Investigation of NOT1 proteins in regulating gene expression in plasmodium falciparum

Liu, Ying

2019

Liu, Y. (2019). Investigation of NOT1 proteins in regulating gene expression in plasmodium falciparum. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/106849>

<https://doi.org/10.32657/10220/49669>

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

**Investigation of NOT1 proteins in regulating gene expression
in *Plasmodium falciparum***

LIU YING

SCHOOL OF BIOLOGICAL SCIENCES

2019

**Investigation of NOT1 proteins in regulating gene expression
in *Plasmodium falciparum***

LIU YING

SCHOOL OF BIOLOGICAL SCIENCES

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the degree of
Doctor of Philosophy

2019

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me except where otherwise stated in this thesis. The thesis work has not been submitted for a degree or professional qualification to any other university or institution. I declare that this thesis is written by myself and is free of plagiarism and of sufficient grammatical clarity to be examined. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

16-01-2019.

Date

Liu Ying / Liu Yj

Name/Signature

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it of sufficient grammatical clarity to be examined. To the best of my knowledge, the thesis is free of plagiarism and the research and writing are those of the candidate's except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

.....
Date
16-01-2019.

.....
Supervisor
Name/Signature

Thesis Title:

Investigation of NOT1 proteins in regulating gene expression in *Plasmodium falciparum*

I have excluded the Authorship Attribution Statement declaration because the thesis does not contain any published material.

Liu Yg
[Student Signature]

Liu Ying
[Student Name]

Date : 17-1-2019

[Supervisor Signature]

Zbyněk Boždek
[Supervisor Name]

Date: 17-1-2019.

Acknowledgment

I express my heartfelt gratitude to Prof. Mark Featherstone for offering me the chance to start my research career during undergraduate studies and then to conduct this project as a post-graduate student. I would like to thank him for his patient guidance, invaluable suggestions and unconditional support.

I would also like to express my gratitude to Prof. Zbynek Bozdech for accepting me in his lab in the last year of my project. His intellectual guide really supports me to finish this project.

I would like to thank the members of my thesis advisory committee, Prof. Zbynek Bozdech and Prof. Francesc Xavier Roca Castella for their valuable suggestions on the design of the projects.

I would also like to thank my lab mates for the support throughout the projects. I would give special thanks to Dr. Ragini Rai, Dr. Sandeep Srivastava, Dr. Uthra Iyer, Dr. Li Xiaolei and Dr. Zhang Changqing for training me in malaria culture and teaching me techniques in molecular biology. I would also thank my lab mates in Prof. Zbynek Bozdech's lab, Dr. Frances Rocamora, Dr. Zhu Lei and Mr. Michael Kucharski for their patient guide in microarray and bioinformatics analysis.

I also thank the School of Biological Sciences (SBS) in Nanyang Technological University for offering me scholarship to pursue this Doctor of Philosophy degree.

Last but not least, I express my sincere gratitude to my family and friends for their unconditional comfort that supports me to overcome struggling time and finally complete this project.

Table of Contents

Acknowledgment	i
Table of Contents	ii
Summary	vii
List of Figures	viii
List of Tables	x
List of Abbreviations.....	xii

Chapter 1 Introduction

1.1 Overview of <i>P. falciparum</i>	1
1.1.1 Life cycle of <i>P. falciparum</i>	1
1.1.2 Morphology of <i>P. falciparum</i> during the IDC.....	2
1.2 Gene regulation in <i>P. falciparum</i>	7
1.2.1 Epigenetic regulation	9
1.2.2 Transcription	11
1.2.3 Post-transcriptional regulation	14
1.3 The CCR4-NOT complex in eukaryotes	18
1.3.1 Components of the CCR4-NOT complex.....	18
1.3.2 Structure of the CCR4-NOT complex	18
1.3.3 Functions of the CCR4-NOT complex.....	20
1.3.4 Functions of NOT1 in eukaryotes.....	23
1.3.5 Functions of other subunits in eukaryotes	25
1.3.6 Proposed mechanisms of the action of CCR4-NOT	28
1.4 The CCR4-NOT complex in <i>P. falciparum</i>	29
1.4.1 Orthologs of the CCR4-NOT subunits in <i>P. falciparum</i>	29
1.4.2 Conserved domains in NOT1.1 and NOT1.2	32
1.4.3 Studies of the CCR4-NOT complex in <i>Plasmodium</i> parasites	38
1.5 Research rationales and objectives	39
1.5.1 To identify potential interacting partners of NOT1.1 and NOT1.2	40
1.5.2 To investigate the effect of <i>Not1</i> mutations on recruitment of RNAPII.....	41
1.5.3 To determine the effect of <i>Not1</i> mutations on transcript abundance	41

