong, M., Jaidee, A., Russell, B., Ginsburg, H., Nosten, F., Day, N. P. J., White, N. J., Carlton, J. M. and Preiser, P. R. (2008) 'The transcriptome of *Plasmodium vivax* reveals divergence and diversity of transcriptional regulation in malaria parasites.', *Proceedings of the National Academy of Sciences of the United States of America*, 105(42): 16290–5.

Bozdech, Z., Zhu, J., Joachimiak, M. P., Cohen, F. E., Pulliam, B. and DeRisi, J. L. (2003) 'Expression profiling of the schizont and trophozoite stages of

Plasmodium falciparum with a long-oligonucleotide microarray', *Genome Biology*, 4(2): R9. doi: 10.1186/gb-2003-4-2-r9.

Braks, J. A. M., Franke-Fayard, B., Kroeze, H., Janse, C. J. and Waters, A. P. (2006) 'Development and application of a positive-negative selectable marker system for use in reverse genetics in *Plasmodium*.', *Nucleic acids research*, 34(5): e39. doi: 10.1093/nar/gnj033.

Brancucci, N. M. B., Bertschi, N. L., Zhu, L., Niederwieser, I., Chin, W. H., Wampfler, R., Freymond, C., Rottmann, M., Felger, I., Bozdech, Z. and Voss, T. S. (2014) 'Heterochromatin protein 1 secures survival and transmission of malaria parasites', *Cell Host and Microbe*, 16(2): 165–176.

Brancucci, N. M. B., Gerdt, J. P., Wang, C., Jiang, H. Y., Clardy, J., Marti, M., De Niz, M., Philip, N., Adapa, S. R., Zhang, M., Hitz, E., Niederwieser, I., Boltryk, S. D., Laffitte, M.-C., Clark, M. A., Grü, C., Ravel, D., Soares, A. B., Demas, A., Bopp, S., Rubio-Ruiz, B. N., Conejo-Garcia, A., Wirth, D. F., Gendaszewska-Darmach, E., Duraisingh, M. T., Adams, J. H., Voss, T. S., Waters, A. P. and Jiang, R. H. Y. (2017) 'Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite *Plasmodium falciparum*', *Cell*, 171: 1–13.

Brancucci, N. M. B., Witmer, K., Schmid, C. D., Flueck, C. and Voss, T. S. (2012) 'Identification of a cis-acting DNA-protein interaction implicated in singular *var* gene choice in *Plasmodium falciparum*', *Cellular Microbiology*, 14(12): 1836–1848.

Brehove, M., Wang, T., North, J., Luo, Y., Dreher, S. J., Shimko, J. C., Ottesen, J. J., Luger, K. and Poirier, M. G. (2015) 'Histone core phosphorylation regulates DNA accessibility', *Journal of Biological Chemistry*, 290(37): 22612–22621.

Brick, K., Watanabe, J., Pizzi, E., Hakimi, M., Roch, K., Zhou, Y., Noble, W. and Roch, K. (2008) 'Core promoters are predicted by their distinct physicochemical properties in the genome of *Plasmodium falciparum*', *Genome Biology* l, 9(12): R178. doi: 10.1186/gb-2008-9-12-r178.

Brind'Amour, J., Liu, S., Hudson, M., Chen, C., Karimi, M. M. and Lorincz, M. C. (2015) 'An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations', *Nature Communications*, 6(1): 6033. doi: 10.1038/ncomms7033.

Briquet, S., Lawson-Hogban, N., Boisson, B., Soares, M. P., Péronet, R., Smith, L., Ménard, R., Huerre, M., Mécheri, S. and Vaquero, C. (2015) ' Disruption of Parasite hmgb2 Gene Attenuates *Plasmodium berghei* ANKA Pathogenicity ', *Infection and Immunity*. doi: 10.1128/iai.03129-14.

Broadbent, K. M., Broadbent, J. C., Ribacke, U., Wirth, D., Rinn, J. L. and Sabeti, P. C. (2015) 'Strand-specific RNA sequencing in *Plasmodium falciparum* malaria identifies developmentally regulated long non-coding RNA and circular RNA', *BMC Genomics*, 16(1). doi: 10.1186/s12864-015-1603-4.

Brown, K. M., Long, S. and Sibley, L. D. (2017) 'Plasma membrane association by N- acylation governs PKG function in Toxoplasma gondii', *mBio*, 8(3). doi: 10.1128/mBio.00375-17.

Bruce, M. C., Alano, P., Duthie, S. and Carter, R. (1990) 'Commitment of the malaria parasite *Plasmodium falciparum* to sexual and asexual development', *Parasitology*, 100(02): 191–200.

Bucci, L. R., Brock, W. A. and Meistrich, M. L. (1982) 'Distribution and synthesis of histone 1 subfractions during spermatogenesis in the rat', *Experimental Cell Research*. doi: 10.1016/0014-4827(82)90162-8.

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. and Greenleaf, W. J. (2013) 'Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position', *Nature Meth*ods, 10(12): 1213–1218.

Bunnik, E. M., Batugedara, G., Saraf, A., Prudhomme, J., Florens, L. and Le Roch, K. G. (2016) 'The mRNA-bound proteome of the human malaria parasite *Plasmodium falciparum*', *Genome Biology*, 17(1). doi: 10.1186/s13059-016-1014-0.

Bunnik, E. M., Chung, D.-W. D., Hamilton, M., Ponts, N., Saraf, A., Prudhomme, J., Florens, L., Le Roch, K. G., Doug Chung, D.-W., Hamilton, M., Ponts, N., Saraf, A., Prudhomme, J., Florens, L., Le Roch, K. G., Chung, D.-W. D., Hamilton, M., Ponts, N., Saraf, A., Prudhomme, J., Florens, L. and Le Roch, K. G. (2013) 'Polysome profiling reveals translational control of gene expression in the human malaria parasite *Plasmodium falciparum*.', *Genome Biology*, 14(11): R128. doi: 10.1186/gb-2013-14-11-r128.

Bunnik, E. M., Polishko, A., Prudhomme, J., Ponts, N., Gill, S. S., Lonardi, S. and Le Roch, K. G. (2014) 'DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite *Plasmodium falciparum*.', *BMC genomics*, 15(1): 347. doi: 10.1186/1471-2164-15-347.

Burda, P.-C., Roelli, M. A., Schaffner, M., Khan, S. M., Janse, C. J. and Heussler, V. T. (2015) 'A *Plasmodium* Phospholipase Is Involved in Disruption of the Liver Stage Parasitophorous Vacuole Membrane', *PLOS Pathogens*, 11(3): e1004760. doi: 10.1371/journal.ppat.1004760.

Bushell, E., Gomes, A. R., Sanderson, T., Anar, B., Girling, G., Herd, C., Metcalf, T., Modrzynska, K., Schwach, F., Martin, R. E., Mather, M. W., McFadden, G. I., Parts, L., Rutledge, G. G., Vaidya, A. B., Wengelnik, K., Rayner, J. C. and Billker, O. (2017) 'Functional Profiling of a *Plasmodium* Genome Reveals an Abundance of Essential Genes', *Cell*, 170(2): 260–272.

Bzik, D. J., Li, W.-B., Horii, T. and Inselburgt, J. (1987) 'Molecular cloning and sequence analysis of the *Plasmodium falciparum* dihydrofolate reductase-thymidylate synthase gene', *Biochemistry*, 84: 8360–8364.

Cabral, F. J., Fotoran, W. L. and Wunderlich, G. (2012) 'Dynamic activation and repression of the *Plasmodium falciparum rif* gene family and their relation

to chromatin modification', *PLoS ONE*, 7(1). doi: 10.1371/journal.pone.0029881.

Calderwood, M. S., Gannoun-Zaki, L., Wellems, T. E. and Deitsch, K. W. (2003) '*Plasmodium falciparum var* Genes Are Regulated by Two Regions with Separate Promoters, One Upstream of the Coding Region and a Second within the Intron', *Journal of Biological Chemistry*, 278(36): 34125–34132.

Calero, G., Wilson, K. F., Ly, T., Rios-Steiner, J. L., Clardy, J. C. and Cerione, R. A. (2002) 'Structural basis of m7GpppG binding to the nuclear cap-binding protein complex', *Nature Structural Biology*, 9(12): 912–917.

Campos, E. I. and Reinberg, D. (2009) 'Histones: Annotating Chromatin', *Annual Review of Genetics*. doi: 10.1146/annurev.genet.032608.103928.

Canning, E. U. and Sinden, R. E. (1973) 'The organization of the ookinete and observations on nuclear division in oocysts of *Plasmodium berghei*.', *Parasitology*, 67(1): 29–40.

Carey, M. F., Peterson, C. L. and Smale, S. T. (2009) 'Chromatin immunoprecipitation (ChIP).', *Cold Spring Harbor protocols*, 2009(9): pdb.prot5279. doi: 10.1101/pdb.prot5279.

Carter, R. and Miller, L. H. (1979) 'Evidence for environmental modulation of gametocytogenesis in *Plasmodium falciparum* in continuous culture.', *Bulletin of the World Health Organization*. World Health Organization, 57 Suppl 1(Suppl), pp. 37–52.

Carter, V., Nacer, A. M. L., Underhill, A., Sinden, R. E. and Hurd, H. (2007) 'Minimum requirements for ookinete to oocyst transformation in *Plasmodium*.', *International journal for parasitology*, 37(11): 1221–32.

Carvalho, T. G., Thiberge, S., Sakamoto, H. and Ménard, R. (2004) 'Conditional mutagenesis using site-specific recombination in *Plasmodium berghei*.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(41): 14931–6.

Carver, T., Harris, S. R., Berriman, M., Parkhill, J. and McQuillan, J. A. (2012) 'Artemis: An integrated platform for visualization and analysis of high-throughput sequence- based experimental data', *Bioinformatics*. doi: 10.1093/bioinformatics/btr703.

Casadio, F., Lu, X., Pollock, S. B., LeRoy, G., Garcia, B. A., Muir, T. W., Roeder, R. G. and Allis, C. D. (2013) 'H3R42me2a is a histone modification with positive transcriptional effects', *Proceedings of the National Academy of Sciences*, 110(37): 14894–14899.

Catherman, A. D., Skinner, O. S. and Kelleher, N. L. (2014) 'Top Down proteomics: Facts and perspectives', *Biochemical and Biophysical Research Communications*, pp. 683–693. doi: 10.1016/j.bbrc.2014.02.041.

Centers for Disease Control and Prevention (CDC).*DPDx Laboratory Diagnosis of Malaria: Plasmodium malariae.* Available at https://www.cdc.gov/dpdx/resources/pdf/benchaids/malaria/pmalariae_bench aidv2.pdf [Accessed 20 July 2017].

Di Cerbo, V., Mohn, F., Ryan, D. P., Montellier, E., Kacem, S., Tropberger, P., Kallis, E., Holzner, M., Hoerner, L., Feldmann, A., Richter, F. M., Bannister, A. J., Mittler, G., Michaelis, J., Khochbin, S., Feil, R., Schuebeler, D., Owen-Hughes, T., Daujat, S. and Schneider, R. (2014) 'Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription', *eLife*. doi: 10.7554/elife.01632.

Chaal, B. K., Gupta, A. P., Wastuwidyaningtyas, B. D., Luah, Y.-H. and Bozdech, Z. (2010) 'Histone deacetylases play a major role in the transcriptional regulation of the *Plasmodium falciparum* life cycle.', *PLoS pathogens*, 6(1): e1000737. doi: 10.1371/journal.ppat.1000737.

Chandra, B. R., Olivieri, A., Silvestrini, F., Alano, P. and Sharma, A. (2005) 'Biochemical characterization of the two nucleosome assembly proteins from *Plasmodium falciparum'*, *Molecular and Biochemical Parasitology*. doi: 10.1016/j.molbiopara.2005.04.006.

Chang, F. T. M., Chan, F. L., McGhie, J. D. R., Udugama, M., Mayne, L., Collas, P., Mann, J. R. and Wong, L. H. (2015) 'CHK1-driven histone H3.3 serine 31 phosphorylation is important for chromatin maintenance and cell survival in human ALT cancer cells', *Nucleic Acids Research*. doi: 10.1093/nar/gkv104.

Chatterjee, N., North, J. A., Dechassa, M. L., Manohar, M., Prasad, R., Luger, K., Ottesen, J. J., Poirier, M. G. and Bartholomew, B. (2015) 'Histone Acetylation near the Nucleosome Dyad Axis Enhances Nucleosome Disassembly by RSC and SWI/SNF', *Molecular and Cellular Biology*, 35(23): 4083–4092.

Chen, P. B., Ding, S., Zanghì, G., Soulard, V., Dimaggio, P. A., Fuchter, M. J., Mecheri, S., Mazier, D., Scherf, A. and Malmquist, N. A. (2016) '*Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites', *Scientific reports*, 6: 21802. doi: 10.1038/srep21802.

Chen, Z. and Riggs, A. D. (2011) 'DNA methylation and demethylation in mammals.', *The Journal of biological chemistry*, 286(21): 18347–53.

Chêne, A., Vembar, S. S., Rivière, L., Lopez-Rubio, J. J., Claes, A., Siegel, T. N., Sakamoto, H., Scheidig-Benatar, C., Hernandez-Rivas, R. and Scherf, A. (2012) 'PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites', *Nucleic Acids Research*, 40(7): 3066–3077.

Cheung, W. L., Turner, F. B., Krishnamoorthy, T., Wolner, B., Ahn, S. H., Foley, M., Dorsey, J. A., Peterson, C. L., Berger, S. L. and Allis, C. D. (2005) 'Phosphorylation of histone H4 serine 1 during DNA damage requires casein kinase II in S. cerevisiae', *Current Biology*, 15(7): 656–660.

Chiu, M. I., Katz, H. and Berlin, V. (1994) 'RAPT1, a mammalian homolog of yeast Tor, interacts with the FKBP12/rapamycin complex.', *Proceedings of the National Academy of Sciences*, 91(26): 12574–12578.

Choi, S.-W., Keyes, M. K. and Horrocks, P. (2006) 'LC/ESI-MS demonstrates the absence of 5-methyl-2′-deoxycytosine in *Plasmodium falciparum* genomic DNA', *Molecular and Biochemical Parasitology*, 150(2): 350–352.

Chookajorn, T., Dzikowski, R., Frank, M., Li, F., Jiwani, A. Z., Hartl, D. L. and Deitsch, K. W. (2007) 'Epigenetic memory at malaria virulence genes.', *Proceedings of the National Academy of Sciences of the United States of America*, 104(3): 899–902.

Chua, M. J., Arnold, M. S. J., Xu, W., Lancelot, J., Lamotte, S., Späth, G. F., Prina, E., Pierce, R. J., Fairlie, D. P., Skinner-Adams, T. S. and Andrews, K. T. (2017) 'Effect of clinically approved HDAC inhibitors on *Plasmodium, Leishmania* and *Schistosoma* parasite growth', *International Journal for Parasitology: Drugs and Drug Resistance*, 7(1): 42–50.

Coetzee, N., Sidoli, S., van Biljon, R., Painter, H., Llinás, M., Garcia, B. A. and Birkholtz, L.-M. (2017) 'Quantitative chromatin proteomics reveals a dynamic histone post- translational modification landscape that defines asexual and sexual *Plasmodium falciparum* parasites.', *Scientific reports*, 7(1): 607. doi: 10.1038/s41598-017- 00687-7.

Cogswell, F. B., Collins, W. E., Krotoski, W. A. and Lowrie, A. C. (1991) 'Hypnozoites of *Plasmodium simiovale*', *American Journal of Tropical Medicine and Hygiene*, 45(2): 211–213.

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., Pradhan, S., Nelson, S. F., Pellegrini, M. and Jacobsen, S. E. (2008) 'Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning', *Nature*, 452(7184): 215–219.

Coleman, B. I., Skillman, K. M., Jiang, R. H. Y., Childs, L. M., Altenhofen, L. M., Ganter, M., Leung, Y., Goldowitz, I., Kafsack, B. F. C., Marti, M., Llinás, M., Buckee, C. O. and Duraisingh, M. T. (2014) 'A *Plasmodium falciparum* Histone Deacetylase regulates antigenic variation and gametocyte conversion', *Cell Host and Microbe*, 16(2): 177–186.

Collas, P. (2010) 'The current state of chromatin immunoprecipitation', *Molecular Biotechnology*. doi: 10.1007/s12033-009-9239-8.

Collins, C. R., Das, S., Wong, E. H., Andenmatten, N., Stallmach, R., Hackett, F., Herman, J.-P. P., Müller, S., Meissner, M. and Blackman, M. J. (2013) 'Robust inducible Cre recombinase activity in the human malaria parasite *Plasmodium falciparum* enables efficient gene deletion within a single asexual erythrocytic growth cycle', *Molecular microbiology*, 88(4): 687–701.

Conway, J. R., Lex, A. and Gehlenborg, N. (2017) 'UpSetR: an R package for the visualization of intersecting sets and their properties', *Bioinformatics*, 33(18): 2938–2940.

Coppi, A., Natarajan, R., Pradel, G., Bennett, B. L., James, E. R., Roggero, M. A., Corradin, G., Persson, C., Tewari, R. and Sinnis, P. (2011) 'The malaria circumsporozoite protein has two functional domains, each with distinct roles as sporozoites journey from mosquito to mammalian host.', *The Journal of experimental medicine*, 208(2): 341–56.

Coulson, R. M. R., Hall, N. and Ouzounis, C. A. (2004) 'Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*.', *Genome research*, 14(8): 1548–54.

Couture, J. F., Collazo, E., Brunzelle, J. S. and Trievel, R. C. (2005) 'Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase', *Genes and Development*, 19(12): 1455–1465.

Cowman, A. F. and Crabb, B. S. (2006) 'Invasion of Red Blood Cells by Malaria Parasites', *Cell*, 124(4): 755–766.

Cowman, A. F., Morry, M. J., Biggs, B. A., Cross, G. A. and Foote, S. J. (1988) 'Amino acid changes linked to pyrimethamine resistance in the dihydrofolate reductase- thymidylate synthase gene of *Plasmodium falciparum*.', *Proceedings of the National Academy of Sciences of the United States of America*, 85(23): 9109–13.

Cox, F. E. G. (2010) 'History of the discovery of the malaria parasites and their vectors', *Parasites & Vectors*, 3(1), p. 5. doi: 10.1186/1756-3305-3-5.

Crabb, B. S., Cooke, B. M., Reeder, J. C., Waller, R. F., Caruana, S. R., Davern, K. M., Wickham, M. E., Brown, G. V, Coppel, R. L. and Cowman, A. F. (1997) 'Targeted gene disruption shows that knobs enable malaria-infected red cells to cytoadhere under physiological shear stress.', *Cell*, 89(2): 287–96.

Craig, A. G., Grau, G. E., Janse, C., Kazura, J. W., Milner, D., Barnwell, J. W., Turner, G. and Langhorne, J. (no date) 'The Role of Animal Models for Research on Severe Malaria'. doi: 10.1371/journal.ppat.1002401.

Cui, L., Fan, Q., Cui, L. and Miao, J. (2008) 'Histone lysine methyltransferases and demethylases in *Plasmodium falciparum*', *International Journal for Parasitology*, 38(10): 1083–1097.

Cui, L., Miao, J., Furuya, T., Li, X., Su, X. Z. and Cui, L. (2007) 'PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in *Plasmodium falciparum*', *Eukaryotic Cell*, 6(7): 1219–1227.

Daher, W., Pierce, R. and Khalife, J. (2007) 'Census, molecular characterization and developmental expression of Leucine-Rich-Repeat proteins in *Plasmodium falciparum*', *Molecular and Biochemical Parasitology*. doi: 10.1016/j.molbiopara.2007.07.001.

Darkin-Rattray, S. J., Gurnett, A. M., Myers, R. W., Dulski, P. M., Crumley, T. M., Allocco, J. J., Cannova, C., Meinke, P. T., Colletti, S. L., Bednarek, M. A., Singh, S. B., Goetz, M. A., Dombrowski, A. W., Polishook, J. D. and Schmatz, D. M. (1996) 'Apicidin: A novel antiprotozoal agent that inhibits

parasite histone deacetylase', *Proceedings of the National Academy of Sciences of the United States of America*, 93(23): 13143–13147.

Dastidar, E. G., Dayer, G., Holland, Z. M., Dorin-Semblat, D., Claes, A., Chêne, A., Sharma, A., Hamelin, R., Moniatte, M., Lopez-Rubio, J. J., Scherf, A. and Doerig, C. (2012) 'Involvement of *Plasmodium falciparum* protein kinase CK2 in the chromatin assembly pathway', *BMC Biology*, 10. doi: 10.1186/1741-7007-10-5.