1.5.4 To outline the function of NOT1 as the scaffold for its interacting partners	42
--	----

Chapter 2 Materials and methods

2.1 Cloning and plasmid constructs.....	43
2.1.1 Restriction digestion and ligation-independent cloning	43
2.1.2 Transformation of competent cells	43
2.1.3 PCR amplification	44
2.1.4 pLNOT1.2-3HA construct	44
2.1.5 pTNOT1.1-FLAG construct	45
2.1.6 pTNOT1.1-3HA construct.....	45
2.1.7 pTNOT4-3HA construct.....	45
2.1.8 pCC1-NOT1.1-KO construct.....	45
2.1.9 pCC1-NOT1.2-KO construct.....	46
2.1.10 pLNOT1.2-HA-ecDDD construct.....	46
2.1.11 pThDHFR construct	46
2.2 Culture, transfection and limiting dilution.....	47
2.2.1 Malaria culture, blood smear and parasite counting	47
2.2.2 Synchronization, cryopreserve and lysis.....	47
2.2.3 Stable transfection.....	48
2.2.4 Drug cycling.....	48
2.2.5 Limiting dilution	49
2.3 DNA work	49
2.3.1 Genomic DNA extraction	49
2.3.2 Southern blot	50
2.4 RNA work	51
2.4.1 Total RNA extraction	51
2.4.2 Reverse transcription.....	52
2.4.3 Real-time PCR.....	52
2.4.4 Northern blot	53
2.5 Microarray.....	54
2.5.1 cDNA preparation by SMART amplification.....	54
2.5.2 Cy-dye labeling.....	55

2.5.3 Hybridization	56
2.5.4 Microarray scanning and processing	57
2.6 Protein work.....	57
2.6.1 Protein extraction	57
2.6.2 Western blot	58
2.6.3 Co-immunoprecipitation.....	60
2.6.4 Mass spectrometry	61
2.7 ChIP-on-chip	62
2.7.1 Crosslinking	62
2.7.2 Chromatin-immunoprecipitation.....	62
2.7.3 DNA extraction and random amplification	64
2.8 Primer list	65

Chapter 3 Co-immunoprecipitation of CCR4-NOT subunits

3.1 NOT1.2 might be a scaffold of the CCR4-NOT complex.....	67
3.1.1 HA-tagging of endogenous <i>Not1.2</i>	67
3.1.2 Expression of HA-tagged NOT1.2	69
3.1.3 Co-immunoprecipitation of HA-tagged NOT1.2	71
3.2 Current progress of tagging endogenous <i>Not1.1</i>	77
3.2.1 Failure to detect FLAG-tagged NOT1.1	77
3.2.2 HA-tagging of endogenous <i>Not1.1</i>	80
3.3 CAF1 is a binding partner of NOT1.1 and NOT1.2	81
3.3.1 HA-tagging of endogenous <i>Caf1</i>	81
3.3.2 Co-immunoprecipitation of HA-tagged CAF1	82
3.4 NOT4 might be a subunit of the CCR4-NOT complex	85
3.4.1 HA-tagging of endogenous <i>Not4</i>	85
3.4.2 Co-immunoprecipitation of HA-tagged NOT4	87