Dastidar, E. G., Dzeyk, K., Krijgsveld, J., Malmquist, N. A., Doerig, C., Scherf, A. and Lopez-Rubio, J. J. (2013) 'Comprehensive Histone Phosphorylation Analysis and Identification of Pf14-3-3 Protein as a Histone H3 Phosphorylation Reader in Malaria Parasites', *PLoS ONE*, 8(1): e53179.

Dawson, M. A., Bannister, A. J., Göttgens, B., Foster, S. D., Bartke, T., Green, A. R. and Kouzarides, T. (2009) 'JAK2 phosphorylates histone H3Y41 and excludes HP1α from chromatin', *Nature*, 461(7265): 819–822.

Dawson, M. A., Foster, S. D., Bannister, A. J., Robson, S. C., Hannah, R., Wang, X., Xhemalce, B., Wood, A. D., Green, A. R., Göttgens, B. and Kouzarides, T. (2012) 'Three Distinct Patterns of Histone H3Y41 Phosphorylation Mark Active Genes', *Cell Reports*. doi: 10.1016/j.celrep.2012.08.016.

Dejung, M., Subota, I., Bucerius, F., Dindar, G., Freiwald, A., Engstler, M., Boshart, M., Butter, F. and Janzen, C. J. (2016) 'Quantitative Proteomics Uncovers Novel Factors Involved in Developmental Differentiation of *Trypanosoma brucei*', *PLoS Pathogens*, 12(2). doi: 10.1371/journal.ppat.1005439.

Delrieu, I., Waller, C. C., Mota, M. M., Grainger, M., Langhorne, J. and Holder, A. A. (2002) 'PSLAP, a protein with multiple adhesive motifs, is expressed in *Plasmodium falciparum* gametocytes', *Molecular and Biochemical Parasitology*, 121(1): 11–20.

Deng, X., Lu, T., Wang, L., Gu, L., Sun, J., Kong, X., Liu, C. and Cao, X. (2016) 'Recruitment of the NineTeen Complex to the activated spliceosome requires AtPRMT5', *Proceedings of the National Academy of Sciences*, 113(19): 5447–5452.

Deshmukh, A. S., Srivastava, S. and Dhar, S. K. (2013) '*Plasmodium falciparum*: Epigenetic Control of *var* Gene Regulation and Disease', *Subcellular Biochemistry*, pp. 659–682. doi: 10.1007/978-94-007-4525-4_28.

Devaiah, B. N., Case-Borden, C., Gegonne, A., Hsu, C. H., Chen, Q., Meerzaman, D., Dey, A., Ozato, K. and Singer, D. S. (2016) 'BRD4 is a histone acetyltransferase that evicts nucleosomes from chromatin', *Nature Structural and Molecular Biology*, 23(6): 540–548.

Dey, K. K., Hsiao, C. J. and Stephens, M. (2017) 'Visualizing the structure of RNA-seq expression data using grade of membership models', *PLoS Genetics*. doi: 10.1371/journal.pgen.1006599.

van Dijk, M. R., Janse, C. J., Thompson, J., Waters, A. P., Braks, J. A. M., Dodemont, H. J., Stunnenberg, H. G., van Gemert, G.-J., Sauerwein, R. W. and Eling, W. (2001) 'A Central Role for P48/45 in Malaria Parasite Male Gamete Fertility', *Cell*, 104(1): 153–164.

van Dijk, M. R., van Schaijk, B. C. L., Khan, S. M., van Dooren, M. W., Ramesar, J., Kaczanowski, S., van Gemert, G.-J., Kroeze, H., Stunnenberg, H. G., Eling, W. M., Sauerwein, R. W., Waters, A. P. and Janse, C. J. (2010) 'Three members of the 6-cys protein family of *Plasmodium* play a role in gamete fertility.', *PLoS pathogens*, 6(4), p. e1000853. doi: 10.1371/journal.ppat.1000853.

van Dijk, M. R., Waters, A. P. and Janse, C. J. (1995) 'Stable transfection of malaria parasite blood stages.', *Science*, 268(5215): 1358–62.

Dixon, M. W. A., Dearnley, M. K., Hanssen, E., Gilberger, T. and Tilley, L. (2012) 'Shape- shifting gametocytes: how and why does *P. falciparum* go banana-shaped?', *Trends in Parasitology*, 28(11): 471–478.

Donald, R. G. K. and Roos, D. S. (1993) 'Stable molecular transformation of *Toxoplasma gondii*: A selectable dihydrofolate reductase-thymidylate synthase marker based on drug-resistance mutations in malaria', *Genetics*, 9024): 11703–11707.

Dondorp, A. M., Kager, P. A., Vreeken, J. and White, N. J. (2000) 'Abnormal Blood Flow and Red Blood Cell Deformability in Severe Malaria', *Parasitology Today*, 16(6): 228–232.

Draker, R., Ng, M. K., Sarcinella, E., Ignatchenko, V., Kislinger, T. and Cheung, P. (2012) 'A Combination of H2A.Z and H4 Acetylation Recruits Brd2 to Chromatin during Transcriptional Activation', *PLoS Genetics*, 8(11). doi: 10.1371/journal.pgen.1003047.

Dreher, K. A. (2006) 'The *Arabidopsis* Aux/IAA Protein Family Has Diversified in Degradation and Auxin Responsiveness', *The Plant Cell Online*, 18(3): 699–714.

Duffy, M. F., Selvarajah, S. A., Josling, G. A. and Petter, M. (2014) 'Epigenetic regulation of the *Plasmodium falciparum* genome', *Briefings in Functional Genomics*, 13(3), pp. 203–216.

Duraisingh, M. T. and Horn, D. (2016) 'Epigenetic Regulation of Virulence Gene Expression in Parasitic Protozoa.', *Cell host & microbe*, 19(5): 629–40.

Duraisingh, M. T., Voss, T. S., Marty, A. J., Duffy, M. F., Good, R. T., Thompson, J. K., Freitas, L. H., Scherf, A., Crabb, B. S. and Cowman, A. F. (2005) 'Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*', *Cell*, 121(1): 13–24.

Duval, L., Fourment, M., Nerrienet, E., Rousset, D., Sadeuh, S. A., Goodman, S. M., Andriaholinirina, N. V., Randrianarivelojosia, M., Paul, R. E., Robert, V., Ayala, F. J. and Ariey, F. (2010) 'African apes as reservoirs of

Plasmodium falciparum and the origin and diversification of the *Laverania* subgenus.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(23): 10561– 10566.

Dvorin, J. D., Martyn, D. C., Patel, S. D., Grimley, J. S., Collins, C. R., Hopp, C. S., Bright, A. T., Westenberger, S., Winzeler, E., Blackman, M. J., Baker, D. A., Wandless, T. J. and Duraisingh, M. T. (2010) 'A plant-like kinase in *Plasmodium falciparum* regulates parasite egress from erythrocytes', *Science*, 328(5980): 910- 912.

Ebine, K., Hirai, M., Sakaguchi, M., Yahata, K., Kaneko, O. and Saito-Nakano, Y. (2016) '*Plasmodium* Rab5b is secreted to the cytoplasmic face of the tubovesicular network in infected red blood cells together with N-acylated adenylate kinase 2', *Malaria Journal*, 15(1). doi: 10.1186/s12936-016-1377-4.

Eckschlager, T., Plch, J., Stiborova, M. and Hrabeta, J. (2017) 'Histone deacetylase inhibitors as anticancer drugs', *International Journal of Molecular Sciences*. doi: 10.3390/ijms18071414.

Egelhofer, T. A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A. A., Cheung, M. S., Day, D. S., Gadel, S., Gorchakov, A. A., Gu, T., Kharchenko, P. V., Kuan, S., Latorre, I., Linder-Basso, D., Luu, Y., Ngo, Q., Perry, M., Rechtsteiner, A., Riddle, N. C., Schwartz, Y. B., Shanower, G. A., Vielle, A., Ahringer, J., Elgin, S. C. R., Kuroda, M. I., Pirrotta, V., Ren, B., Strome, S., Park, P. J., Karpen, G. H., Hawkins, R. D. and Lieb, J. D. (2011) 'An assessment of histone-modification antibody quality', *Nature Structural and Molecular Biology*, 18(1): 91–94.

Eichner, M., Diebner, H. H., Molineaux, L., Collins, W. E., Jeffery, G. M. and Dietz, K. (2001) 'Genesis, sequestration and survival of *Plasmodium falciparum* gametocytes: parameter estimates from fitting a model to malariatherapy data', *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 95(5): 497– 501.

Escalante, A. A. and Ayala, F. J. (1995) 'Evolutionary origin of Plasmodium and other Apicomplexa based on rRNA genes', *Evolution*, 92: 5793–5797.

Evertts, A. G., Manning, A. L., Wang, X., Dyson, N. J., Garcia, B. A. and Coller, H. A. (2013) 'H4K20 methylation regulates quiescence and chromatin compaction', *Molecular Biology of the Cell*. doi: 10.1091/mbc.e12-07-0529.

Falae, A., Combe, A., Amaladoss, A., Carvalho, T., Menard, R. and Bhanot, P. (2010) 'Role of *Plasmodium berghei* cGMP-dependent protein kinase in late liver stage development.', *The Journal of biological chemistry*, 285(5): 3282–8.

Fan, Q., An, L. and Cui, L. (2004) '*Plasmodium falciparum* histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling', *Eukaryotic Cell*, 3(2): 264–276.

Fang, H., Klages, N., Baechler, B., Hillner, E., Yu, L., Pardo, M., Choudhary, J. and Brochet, M. (2017) 'Multiple short windows of calcium-dependent protein kinase 4 activity coordinate distinct cell cycle events during *Plasmodium* gametogenesis', *eLife*, 6. doi: 10.7554/eLife.26524.

Farooq, Z., Banday, S., Pandita, T. K. and Altaf, M. (2016) 'The many faces of histone H3K79 methylation', *Mutation Research - Reviews in Mutation Research*, pp. 46– 52. doi: 10.1016/j.mrrev.2016.03.005.

Fast, N. M., Kissinger, J. C., Roos, D. S. and Keeling, P. J. (2001) 'Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids.', *Molecular biology and evolution*, 18(3): 418–426.

Feng, Y., Wang, J., Asher, S., Hoang, L., Guardiani, C., Ivanov, I. and Zheng, Y. G. (2011) 'Histone H4 acetylation differentially modulates arginine methylation by an in Cis mechanism.', *The Journal of biological chemistry*, 286(23): 20323–34.

Fennell, C., Babbitt, S., Russo, I., Wilkes, J., Ranford-Cartwright, L., Goldberg, D. E. and Doerig, C. (2009) 'PfeIK1, a eukaryotic initiation factor 2alpha kinase of the human malaria parasite *Plasmodium falciparum*, regulates stress-response to amino-acid starvation.', *Malaria journal*, 8, p. 99. doi: 10.1186/1475-2875-8-99.

Ferreira, H., Somers, J., Webster, R., Flaus, A. and Owen-Hughes, T. (2007) 'Histone Tails and the H3 N Helix Regulate Nucleosome Mobility and Stability', *Molecular and Cellular Biology*. doi: 10.1128/mcb.02229-06.

Filarsky, M., Fraschka, S. A., Niederwieser, I., Brancucci, N. M. B. B., Carrington, E., Carrió, E., Moes, S., Jenoe, P., Bártfai, R. and Voss, T. S. (2018) 'GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing', *Science*, 359(6381): 1259-1263.

Fingerman, I. M., Du, H. N. and Briggs, S. D. (2008) 'In vitro histone methyltransferase assay', *Cold Spring Harbor Protocols*. doi: 10.1101/pdb.prot4939.

Flannery, E. L., Chatterjee, A. K. and Winzeler, E. A. (2013) 'Antimalarial drug discovery - approaches and progress towards new medicines.', *Nature reviews. Microbiology*, 11(12): 849–62.

Flensburg, C., Kinkel, S. A., Keniry, A., Blewitt, M. E. and Oshlack, A. (2014) 'A comparison of control samples for ChIP-seq of histone modifications.', *Frontiers in genetics*. Frontiers Media SA, 5, p. 329. doi: 10.3389/fgene.2014.00329.

Flueck, C., Bartfai, R., Niederwieser, I., Witmer, K., Alako, B. T. F., Moes, S., Bozdech, Z., Jenoe, P., Stunnenberg, H. G. and Voss, T. S. (2010) 'A major role for the *Plasmodium falciparum* ApiAP2 protein PfSIP2 in chromosome end biology', *PLoS Pathogens*, 6(2). doi: 10.1371/journal.ppat.1000784.

Fonager, J., Franke-Fayard, B. M., Adams, J. H., Ramesar, J., Klop, O., Khan, S. M., Janse, C. J. and Waters, A. P. (2011) 'Development of the *piggyBac* transposable system for *Plasmodium berghei* and its application for random mutagenesis in malaria parasites', *BMC Genomics*, 12(1), p. 155. doi: 10.1186/1471-2164-12-155.

Foote, S. J., Galatis, D. and Cowman, A. F. (1990) 'Amino acids in the dihydrofolate reductase-thymidylate synthase gene of *Plasmodium falciparum* involved in cycloguanil resistance differ from those involved in pyrimethamine resistance', *Proceedings of the National Academy of Sciences of the United States of America*, 87(8): 3014–3017.

Foth, B. J., Zhang, N., Chaal, B. K., Sze, S. K., Preiser, P. R. and Bozdech, Z. (2011) 'Quantitative time-course profiling of parasite and host cell proteins in the human malaria parasite *Plasmodium falciparum*.', *Molecular & cellular proteomics*, 10(8), p. M110.006411. doi: 10.1074/mcp.M110.006411.

Fougère, A., Jackson, A. P., Paraskevi Bechtsi, D., Braks, J. A. M., Annoura, T., Fonager, J., Spaccapelo, R., Ramesar, J., Chevalley-Maurel, S., Klop, O., van der Laan, A. M. A., Tanke, H. J., Kocken, C. H. M., Pasini, E. M., Khan, S. M., Böhme, U., van Ooij, C., Otto, T. D., Janse, C. J. and Franke-Fayard, B. (2016) 'Variant Exported Blood-Stage Proteins Encoded by *Plasmodium* Multigene Families Are Expressed in Liver Stages Where They Are Exported into the Parasitophorous Vacuole', *PLoS Pathogens*, 12(11). doi: 10.1371/journal.ppat.1005917.

Frank, M., Dzikowski, R., Amulic, B. and Deitsch, K. (2007) 'Variable switching rates of malaria virulence genes are associated with chromosomal position', *Molecular Microbiology*, 64(6): 1486–1498.

Franke-Fayard, B., Trueman, H., Ramesar, J., Mendoza, J., Van Der Keur, M., Van Der Linden, R., Sinden, R. E., Waters, A. P. and Janse, C. J. (2004) 'A *Plasmodium berghei* reference line that constitutively expresses GFP at a high level throughout the complete life cycle', *Molecular and Biochemical Parasitology*, 137(1): 23–33.

Fraschka, S. A., Filarsky, M., Hoo, R., Niederwieser, I., Yam, X. Y., Brancucci, N. M. B., Mohring, F., Mushunje, A. T., Huang, X., Christensen, P. R., Nosten, F., Bozdech, Z., Russell, B., Moon, R. W., Marti, M., Preiser, P. R., Bártfai, R. and Voss, T. S. (2018) 'Comparative Heterochromatin Profiling Reveals Conserved and Unique Epigenome Signatures Linked to Adaptation and Development of Malaria Parasites', *Cell Host & Microbe*, 23(3): 407–420.

Fraschka, S. A. K., Henderson, R. W. M. and Bártfai, R. (2016) 'H3.3 demarcates GC-rich coding and subtelomeric regions and serves as potential memory mark for virulence gene expression in *Plasmodium falciparum*', *Scientific Reports*. doi:10.1038/srep31965.

Freitas-Junior, L. H., Bottius, E., Pirrit, L. A., Deitsch, K. W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T. E. and Scherf, A. (2000) 'Frequent ectopic

recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*', *Nature*, 407(6807): 1018–1022.

Frischknecht, F., Baldacci, P., Martin, B., Zimmer, C., Thiberge, S., Olivo-Marin, J.-C., Shorte, S. L. and Ménard, R. (2004) 'Imaging movement of malaria parasites during transmission by *Anopheles* mosquitoes', *Cellular Microbiology*, 6(7): 687–694.

Funes, S., Davidson, E., Reyes-Prieto, A., Magallón, S., Herion, P., King, M. P. and González-Halphen, D. (2002) 'A Green Algal Apicoplast Ancestor', *Science,* 298(5601): 2155.

Furey, T. S. (2012) 'ChIP-seq and beyond: New and improved methodologies to detect and characterize protein-DNA interactions', *Nature Reviews Genetics*, pp. 840– 852. doi: 10.1038/nrg3306.

Fussner, E., Ching, R. W. and Bazett-Jones, D. P. (2011) 'Living without 30 nm chromatin fibers', *Trends in Biochemical Sciences*, 36(1): 1–6.

Gallo, S. and Manfrini, N. (2015) 'Working hard at the nexus between cell signaling and the ribosomal machinery: An insight into the roles of RACK1 in translational regulation', *Translation*, 3(2), p. e1120382. doi: 10.1080/21690731.2015.1120382.

Gana, R., Rao, S., Huang, H., Wu, C. and Vasudevan, S. (2013) 'Structural and functional studies of S-adenosyl-L-methionine binding proteins: A ligand-centric approach', *BMC Structural Biology*, 13(1). doi: 10.1186/1472-6807-13-6.

Garcia, B. A., Mollah, S., Ueberheide, B. M., Busby, S. A., Muratore, T. L., Shabanowitz, J. and Hunt, D. F. (2007) 'Chemical derivatization of histones for facilitated analysis by mass spectrometry', *Nature Protocols*, 2(4): 933–938.

Garnham, P.C.C (1966) Malaria Parasites and Other Haemosporidia Blackwell Scientific Publications, Oxford, United Kingdom.

Garten, M., Nasamu, A. S., Niles, J. C., Zimmerberg, J., Goldberg, D. E. and Beck, J. R. (2018) 'EXP2 is a nutrient-permeable channel in the vacuolar membrane of *Plasmodium* and is essential for protein export via PTEX', *Nature Microbiology*. doi: 10.1038/s41564-018-0222-7.

Gates, L. A., Shi, J., Rohira, A. D., Feng, Q., Zhu, B., Bedford, M. T., Sagum, C. A., Jung, S. Y., Qin, J., Tsai, M. J., Tsai, S. Y., Li, W., Foulds, C. E. and O'Malley, B. W. (2017) 'Acetylation on histone H3 lysine 9 mediates a switch from transcription initiation to elongation', *Journal of Biological Chemistry*, 292(35): 14456–14472.

Gbotosho, G. O., Sowunmi, A., Happi, C. T. and Okuboyejo, T. M. (2011) '*Plasmodium falciparum* gametocyte carriage, sex ratios and asexual parasite rates in Nigerian children before and after a treatment protocol

policy change instituting the use of artemisinin-based combination therapies', *Memórias do Instituto Oswaldo Cruz*, 106(6): 685–690.

Gerald, N., Mahajan, B. and Kumar, S. (2011) 'Mitosis in the human malaria parasite *Plasmodium falciparum*.', *Eukaryotic cell*, 10(4): 474–82.

Gething, P. W., Elyazar, I. R. F., Moyes, C. L., Smith, D. L., Battle, K. E., Guerra, C. A., Patil, A. P., Tatem, A. J., Howes, R. E., Myers, M. F., George, D. B., Horby, P., Wertheim, H. F. L., Price, R. N., Müeller, I., Baird, J. K. and Hay, S. I. (2012) 'A Long Neglected World Malaria Map: *Plasmodium vivax* Endemicity in 2010', *PLoS Neglected Tropical Diseases*, 6(9), p. e1814. doi: 10.1371/journal.pntd.0001814.

Gilfillan, G. D., Hughes, T., Sheng, Y., Hjorthaug, H. S., Straub, T., Gervin, K., Harris, J. R., Undlien, D. E. and Lyle, R. (2012) 'Limitations and possibilities of low cell number ChIP-seq.', *BMC genomics*, 13, p. 645. doi: 10.1186/1471-2164-13-645.

Gill, J., Kumar, A., Yogavel, M., Belrhali, H., Jain, S. K., Rug, M., Brown, M., Maier, A. G. and Sharma, A. (2010) 'Structure, localization and histone binding properties of nuclear-associated nucleosome assembly protein from *Plasmodium falciparum*.', *Malaria journal*, 9, p. 90. doi: 10.1186/1475-2875-9-90.