Chapter 4 Functional analysis of NOT1.1 and NOT1.2

4.1 Creation of the <i>Not1.1</i> mutant	92
4.1.1 Disruption of endogenous <i>Not1.1</i> by a selectable marker	92

4.1.2 Validation of the knockout of <i>Not1.1</i>	94
4.2 Phenotypes of the <i>Not1.1</i> mutant	99
4.3 Differential mRNA abundance in the <i>Not1.1</i> mutant	104
4.3.1 Knockout of <i>Not1.1</i> affects the expression of 715 genes.....	104
4.3.2 Functional analysis of the genes influenced by NOT1.1.....	107
4.3.3 Validation of microarray data by RT-qPCR in the <i>Not1.1</i> mutant	110
4.3.4 NOT1.1 may be required for setting the mRNA abundance of egress- and invasion-related genes	111
4.4 Knockout of <i>Not1.1</i> affects recruitment of RNAPII.....	113
4.5 Creation of the <i>Not1.2</i> mutant	118
4.5.1 Knockout of endogenous <i>Not1.2</i> gene was not successful	118
4.5.2 Conditional knockdown system	119
4.5.3 Endogenous <i>Not1.2</i> was tagged with HA epitopes and destabilizing domain ..	121
4.5.4 Knockdown of ecDDD-tagged NOT1.2 in the absence of stabilizing ligand	124
4.6 Phenotype of the <i>Not1.2</i> mutant	127
4.7 Differentially expressed genes in the <i>Not1.2</i> mutant	131
4.7.1 Knockdown of NOT1.2 affects the expression of 202 genes	131
4.7.2 Validation of microarray data by RT-qPCR in the <i>Not1.2</i> mutant	134

Chapter 5 Discussions

5.1 Proposed architecture of the CCR4-NOT complex in <i>P. falciparum</i>	136
5.1.1 Components of the CCR4-NOT complex in <i>P. falciparum</i>	136
5.1.2 Architecture of the CCR4-NOT complex in <i>P. falciparum</i>	138
5.1.3 Candidate interacting partners and potential functions of the CCR4-NOT complex in <i>P. falciparum</i>	139
5.2 Hypothesized functions of NOT1 paralogs in gene regulation in <i>P. falciparum</i>	143
5.3 Functional correlation of NOT1.1 and NOT1.2	145
5.3.1 NOT1.1 and NOT1.2 have opposite functions in regulating mRNA abundance	145
5.3.2 NOT1.1 and NOT1.2 may have opposite functions in gametocytogenesis	149
5.4 NOT1.1 and NOT1.2 are functionally correlated with their interacting partners	151
5.4.1 Functional correlation with CAF1	151

5.4.2 Functional correlation with ALBA1	153
5.5 Proposed mechanism of NOT1.1- and NOT1.2-mediated gene regulation	155
5.6 Improvement and future perspectives	157
 Chapter 6 Conclusion	 161
Appendix	162
Reference	224

Summary

Plasmodium falciparum is the most deadly form of malaria that threatens millions of people around the world. The CCR4-NOT complex is a highly conserved complex that regulates gene expression across the eukaryotic kingdom and NOT1 is the scaffold of the complex. *P. falciparum* encodes two paralogs of NOT1, NOT1.1 (PF3D7_1103800) and NOT1.2 (PF3D7_1417200). My study investigates the function of two NOT1 proteins in gene regulation during the 48-hour intraerythrocytic developmental cycle. I found that NOT1.1 and NOT1.2 are probably the scaffold of the complex in *P. falciparum*, interacting with CAF1, CCR4, CAF40 as well as NOT4 subunits. The independent mutation of *Not1.1* and *Not1.2* leads to altered mRNA abundance of many genes, including many egress- and invasion-related genes, which may result in defective growth and invasion rate. The knockout of *Not1.1* affects recruitment of the elongating form of RNA Polymerase II, which however is not significantly correlated to differential mRNA abundance, suggesting the role of NOT1.1 in post-transcriptional regulation. In addition, comparison between the differential mRNA abundance in the two mutants reveals a negative correlation, demonstrating opposite functions of NOT1.1 and NOT1.2 in gene regulation. Furthermore, I found that differential mRNA abundance in the two mutants is correlated to that in *Caf1* mutant and *Alba1* mutant. I propose that NOT1.1 and NOT1.2 compete to be the scaffold of the CCR4-NOT complex and regulate gene expression post-transcriptionally by modulating the functions of their interacting proteins, including CAF1 and ALBA1, which are key players in mRNA metabolism.