Gilson, P . R. and Crabb, B. S. (2009) 'Morphology and kinetics of the three distinct phases of red blood cell invasion by *Plasmodium falciparum* merozoites', *International Journal for Parasitology*, 39(1): 91–96.

Gissot, M., Briquet, S., Refour, P., Boschet, C. and Vaquero, C. (2005) 'PfMyb1, a *Plasmodium falciparum* transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation', *Journal of Molecular Biology*, 346(1): 29–42.

Gissot, M., Ting, L. M., Daly, T. M., Bergman, L. W., Sinnis, P. and Kim, K. (2008) 'High mobility group protein HMGB2 is a critical regulator of *Plasmodium* oocyst development', *Journal of Biological Chemistry*, 283(25): 17030–17038.

Godiska, R., Mead, D., Dhodda, V., Wu, C., Hochstein, R., Karsi, A., Usdin, K., Entezam, A. and Ravin, N. (2009) 'Linear plasmid vector for cloning of repetitive or unstable sequences in *Escherichia coli*', *Nucleic Acids Research*, 38(6). doi: 10.1093/nar/gkp1181.

Goldberg, D. E., Slater, A. F., Cerami, A. and Henderson, G. B. (1990) 'Hemoglobin degradation in the malaria parasite *Plasmodium falciparum*: an ordered process in a unique organelle.', *Proceedings of the National Academy of Sciences of the United States of America*, 87(8): 2931–2935.

Gomes-Santos, C. S. S., Braks, J., Prudêncio, M., Carret, C., Gomes, A. R., Pain, A., Feltwell, T., Khan, S., Waters, A., Janse, C., Mair, G. R. and Mota, M. M. (2011) 'Transition of *Plasmodium* sporozoites into liver stage-like forms

is regulated by the RNA binding protein Pumilio', *PLoS Pathogens*, 7(5). doi: 10.1371/journal.ppat.1002046.

Gómez-Díaz, E., Rivero, A., Chandre, F. and Corces, V. G. (2014) 'Insights into the epigenomic landscape of the human malaria vector *Anopheles gambiae*.', *Frontiers in genetics*, 5, p. 277. doi: 10.3389/fgene.2014.00277.

Goonewardene, R., Daily, J., Kaslow, D., Sullivan, T. J., Duffyt, P., Carter, R., Mendist, K. and Wirth, D. (1993) 'Transfection of the malaria parasite and expression of firefly luciferase', *Biochemistry*, 90: 5234–5236.

Gordon, J. L. and Sibley, L. D. (2005) 'Comparative genome analysis reveals a conserved family of actin-like proteins in apicomplexan parasites', *BMC Genomics*, 6, p. 179. doi: 10.1186/1471-2164-6-179.

Gottschling, D. E., Aparicio, O. M., Billington, B. L. and Zakian, V. A. (1990) 'Position effect at S. cerevisiae telomeres: Reversible repression of Pol II transcription', *Cell*, 63(4): 751–762.

Govin, J., Schug, J., Krishnamoorthy, T., Dorsey, J., Khochbin, S. and Berger, S. L. (2010) 'Genome-wide mapping of histone H4 serine-1 phosphorylation during sporulation in *Saccharomyces cerevisiae'*, *Nucleic Acids Research*, 38(14): 4599–4606.

Graves, H. K., Wang, P., Lagarde, M., Chen, Z. and Tyler, J. K. (2016) 'Mutations that prevent or mimic persistent post-translational modifications of the histone H3 globular domain cause lethality and growth defects in *Drosophila*', *Epigenetics and Chromatin*, 9(1). doi: 10.1186/s13072-016-0059-3.

Grüring, C., Heiber, A., Kruse, F., Ungefehr, J., Gilberger, T.-W. and Spielmann, T. (2011) 'Development and host cell modifications of *Plasmodium falciparum* blood stages in four dimensions', *Nature Communications*, 2, p. 165. doi: 10.1038/ncomms1169.

Guerreiro, A., Deligianni, E., Santos, J. M., Silva, P. A. G. C., Louis, C., Pain, A., Janse, C. J., Franke-Fayard, B., Carret, C. K., Siden-Kiamos, I. and Mair, G. R. (2014) 'Genome-wide RIP-Chip analysis of translational repressor-bound mRNAs in the *Plasmodium* gametocyte', *Genome Biology*, 15(11). doi: 10.1186/s13059-014- 0493-0.

Gupta, A. P., Chin, W. H., Zhu, L., Mok, S., Luah, Y. H., Lim, E.-H. H. and Bozdech, Z. (2013) 'Dynamic Epigenetic Regulation of Gene Expression during the Life Cycle of Malaria Parasite *Plasmodium falciparum*', *PLoS Pathog*, 9(2), p. e1003170. doi: 10.1371/journal.ppat.1003170.

Gupta, A. P., Zhu, L., Tripathi, J., Kucharski, M., Patra, A. and Bozdech, Z. (2017) 'Histone 4 lysine 8 acetylation regulates proliferation and host–pathogen interaction in *Plasmodium falciparum*', *Epigenetics & Chromatin*, 10(1), p. 40. doi: 10.1186/s13072-017-0147-z.

Gupta, A., Thiruvengadam, G. and Desai, S. A. (2015) 'The conserved clag multigene family of malaria parasites: Essential roles in host-pathogen interaction', *Drug Resistance Updates*, 18, pp. 47–54. doi: 10.1016/j.drup.2014.10.004.

Guttery, D. S., Holder, A. A., Tewari, R., Franke-Fayard, B. and Carret, C. (2012) 'Sexual Development in *Plasmodium*: Lessons from Functional Analyses', *PLoS Pathogens*, 8(1), p. e1002404. doi: 10.1371/journal.ppat.1002404.

Hainer, S. J. and Martens, J. A. (2011) 'Identification of Histone Mutants That Are Defective for Transcription-Coupled Nucleosome Occupancy', *Molecular and Cellular Biology*, 31(17): 3557–3568.

Hake, S. B., Garcia, B. A., Kauer, M., Baker, S. P., Shabanowitz, J., Hunt, D. F. and Allis, C. D. (2005) 'Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes', *Proceedings of the National Academy of Sciences*, 102(18): 6344–6349.

Hall, N., Karras, M., Raine, J. D., Carlton, J. M., Kooij, T. W. A., Berriman, M., Florens, L., Janssen, C. S., Pain, A., Christophides, G. K., James, K., Rutherford, K., Harris, B., Harris, D., Churcher, C., Quail, M. A., Ormond, D., Doggett, J., Trueman, H. E., Mendoza, J., Bidwell, S. L., Rajandream, M.-A., Carucci, D. J., Yates, J. R., Kafatos, F. C., Janse, C. J., Barrell, B., Turner, C. M. R., Waters, A. P. and Sinden, R. E. (2005) 'A Comprehensive Survey of the *Plasmodium* Life Cycle by Genomic, Transcriptomic, and Proteomic Analyses', *Science*, 307(5706): 82–6.

Hammond, C. M., Strømme, C. B., Huang, H., Patel, D. J. and Groth, A. (2017) 'Histone chaperone networks shaping chromatin function', *Nature Reviews Molecular Cell Biology*, pp. 141–158. doi: 10.1038/nrm.2016.159.

Hansen, F. K., Sumanadasa, S. D. M., Stenzel, K., Duffy, S., Meister, S., Marek, L., Schmetter, R., Kuna, K., Hamacher, A., Mordmüller, B., Kassack, M. U., Winzeler, E. A., Avery, V. M., Andrews, K. T. and Kurz, T. (2014) 'Discovery of HDAC inhibitors with potent activity against multiple malaria parasite life cycle stages.', *European journal of medicinal chemistry*, 82, pp. 204–13. doi:10.1016/j.ejmech.2014.05.050.

Haring, M., Offermann, S., Danker, T., Horst, I., Peterhansel, C. and Stam, M. (2007) 'Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization.', *Plant methods*, 3, p. 11. doi: 10.1186/1746-4811-3-11.

Hatanaka, Y., Tsusaka, T., Shimizu, N., Morita, K., Suzuki, T., Machida, S., Satoh, M., Honda, A., Hirose, M., Kamimura, S., Ogonuki, N., Nakamura, T., Inoue, K., Hosoi, Y., Dohmae, N., Nakano, T., Kurumizaka, H., Matsumoto, K., Shinkai, Y. and Ogura, A. (2017) 'Histone H3 Methylated at Arginine 17 Is Essential for Reprogramming the Paternal Genome in Zygotes', *Cell Reports*, 20(12): 2756– 2765.

Hay, S. I., Guerra, C. A., Tatem, A. J., Noor, A. M. and Snow, R. W. (2004) 'The global distribution and population at risk of malaria: past, present, and future', *Lancet Infect Dis.*, 4(6): 327–336.

Hayakawa, T., Culleton, R., Otani, H., Horii, T. and Tanabe, K. (2008) 'Big Bang in the Evolution of Extant Malaria Parasites', *Molecular Biology and Evolution*, 25(10): 2233–2239.

Heinig, M., Colomé-Tatché, M., Taudt, A., Rintisch, C., Schafer, S., Pravenec, M., Hubner, N., Vingron, M. and Johannes, F. (2015) 'histoneHMM: Differential analysis of histone modifications with broad genomic footprints', *BMC Bioinformatics*. doi: 10.1186/s12859-015-0491-6.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H. and Glass, C. K. (2010) 'Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities.', *Molecular cell*, 38(4): 576–89.

Hirai, M., Arai, M., Mori, T., Miyagishima, S., Kawai, S., Kita, K., Kuroiwa, T., Terenius, O. and Matsuoka, H. (2008) 'Male Fertility of Malaria Parasites Is Determined by GCS1, a Plant-Type Reproduction Factor', *Current Biology*. doi: 10.1016/j.cub.2008.03.045.

Hirota, T., Lipp, J. J., Toh, B.-H. H. and Peters, J.-M. M. (2005) 'Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin', *Nature*, 438(7071):1176–1180.

Hliscs, M., Millet, C., Dixon, M. W., Siden-Kiamos, I., McMillan, P. and Tilley, L. (2015) 'Organization and function of an actin cytoskeleton in *Plasmodium falciparum* gametocytes', *Cellular Microbiology*, 17(2): 207–225.

Hoeijmakers, W. A. M., Flueck, C., Françoijs, K. J., Smits, A. H., Wetzel, J., Volz, J. C., Cowman, A. F., Voss, T., Stunnenberg, H. G. and Bártfai, R. (2012) '*Plasmodium falciparum* centromeres display a unique epigenetic makeup and cluster prior to and during schizogony', *Cellular Microbiology*, 14(9): 1391–1401.

Hoeijmakers, W. A. M., Salcedo-Amaya, A. M., Smits, A. H., Françoijs, K.-J. J., Treeck, M., Gilberger, T.-W. W., Stunnenberg, H. G. and Bártfai, R. (2013) 'H2A.Z/H2B.Z double-variant nucleosomes inhabit the AT-rich promoter regions of the *Plasmodium falciparum* genome.', *Molecular Microbiology*, 87(5): 1061–73.

Hoffman, E. A., Frey, B. L., Smith, L. M. and Auble, D. T. (2015) 'Formaldehyde crosslinking: a tool for the study of chromatin complexes.', *The Journal of biological chemistry*, 290(44): 26404–11.

Holland, Z., Prudent, R., Reiser, J.-B., Cochet, C. and Doerig, C. (2009) 'Functional analysis of protein kinase CK2 of the human malaria parasite *Plasmodium falciparum*.', *Eukaryotic cell*, 8(3): 388–397.

Hong, D., Radke, J. B. and White, M. W. (2017) 'Opposing Transcriptional Mechanisms Regulate *Toxoplasma* Development', *mSphere*, 2(1): e00347-16. doi: 10.1128/mSphere.00347-16.

Hughes, K. R. and Waters, A. P. (2017) 'Rapid inducible protein displacement in *Plasmodium in vivo* and *in vitro* using knocksideways technology.', *Wellcome open research*, 2: 18. doi: 10.12688/wellcomeopenres.11005.1.

Hyland, E. M., Cosgrove, M. S., Molina, H., Wang, D., Pandey, A., Cotter, R. J. and Boeke, J. D. (2005) 'Insights into the Role of Histone H3 and Histone H4 Core Modifiable Residues in *Saccharomyces cerevisiae*', *Molecular and Cellular Biology*. doi: 10.1128/mcb.25.24.11193.2005.

Hyllus, D., Stein, C., Schnabel, K., Schiltz, E., Imhof, A., Dou, Y., Hsieh, J. and Bauer, U. M. (2007) 'PRMT6-mediated methylation of R2 in histone H3 antagonizes H3 K4 trimethylation', *Genes and Development*, 21(24): 3369–3380.

Ikura, T., Tashiro, S., Kakino, A., Shima, H., Jacob, N., Amunugama, R., Yoder, K., Izumi, S., Kuraoka, I., Tanaka, K., Kimura, H., Ikura, M., Nishikubo, S., Ito, T., Muto, A., Miyagawa, K., Takeda, S., Fishel, R., Igarashi, K. and Kamiya, K. (2007) 'DNA Damage-Dependent Acetylation and Ubiquitination of H2AX Enhances Chromatin Dynamics', *Molecular and Cellular Biology*. doi: 10.1128/mcb.00579-07.

Invergo, B. M., Brochet, M., Yu, L., Choudhary, J., Beltrao, P., Correspondence, O. B. and Billker, O. (2017) 'Sub-minute Phosphoregulation of Cell Cycle Systems during *Plasmodium* Gamete Formation', *Cell Reports*, 21(7): 2017–2029.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J. and Mann, M. (2005) 'Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein', *Molecular & Cellular Proteomics*. doi: 10.1074/mcp.M500061-MCP200.

Iwanaga, S., Kaneko, I., Kato, T. and Yuda, M. (2012) 'Identification of an AP2-family Protein That Is Critical for Malaria Liver Stage Development', *PLoS ONE*, 7(11). doi: 10.1371/journal.pone.0047557.

Iwasaki, W., Miya, Y., Horikoshi, N., Osakabe, A., Taguchi, H., Tachiwana, H., Shibata, T., Kagawa, W. and Kurumizaka, H. (2013) 'Contribution of histone N-terminal tails to the structure and stability of nucleosomes', *FEBS Open Bio*, 3: 363–369.

Izzo, A. and Schneider, R. (2010) 'Chatting histone modifications in mammals', *Briefings in Functional Genomics*, 9(5–6): 429–443.

Jackson, K. E., Habib, S., Frugier, M., Hoen, R., Khan, S., Pham, J. S., Pouplana, L. R. de, Royo, M., Santos, M. A. S., Sharma, A. and Ralph, S. A. (2011) 'Protein translation in *Plasmodium* parasites', *Trends in Parasitology*, pp. 467–476. doi: 10.1016/j.pt.2011.05.005.

Jaijyan, D. K., Singh, H. and Singh, A. P. (2015) 'A Sporozoite- and Liver Stage-expressed Tryptophan-rich Protein Plays an Auxiliary Role in *Plasmodium* Liver Stage Development and Is a Potential Vaccine Candidate.', *The Journal of biological chemistry*, 290(32): 19496–511.

Jain, M., Koren, S., Miga, K. H., Quick, J., Rand, A. C., Sasani, T. A., Tyson, J. R., Beggs, A. D., Dilthey, A. T., Fiddes, I. T., Malla, S., Marriott, H., Nieto, T., O'Grady, J., Olsen, H. E., Pedersen, B. S., Rhie, A., Richardson, H., Quinlan, A. R., Snutch, T. P., Tee, L., Paten, B., Phillippy, A. M., Simpson, J. T., Loman, N. J. and Loose, M. (2018) 'Nanopore sequencing and assembly of a human genome with ultra-long reads', *Nature Biotechnology*, 36(4): 338–345.

Jang, S. M., Azebi, S., Soubigou, G. and Muchardt, C. (2014) 'DYRK1A phoshorylates histone H3 to differentially regulate the binding of HP1 isoforms and antagonize HP1-mediated transcriptional repression', *EMBO reports*. doi: 10.15252/embr .201338356.

Janouskovec, J., Horák, A., Oborník, M., Lukes, J. and Keeling, P. J. (2010) 'A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids.', *Proceedings of the National Academy of Sciences of the United States of America*, 107(24): 10949–54.

Janse, C. J., van der Klooster, P. F. J., van der Kaay, H. J., van der Ploeg, M. and Prosper Overdulve, J. (1986) 'DNA synthesis in *Plasmodium berghei* during asexual and sexual development', *Molecular and Biochemical Parasitology*, 20(2): 173–182.

Janse, C. J., Ramesar, J. and Waters, A. P. (2006) 'High-efficiency transfection and drug selection of genetically transformed blood stages of the rodent malaria parasite *Plasmodium berghei*', *Nat. Protocols*, 1(1): 346–356.

Janzen, C. J., Hake, S. B., Lowell, J. E. and Cross, G. A. M. (2006) 'Selective Di- or Trimethylation of Histone H3 Lysine 76 by Two DOT1 Homologs Is Important for Cell Cycle Regulation in *Trypanosoma brucei'*, *Molecular Cell*, 23(4): 497–507.

Jenuwein, T., Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z.-W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P. and Allis, C. D. (2000) 'Regulation of chromatin structure by site-specific histone H3 methyltransferases.', *Nature*, 406(6796): 593–599.

Jiang, L., Mu, J., Zhang, Q., Ni, T., Srinivasan, P., Rayavara, K., Yang, W., Turner, L., Lavstsen, T., Theander, T. G., Peng, W., Wei, G., Jing, Q., Wakabayashi, Y., Bansal, A., Luo, Y., Ribeiro, J. M. C., Scherf, A., Aravind, L., Zhu, J., Zhao, K. and Miller, L. H. (2013) 'PfSETvs methylation of histone H3K36 represses virulence genes in *Plasmodium falciparum*.', *Nature*, 499(7457): 223–7.

Jimenez-Ruiz, E., Morlon-Guyot, J., Daher, W. and Meissner, M. (2016) 'Vacuolar protein sorting mechanisms in apicomplexan parasites', *Molecular and Biochemical Parasitology*, 209(1–2): 18–25.

Jørgensen, S., Schotta, G. and Sørensen, C. S. (2013) 'Histone H4 Lysine 20 methylation: Key player in epigenetic regulation of genomic integrity', *Nucleic Acids Research*. doi: 10.1093/nar/gkt012.

Josling, G. A. and Llinas, M. (2015) 'Sexual development in *Plasmodium* parasites: knowing when it's time to commit', *Nat Rev Micro.*, 13(9): 573–587.

Josling, G. A., Williamson, K. C. and Llinás, M. (2018) 'Regulation of Sexual Commitment and Gametocytogenesis in Malaria Parasites', *Annual Review of Microbiology*. doi: 10.1146/annurev-micro-090817-062712.

Jullien, N., Goddard, I., Selmi-Ruby, S., Fina, J. L., Cremer, H. and Herman, J. P. (2007) 'Conditional transgenesis using Dimerizable Cre (DiCre)', *PLoS ONE*, 2(12). doi: 10.1371/journal.pone.0001355.

Kadekoppala, M. and Holder, A. A. (2010) 'Merozoite surface proteins of the malaria parasite: The MSP1 complex and the MSP7 family', *International Journal for Parasitology*, pp. 1155–1161. doi: 10.1016/j.ijpara.2010.04.008.

Kadota, K., Ishino, T., Matsuyama, T., Chinzei, Y. and Yuda, M. (2004) 'Essential role of membrane-attack protein in malarial transmission to mosquito host.', *Proceedings of the National Academy of Sciences of the United States of America*, 101(46),:16310–5.

Kaestli, M., Cockburn, I. a, Cortés, A., Baea, K., Rowe, J. A. and Beck, H.-P. (2006) 'Virulence of malaria is associated with differential expression of *Plasmodium falciparum var* gene subgroups in a case-control study.', *The Journal of infectious diseases*, 193(11): 1567–1574.

Kafsack, B. F. C., Rovira-Graells, N., Clark, T. G., Bancells, C., Crowley, V. M., Campino, S. G., Williams, A. E., Drought, L. G., Kwiatkowski, D. P., Baker, D. A., Cortes, A. and Llinas, M. (2014) 'A transcriptional switch underlies commitment to sexual development in malaria parasites', *Nature*, 507(7491): 248–252.