List of Figures

Figure 1.1 Life cycle of <i>P. falciparum</i>	1
Figure 1.2 Three-dimensional organization of the merozoite	3
Figure 1.3 Three-dimensional organization of the ring	5
Figure 1.4 Three-dimensional organization of the trophozoite	6
Figure 1.5 Three-dimensional organization of the schizont stage	7
Figure 1.6 The phaseogram of transcriptional profiling in <i>P. falciparum</i>	8
Figure 1.7 Architecture of the CCR4-NOT complex in eukaryotes	19
Figure 1.8 Expression of <i>Not1.1</i> and <i>Not1.2</i> during one IDC	30
Figure 1.9 Alignment of protein sequences of the MIF4G domain in NOT1	34
Figure 1.10 Alignment of protein sequences of the DUF3819 domain in NOT1	35
Figure 1.11 Alignment of protein sequences of the NOT1 domain in NOT1	36
Figure 1.12 Conserved domains of yeast NOT1, NOT1.1 and NOT1.2	37
Figure 3.1 HA-tagging of endogenous <i>Not1.2</i>	68
Figure 3.2 Expression of tagged <i>Not1.2</i> at the transcript level	69
Figure 3.3 Expression of HA-tagged NOT1.2 protein	70
Figure 3.4 Immunoprecipitation of HA-tagged NOT1.2	72
Figure 3.5 FLAG-tagging of endogenous <i>Not1.1</i>	77
Figure 3.6 Unspecific detection of the FLAG tag in the T996 strain	78
Figure 3.7 Unspecific immunoprecipitation of the FLAG tag	79
Figure 3.8 HA-tagging of endogenous <i>Not1.1</i>	80
Figure 3.9. HA-tagging of endogenous <i>Caf1</i>	81
Figure 3.10 Immunoprecipitation of HA-tagged CAF1	83
Figure 3.11 HA-tagging of endogenous <i>Not4</i>	86
Figure 3.12 Expression of HA-tagged NOT4	87
Figure 3.13 Immunoprecipitation of HA-tagged NOT4	88
Figure 4.1 Disruption of endogenous <i>Not1.1</i>	93
Figure 4.2 Southern blot on the <i>Not1.1</i> mutant	95
Figure 4.3 Northern blot on the <i>Not1.1</i> mutant	97
Figure 4.4 Northern blot with the control probe on the <i>Not1.1</i> mutant	98
Figure 4.5 Proliferation of the <i>Not1.1</i> mutant	100
Figure 4.6 Invasion rate of the <i>Not1.1</i> mutant	101
Figure 4.7 Merozoite number per schizont in the <i>Not1.1</i> mutant	102
Figure 4.8 Progression of parasites in the <i>Not1.1</i> mutant	103
Figure 4.9 Significantly differential mRNA abundance in the <i>Not1.1</i> mutant	106

Figure 4.10 Validation of differential mRNA abundance in the <i>Not1.1</i> mutant.....	110
Figure 4.11 Egress- and invasion-related genes with significantly differential mRNA abundance in the <i>Not1.1</i> mutant	112
Figure 4.12 Chromatin-immunoprecipitation of RNAPII tested by Western blot	114
Figure 4.13 Correlation of RNAPII recruitment and mRNA abundance in the <i>Not1.1</i> mutant	117
Figure 4.14. Failure to disrupt endogenous <i>Not1.2</i>	118
Figure 4.15 Design of conditional knockdown of NOT1.2	120
Figure 4.16 Tagging of endogenous <i>Not1.2</i> with the destabilizing domain	122
Figure 4.17 Expression of tagged <i>Not1.2</i> gene at the transcript level	123
Figure 4.18 Expression of NOT1.2-HA-ecDDD in the presence or absence of TMP ...	125
Figure 4.19 Proliferation of the <i>Not1.2</i> mutant.....	128
Figure 4.20 Invasion rate of the <i>Not1.2</i> mutant	129
Figure 4.21 Merozoite number per schizont in the <i>Not1.2</i> mutant	130
Figure 4.22 Progression of parasites in the <i>Not1.2</i> mutant	131
Figure 4.23 Significantly differential mRNA abundance in <i>Not1.2</i> mutant	133
Figure 4.24 Validation of differential mRNA abundance in the <i>Not1.2</i> mutant.....	135
Figure 5.1 Proposed architecture of the CCR4-NOT complexes in <i>P. falciparum</i>	138
Figure 5.2 Overlap of genes with differential mRNA abundance in the <i>Not1.1</i> mutant and the <i>Not1.2</i> mutant	146
Figure 5.3 Opposite functions of NOT1.1 and NOT1.2 in regulating mRNA abundance	148
Figure 5.4 Correlation between proportion of gametocyte and expression of <i>Not1.1</i> and <i>Not1.2</i>	150
Figure 5.5 Correlation of NOT1.1, CAF1 and NOT1.2 in regulating mRNA abundance	152
Figure 5.6 Correlation of NOT1.1, ALBA1 and NOT1.2 in regulating mRNA abundance	154
Figure 5.7 Proposed mechanism of NOT1.1- and NOT1.2-mediated gene regulation	155

List of Tables

Table 1.1 Predicted CCR4-NOT subunits in <i>P. falciparum</i>	29
Table 3.1 Enriched functional pathways of NOT1.2 potential interacting partners	73
Table 3.2 Enriched functional pathways of NOT4 potential interacting partners	90
Table 4.1 Number of differential mRNA abundance in the <i>Not1.1</i> mutant	105
Table 4.2 Number of differential RNAPII recruitment in the <i>Not1.1</i> mutant.....	116
Table 4.3 Number of differential mRNA abundance in the <i>Not1.2</i> mutant	132
Table 5.1 Interactions between the core subunits of the CCR4-NOT complex.....	136
Table 5.2 Potential interacting proteins of the CCR4-NOT complex	139
Supplementary Table 1 Similarity between the conserved domains of NOT1 in yeast, human, <i>Drosophila</i> and <i>P. falciparum</i>	162
Supplementary Table 2 20 NOT1.2 potential interacting proteins at schizont stage	163
Supplementary Table 3 244 CAF1 potential interacting proteins at schizont stage ..	164
Supplementary Table 4 Enriched functions of CAF1 potential interacting proteins..	174
Supplementary Table 5 45 NOT4 potential interacting proteins at schizont stage ...	178
Supplementary Table 6 Proliferation of the <i>Not1.1</i> mutant	180
Supplementary Table 7 Invasion rate of the <i>Not1.1</i> mutant.....	181
Supplementary Table 8 Merozoite number per schizont in the <i>Not1.1</i> mutant.....	182
Supplementary Table 9 Estimated age of parasites harvested at 6 timepoints across one IDC in the <i>Not1.1</i> mutant.....	183
Supplementary Table 10 715 genes with significantly differential mRNA abundance in the <i>Not1.1</i> mutant	184
Supplementary Table 11 Differential mRNA abundance assessed by microarray and RT-qPCR in the <i>Not1.1</i> mutant	202
Supplementary Table 12 150 genes with significantly differential RNAPII recruitment in <i>Not1.1</i> mutant.....	203
Supplementary Table 13 Proliferation of the <i>Not1.2</i> mutant	208
Supplementary Table 14 Invasion rate of the <i>Not1.2</i> mutant.....	209
Supplementary Table 15 Merozoite number per schizont in the <i>Not1.2</i> mutant.....	210

Supplementary Table 16 Estimated age of parasites harvested at 6 timepoints across one IDC in the <i>Not1.2</i> mutant.....	211
Supplementary Table 17 202 genes with significantly differential mRNA abundance in the <i>Not1.2</i> mutant	212
Supplementary Table 18 Differential mRNA abundance assessed by microarray and RT-qPCR in the <i>Not1.2</i> mutant	220
Supplementary Table 19 114 genes with significantly differential mRNA abundance at the same timepoint in the <i>Not1.1</i> mutant and the <i>Not1.2</i> mutant.	221

List of Abbreviations

Abbreviation	Name
5-FC	5-fluorocytosine
ACT	Artemisinin-based combination therapy
ALBA	Acetylation lowers binding affinity
AMA1	Apical membrane antigen-1
Bam	Bag-of-marbles
CDPK5	Calcium-dependent protein kinase 5
CELF	CUG-BP, Elav-like family
ChIP	Chromatin-immunoprecipitation
CITH	CAR-I/Trailer Hitch Homolog
CK2 β 2	Casein kinase II beta chain-2
CPSF3	Cleavage and polyadenylation specificity factor-3
CTD	C-terminal domain of RPB1
DCP2	mRNA-decapping enzyme-2
DHFR	Dihydrofolate reductase
EBAs	Erythrocyte binding antigens
EBL1	Erythrocyte binding like protein-1
ecDDD	<i>E.coli</i> DHFR destabilizing domain
EEP	Exonuclease-endonuclease-phosphatase
eIF	Eukaryotic translation initiation factor
EMP1	Erythrocyte membrane protein-1
ER	Endoplasmic reticulum
GAP	Glideosome-associated protein
GTF	General transcription factor
HFD	Histone fold domain
Hi-C	Chromosome conformation capture
hnRNP	Heterogeneous nuclear ribonucleoprotein
HP1	Heterochromatin protein-1
hpi	Hour post-invasion