Kaneko, I., Iwanaga, S., Kato, T., Kobayashi, I. and Yuda, M. (2015) 'Genome-Wide Identification of the Target Genes of AP2-O, a *Plasmodium* AP2-Family Transcription Factor', *PLOS Pathogens*, 11(5): e1004905. doi: 10.1371/journal.ppat.1004905.

Kanyal, A., Rawat, M., Gurung, P., Choubey, D., Anamika, K. and Karmodiya, K. (2018) 'Genome-wide survey and phylogenetic analysis of histone acetyltransferases and histone deacetylases of *Plasmodium falciparum*', *The FEBS Journal*. doi: 10.1111/febs.14376.

Kappe, S. H. I., Buscaglia, C. A. and Nussenzweig, V. (2004) '*Plasmodium* Sporozoite Molecular Cell Biology', *Annual Review of Cell and Developmental Biology*, 20(1): 29–59.

Karch, K. R., Sidoli, S. and Garcia, B. A. (2016) 'Chapter One – Identification and Quantification of Histone PTMs Using High-Resolution Mass Spectrometry', in *Methods in Enzymology*, pp. 3–29. doi: 10.1016/bs.mie.2015.12.007.

Karmodiya, K., Pradhan, S. J., Joshi, B., Jangid, R., Reddy, P. C. and Galande, S. (2015) 'A comprehensive epigenome map of *Plasmodium falciparum* reveals unique mechanisms of transcriptional regulation and identifies H3K36me2 as a global mark of gene suppression.', *Epigenetics & chromatin*, 8(1), p. 32. doi: 10.1186/s13072-015-0029-1.

Kaushansky, A. and Kappe, S. H. (2015) 'Selection and refinement: the malaria parasite's infection and exploitation of host hepatocytes.', *Current opinion in microbiology*, 26: 71–8.

Kawamoto, F., Alejo-Blanco, R., Fleck, S. L., Kawamoto, Y. and Sinden, R. E. (1990) 'Possible roles of Ca2+ and cGMP as mediators of the exflagellation of *Plasmodium berghei* and *Plasmodium falciparum*', *Molecular and Biochemical Parasitology*, 42(1): 101–108.

Kehrer, J., Frischknecht, F. and Mair, G. R. (2016) 'Proteomic Analysis of the *Plasmodium berghei* Gametocyte Egressome and Vesicular bioID of Osmiophilic Body Proteins Identifies Merozoite TRAP-like Protein (MTRAP) as an Essential Factor for Parasite Transmission', *Molecular & Cellular Proteomics*, 15(9): 2852–2862.

Kensche, P. R., Hoeijmakers, W. A. M., Toenhake, C. G., Bras, M., Chappell, L., Berriman, M. and Bártfai, R. (2015) 'The nucleosome landscape of *Plasmodium falciparum* reveals chromatin architecture and dynamics of regulatory sequences', *Nucleic Acids Research*, 44(5): 2110–2124.

Kent, R. S., Modrzynska, K. K., Cameron, R., Philip, N., Billker, O. and Waters, A. P. (2018) 'Inducible developmental reprogramming redefines commitment to sexual development in the malaria parasite *Plasmodium berghei'*, *Nature Microbiology*. doi: 10.1038/s41564-018-0223-6.

Keogh, M. C., Mennella, T. A., Sawa, C., Berthelet, S., Krogan, N. J., Wolek, A., Podolny, V., Carpenter, L. R., Greenblatt, J. F., Baetz, K. and Buratowski, S. (2006) 'The *Saccharomyces cerevisiae* histone H2A variant Htz1 is acetylated by NuA4', *Genes and Development*. doi: 10.1101/gad.1388106.

Kerlin, D. H., Gatton, M. L., Roberts, J., Munro, J. and Carlton, J. (2013) 'Preferential Invasion by *Plasmodium* Merozoites and the Self-Regulation of Parasite Burden', *PLoS ONE*, 8(2): e57434. doi: 10.1371/journal.pone.0057434.

Khan, S. M., Franke-Fayard, B., Mair, G. R., Lasonder, E., Janse, C. J., Mann, M. and Waters, A. P. (2005) 'Proteome Analysis of Separated Male and Female Gametocytes Reveals Novel Sex-Specific *Plasmodium* Biology', *Cell*, 121(5): 675– 687.

Khan, S. M., Kroeze, H., Franke-Fayard, B. and Janse, C. J. (2013) 'Standardization in generating and reporting genetically modified rodent malaria parasites: The RMgmDB database', *Methods in Molecular Biology*, 923: 139–150.

Khomtchouk, B. B., Van Booven, D. J. and Wahlestedt, C. (2017) 'SUPERmerge: ChIP-Seq Coverage Island Analysis Algorithm For Broad Histone Marks', *bioRxiv*. doi: 10.1101/121897.

Kim, D., Pertea, M., Kim, D., Pertea, G. M., Leek, J. T. and Salzberg, S. L. (2016) 'Transcript-level expression analysis of RNA- seq experiments with HISAT , StringTie and Transcript-level expression analysis of RNA-seq experiments with HISAT , StringTie and Ballgown', *Nature Protocols*, 11(9): 1650–1667.

Klinker, H., Haas, C., Harrer, N., Becker, P. B. and Mueller-Planitz, F. (2014) 'Rapid purification of recombinant histones', *PLoS ONE*. doi: 10.1371/journal.pone.0104029.

Knoll, A. H., Walter, M. R., Narbonne, G. M. and Christie-Blick, N. (2004) 'A New Period for the Geologic Time Scale', *Science*, 305(5684): 621-622.

Komaki-Yasuda, K., Okuwaki, M., Nagata, K., Kawazu, S. ichiro and Kano, S. (2013) 'Identification of a Novel and Unique Transcription Factor in the Intraerythrocytic Stage of *Plasmodium falciparum*', *PLoS ONE*, 8(9). doi: 10.1371/journal.pone.0074701.

de Koning-Ward, T. F., Fidock, D. A., Thathy, V., Menard, R., van Spaendonk, R. M. L., Waters, A. P. and Janse, C. J. (2000) 'The selectable marker human dihydrofolate reductase enables sequential genetic manipulation of the *Plasmodium berghei* genome', *Molecular and Biochemical Parasitology*, 106(2): 199–212.

de Koning-Ward, T. F., Gilson, P. R. and Crabb, B. S. (2015) 'Advances in molecular genetic systems in malaria', *Nature Reviews Microbiology*, 13(6): 373–387.

de Koning-Ward, T. F., Janse, C. J. and Waters, A. P. (2000) 'The development of genetic tools for dissecting the biology of malaria parasites.', *Annual review of microbiology*, 54: 157–85.

De Koning-Ward, T. F., Olivieri, A., Bertuccini, L., Hood, A., Silvestrini, F., Charvalias, K., Berzosa Díaz, P., Camarda, G., McElwain, T. F., Papenfuss, T., Healer, J., Baldassarri, L., Crabb, B. S., Alano, P. and Ranford-Cartwright, L. C. (2007) 'The role of osmiophilic bodies and Pfg377 expression in female gametocyte emergence and mosquito infectivity in the human malaria parasite *Plasmodium falciparum*', *Molecular Microbiology*, 67(2): 278–290.

de Koning-Ward, T. F., Sperança, M. A., Waters, A. P. and Janse, C. J. (1999) 'Analysis of stage specificity of promoters in *Plasmodium berghei* using luciferase as a reporter', *Molecular and Biochemical Parasitology*, 100(1): 141–146.

de Koning-Ward, T. F., Thomas, A. W., Waters, A. P. and Janse, C. J. (1998) 'Stable expression of green fluorescent protein in blood and mosquito stages of *Plasmodium berghei*', *Molecular and Biochemical Parasitology*, 97(1–2): 247–252.

Kooij, T. W. A. A., Carlton, J. M., Bidwell, S. L., Hall, N., Ramesar, J., Janse, C. J. and Waters, A. P. (2005) 'A Plasmodium Whole-Genome Synteny Map: Indels and Synteny Breakpoints as Foci for Species-Specific Genes', *PLoS Pathogens*, 1(4): 349–361.

Kouzarides, T. (2007) 'Chromatin Modifications and Their Function', *Cell*, 128(4): 693– 705.

Kreidenweiss, A., Hopkins, A. V. and Mordmüller, B. (2013) '2A and the auxin-based degron system facilitate control of protein levels in *Plasmodium falciparum*', *PLoS ONE*, 8(11). doi: 10.1371/journal.pone.0078661.

Krishnamoorthy, T., Chen, X., Govin, J., Cheung, W. L., Dorsey, J., Schindler, K., Winter, E., Allis, C. D., Guacci, V., Khochbin, S., Fuller, M. T. and Berger, S. L. (2006) 'Phosphorylation of histone H4 Ser1 regulates sporulation in yeast and is conserved in fly and mouse spermatogenesis', *Genes and Development*, 20(18): 2580–2592.

Kuehn, A. and Pradel, G. (2010) 'The coming-out of malaria gametocytes.', *Journal of biomedicine & biotechnology*, 2010: 976827. doi: 10.1155/2010/976827.

Kühn, R. and Torres, R. M. (2002) 'Cre/loxP Recombination System and Gene Targeting', *Transgenesis Techniques: Principles and Protocols*, 180: 175–204. doi: 10.1385/1- 59259-178-7:175.

Kumar, A., Vaid, A., Syin, C. and Sharma, P. (2004) 'PfPKB, a novel protein kinase B-like enzyme from Plasmodium falciparum. I. Identification, characterization, and possible role in parasite development', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M312855200.

Kumar, K., Singal, A., Rizvi, M. M. A. and Chauhan, V. S. (2008) 'High mobility group box (HMGB) proteins of *Plasmodium falciparum*: DNA binding proteins with pro- inflammatory activity', *Parasitology International*, 57(2): 150–157.

Kyes, S. A., Christodoulou, Z., Raza, A., Horrocks, P., Pinches, R., Rowe, J. A. and Newbold, C. I. (2003) 'A well-conserved *Plasmodium falciparum var* gene shows an unusual stage-specific transcript pattern', *Molecular Microbiology*, 48(5): 1339–1348.

Lacroix, C., Giovannini, D., Combe, A., Bargieri, D. Y., Späth, S., Panchal, D., Tawk, L., Thiberge, S., Carvalho, T. G., Barale, J.-C., Bhanot, P. and Ménard, R. (2011) 'FLP/FRT-mediated conditional mutagenesis in pre-erythrocytic stages of *Plasmodium berghei*', *Nature Protocols*, 6(9): 1412.

Lal, K., Delves, M. J., Bromley, E., Wastling, J. M., Tomley, F. M. and Sinden, R. E. (2009) '*Plasmodium* male development gene-1 (mdv-1) is important for female sexual development and identifies a polarised plasma membrane during zygote development', *International Journal for Parasitology*, 39(7): 755–761.

Lamarque, M., Besteiro, S., Papoin, J., Roques, M., Vulliez-Le Normand, B., Morlon- Guyot, J., Dubremetz, J.-F., Fauquenoy, S., Tomavo, S., Faber, B. W., Kocken, C. H., Thomas, A. W., Boulanger, M. J., Bentley, G. A. and Lebrun, M. (2011) 'The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites.', *PLoS pathogens*, 7(2): e1001276. doi:10.1371/journal.ppat.1001276.

Lamb, T. J. and Langhorne, J. (2008) 'The severity of malarial anaemia in *Plasmodium chabaudi* infections of BALB/c mice is determined independently of the number of circulating parasites', *Malaria Journal*, 7(1), p. 68. doi: 10.1186/1475-2875-7- 68.

Langhorne, J., Buffet, P., Galinski, M., Good, M., Harty, J., Leroy, D., Mota, M. M., Pasini, E., Renia, L., Riley, E., Stins, M. and Duffy, P. (2011) 'The relevance of non-human primate and rodent malaria models for humans', *Malaria Journal*, 10(1): 23. doi: 10.1186/1475-2875-10-23.

Langmead, B. and Salzberg, S. L. (2012) 'Fast gapped-read alignment with Bowtie 2', *Nature Methods*, 9(4): 357–359.

Lasonder, E., Rijpma, S. R., van Schaijk, B. C. L., Hoeijmakers, W. A. M., Kensche, P. R., Gresnigt, M. S., Italiaander, A., Vos, M. W., Woestenenk, R., Bousema, T., Mair, G. R., Khan, S. M., Janse, C. J., Bártfai, R. and Sauerwein, R. W. (2016) 'Integrated transcriptomic and proteomic analyses of *P. falciparum* gametocytes: molecular insight into sex-specific processes and translational repression', *Nucleic Acids Research*, 44(13): 6087–6101.

Lee, J. H., Kang, B. H., Jang, H., Kim, T. W., Choi, J., Kwak, S., Han, J., Cho, E. J. and Youn, H. D. (2015) 'AKT phosphorylates H3-threonine 45 to facilitate termination of gene transcription in response to DNA damage', *Nucleic Acids Research*, 43(9): 4505–4516.

Lee, K.-S., Cox-Singh, J. and Singh, B. (2009) 'Morphological features and differential counts of *Plasmodium knowlesi* parasites in naturally acquired human infections.', *Malaria journal*, 8(1): 73. doi: 10.1186/1475-2875-8-73.

Lee, R. S., Waters, A. P. and Brewer, J. M. (2018) 'A cryptic cycle in haematopoietic niches promotes initiation of malaria transmission and evasion of chemotherapy', *Nature Communications*, 9(1): 1689. doi: 10.1038/s41467-018-04108-9.

Van Leeuwen, F., Gafken, P. R. and Gottschling, D. E. (2002) 'Dot1p modulates silencing in yeast by methylation of the nucleosome core', *Cell*, 109(6): 745–756.

Lenhard, B., Sandelin, A. and Carninci, P. (2012) 'Metazoan promoters: emerging characteristics and insights into transcriptional regulation', *Nature Reviews Genetics*, 13(4): 233. doi: 10.1038/nrg3163.

Lensen, A., Bril, A., Van De Vegte, M., Van Gemert, G. J., Eling, W. and Sauerwein, R. (1999) '*Plasmodium falciparum*: Infectivity of Cultured,

Synchronized Gametocytes to Mosquitoes', *Experimental Parasitology*, 914354: 101–103.

LeRoux, M., Lakshmanan, V. and Daily, J. P. (2009) '*Plasmodium falciparum* biology: analysis of *in vitro* versus *in vivo* growth conditions', *Trends in Parasitology*, 25(10): 474–481.

Levine, N. D. (1988) 'Progress in taxonomy of the Apicomplexan protozoa.', *The Journal of protozoology*, 35(4): 518–20.

Lew, V. L., Tiffert, T. and Ginsburg, H. (2003) 'Excess hemoglobin digestion and the osmotic stability of *Plasmodium falciparum*-infected red blood cells', *Blood*, 101(10): 4189-94.

Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R. and Pfister, H. (2014) 'UpSet: Visualization of Intersecting Sets.', *IEEE transactions on visualization and computer graphics*, 20(12): 1983–92.

Li, E. and Zhang, Y. (2014) 'DNA methylation in mammals.', *Cold Spring Harbor perspectives in biology*, 6(5): a019133. doi: 10.1101/cshperspect.a019133.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and 1000 Genome Project Data Processing Subgroup, 1000 Genome Project Data Processing (2009) 'The Sequence Alignment/Map format and SAMtools.', *Bioinformatics*, 25(16): 2078–9.

Li, J., Tao, Z., Li, Q., Brashear, A., Wang, Y., Xia, H., Fang, Q. and Cui, L. (2017) 'Further evaluation of the NWF filter for the purification of *Plasmodium vivax*-infected erythrocytes', *Malaria Journal*, 16(1): 201. doi: 10.1186/s12936-017-1855-3.

Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A. and Zhang, Z. (2008) 'Acetylation of Histone H3 Lysine 56 Regulates Replication-Coupled Nucleosome Assembly', *Cell*. doi: 10.1016/j.cell.2008.06.018.

Li, X., Seidel, C. W., Szerszen, L. T., Lange, J. J., Workman, J. L. and Abmayr, S. M. (2017) 'Enzymatic modules of the SAGA chromatin-modifying complex play distinct roles in *Drosophila* gene expression and development', *Genes and Development*. doi: 10.1101/gad.300988.117.

Li, Y. and Seto, E. (2016) 'HDACs and HDAC inhibitors in cancer development and therapy', *Cold Spring Harbor Perspectives in Medicine*, 6(10). doi: 10.1101/cshperspect.a026831.

Liang, X., Hart, K. J., Dong, G., Siddiqui, F. A., Sebastian, A., Li, X., Albert, I., Miao, J., Lindner, S. E. and Cui, L. (2018) 'Puf3 participates in ribosomal biogenesis in malaria parasites.', *Journal of cell science*, p. jcs.212597. doi: 10.1242/jcs.212597.

Liang, X., Shan, S., Pan, L., Zhao, J., Ranjan, A., Wang, F., Zhang, Z., Huang, Y., Feng, H., Wei, D., Huang, L., Liu, X., Zhong, Q., Lou, J., Li, G., Wu, C. and Zhou, Z. (2016) 'Structural basis of H2A.Z recognition by SRCAP chromatin-

remodeling subunit YL1', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.3190.

Liao,Q., Shen,J., Liu,J., Sun,X., Zhao,G., Chang,Y., Xu,L., Li,X., Zhao,Y., Zheng,H. Zhao, Y. and Wu, Z. (2014) 'Genome-wide identification and functional annotation of *Plasmodium falciparum* long noncoding RNAs from RNA-seq data', *Parasitology Research*, 113(4): 1269–1281.

Lin, J., Annoura, T., Sajid, M., Chevalley-Maurel, S., Ramesar, J., Klop, O., Franke- Fayard, B. M. D., Janse, C. J. and Khan, S. M. (2011) 'A Novel "Gene Insertion/Marker Out" (GIMO) Method for Transgene Expression and Gene Complementation in Rodent Malaria Parasites', *PLoS ONE*, 6(12): e29289. doi: 10.1371/journal.pone.0029289.

Lindner, S. E., Mikolajczak, S. A., Vaughan, A. M., Moon, W., Joyce, B. R., Sullivan, W. J. and Kappe, S. H. I. (2013) 'Perturbations of *Plasmodium* Puf2 expression and RNA-seq of Puf2-deficient sporozoites reveal a critical role in maintaining RNA homeostasis and parasite transmissibility', *Cellular Microbiology*, 15(7): 1266– 1283.

Liokatis, S., Stützer, A., Elsässer, S. J., Theillet, F. X., Klingberg, R., Van Rossum, B., Schwarzer, D., Allis, C. D., Fischle, W. and Selenko, P. (2012) 'Phosphorylation of histone H3 Ser10 establishes a hierarchy for subsequent intramolecular modification events', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb.2310.

Liu, Y., Tewari, R., Ning, J., Blagborough, A. M., Garbom, S., Pei, J., Grishin, N. V, Steele, R. E., Sinden, R. E., Snell, W. J. and Billker, O. (2008) 'The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in *Chlamydomonas* and *Plasmodium* gametes.', *Genes & development*, 22(8): 1051– 68.

Longhurst, H. J. and Holder, A. A. (1997) 'The histones of *Plasmodium falciparum:* identification, purification and a possible role in the pathology of malaria.', *Parasitology*, 114(5): 413–9.

Lopez-Rubio, J.-J. J., Mancio-Silva, L. and Scherf, A. (2009) 'Genome-wide Analysis of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear Repressive Centers in Malaria Parasites', *Cell Host & Microbe*, 5(2): 179–190.

Lopez-Rubio, J.-J., Siegel, T. N. and Scherf, A. (2012) 'Genome-wide Chromatin Immunoprecipitation-Sequencing in *Plasmodium*', *Methods In Molecular Biology,* pp. 321–333. doi: 10.1007/978-1-62703-026-7_23.

Lopez-Rubio, J. J., Gontijo, A. M., Nunes, M. C., Issar, N., Hernandez Rivas, R. and Scherf, A. (2007) '5' flanking region of *var* genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites', *Molecular Microbiology*, 66(6): 1296–1305.

Di Lorenzo, A. and Bedford, M. T. (2011) 'Histone arginine methylation', *FEBS Letters*, pp. 2024–2031. doi: 10.1016/j.febslet.2010.11.010.

Lu, X. M., Bunnik, E. M., Pokhriyal, N., Nasseri, S., Lonardi, S. and Le Roch, K. G. (2015) 'Analysis of nucleosome positioning landscapes enables gene discovery in the human malaria parasite *Plasmodium falciparum*.', *BMC genomics*, 16(1): 1005. doi: 10.1186/s12864-015-2214-9.

Luah, Y.-H., Chaal, B. K., Ong, E. Z. and Bozdech, Z. (2010) 'A moonlighting function of *Plasmodium falciparum* histone 3, mono-methylated at lysine 9?', *PloS one*, 5(4), p. e10252. doi: 10.1371/journal.pone.0010252.

Luger, K., Dechassa, M. L. and Tremethick, D. J. (2012) 'New insights into nucleosome and chromatin structure: an ordered state or a disordered affair?', *Nature Reviews Molecular Cell Biology*, 13(7): 436–447.

Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997) 'Crystal structure of the nucleosome core particle at 2.8 Å resolution', *Nature*, 389(6648): 251–260.

Madeira, L., DeMarco, R., Gazarini, M. L., Verjovski-Almeida, S. and Garcia, C. R. S. (2003) 'Human malaria parasites display a receptor for activated C kinase ortholog', *Biochemical and Biophysical Research Communications*, 306(4): 995– 1001.

Maheshwari, S., Ishii, T., Brown, C. T., Houben, A. and Comai, L. (2017) 'Centromere location in *Arabidopsis* is unaltered by extreme divergence in CENH3 protein sequence', *Genome Research*. doi: 10.1101/gr.214619.116.

Maier, A. G., Braks, J. A. M. M., Waters, A. P. and Cowman, A. F. (2006) 'Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in *Plasmodium falciparum* for targeted gene deletion by double crossover recombination', *Molecular & Biochemical Parasitology*, 150(1): 118–121.

Maier, A. G., Cooke, B. M., Cowman, A. F. and Tilley, L. (2009) 'Malaria parasite proteins that remodel the host erythrocyte', *Nature Reviews Microbiology*, 7(5): 341–354.

Mair, G. R., Braks, J. A. M., Garver, L. S., Dimopoulos, G., Hall, N., Wiegant, J. C. A. G., Dirks, R. W., Khan, S. M., Janse, C. J. and Waters, A. P. (2006) 'Translational Repression is essential for *Plasmodium* sexual development and mediated by a DDX6-type RNA helicase', *Methods*. doi: 10.1038/nature09421.Oxidative.

Mair, G. R., Braks, J. A. M., Garver, L. S., Wiegant, J. C. A. G., Hall, N., Dirks, R. W., Khan, S. M., Dimopoulos, G., Janse, C. J. and Waters, A. P. (2006) 'Regulation of sexual development of *Plasmodium* by translational repression', *Science*, 313(5787): 667–669.

Mair, G. R., Lasonder, E., Garver, L. S., Franke-Fayard, B. M. D. D., Carret, C. K., Wiegant, J. C. A. G. A. G., Dirks, R. W., Dimopoulos, G., Janse, C. J. and Waters, A. P. (2010a) 'Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development', *PLoS Pathogens*, 6(2), p. e1000767. doi: 10.1371/journal.ppat.1000767.

Mair, G. R., Lasonder, E., Garver, L. S., Franke-Fayard, B. M. D. D., Carret, C. K., Wiegant, J. C. A. G. A. G., Dirks, R. W., Dimopoulos, G., Janse, C. J. and Waters, A. P. (2010b) 'Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development', *PLoS Pathogens*, 6(2): e1000767. doi: 10.1371/journal.ppat.1000767.

Malmquist, N. A., Sundriyal, S., Caron, J., Chen, P., Witkowski, B., Menard, D., Suwanarusk, R., Renia, L., Nosten, F., Jimenez-Diaz, M. B., Angulo-Barturen, I., Santos Martinez, M., Ferrer, S., Sanz, L. M., Gamo, F. J., Wittlin, S., Duffy, S., Avery, V. M., Ruecker, A., Delves, M. J., Sinden, R. E., Fuchter, M. J. and Scherf, A. (2015) 'Histone methyltransferase inhibitors are orally bioavailable, fast-acting molecules with activity against different species causing malaria in humans', *Antimicrobial Agents and Chemotherapy*, 59(2): 950–959.

Mancio-Silva, L., Lopez-Rubio, J. J., Claes, A. and Scherf, A. (2013) 'Sir2a regulates rDNA transcription and multiplication rate in the human malaria parasite *Plasmodium falciparum*.', *Nature communications*, 4, p. 1530. doi: 10.1038/ncomms2539.

Manzoni, G., Briquet, S., Risco-Castillo, V., Gaultier, C., Topçu, S., Ivaňescu, M. L., Franetich, J.-F., Hoareau-Coudert, B., Mazier, D. and Silvie, O. (2015) 'A rapid and robust selection procedure for generating drug-selectable marker-free recombinant malaria parasites', *Scientific Reports*, 4(1): 4760. doi: 10.1038/srep04760.

Marfella, C. G. A. and Imbalzano, A. N. (2007) 'The Chd family of chromatin remodelers.', *Mutation research*, 618(1–2): 30–40.

Martín-Jaular, L., Elizalde-Torrent, A., Thomson-Luque, R., Ferrer, M., Segovia, J. C., Herreros-Aviles, E., Fernández-Becerra, C. and del Portillo, H. A. (2013) 'Reticulocyte-prone malaria parasites predominantly invade CD71hi immature cells: implications for the development of an in vitro culture for Plasmodium vivax', *Malaria Journal*, 12(1): 434. doi: 10.1186/1475-2875-12-434.

Martin, S. K., Jett, M. and Schneider, I. (1994) 'Correlation of Phosphoinositide Hydrolysis with Exflagellation in the Malaria Microgametocyte', *The Journal of Parasitology*, 80(3): 371. doi: 10.2307/3283406.

Martins, R. M., Macpherson, C. R., Claes, A., Scheidig-Benatar, C., Sakamoto, H., Yam, X. Y., Preiser, P., Goel, S., Wahlgren, M., Sismeiro, O., Coppée, J. Y. and Scherf, A. (2017) 'An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite *Plasmodium falciparum*', *Scientific Reports*, 7(1). doi: 10.1038/s41598-017-12578-y .

Martinsen, E. S., Perkins, S. L. and Schall, J. J. (2007) 'A three-genome phylogeny of malaria parasites (*Plasmodium* and closely related genera): Evolution of life-history traits and host switches', *Molecular Phylogenetics and Evolution*, 47(1): 261–273.

Marty, A. J., Thompson, J. K., Duffy, M. F., Voss, T. S., Cowman, A. F. and Crabb, B. S. (2006) 'Evidence that *Plasmodium falciparum* chromosome end clusters are cross- linked by protein and are the sites of both virulence gene silencing and activation', *Molecular Microbiology*, 62(1): 72–83.

Matz, J. M. and Kooij, T. W. a (2015) 'Towards genome-wide experimental genetics in the *in vivo* malaria model parasite *Plasmodium berghei*.', *Pathogens and global health*, 109(2): 46–60.

McAndrew, M. B., Read, M., Sims, P. F. G. and Hyde, J. E. (1993) 'Characterisation of the gene encoding an unusually divergent TATA-Binding Protein (TBP) from the extremely A+T-rich human malaria parasite *Plasmodium falciparum*', *Gene*, 124(2): 165–171.

McCutchan, T. F., Li, J., McConkey, G. A., Rogers, M. J. and Waters, A. P. (1995) 'The cytoplasmic ribosomal RNAs of *Plasmodium* spp', *Parasitology Today*, 11(4): 134– 138.

McInroy, G. R., Beraldi, D., Raiber, E.-A., Modrzynska, K., van Delft, P., Billker, O. and Balasubramanian, S. (2016) 'Enhanced Methylation Analysis by Recovery of Unsequenceable Fragments', *PLOS ONE*, 11(3): e0152322. doi: 10.1371/journal.pone.0152322.

McKittrick, E., Gafken, P. R., Ahmad, K. and Henikoff, S. (2004) 'From The Cover: Histone H3.3 is enriched in covalent modifications associated with active chromatin', *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.0308092100.

McRobert, L., Taylor, C. J., Deng, W., Fivelman, Q. L., Cummings, R. M., Polley, S. D., Billker, O. and Baker, D. A. (2008) 'Gametogenesis in Malaria Parasites Is Mediated by the cGMP-Dependent Protein Kinase', *PLoS Biology*, 6(6): e139. doi: 10.1371/journal.pbio.0060139.

Medica, D. L. and Sinnis, P. (2005) 'Quantitative dynamics of Plasmodium yoelii sporozoite transmission by infected anopheline mosquitoes.', *Infection and immunity*, 73(7): 4363–9.

Mehta, M., Braberg, H., Wang, S., Lozsa, A., Shales, M., Solache, A., Krogan, N. J. and Keogh, M. C. (2010) 'Individual lysine acetylations on the N terminus of *Saccharomyces cerevisiae* H2A.Z are highly but not differentially regulated', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M110.185967.

Ménard, R., Sultan, A. A., Cortes, C., Altszuler, R., van Dijk, M. R., Janse, C. J., Waters, A. P., Nussenzweig, R. S., Nussenzweig, V., Menard, R., Sultan, A. A., Cortes, C., Altszuler, R., van Dijk, M. R., Janse, C. J., Waters, A. P., Nussenzweig, R. S. and Nussenzweig, V. (1997) 'Circumsporozoite protein is required for development of malaria sporozoites in mosquitoes', *Nature*, 385(6614): 336–340.

Merrick, C. J., Jiang, R. H. Y., Skillman, K. M., Samarakoon, U., Moore, R. M., Dzikowski, R., Ferdig, M. T. and Duraisingh, M. T. (2015) 'Functional

analysis of sirtuin genes in multiple *Plasmodium falciparum* strains', *PLoS ONE*, 10(3). doi: 10.1371/journal.pone.0118865.

Mersfelder, E. L. and Parthun, M. R. (2006) 'The tale beyond the tail: Histone core domain modifications and the regulation of chromatin structure', *Nucleic Acids Research*, pp. 2653–2662. doi: 10.1093/nar/gkl338.

Miao, J., Fan, Q., Cui, L., Li, J., Li, J. and Cui, L. (2006) 'The malaria parasite *Plasmodium falciparum* histones: Organization, expression, and acetylation', *Gene*, 369(1–2): 53–65.

Miao, J., Fan, Q., Cui, L., Li, X., Wang, H., Ning, G., Reese, J. C. and Cui, L. (2010) 'The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite *Plasmodium falciparum*', *Molecular Microbiology*, 78(4): 883– 902.

Miao, J., Fan, Q., Parker, D., Li, X., Li, J. and Cui, L. (2013) 'Puf Mediates Translation Repression of Transmission-Blocking Vaccine Candidates in Malaria Parasites', *PLoS Pathogens*, 9(4). doi: 10.1371/journal.ppat.1003268.

Miao, J., Li, J., Fan, Q., Li, X., Li, X. and Cui, L. (2010) 'The Puf-family RNA-binding protein PfPuf2 regulates sexual development and sex differentiation in the malaria parasite *Plasmodium falciparum*.', *Journal of cell science*, 123(7): 1039– 1049.

Miller, L. H., Ackerman, H. C., Su, X. and Wellems, T. E. (2013) 'Malaria biology and disease pathogenesis: insights for new treatments', *Nature Medicine*, 19(2): 156– 167.

Miller, L. H., Baruch, D. I., Marsh, K. and Doumbo, O. K. (2002) 'The pathogenic basis of malaria', *Nature*, 415(6872): 673–679.

Miller, M. A. and Olivas, W. M. (2011) 'Roles of Puf proteins in mRNA degradation and translation', *Wiley Interdisciplinary Reviews: RNA*, pp. 471–492. doi: 10.1002/wrna.69.

Min, J., Feng, Q., Li, Z., Zhang, Y. and Xu, R. M. (2003) 'Structure of the catalytic domain of human Dot1L, a non-SET domain nucleosomal histone methyltransferase', *Cell*, 112(5): 711–723.

Minns, A. M., Hart, K. J., Subramanian, S., Hafenstein, S. and Lindner, S. E. (2018) 'Nuclear, Cytosolic, and Surface-Localized Poly(A)-Binding Proteins of *Plasmodium yoelii*', *mSphere*, 3(1). doi: 10.1128/mSphere.00435-17.

Modrzynska, K., Pfander, C., Chappell, L., Yu, L., Suarez, C., Dundas, K., Gomes, A. R., Goulding, D., Rayner, J. C., Choudhary, J. and Billker, O. (2017) 'A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the *Plasmodium* Life Cycle', *Cell Host and Microbe*, 21(1): 11–22.

Mohrle, J. J., Zhao, Y., Wernli, B., Franklin, R. M. and Kappes, B. (1997) 'Molecular cloning, characterization and localization of PfPK4, an eIF-2alpha kinase-related enzyme from the malarial parasite *Plasmodium falciparum*', *Biochem J*, 328(2): 677–687.

Moll, K., Kaneko, A., Scherf, A. and Wahlgren, M. (2013) 'Methods in Malaria Research'. Available at: https://www.beiresources.org/portals/2/MR4/Methods_In_Malaria_Research- 6th_edition.pdf (Accessed: 31 July 2017).

Morrison, D. A. and Ellis, J. T. (1997) 'Effects of nucleotide sequence alignment on phylogeny estimation: a case study of 18S rDNAs of apicomplexa.', *Molecular biology and evolution*, 14(4): 428–441.

Morrissette, N. S. and Sibley, L. D. (2002) 'Cytoskeleton of apicomplexan parasites.', *Microbiology and molecular biology reviews : MMBR*, 66(1): 21–38.

Mota, M. M., Pradel, G., Vanderberg, J. P., Hafalla, J. C. R., Frevert, U., Nussenzweig, R. S., Nussenzweig, V. and Rodríguez, A. (2001) 'Migration of *Plasmodium* Sporozoites Through Cells Before Infection', *Science*, 291(5501): 141-4.

Mueller, I., Zimmerman, P. A. and Reeder, J. C. (2007) '*Plasmodium malariae* and *Plasmodium ovale* - the "bashful" malaria parasites', *Trends in Parasitology*, pp. 278–283. doi: 10.1016/j.pt.2007.04.009.

Müller, K., Matuschewski, K. and Silvie, O. (2011) 'The Puf-family RNA-binding protein Puf2 controls sporozoite conversion to liver stages in the malaria parasite', *PLoS ONE*, 6(5). doi: 10.1371/journal.pone.0019860.

Musselman, C. A., Gibson, M. D., Hartwick, E. W., North, J. A., Gatchalian, J., Poirier, M. G. and Kutateladze, T. G. (2013) 'Binding of PHF1 Tudor to H3K36me3 enhances nucleosome accessibility', *Nature Communications*. doi: 10.1038/ncomms3969.

Näär, A. M., Lemon, B. D. and Tjian, R. (2001) 'Transcriptional Coactivator Complexes', *Annual Review of Biochemistry*, 70(1): 475–501.

Nacer, A., Walker, K. and Hurd, H. (2008) 'Localisation of laminin within *Plasmodium berghei* oocysts and the midgut epithelial cells of *Anopheles stephensi*', *Parasites & Vectors*, 1(1), p. 33. doi: 10.1186/1756-3305-1-33.

Navadgi, V. M., Chandra, B. R., Mishra, P. C. and Sharma, A. (2006) 'The two *Plasmodium falciparum* nucleosome assembly proteins play distinct roles in histone transport and chromatin assembly', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M602243200.

Nijhout, M. M. (1979) '*Plasmodium gallinaceum*: Exflagellation stimulated by a mosquito factor', *Experimental Parasitology*, 48(1): 75–80.

Nijhout, M. M. and Carter, R. (1978) 'Gamete development in malaria parasites: bicarbonate-dependent stimulation by pH in vitro.', *Parasitology*, 76(1): 39–53.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. and Kanemaki, M. (2009) 'An auxin-based degron system for the rapid depletion of proteins in nonplant cells', *Nature Methods*, 6(12): 917–922.

Nishimura, K. and Kanemaki, M. T. (2014) 'Rapid Depletion of Budding Yeast Proteins via the Fusion of an Auxin-Inducible Degron (AID)', in *Current Protocols in Cell Biology*, p. 20.9.1-20.9.16. doi: 10.1002/0471143030.cb2009s64.

van Nuland, R. and Gozani, O. (2016) 'Histone H4 Lysine 20 (H4K20) Methylation, Expanding the Signaling Potential of the Proteome One Methyl Moiety at a Time', *Molecular & Cellular Proteomics*, 15(3): 755–764.

Oberdoerffer, P. (2003) 'Unidirectional Cre-mediated genetic inversion in mice using the mutant loxP pair lox66/lox71', *Nucleic Acids Research*, 31(22): e140.

Oberstaller, J., Pumpalova, Y., Schieler, A., Llinás, M. and Kissinger, J. C. (2014) 'The Cryptosporidium parvum ApiAP2 gene family: Insights into the evolution of apicomplexan AP2 regulatory systems', *Nucleic Acids Research*, 42(13): 8271– 8284.

Ogwan'g, R., Mwangi, J., Gachihi, G., Nwachukwu, A., Roberts, C. R. and Martin, S. K. (1993) 'Use of pharmacological agents to implicate a role for phosphoinositide hydrolysis products in malaria gamete formation', *Biochemical Pharmacology*, 46(9): 1601–1606.

Olivieri, A., Bertuccini, L., Deligianni, E., Franke-Fayard, B., Currà, C., Siden-Kiamos, I., Hanssen, E., Grasso, F., Superti, F., Pace, T., Fratini, F., Janse, C. J. and Ponzi, M. (2015) 'Distinct properties of the egress-related osmiophilic bodies in male and female gametocytes of the rodent malaria parasite *Plasmodium berghei*', *Cellular Microbiology*, 17(3): 355–368.

Olliaro, P. (2008) '*Editorial Commentary:* Mortality Associated with Severe *Plasmodium falciparum* Malaria Increases with Age', *Clinical Infectious Diseases*, 47(2): 158– 160.

Orr, R. Y., Philip, N. and Waters, A. P. (2012) 'Improved negative selection protocol for *Plasmodium berghei* in the rodent malarial model.', *Malaria journal*, 11(1): 103. doi: 10.1186/1475-2875-11-103.

Otto, T. D., Böhme, U., Jackson, A. P., Hunt, M., Franke-Fayard, B., Hoeijmakers, W. A. M., Religa, A. A., Robertson, L., Sanders, M., Ogun, S. A., Cunningham, D., Erhart, A., Billker, O., Khan, S. M., Stunnenberg, H. G., Langhorne, J., Holder, A. A., Waters, A. P., Newbold, C. I., Pain, A., Berriman, M. and Janse, C. J. (2014) 'A comprehensive evaluation of rodent malaria parasite genomes and gene expression', *BMC Biology*, 12(1), p. 86. doi: 10.1186/s12915-014-0086-0.

Otto, T. D., Rayner, J. C., Böhme, U., Pain, A., Spottiswoode, N., Sanders, M., Quail, M., Ollomo, B., Renaud, F., Thomas, A. W., Prugnolle, F., Conway, D. J., Newbold, C. and Berriman, M. (2014) 'Genome sequencing of chimpanzee malaria parasites reveals possible pathways of adaptation to human hosts'. *Nature Communications,* 9(5): 4754.

Painter, H. J., Carrasquilla, M. and Llinás, M. (2017) 'Capturing in vivo RNA transcriptional dynamics from the malaria parasite *Plasmodium falciparum*.', *Genome research*, 27(6): 1074–1086.

Painter, H. J., Chung, N. C., Sebastian, A., Albert, I., Storey, J. D. and Llinás, M. (2018) 'Genome-wide real-time in vivo transcriptional dynamics during *Plasmodium falciparum* blood-stage development', *Nature Communications*, 9(1): 2656. doi: 10.1038/s41467-018-04966-3.

Papazyan, R., Voronina, E., Chapman, J. R., Luperchio, T. R., Gilbert, T. M., Meier, E., Mackintosh, S. G., Shabanowitz, J., Tackett, A. J., Reddy, K. L., Coyne, R. S., Hunt, D. F., Liu, Y. and Taverna, S. D. (2014) 'Methylation of histone H3K23 blocks DNA damage in pericentric heterochromatin during meiosis', *eLife*. doi: 10.7554/eLife.02996.

Parra, M. A., Kerr, D., Fahy, D., Pouchnik, D. J. and Wyrick, J. J. (2006) 'Deciphering the roles of the histone H2B N-terminal domain in genome-wide transcription.', *Molecular and cellular biology*, 26(10): 3842–52.

Pasini, E. M., Böhme, U., Rutledge, G. G., Voorberg-Van der Wel, A., Sanders, M., Berriman, M., Kocken, C. H. and Otto, T. D. (2017) 'An improved *Plasmodium cynomolgi* genome assembly reveals an unexpected methyltransferase gene expansion', *Wellcome Open Research*, 2, p. 42. doi: 10.12688/wellcomeopenres.11864.1.

Paul, R. E. L., Brey, P. T. and Robert, V. (2002) '*Plasmodium* sex determination and transmission to mosquitoes', *Trends in Parasitology*, 18(1): 32–38.

Peck, S. C. (2006) 'Analysis of protein phosphorylation: Methods and strategies for studying kinases and substrates', *Plant Journal*. doi: 10.1111/j.1365-313X.2005.02613.x.

Perkins, D. N., Pappin, D. J., Creasy, D. M. and Cottrell, J. S. (1999) 'Probability-based protein identification by searching sequence databases using mass spectrometry data.', *Electrophoresis*, 20(18): 3551–67.

Perreault, A. A. and Venters, B. J. (2016) 'The ChIP-exo Method: Identifying Protein-DNA Interactions with Near Base Pair Precision', *Journal of Visualized Experiments*, (118), pp. e55016–e55016. doi: 10.3791/55016.

Peterson, D. S., Milhous, W. K. and Wellems, T. E. (1990) 'Molecular basis of differential resistance to cycloguanil and pyrimethamine in *Plasmodium falciparum* malaria', *Proceddings of the National Academy of Sciences of the United States of America*, 87(8): 3018–3022.

Peterson, D. S., Walliker, D. and Wellems, T. E. (1988) 'Evidence that a point mutation in dihydrofolate reductase-thymidylate synthase confers resistance to pyrimethamine in *falciparum* malaria.', *Proceedings of the National Academy of Sciences of the United States of America*, 85(23): 9114–8.

Petter, M., Lee, C. C., Byrne, T. J., Boysen, K. E., Volz, J., Ralph, S. A., Cowman, A. F., Brown, G. V. and Duffy, M. F. (2011) 'Expression of *P. falciparum var* genes involves exchange of the histone variant H2A.Z at the promoter', *PLoS Pathogens*, 7(2), p. e1001292. doi: 10.1371/journal.ppat.1001292.

Petter, M., Selvarajah, S. A., Lee, C. C., Chin, W. H., Gupta, A. P., Bozdech, Z., Brown, G. V. and Duffy, M. F. (2013) 'H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy var gene promoters in the malaria parasite *Plasmodium falciparum*', *Molecular Microbiology*, 87(6): 1167–1182.

Pfander, C., Anar, B., Brochet, M., Rayner, J. C. and Billker, O. (2013) 'Recombination- mediated genetic engineering of *Plasmodium berghei* DNA', *Methods in Molecular Biology*, 923, pp. 127–138. doi: 10.1007/978-1-62703-26-7_8.

Pfander, C., Anar, B., Schwach, F., Otto, T. D., Brochet, M., Volkmann, K., Quail, M. A., Pain, A., Rosen, B., Skarnes, W., Rayner, J. C. and Billker, O. (2011) 'A scalable pipeline for highly effective genetic modification of a malaria parasite', *Nature methods*, 8(12): 1078–1084.

Philip, N. and Waters, A. P. (2015) 'Conditional Degradation of *Plasmodium* Calcineurin Reveals Functions in Parasite Colonization of both Host and Vector.', *Cell host & microbe*, 18(1): 122–31.

Picelli, S., Björklund, A. K., Reinius, B., Sagasser, S., Winberg, G. and Sandberg, R. (2014) 'Tn5 transposase and tagmentation procedures for massively scaled sequencing projects', *Genome Research*, 24(12): 2033–2040.

Pino, P., Sebastian, S., Kim, E. A., Bush, E., Brochet, M., Volkmann, K., Kozlowski, E., Llinás, M., Billker, O., Soldati-Favre, D. and al., et (2012) 'A tetracycline- repressible transactivator system to study essential genes in malaria parasites.', *Cell host & microbe*, 12(6): 824–34.

*Plasmo*GEM (2018) The *Plasmodium* Genetic Modification Project (*Plasmo*GEM) [online] Available at: http://plasmogem.sanger.ac.uk [Accessed 01 Feb. 2018]

Pollack, Y., Kogan, N. and Golenser, J. (1991) '*Plasmodium falciparum*: evidence for a DNA methylation pattern.', *Experimental parasitology*, 72(4): 339–344.

Ponts, N., Fu, L., Harris, E. Y., Zhang, J., Chung, D.-W. W. D., Cervantes, M. C., Prudhomme, J., Atanasova-Penichon, V., Zehraoui, E., Bunnik, E. M., Rodrigues, E. M., Lonardi, S., Hicks, G. R., Wang, Y. and Le Roch, K. G.

(2013) 'Genome-wide mapping of DNA methylation in the human malaria parasite *Plasmodium falciparum*.', *Cell host & microbe*, 14(6): 696–706.

Ponts, N., Harris, E. Y., Prudhomme, J., Wick, I., Eckhardt-Ludka, C., Hicks, G. R., Hardiman, G., Lonardi, S. and Le Roch, K. G. (2010) 'Nucleosome landscape and control of transcription in the human malaria parasite', *Genome Research*, 20(2):228–238.

Ponzi, M., Sidén-Kiamos, I., Bertuccini, L., Currà, C., Kroeze, H., Camarda, G., Pace, T., Franke-Fayard, B., Laurentino, E. C., Louis, C., Waters, A. P., Janse, C. J. and Alano, P. (2009) 'Egress of *Plasmodium berghei* gametes from their host erythrocyte is mediated by the MDV-1/PEG3 protein', *Cellular Microbiology*, 11(8): 1272–1288.

Poran, A., Nötzel, C., Aly, O., Mencia-Trinchant, N., Harris, C. T., Guzman, M. L., Hassane, D. C., Elemento, O. and Kafsack, B. F. C. (2017) 'Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites', *Nature*, 551(7678): 95–99.

Pradeepa, M. M., Grimes, G. R., Kumar, Y., Olley, G., Taylor, G. C. A., Schneider, R. and Bickmore, W. A. (2016) 'Histone H3 globular domain acetylation identifies a new class of enhancers', *Nature Genetics*, 48(6): 681–686.

Pradel, G., Hayton, K., Aravind, L., Iyer, L. M., Abrahamsen, M. S., Bonawitz, A., Mejia, C. and Templeton, T. J. (2004) 'A Multidomain Adhesion Protein Family Expressed in *Plasmodium falciparum* Is Essential for Transmission to the Mosquito', *Journal of Experimental Medicine*, 199(11): 1533-1544.

Prakash, K. and Fournier, D. (2017) 'Deciphering the histone code to build the genome structure', *bioRxiv*, p. 217190. doi: 10.1101/217190.

Prasanth, S. G., Prasanth, K. V., Siddiqui, K., Spector, D. L. and Stillman, B. (2004) 'Human Orc2 localizes to centrosomes, centromeres and heterochromatin during chromosome inheritance', *EMBO Journal*, 23(13): 2651–2663.

Prudêncio, M., Rodriguez, A. and Mota, M. M. (2006) 'The silent path to thousands of merozoites: the *Plasmodium* liver stage', *Nature Reviews Microbiology*, 4(11): 849–856.

Pumpuni, C. B. and Beier, J. C. (1995) 'Long-term survival of *Plasmodium* sporozoites in vitro', *Parasitology Res*earch 81(2): 178–180.

Qiu, Y., Liu, L., Zhao, C., Han, C., Li, F., Zhang, J., Wang, Y., Li, G., Mei, Y., Wu, M., Wu, J. and Shi, Y. (2012) 'Combinatorial readout of unmodified H3R2 and acetylated H3K14 by the tandem PHD finger of MOZ reveals a regulatory mechanism for HOXA9 transcription', *Genes and Development*, 26(12): 1376–1391.

Rajagopal, N., Ernst, J., Ray, P., Wu, J., Zhang, M., Kellis, M. and Ren, B. (2014) 'Distinct and Predictive Histone Lysine Acetylation Patterns at Promoters, Enhancers, and Gene Bodies', *Genes,Genomes,Genetics* (G3). doi: 10.1534/g3.114.013565.

Ram, E. V. S. R., Naik, R., Ganguli, M. and Habib, S. (2008) 'DNA organization by the apicoplast-targeted bacterial histone-like protein of *Plasmodium falciparum*.', *Nucleic acids research*, 36(15): 5061–73.

Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F. and Manke, T. (2016) 'deepTools2: a next generation web server for deep-sequencing data analysis', *Nucleic acids research*, 44(W1), pp. W160– W165. doi: 10.1093/nar/gkw257.

Rawlings, D. J., Fujioka, H., Fried, M., Keister, D. B., Aikawa, M. and Kaslow, D. C. (1992) 'Alpha-tubulin II is a male-specific protein in *Plasmodium falciparum*.', *Molecular and biochemical parasitology*, 56(2): 239–50.

Reece, S. E., Drew, D. R. and Gardner, A. (2008) 'Sex ratio adjustment and kin discrimination in malaria parasites', *Nature*, 453(7195): 609–614.

Reeves, R. (2015) 'High mobility group (HMG) proteins: Modulators of chromatin structure and DNA repair in mammalian cells', *DNA Repair*, pp. 122–136. doi:10.1016/j.dnarep.2015.09.015.

Reininger, L., Billker, O., Tewari, R., Mukhopadhyay, A., Fennell, C., Dorin-Semblat, D., Doerig, C. C., Goldring, D., Harmse, L., Ranford-Cartwright, L., Packer, J. and Doerig, C. C. (2005) 'A NIMA-related protein kinase is essential for completion of the sexual cycle of malaria parasites', *Journal of Biological Chemistry*, 280(36): 31957–64.

Reininger, L., Tewari, R., Fennell, C., Holland, Z., Goldring, D., Ranford-Cartwright, L., Billker, O. and Doerig, C. (2009) 'An essential role for the *Plasmodium* Nek-2 Nima- related protein kinase in the sexual development of malaria parasites', *Journal of Biological Chemistry*, 284(31): 20858–68.

Rennenberg, A., Lehmann, C., Heitmann, A., Witt, T., Hansen, G., Nagarajan, K., Deschermeier , C., T urk, V ., Hilgenfeld, R. and Heussler , V . T . (2010) 'Exoerythrocytic *Plasmodium* Parasites Secrete a Cysteine Protease Inhibitor Involved in Sporozoite Invasion and Capable of Blocking Cell Death of Host Hepatocytes', *PLoS Pathogens.,* 6(3): e1000825. Doi:10.1371/journal.ppat.1000825.

Rhoads, A. and Au, K. F. (2015) 'PacBio Sequencing and Its Applications', *Genomics, Proteomics & Bioinformatics*, 13(5): 278–289.

Ribacke, U., Moll, K., Albrecht, L., Ahmed Ismail, H., Normark, J., Flaberg, E., Szekely, L., Hultenby, K., Persson, K. E. M., Egwang, T. G. and Wahlgren, M. (2013) 'Improved *In Vitro* Culture of *Plasmodium falciparum* Permits Establishment of Clinical Isolates with Preserved Multiplication, Invasion and Rosetting Phenotypes', *PLoS ONE*, 8(7): e69781. doi: 10.1371/journal.pone.0069781.

Ribaut, C., Berry, A., Chevalley, S., Reybier, K., Morlais, I., Parzy, D., Nepveu, F., Benoit-Vical, F. and Valentin, A. (2008) 'Concentration and purification by magnetic separation of the erythrocytic stages of all human Plasmodium species', *Malaria Journal*, 7(1), p. 45. doi: 10.1186/1475-2875-7-45.

Rintisch, C., Heinig, M., Bauerfeind, A., Schafer, S., Mieth, C., Patone, G., Hummel, O., Chen, W., Cook, S., Cuppen, E., Colomé-Tatché, M., Johannes, F., Jansen, R. C., Neil, H., Werner, M., Pravenec, M., Vingron, M. and Hubner, N. (2014) 'Natural variation of histone modification and its impact on gene expression in the rat genome', *Genome Research*. doi: 10.1101/gr.169029.113.

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. and Smyth, G. K. (2015) 'Limma powers differential expression analyses for RNA-sequencing and microarray studies', *Nucleic Acids Research*, 43(7), p. e47. doi: 10.1093/nar/gkv007.

Robert, V., Sokhna, C. S., Rogier, C., Ariey, F. and Trape, J. F. (2003) 'Sex ratio of *Plasmodium falciparum* gametocytes in inhabitants of Dielmo, Senegal.', *Parasitology*, 127(1): 1–8.

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G. and Mesirov, J. P. (2011) 'Integrative genomics viewer', *Nature Biotechnology*, pp. 24–26. doi: 10.1038/nbt.1754.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010) 'edgeR: a Bioconductor package for differential expression analysis of digital gene expression data.', *Bioinformatics*, 26(1): 139–40.

Rodriguez-Collazo, P., Leuba, S. H. and Zlatanova, J. (2009) 'Robust methods for purification of histones from cultured mammalian cells with the preservation of their native modifications', *Nucleic Acids Research*, 37(11). doi: 10.1093/nar/gkp273.

Rodríguez, M. A., Gomez, C., Esther Ramirez, M., Calixto-Galvez, M. and Medel, O. (2010) 'Regulation of gene expression in protozoa parasites', *Journal of Biomedicine and Biotechnology*. doi: 10.1155/2010/726045.

Roobsoong, W., Tharinjaroen, C. S., Rachaphaew, N., Chobson, P., Schofield, L., Cui, L., Adams, J. H. and Sattabongkot, J. (2015) 'Improvement of culture conditions for long-term in vitro culture of *Plasmodium vivax*', *Malaria Journal*, 14(1), p. 297. doi: 10.1186/s12936-015-0815-z.

Rotem, A., Ram, O., Shoresh, N., Sperling, R. A., Goren, A., Weitz, D. A. and Bernstein, B. E. (2015) 'Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state', *Nature Biotechnology*, 33(11): 1165–1172.

Rovira-Graells, N., Gupta, A. P., Planet, E., Crowley, V. M., Mok, S., Ribas de Pouplana, L., Preiser, P. R., Bozdech, Z., Cortés, A., De Pouplana, L. R., Preiser, P. R., Bozdech, Z. and Cortés, A. (2012) 'Transcriptional variation in the malaria parasite *Plasmodium falciparum*.', *Genome Research*, 22(5): 925–938.

Ruan, K., Yamamoto, T. G., Asakawa, H., Chikashige, Y., Kimura, H., Masukata, H., Haraguchi, T. and Hiraoka, Y. (2015) 'Histone H4 acetylation required for chromatin decompaction during DNA replication', *Scientific Reports*. doi: 10.1038/srep12720.

Ruegger, M., Dewey, E., Gray, W. M., Hobbie, L., Turner, J. and Estelle, M. (1998) 'The *TIR1* protein of *Arabidopsis* functions in auxin response and is related to human *SKP2* and yeast *Grr1p*', *Genes & Development*, 12(2): 198–207.

Ruiz-Carrillo, D., Lin, J., El Sahili, A., Wei, M., Sze, S. K., Cheung, P. C. F., Doerig, C. and Lescar, J. (2018) 'The protein kinase CK2 catalytic domain from *Plasmodium falciparum*: crystal structure, tyrosine kinase activity and inhibition', *Scientific Reports*, 8(1): 7365. doi: 10.1038/s41598-018-25738-5.

Ruiz, J. L., Tena, J. J., Bancells, C., Cortés, A., Gómez-Skarmeta, J. L. and Gómez-Díaz, E. (2018) 'Characterization of the accessible genome in the human malaria parasite *Plasmodium falciparum*', *Nucleic Acids Research*. doi: 10.1093/nar/gky643.

Russell, D. G. and Burns, R. G. (1984) 'The polar ring of coccidian sporozoites: a unique microtubule-organizing centre', *Journal of Cell Science*, 65(1): 193-207.

Sainsbury, S., Bernecky, C. and Cramer, P. (2015) 'Structural basis of transcription initiation by RNA polymerase II', *Nature Reviews Molecular Cell Biology*, pp. 129– 143. doi: 10.1038/nrm3952.

Salcedo-Amaya, A. M., van Driel, M. A., Alako, B. T., Trelle, M. B., van den Elzen, A. M. G., Cohen, A. M., Janssen-Megens, E. M., van de Vegte-Bolmer, M., Selzer, R. R., Iniguez, A. L., Green, R. D., Sauerwein, R. W., Jensen, O. N. and Stunnenberg, H. G. (2009) 'Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*.', *Proceedings of the National Academy of Sciences of the United States of America*, 106(24): 9655–60.

Sanders, P. R., Gilson, P. R., Cantin, G. T., Greenbaum, D. C., Nebl, T., Carucci, D. J., McConville, M. J., Schofield, L., Hodder, A. N., Yates, J. R. and Crabb, B. S. (2005) 'Distinct protein classes including novel merozoite surface antigens in Raft-like membranes of *Plasmodium falciparum*.', *The Journal of biological chemistry*, 280(48): 40169–76.

Sanderson, T. and Rayner, J. C. (2017) 'PhenoPlasm: a database of disruption phenotypes for malaria parasite genes', *Wellcome Open Research*, 2, p. 45. doi: 10.12688/wellcomeopenres.11896.1.

Santos, J. M., Kehrer, J., Franke-Fayard, B., Frischknecht, F., Janse, C. J. and Mair, G. R. (2015) 'The *Plasmodium* palmitoyl-S-acyl-transferase DHHC2 is essential for ookinete morphogenesis and malaria transmission', *Scientific Reports*, 5. doi: 10.1038/srep16034.

Sasi, P., Burns, S. P., Waruiru, C., English, M., Hobson, C. L., King, C. G., Mosobo, M., Beech, J. S., Iles, R. A., Boucher, B. J. and Cohen, R. D. (2007) 'Metabolic Acidosis and Other Determinants of Hemoglobin-Oxygen Dissociation in Severe Childhood *Plasmodium falciparum* Malaria', *The American Journal of Tropical Hygiene and Medicine,* 77(2): 256-60.

Sautel, C. F., Cannella, D., Bastien, O., Kieffer, S., Aldebert, D., Garin, J., Tardieux, I., Belrhali, H. and Hakimi, M.-A. (2007) 'SET8-Mediated Methylations of Histone H4 Lysine 20 Mark Silent Heterochromatic Domains in Apicomplexan Genomes', *Molecular and Cellular Biology*. doi: 10.1128/mcb.00482-07.

Schaer, J., Perkins, S. L., Decher, J., Leendertz, F. H., Fahr, J., Weber, N. and Matuschewski, K. (2013) 'High diversity of West African bat malaria parasites and a tight link with rodent *Plasmodium* taxa', *Proceedings of the National Academy of Sciences*, 110(43): 17415–17419.

van Schaijk, B. C. L., van Dijk, M. R., van de Vegte-Bolmer, M., van Gemert, G.-J., van Dooren, M. W., Eksi, S., Roeffen, W. F. G., Janse, C. J., Waters, A. P. and Sauerwein, R. W. (2006) 'Pfs47, paralog of the male fertility factor Pfs48/45, is a female specific surface protein in *Plasmodium falciparum*', *Molecular and Biochemical Parasitology*, 149(2): 216–222.

Scherf, A., Hernandez-Rivas, R., Buffet, P., Bottius, E., Benatar, C., Pouvelle, B., Gysin, J. and Lanzer, M. (1998) 'Antigenic variation in malaria: In situ switching, relaxed and mutually exclusive transcription of *var* genes during intra-erythrocytic development in *Plasmodium falciparum*', *EMBO Journal*, 17(18): 5418–5426.

Schiza, V., Molina-Serrano, D., Kyriakou, D., Hadjiantoniou, A. and Kirmizis, A. (2013) 'N-alpha-terminal Acetylation of Histone H4 Regulates Arginine Methylation and Ribosomal DNA Silencing', *PLoS Genetics*, 9(9). doi: 10.1371/journal.pgen.1003805.

Schmidl, C., Rendeiro, A. F., Sheffield, N. C. and Bock, C. (2015) 'ChIPmentation: fast, robust, low-input ChIP-seq for histones and transcription factors', *Nat Meth*ods, 12(10): 963–965.

Schuster, F. L. (2002) 'Cultivation of plasmodium spp.', *Clinical microbiology reviews*, 15(3): 355–64.

Schwach, F., Bushell, E., Gomes, A. R., Anar, B., Girling, G., Herd, C., Rayner, J. C. and Billker, O. (2015) 'PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites', *Nucleic Acids Research*, 43(D1), pp. D1176–D1182. doi: 10.1093/nar/gku1143.

Schwämmle, V., Sidoli, S., Ruminowicz, C., Wu, X., Lee, C.-F., Helin, K. and Jensen, O. N. (2016) 'Systems Level Analysis of Histone H3 Post-translational Modifications (PTMs) Reveals Features of PTM Crosstalk in Chromatin Regulation', *Molecular & Cellular Proteomics*, 15(8): 2715–2729.

Schwank, S., Sutherland, C. J., Drakeley, C. J., Goman, M. and Holloway, S. (2010) 'Promiscuous Expression of α-Tubulin II in Maturing Male and Female *Plasmodium falciparum* Gametocytes', *PLoS ONE*, 5(12): e14470. doi: 10.1371/journal.pone.0014470.

Sebastian, S., Brochet, M., Collins, M. O., Schwach, F., Jones, M. L., Goulding, D., Rayner, J. C., Choudhary, J. S. and Billker, O. (2012) 'A *Plasmodium* Calcium- Dependent Protein Kinase Controls Zygote Development and Transmission by Translationally Activating Repressed mRNAs', *Cell Host & Microbe*, 12(1): 9–19.

Sen, U., Saxena, H., Khurana, J., Nayak, A. and Gupta, A. (2018) '*Plasmodium falciparum* RUVBL3 protein: a novel DNA modifying enzyme and an interacting partner of essential HAT protein MYST', *Scientific Reports*. doi: 10.1038/s41598-018-29137- 8.

Sengupta, J., Nilsson, J., Gursky, R., Spahn, C. M. T., Nissen, P. and Frank, J. (2004) 'Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM', *Nature Structural and Molecular Biology*, 11(10): 957–962.

Seow, F., Sato, S., Janssen, C. S., Riehle, M. O., Mukhopadhyay, A., Phillips, R. S., Wilson, R. J. M. and Barrett, M. P. (2005) 'The plastidic DNA replication enzyme complex of *Plasmodium falciparum*', *Molecular and Biochemical Parasitology*, 141(2): 145–153.

Severini, C., Silvestrini, F., Sannella, A., Barca, S., Gradoni, L. and Alano, P. (1999) 'The production of the osmiophilic body protein Pfg377 is associated with stage of maturation and sex in *Plasmodium falciparum* gametocytes', *Molecular and Biochemical Parasitology*, 100(2): 247–252.

Sharma, R., Sharma, B., Gupta, A. and Dhar, S. K. (2018) 'Identification of a novel trafficking pathway exporting a replication protein, Orc2 to nucleus via classical secretory pathway in *Plasmodium falciparum*', *Biochimica et Biophysica Acta - Molecular Cell Research*. doi: 10.1016/j.bbamcr.2018.03.003.

Shaw, P. J., Ponmee, N., Karoonuthaisiri, N., Kamchonwongpaisan, S. and Yuthavong, Y. (2007) 'Characterization of human malaria parasite *Plasmodium falciparum* eIF4E homologue and mRNA 5′ cap status', *Molecular and Biochemical Parasitology*, 155(2): 146–155.

Shechter, D., Dormann, H. L., Allis, C. D. and Hake, S. B. (2007) 'Extraction, purification and analysis of histones', *Nature Protocols*, 2(6): 1445–1457.

Shoaib, M., Walter, D., Gillespie, P. J., Izard, F., Fahrenkrog, B., Lleres, D., Lerdrup, M., Johansen, J. V., Hansen, K., Julien, E., Blow, J. J. and Sørensen, C. S. (2018) 'Histone H4K20 methylation mediated chromatin compaction threshold ensures genome integrity by limiting DNA replication licensing', *Nature Communications*. doi: 10.1038/s41467-018-06066-8.

Shrestha, S., Li, X., Ning, G., Miao, J. and Cui, L. (2016) 'The RNA-binding protein PfPuf1 functions in the maintenance of gametocytes in *Plasmodium falciparum*', *Journal of Cell Science*, p. jcs.186908. doi: 10.1242/jcs.186908.

Sidoli, S., Bhanu, N. V., Karch, K. R., Wang, X. and Garcia, B. A. (2016) 'Complete Workflow for Analysis of Histone Post-translational Modifications Using Bottom-up Mass Spectrometry: From Histone Extraction to Data Analysis', *Journal of Visualized Experiments*, (111). doi: 10.3791/54112.

Sidoli, S., Cheng, L. and Jensen, O. N. (2012) 'Proteomics in chromatin biology and epigenetics: Elucidation of post-translational modifications of histone proteins by mass spectrometry', *Journal of Proteomics*, pp. 3419–3433. doi: 10.1016/j.jprot.2011.12.029.

Sidoli, S., Yuan, Z. F., Lin, S., Karch, K., Wang, X., Bhanu, N., Arnaudo, A. M., Britton, L. M., Cao, X. J., Gonzales-Cope, M., Han, Y., Liu, S., Molden, R. C., Wein, S., Afjehi-Sadat, L. and Garcia, B. A. (2015) 'Drawbacks in the use of unconventional hydrophobic anhydrides for histone derivatization in bottom-up proteomics PTM analysis', *Proteomics*, 15(9): 1459–1469.

Siegel, T. N., Hon, C., Zhang, Q., Lopez-Rubio, J.-J., Scheidig-Benatar, C., Martins, R. M., Sismeiro, O., Coppée, J. and Scherf, A. (2014) 'Strand-specific RNA-Seq reveals widespread and developmentally regulated transcription of natural antisense transcripts in *Plasmodium falciparum*.', *BMC genomics*, 15(1), p. 150. doi: 10.1186/1471-2164-15-150.

Siekierka, J. J., Hung, S. H. Y., Poe, M., Lin, C. S. and Sigal, N. H. (1989) 'A cytosolic binding protein for the immunosuppressant FK506 has peptidyl-prolyl isomerase activity but is distinct from cyclophilin', *Nature*, 341(6244): 755–757.

Sierra-Miranda, M., Vembar, S. S., Delgadillo, D. M., Ávila-López, P. A., Herrera-Solorio, A. M., Lozano Amado, D., Vargas, M. and Hernandez-Rivas, R. (2017) 'PfAP2Tel, harbouring a non-canonical DNA-binding AP2 domain, binds to *Plasmodium falciparum* telomeres', *Cellular Microbiology*, 19(9). doi: 10.1111/cmi.12742.

Sievers, F. and Higgins, D. G. (2014) 'Clustal Omega, Accurate Alignment of Very Large Numbers of Sequences', in. Humana Press, Totowa, NJ, pp. 105–116. doi: 10.1007/978-1-62703-646-7_6.

Silberhorn, E., Schwartz, U., Löffler, P., Schmitz, S., Symelka, A., de Koning-Ward, T., Merkl, R. and Längst, G. (2016) '*Plasmodium falciparum* Nucleosomes Exhibit Reduced Stability and Lost Sequence Dependent Nucleosome Positioning', *PLOS Pathogens*, 12(12), p. e1006080. doi: 10.1371/journal.ppat.1006080.

Silva, P. A., Guerreiro, A., Santos, J. M., Braks, J. A. M., Janse, C. J. and Mair, G. R. (2016) 'Translational control of UIS4 protein of the host-parasite interface is mediated by the RNA binding protein Puf2 in *Plasmodium berghei* sporozoites', *PLoS ONE*, 11(1). doi: 10.1371/journal.pone.0147940.

Silvestrini, F., Alano, P. And Williams, J. L. (2000) 'Commitment to the production of male and female gametocytes in the human malaria parasite *Plasmodium falciparum*', *Parasitology*, 121(05): 465–471.

Simonetta Friso, *,†, Sang-Woon Choi, †, Gregory G. Dolnikowski, ‡ and and Selhub†, J. (2002) 'A Method to Assess Genomic DNA Methylation Using High-Performance Liquid Chromatography/Electrospray Ionization Mass Spectrometry'. American Chemical Society . doi: 10.1021/AC020050H.

Sinden, R. E. (1983) 'Sexual Development of Malarial Parasites', in, pp. 153–216. doi: 10.1016/S0065-308X(08)60462-5.

Sinden, R. E., Butcher, G. A., Billker, O. and Fleck, S. L. (1996) 'Regulation of Infectivity of *Plasmodium* to the Mosquito Vector', in, pp. 53–117. doi: 10.1016/S0065- 308X(08)60033-0.

Sindikubwabo, F., Ding, S., Hussain, T., Ortet, P., Barakat, M., Baumgarten, S., Cannella, D., Palencia, A., Bougdour, A., Belmudes, L., Couté, Y., Tardieux, I., Botté, C. Y., Scherf, A. and Hakimi, M. (2017) 'Modifications at K31 on the lateral surface of histone H4 contribute to genome structure and expression in apicomplexan parasites', *eLife*, 6, p. e29391. doi: 10.7554/eLife.29391.

Singh, A. P., Buscaglia, C. A., Wang, Q., Levay, A., Nussenzweig, D. R., Walker, J. R., Winzeler, E. A., Fujii, H., Fontoura, B. M. A. and Nussenzweig, V. (2007) '*Plasmodium* Circumsporozoite Protein Promotes the Development of the Liver Stages of the Parasite', *Cell*, 131(3): 492–504.

Singh, B. and Daneshvar, C. (2013) 'Human infections and detection of *Plasmodium knowlesi*.', *Clinical microbiology reviews*, 26(2): 165–84.

Sinha, A., Hughes, K. R., Modrzynska, K. K., Otto, T. D., Pfander, C., Dickens, N. J., Religa, A. A., Bushell, E., Graham, A. L., Cameron, R., Kafsack, B. F. C., Williams, A. E., Llinas, M., Berriman, M., Billker, O. and Waters, A. P. (2014) 'A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*', *Nature*, 507(7491): 253–257.

Sirawaraporn, W., Sirawaraporn, R., Cowman, A. F., Yuthavong, Y. and Santi, D. V (1990) 'Heterologous expression of active thymidylate synthase-dihydrofolate reductase from *Plasmodium falciparum*.', *Biochemistry*, 29(48): 10779–85.

Smith, T. G., Lourenço, P., Carter, R., Walliker, D. and Ranford-Cartwright, L. C. (2000) 'Commitment to sexual differentiation in the human malaria parasite, *Plasmodium falciparum*.', *Parasitology*, 121, pp. 127–133. doi: 10.1017/S0031182099006265.

Sologub, L., Kuehn, A., Kern, S., Przyborski, J., Schillig, R. and Pradel, G. (2011) 'Malaria proteases mediate inside-out egress of gametocytes from red blood cells following parasite transmission to the mosquito', *Cellular Microbiology*, 13(6): 897–912.

Solyakov, L., Halbert, J., Alam, M. M., Semblat, J.-P., Dorin-Semblat, D., Reininger, L., Bottrill, A. R., Mistry, S., Abdi, A., Fennell, C., Holland, Z., Demarta, C., Bouza, Y., Sicard, A., Nivez, M.-P., Eschenlauer, S., Lama, T., Thomas, D. C., Sharma, P., Agarwal, S., Kern, S., Pradel, G., Graciotti, M.,

Tobin, A. B. and Doerig, C. (2011) 'Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*', *Nature Communications*, 2, p. 565. doi: 10.1038/ncomms1558.

Sorber, K., Dimon, M. T. and Derisi, J. L. (2011) 'RNA-Seq analysis of splicing in *Plasmodium falciparum* uncovers new splice junctions, alternative splicing and splicing of antisense transcripts', *Nucleic Acids Research*, 39(9): 3820–3835.

van Spaendonk, R. M., Ramesar, J., van Wigcheren, a, Eling, W., Beetsma, a L., van Gemert, G. J., Hooghof, J., Janse, C. J. and Waters, a P. (2001) 'Functional equivalence of structurally distinct ribosomes in the malaria parasite, *Plasmodium berghei*.', *The Journal of biological chemistry*, 276(25): 22638–47.

Srinivasan, P., Fujioka, H. and Jacobs-Lorena, M. (2008) 'PbCap380, a novel oocyst capsule protein, is essential for malaria parasite survival in the mosquito', *Cellular Microbiology*, 10(6): 1304–1312.

Srivastava, S., Bhowmick, K., Chatterjee, S., Basha, J., Kundu, T. K. and Dhar, S. K. (2014) 'Histone H3K9 acetylation level modulates gene expression and may affect parasite growth in human malaria parasite *Plasmodium falciparum*', *FEBS Journal*, 281(23): 5265–5278.

Staby, L., Shea, C., Willemoës, M., Theisen, F., Kragelund, B. B. and Skriver, K. (2017) 'Eukaryotic transcription factors: paradigms of protein intrinsic disorder', *Biochemical Journal*, 474(15): 2509–2532.

Starmer, J. and Magnuson, T. (2016) 'Detecting broad domains and narrow peaks in ChIP- seq data with hiddenDomains', *BMC Bioinformatics*. doi: 10.1186/s12859-016- 0991-z.

Steinhauser, S., Kurzawa, N., Eils, R. and Herrmann, C. (2016) 'A comprehensive comparison of tools for differential ChIP-seq analysis', *Briefings in Bioinformatics*. doi: 10.1093/bib/bbv110.

Stejskal, S., Stepka, K., Tesarova, L., Stejskal, K., Matejkova, M., Simara, P., Zdrahal, Z. and Koutna, I. (2015) 'Cell cycle-dependent changes in H3K56ac in human cells', *Cell Cycle*. doi: 10.1080/15384101.2015.1106760.

Straschil, U., Talman, A. M., Ferguson, D. J. P., Bunting, K. A., Xu, Z., Bailes, E., Sinden, R. E., Holder, A. A., Smith, E. F., Coates, J. C. and Rita Tewari, R. (2010) 'The Armadillo Repeat Protein PF16 Is Essential for Flagellar Structure and Function in *Plasmodium* Male Gametes', *PLoS ONE.*, 5(9), p. e12901. doi: 10.1371/journal.pone.0012901.

Sturm, A., Amino, R., van de Sand, C., Regen, T., Retzlaff, S., Rennenberg, A., Krueger, A., Pollok, J.-M., Menard, R. and Heussler, V. T. (2006) 'Manipulation of Host Hepatocytes by the Malaria Parasite for Delivery into Liver Sinusoids', *Science*, 313(5791): 1287-90.

Su, S., Law, C. W., Ah-Cann, C., Asselin-Labat, M.-L., Blewitt, M. E. and Ritchie, M. E. (2017) 'Glimma: Interactive graphics for gene expression analysis', *Bioinformatics*, 33(13). doi: 10.1093/bioinformatics/btx094.

Su, X.-R., Heatwoie, V. M., Wertheimer, S. P., Guinet, F., Hertfeldt, J. A., Peterson, D. S., Ravetch, J. A. and Weilems ', T. E. (1995) 'The Large Diverse Gene Family w Encodes Proteins Involved in Cytoadherence and Antigenic Variation of *Plasmodium falciparum*-Infected Erythrocytes', *Cell*, 82(1): 89–100.

Sultan, A. A., Thathy, V., de Koning-Ward, T. F. and Nussenzweig, V. (2001) 'Complementation of *Plasmodium berghei* TRAP knockout parasites using human dihydrofolate reductase gene as a selectable marker', *Molecular and Biochemical Parasitology*, 113(1): 151–156.

Sumanadasa, S. D. M., Goodman, C. D., Lucke, A. J., Skinner-Adams, T., Saham, I., Haque, A., Do, T. A., McFadden, G. I., Fairlie, D. P. and Andrews, K. T. (2012) 'Antimalarial activity of the anticancer histone deacetylase inhibitor SB939', *Antimicrobial Agents and Chemotherapy*, 56(7): 3849–3856.

Sutherland, C. J., Tanomsing, N., Nolder, D., Oguike, M., Jennison, C., Pukrittayakamee, S., Dolecek, C., Hien, T. T., Do Rosá Rio, V. E., Arez, A. P., Pinto, J. O., Michon, P., Escalante, A. A., Nosten, F., Burke, M., Lee, R., Blaze, M., Otto, T. D., Barnwell, J. W., Pain, A., Williams, J., White, N. J., Day, N. P. J., Snounou, G., Lockhart, P. J., Chiodini, P. L., Imwong, M., Polley, S. D. and Sutherland, C. (2010) 'Two Nonrecombining Sympatric Forms of the Human Malaria Parasite *Plasmodium ovale* Occur Globally', *The Journal of Infectious Diseases*, 201(10): 1544–1550.

Suzuki, S., Kato, H., Suzuki, Y., Chikashige, Y., Hiraoka, Y., Kimura, H., Nagao, K., Obuse, C., Takahata, S. and Murakami, Y. (2016) 'Histone H3K36 trimethylation is essential for multiple silencing mechanisms in fission yeast', *Nucleic Acids Research*. doi: 10.1093/nar/gkw008.

Ta, T. H., Hisam, S., Lanza, M., Jiram, A. I., Ismail, N. and Rubio, J. M. (2014) 'First case of a naturally acquired human infection with *Plasmodium cynomolgi*.', *Malaria journal*. BioMed Central, 13, p. 68. doi: 10.1186/1475-2875-13-68.

Tafrova, J. I. and Tafrov, S. T. (2014) 'Human histone acetyltransferase 1 (Hat1) acetylates lysine 5 of histone H2A in vivo', *Molecular and Cellular Biochemistry*. doi: 10.1007/s11010-014-2036-0.

Takebe, S., Witola, W. H., Schimanski, B., Günzl, A. and Ben Mamoun, C. (2007) 'Purification of components of the translation elongation factor complex of *Plasmodium falciparum* by tandem affinity purification', *Eukaryotic Cell*, 6(4): 584–591.

Taleb, O., Maammar, M., Brumaru, D., Bourguignon, J.-J., Schmitt, M., Klein, C., Kemmel, V., Maitre, M. and Mensah-Nyagan, A. G. (2012) 'Xanthurenic Acid Binds to Neuronal G-Protein-Coupled Receptors That Secondarily Activate

Cationic Channels in the Cell Line NCB-20', *PLoS ONE*, 7(11), p. e48553. doi: 10.1371/journal.pone.0048553.

Talman, A. M., Domarle, O., McKenzie, F. E., Ariey, F. and Robert, V. (2004) 'Gametocytogenesis : the puberty of *Plasmodium falciparum*', *Malaria Journal*, 3(1), p. 24. doi: 10.1186/1475-2875-3-24.

Talman, A. M., Lacroix, C., Marques, S. R., Blagborough, A. M., Carzaniga, R., Ménard, R. and Sinden, R. E. (2011) 'PbGEST mediates malaria transmission to both mosquito and vertebrate host', *Molecular Microbiology*, 82(2): 462–474.

Tan, X., Calderon-Villalobos, L. I. A., Sharon, M., Zheng, C., Robinson, C. V., Estelle, M. and Zheng, N. (2007) 'Mechanism of auxin perception by the TIR1 ubiquitin ligase', *Nature*, 446(7136): 640–645.

Tanaka, K. and Okamoto, A. (2007) 'Degradation of DNA by bisulfite treatment', *Bioorganic & Medicinal Chemistry Letters*, 17(7): 1912–1915.

Tarr, S. J., Díaz-Ingelmo, O., Stewart, L. B., Hocking, S. E., Murray, L., Duffy, C. W., Otto, T. D., Chappell, L., Rayner, J. C., Awandare, G. A. and Conway, D. J. (2018) '*Plasmodium falciparum* mature schizont transcriptome variation among clinical isolates and laboratory-adapted clones', *bioRxiv*, p. 329532. doi: 10.1101/329532.

Tavares, J., Formaglio, P., Thiberge, S., Mordelet, E., Van Rooijen, N., Medvinsky, A., Ménard, R. and Amino, R. (2013) 'Role of host cell traversal by the malaria sporozoite during liver infection.', *The Journal of experimental medicine*, 210(5): 905–15.

Teale, W. D., Paponov, I. A. and Palme, K. (2006) 'Auxin in action: Signalling, transport and the control of plant growth and development', *Nature Reviews Molecular Cell Biology*, pp. 847–859. doi: 10.1038/nrm2020.

Terzakis, J. A., Sprinz, H. and Ward, R. A. (1967) 'The transformation of the *Plasmodium gallinaceum* oocyst in Aedes aegypti mosquitoes.', *The Journal of cell biology*, 34(1): 311–26.

Tessarz, P. and Kouzarides, T. (2014) 'Histone core modifications regulating nucleosome structure and dynamics', *Nature Reviews Molecular Cell Biology*, 15(11): 703–708.

Tewari, R., Straschil, U., Bateman, A., Böhme, U., Cherevach, I., Gong, P., Pain, A. and Billker, O. (2010) 'The systematic functional analysis of *Plasmodium* protein kinases identifies essential regulators of mosquito transmission', *Cell Host and Microbe*, 8(4): 377–387.

The EuPathDB Project Team (2018) PlasmoDB: Plasmodium Genomics Resource [online] Available at: http://plasmodb.org/plasmo/showApplication.do [Accessed 23 Jan. 2018]

The Gene Ontology Consortium (2017) 'Expansion of the Gene Ontology knowledgebase and resources.', *Nucleic acids research*, 45(D1), pp. D331–D338. doi: 10.1093/nar/gkw1108.

Thompson, J., Janse, C. J. and Waters, A. P. (2001) 'Comparative genomics in *Plasmodium*: a tool for the identification of genes and functional analysis', *Molecular and Biochemical Parasitology*, 118(2): 147–154.

Tibú Rcio, M., Sauerwein, R., Lavazec, C. and Alano, P. (2015) 'Erythrocyte remodeling by *Plasmodium falciparum* gametocytes in the human host interplay', *Trends in Parasitology*, 31(6): 270–278.

Toenhake, C. G., Fraschka, S. A.-K., Vijayabaskar, M. S., Westhead, D. R., van Heeringen, S. J. and Bártfai, R. (2018) 'Chromatin Accessibility-Based Characterization of the Gene Regulatory Network Underlying *Plasmodium falciparum* Blood-Stage Development', *Cell Host & Microbe*, 23(4): 557–569.

Tonkin, C. J., Carret, C. K., Duraisingh, M. T., Voss, T. S., Ralph, S. A., Hommel, M., Duffy, M. F., Silva, L. M. da, Scherf, A., Ivens, A., Speed, T. P., Beeson, J. G., Cowman, A. F., Da Silva, L. M., Scherf, A., Ivens, A., Speed, T. P., Beeson, J. G. and Cowman, A. F. (2009) 'Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in *Plasmodium falciparum*', *PLoS Biology*, 7(4): 771–788.

Topolska, A. E., Lidgett, A., Truman, D., Fujioka, H. and Coppel, R. L. (2004) 'Characterization of a Membrane-associated Rhoptry Protein of *Plasmodium falciparum*', *Journal of Biological Chemistry*, 279(6): 4648–4656.

Trelle, M. B., Salcedo-Amaya, A. M., Cohen, A. M., Stunnenberg, H. G. and Jensen, O. N. (2009) 'Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite *Plasmodium falciparum*', *Journal of Proteome Research*, 8(7): 3439–3450.

Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E. T., Bühler, M., Margueron, R. and Schneider, R. (2013) 'Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer', *Cell*, 152(4): 859–872.

Tropberger, P. and Schneider, R. (2010) 'Going global: Novel histone modifications in the globular domain of H3', *Epigenetics*. doi: 10.4161/epi.5.2.11075.

Trost, M., Blattner, A. C. and Lehner, C. F. (2016) 'Regulated protein depletion by the auxin-inducible degradation system in *Drosophila melanogaster*', *Fly*, 10(1): 35– 46.

Tu, S. and Shao, Z. (2017) 'An introduction to computational tools for differential binding analysis with ChIP-seq data', *Quantitative Biology*. doi: 10.1007/s40484-017-0111- 8.

Turner, B. M. (2005) 'Reading signals on the nucleosome with a new nomenclature for modified histones', *Nature Structural and Molecular Biology*. doi: 10.1038/nsmb0205-110.

Tuteja, R., Ansari, A. and Chauhan, V. S. (2011) 'Emerging functions of transcription factors in malaria parasite', *Journal of Biomedicine and Biotechnology*. doi: 10.1155/2011/461979.

Tuteja, R. and Pradhan, A. (2009) 'Isolation and functional characterization of eIF4F components and poly(A)-binding protein from *Plasmodium falciparum*', *Parasitology International*, 58(4): 481–485.

Tvardovskiy, A., Schwämmle, V., Kempf, S. J., Rogowska-Wrzesinska, A. and Jensen, O. N. (2017) 'Accumulation of histone variant H3.3 with age is associated with profound changes in the histone methylation landscape', *Nucleic Acids Research*, 45(16): 9272–9289.

Udomsangpetch, R., Kaneko, O., Chotivanich, K. and Sattabongkot, J. (2008) 'Cultivation of Plasmodium vivax', *Trends in Parasitology*, 24(2): 85–88.

Vaidya, A. B., Lashgari, M. S., Pologe, L. G. and Morrisey, J. (1993) 'Structural features of *Plasmodium* cytochrome b that may underlie susceptibility to 8- aminoquinolines and hydroxynaphthoquinones', *Molecular and Biochemical Parasitology*, 58(1): 33–42.

Vembar, S. S., Droll, D. and Scherf, A. (2016) 'Translational regulation in blood stages of the malaria parasite *Plasmodium* spp.: systems-wide studies pave the way', *Wiley Interdisciplinary Reviews: RNA*, 7(6): 772–792.

Vembar, S. S., Macpherson, C. R., Sismeiro, O., Coppée, J. Y. and Scherf, A. (2015) 'The PfAlba1 RNA-binding protein is an important regulator of translational timing in *Plasmodium falciparum* blood stages', *Genome Biology*, 16(1). doi: 10.1186/s13059-015-0771-5.

Verma, G. and Surolia, N. (2013) '*Plasmodium falciparum* CENH3 is able to functionally complement Cse4p and its, C-terminus is essential for centromere function', *Molecular and Biochemical Parasitology*, 192(1–2): 21–29.

Vincke, I. H. and Lips, M. (1948) 'Un nouveau plasmodium d'un rongeur sauvage du Congo *Plasmodium berghei* n. sp.', *Annales de la Societe belge de medecine tropicale (1920)*, 28(1): 97–104.

Vinetz, J. M., Valenzuela, J. G., Specht, C. A., Aravind, L., Langer, R. C., Ribeiro, J. M. and Kaslow, D. C. (2000) 'Chitinases of the avian malaria parasite *Plasmodium gallinaceum*, a class of enzymes necessary for parasite invasion of the mosquito midgut.', *The Journal of biological chemistry*, 275(14): 10331–41.

Vlachou, D., Schlegelmilch, T., Runn, E., Mendes, A. and Kafatos, F. C. (2006) 'The developmental migration of *Plasmodium* in mosquitoes', *Current Opinion in Genetics & Development*, 16(4): 384–391.

Volz, J. C. C., Bártfai, R., Petter, M., Langer, C., Josling, G. A. A., Tsuboi, T., Schwach, F., Baum, J., Rayner, J. C. C., Stunnenberg, H. G. G., Duffy, M. F. F. and Cowman, A. F. F. (2012) 'PfSET10, a *Plasmodium falciparum*

Methyltransferase, Maintains the Active *var* Gene in a Poised State during Parasite Division', *Cell Host & Microbe*, 11(1): 7–18.

Volz, J., Carvalho, T. G., Ralph, S. A., Gilson, P., Thompson, J., Tonkin, C. J., Langer, C., Crabb, B. S. and Cowman, A. F. (2010) 'Potential epigenetic regulatory proteins localise to distinct nuclear sub-compartments in *Plasmodium falciparum*', *International Journal for Parasitology*, 40(1): 109–121.

Voss, T. S., Healer, J., Marty, A. J., Duffy, M. F., Thompson, J. K., Beeson, J. G., Reeder, J. C., Crabb, B. S. and Cowman, A. F. (2006) 'A *var* gene promoter controls allelic exclusion of virulence genes in *Plasmodium falciparum* malaria.', *Nature*, 439(7079): 1004–1008.

Voss, T. S., Kaestli, M., Vogel, D., Bopp, S. and Beck, H.-P. P. (2003) 'Identification of nuclear proteins that interact differentially with *Plasmodium falciparum var* gene promoters', *Molecular Microbiology*, 48(6): 1593–1607.

Voss, T. S., Tonkin, C. J., Marty, A. J., Thompson, J. K., Healer, J., Crabb, B. S. and Cowman, A. F. (2007) 'Alterations in local chromatin environment are involved in silencing and activation of subtelomeric *var* genes in *Plasmodium falciparum*', *Molecular Microbiology*, 66(1): 139–150.

Wagner, E. J. and Carpenter, P. B. (2012) 'Understanding the language of Lys36 methylation at histone H3', *Nature Reviews Molecular Cell Biology*. doi: 10.1038/nrm3274.

Wagner, J. C., Platt, R. J., Goldfless, S. J., Zhang, F. and Niles, J. C. (2014) 'Efficient CRISPR-Cas9-mediated genome editing in *Plasmodium falciparum*', *Nat Methods*, 11(9): 915–918.

Wang, J., Lunyak, V. V. and King Jordan, I. (2013) 'BroadPeak: A novel algorithm for identifying broad peaks in diffuse ChIP-seq datasets', *Bioinformatics*. doi: 10.1093/bioinformatics/bts722.

Wang, J., Sarov, M., Rientjes, J., Fu, J., Hollak, H., Kranz, H., Xie, W., Stewart, A. F. and Zhang, Y. (2006) 'An Improved Recombineering Approach by Adding RecA to λ Red Recombination', *Molecular Biotechnology*, 32(1): 43–54.

Wang, W. L., Anderson, L. C., Nicklay, J. J., Chen, H., Gamble, M. J., Shabanowitz, J., Hunt, D. F. and Shechter, D. (2014) 'Phosphorylation and arginine methylation mark histone H2A prior to deposition during *Xenopus laevis* development', *Epigenetics and Chromatin*. doi: 10.1186/1756-8935-7-22.

Watanabe, J., Sasaki, M., Suzuki, Y. and Sugano, S. (2002) 'Analysis of transcriptomes of human malaria parasite *Plasmodium falciparum* using full-length enriched library: Identification of novel genes and diverse transcription start sites of messenger RNAs', *Gene*, 291(1–2):105–113.

Wei, H., Jiang, S., Chen, L., He, C., Wu, S. and Peng, H. (2017) 'Characterization of Cytosine Methylation and the DNA Methyltransferases of *Toxoplasma gondii*.', *International journal of biological sciences*, 13(4): 458–470.

Weirich, S., Kusevic, D., Kudithipudi, S. and Jeltsch, A. (2015) 'Investigation of the methylation of Numb by the SET8 protein lysine methyltransferase', *Scientific Reports*, 5. doi: 10.1038/srep13813.

Westenberger, S. J., Cui, L. L., Dharia, N., Winzeler, E. and Cui, L. L. (2009) 'Genome- wide nucleosome mapping of *Plasmodium falciparum* reveals histone-rich coding and histone-poor intergenic regions and chromatin remodeling of core and subtelomeric genes', *BMC Genomics*. BioMed Central, 10(1), p. 610. doi: 10.1186/1471-2164-10-610.

Wheatley, N. C., Andrews, K. T., Tran, T. L., Lucke, A. J., Reid, R. C. and Fairlie, D. P. (2010) 'Antimalarial histone deacetylase inhibitors containing cinnamate or NSAID components', *Bioorganic and Medicinal Chemistry Letters*, 20(23): 7080–7084.

White, N. J., Turner, G. D. H., Medana, I. M., Dondorp, A. M. and Day, N. P. J. (2010) 'The murine cerebral malaria phenomenon', *Trends in Parasitology*, 26(1): 11–15.

World Health Organisation (2018) World Malaria Report 2018. Geneva: WHO Press.

Wiles, E. T. and Selker, E. U. (2017) 'H3K27 methylation: a promiscuous repressive chromatin mark', *Current Opinion in Genetics and Development*, pp. 31–37. doi: 10.1016/j.gde.2016.11.001.

Williams, J. L. (1999) 'Stimulation Of *Plasmodium falciparum* Gametocytogenesis By Conditioned Medium From Parasite Cultures', *Am. J. Trop. Med. Hyg*, 60(1), pp. 7–13. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.608.1384&rep=rep1 &type=pdf (Accessed: 12 January 2018).

Wong, W., Bai, X. chen, Brown, A., Fernandez, I. S., Hanssen, E., Condron, M., Tan, Y. H. on., Baum, J. and Scheres, S. H. W. (2014) 'Cryo-EM structure of the *Plasmodium falciparum* 80S ribosome bound to the anti-protozoan drug emetine', *eLife*, 3. doi: 10.7554/eLife.03080.

Wongsombat, C., Aroonsri, A., Kamchonwongpaisan, S., Morgan, H. P., Walkinshaw, M. D., Yuthavong, Y. and Shaw, P. J. (2014) 'Molecular characterization of *Plasmodium falciparum* Bruno/CELF RNA binding proteins', *Molecular and Biochemical Parasitology*, 198(1): 1–10.

Wu, J., Cui, N., Wang, R., Li, J. and Wong, J. (2012) 'A role for CARM1-mediated histone H3 arginine methylation in protecting histone acetylation by releasing corepressors from chromatin', *PLoS ONE*, 7(6). doi: 10.1371/journal.pone.0034692.

Wykes, M. N. and Good, M. F. (no date) 'Infectious disease: Malaria What have we learnt from mouse models for the study of malaria?' doi: 10.1002/eji.200939552.

Xu, S., Grullon, S., Ge, K. and Peng, W. (2014) 'Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells.', *Methods in molecular biology (Clifton, N.J.)*. NIH Public Access, 1150, pp. 97–111. doi: 10.1007/978-1-4939-0512-6_5.

Yam, X. Y., Brugat, T., Siau, A., Lawton, J., Wong, D. S., Farah, A., Twang, J. S., Gao, X., Langhorne, J. and Preiser, P. R. (2016) 'Characterization of the *Plasmodium* Interspersed Repeats (PIR) proteins of *Plasmodium chabaudi* indicates functional diversity', *Scientific Reports*, 6. doi: 10.1038/srep23449.

Yang, A. S. P., Lopaticki, S., O'Neill, M. T., Erickson, S. M., Douglas, D. N., Kneteman, N. M. and Boddey, J. A. (2017) 'AMA1 and MAEBL are important for *Plasmodium falciparum* sporozoite infection of the liver', *Cellular Microbiology*, 19(9). doi: 10.1111/cmi.12745.

Yap, A., Azevedo, M. F., Gilson, P. R., Weiss, G. E., O'Neill, M. T., Wilson, D. W., Crabb, B. S. and Cowman, A. F. (2014) 'Conditional expression of apical membrane antigen 1 in *Plasmodium falciparum* shows it is required for erythrocyte invasion by merozoites', *Cellular Microbiology*, 16(5): 642–656.

Yeoh, L. M., Goodman, C. D., Mollard, V., McFadden, G. I. and Ralph, S. A. (2017) 'Comparative transcriptomics of female and male gametocytes in *Plasmodium berghei* and the evolution of sex in alveolates.', *BMC genomics*. BioMed Central, 18(1), p. 734. doi: 10.1186/s12864-017-4100-0.

Young, J. A., Fivelman, Q. L., Blair, P. L., De La Vega, P., Le Roch, K. G., Zhou, Y., Carucci, D. J., Baker, D. A. and Winzeler, E. A. (2005) 'The *Plasmodium falciparum* sexual development transcriptome: A microarray analysis using ontology-based pattern identification', *Molecular & Biochemical Parasitology*, 143, pp. 67–79. doi: 10.1016/j.molbiopara.2005.05.007.

Yu, Q., Olsen, L., Zhang, X., Boeke, J. D. and Bi, X. (2011) 'Differential contributions of histone H3 and H4 residues to heterochromatin structure', *Genetics*. doi: 10.1534/genetics.111.127886.

Yuan, Z.-F., Lin, S., Molden, R. C., Cao, X.-J., Bhanu, N. V., Wang, X., Sidoli, S., Liu, S. and Garcia, B. A. (2015) 'EpiProfile Quantifies Histone Peptides With Modifications by Extracting Retention Time and Intensity in High-resolution Mass Spectra', *Molecular & Cellular Proteomics*, 14(6), pp. 1696–1707. doi: 10.1074/mcp.M114.046011.

Yuan, Z. F., Lin, S., Molden, R. C. and Garcia, B. A. (2014) 'Evaluation of proteomic search engines for the analysis of histone modifications', *Journal of Proteome Research*, 13(10), pp. 4470–4478. doi: 10.1021/pr5008015.

Yuda, M., Iwanaga, S., Kaneko, I. and Kato, T. (2015) 'Global transcriptional repression: An initial and essential step for *Plasmodium* sexual development.', *Proceedings of the National Academy of Sciences of the United States of America*, 112(41): 12824–9.

Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T. and Kaneko, I. (2010) 'Transcription factor AP2-Sp and its target genes in malarial sporozoites', *Molecular Microbiology*, 75(4), pp. 854–863. doi: 10.1111/j.1365-2958.2009.07005.x.

Yuda, M., Iwanaga, S., Shigenobu, S., Mair, G. R., Janse, C. J., Waters, A. P., Kato, T. and Kaneko, I. (2009) 'Identification of a transcription factor in the mosquito- invasive stage of malaria parasites', *Molecular Microbiology*, 71(6):1402–1414.

Zeng, W., Ball, A. R., Yokomori, K. and Yokomori, K. (2010) 'HP1: heterochromatin binding proteins working the genome.', *Epigenetics*, 5(4): 287–92.

Zhang, K. and Rathod, P. K. (2002) 'Divergent regulation of dihydrofolate reductase between malaria parasite and human host', *Science*, 296(5567): 545–547.

Zhang, M., Fennell, C., Ranford-Cartwright, L., Sakthivel, R., Gueirard, P., Meister, S., Caspi, A., Doerig, C., Nussenzweig, R. S., Tuteja, R., Sullivan, W. J., Roos, D. S., Fontoura, B. M. A., Ménard, R., Winzeler, E. A. and Nussenzweig, V. (2010) 'The *Plasmodium* eukaryotic initiation factor-2α kinase IK2 controls the latency of sporozoites in the mosquito salivary glands', *The Journal of Experimental Medicine*, 207(7): 1465–1474.

Zhang, M., Joyce, B. R., Sullivan, W. J. and Nussenzweig, V. (2013) 'Translational control in *Plasmodium* and *Toxoplasma* parasites', *Eukaryotic Cell*, pp. 161–167. doi: 10.1128/EC.00296-12.

Zhang, M., Wang, C., Otto, T. D., Oberstaller, J., Liao, X., Adapa, S. R., Udenze, K., Bronner, I. F., Casandra, D., Mayho, M., Brown, J., Li, S., Swanson, J., Rayner, J. C., Jiang, R. H. Y. and Adams, J. H. (2018) 'Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis.', *Science*, 360(6388), p. eaap7847. doi: 10.1126/science.aap7847.

Zhang, Q., Ma, C., Oberli, A., Zinz, A., Engels, S. and Przyborski, J. M. (2017) 'Proteomic analysis of exported chaperone/co-chaperone complexes of *P. falciparum* reveals an array of complex protein-protein interactions', *Scientific Reports*, 7. doi: 10.1038/srep42188.

Zhang, Q., Siegel, T. N., Martins, R. M., Wang, F., Cao, J., Gao, Q., Cheng, X., Jiang, L., Hon, C.-C., Scheidig-Benatar, C., Sakamoto, H., Turner, L., Jensen, A. T. R., Claes, A., Guizetti, J., Malmquist, N. A. and Scherf, A. (2014) 'Exonuclease- mediated degradation of nascent RNA silences genes linked to severe malaria', *Nature*, 513(7518): 431–435.

Zhang, Y., Griffin, K., Mondal, N. and Parvin, J. D. (2004) 'Phosphorylation of histone H2A inhibits transcription chromatin templates', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M400099200.

Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W. and Liu, X. S. (2008) 'Model-based analysis of ChIP-Seq (MACS).', *Genome biology*. BioMed Central, 9(9), p. R137. doi: 10.1186/gb-2008-9-9-r137.

Zhang, Z. and Lutz, B. (2002) 'Cre recombinase-mediated inversion using lox66 and lox71: method to introduce conditional point mutations into the CREB-binding protein', *Nucleic Acids Research*, 30(17): e90. doi: 10.1093/nar/gnf089.

Zhao, J., Bhanot, P., Hu, J. and Wang, Q. (2016) 'A comprehensive analysis of *Plasmodium* circumsporozoite protein binding to hepatocytes', *PLoS ONE*, 11(8). doi: 10.1371/journal.pone.0161607.

Zheng, W., Kou, X., Du, Y., Liu, F., Yu, C., Tsuboi, T., Fan, Q., Luo, E., Cao, Y. and Cui, L. (2016) 'Identification of three ookinete-specific genes and evaluation of their transmission-blocking potentials in *Plasmodium berghei'*, *Vaccine*, 34(23): 2570– 2578.

Zolg, J. W., Plitt, J. R., Chen, G. X. and Palmer, S. (1989) 'Point mutations in the dihydrofolate reductase-thymidylate synthase gene as the molecular basis for pyrimethamine resistance in *Plasmodium falciparum*.', *Molecular and biochemical parasitology*, 36(3): 253–62.