HRP	Horseradish peroxidase
HSP	Heat shock protein
IDC	Intraerythrocytic developmental cycle
iRBC	Infected RBC
KAHRP	Knob-associated histidine-rich protein
LIC	Ligation-independent cloning
lncRNA	Long non-coding RNA
MIF4G	Middle domain of eIF4G
MS	Mass spectrometry
MSP	Merozoite surface protein
NAC	Nascent associated polypeptide-associated complex
NTF	Nuclear transport factor
ORF	Open reading frames
PABP	Polyadenylate-binding protein
PAF1	RNAPII-associated factor-1
PAIP1	PABP-interacting protein-1
PIC	Pre-initiation complex
PKG	Protein kinase G
PM	Plasma membrane
PRP	Pre-mRNA processing protein
PTEX	<i>Plasmodium</i> translocon of exported proteins
PUF	Pumilio and fem-3 binding factor
PUS	RNA pseudouridylate synthase
PV	Parasitophorous vacuole
PVM	Parasitophorous vacuole membrane
RBC	Red blood cell
RH	Reticulocyte binding-like homologues
RNAPII	RNA polymerase II
RON	Rhoptry neck protein
RT-qPCR	Reverse Transcription Quantitative PCR
ScFCU	<i>S.cerevisiae</i> cytosine deaminase/uracil phosphoribosyltransferase

SERA	Serine repeat antigen
S/N	Supernatant
SR	bona fide Ser/Arg-rich protein
SUB1	Subtilisin-like protease 1
TAF	TBP-associated factors
TAP	Transcription activating proteins
TBP	TATA-binding protein
TFII	Transcription Factor II
TMP	Trimethoprim
TP	Timepoint
TSS	Transcription start site
TTP	Tristetraprolin
uORF	Upstream open reading frame
UTR	Untranslated regions

Chapter 1 Introduction

Malaria is one of the most severe diseases in the world and causes global health concerns. According to the latest estimate, 216 million cases were reported in 2016, leading to 445,000 deaths [1]. The artemisinin-based combination therapy (ACT) is used worldwide to eradicate this leading cause of childhood mortality [2]. However, this first-line treatment is now threatened by the emergence of artemisinin resistance in many countries [3]. Unfortunately, no alternative and fully effective therapies are currently available to replace the ACT.

1.1 Overview of *P. falciparum*

1.1.1 Life cycle of *P. falciparum*

Malaria is caused by infection with the unicellular *Apicomplexan* parasites in the genus *Plasmodium*. Of five species that infect humans, *Plasmodium falciparum* is the most lethal form. The life cycle of *P. falciparum* consists of two stages, a sexual stage in *Anopheles spp.* mosquitoes and an asexual stage in humans (**Figure 1.1**, adapted from Pasvol, 2010 [4], copyright license number:4504600019467).

Figure 1.1. Life cycle of *P. falciparum* [4]. The parasite undergoes repeated asexual cycles in RBCs, giving rise to clinical symptoms. A portion of parasites develops into gametocytes, taken up by mosquitoes, in which the sexual stage is completed to form sporozoites, ready to inject into a human host via a blood meal.

The sporozoite form is injected into a human host during a blood meal. Sporozoites migrate to the liver via the bloodstream and infect hepatocytes, where they divide into thousands of merozoites. These merozoites are then released to invade red blood cells (RBCs), thus beginning the asexual intraerythrocytic developmental cycle (IDC). The parasite passes through ring, trophozoite and schizont stages in a period of 48 hours and eventually releases up to 32 new merozoites, ready for another round of RBC invasion. All clinical symptoms arise during the asexual stage [5].

A fraction of parasites in the asexual stage develops into male and female gametocytes and is then taken up by a mosquito during a blood meal. The gametocytes develop into gametes in the insect gut and fuse to form the diploid zygote. The zygote then develops into the invasive ookinete form, which traverses the midgut and transforms into an oocyst from which multiple haploid sporozoites are released. These sporozoites migrate to the salivary glands for injection into a human host [5].

1.1.2 Morphology of *P. falciparum* during the IDC

Four stages, reflected in the changing morphology of the parasite, traverse in the IDC cycle: merozoites, ring, trophozoite and schizont. The transformation between morphologies is a dynamic process across the 48-hour IDC. The ultrastructure of the parasite in the four stages was reported in 2000 [6], from which **Figure 1.2-1.5** are adapted (copyright license number: 4504611206050).

Merozoite

The morphology of the merozoite is shown in **Figure 1.2**. The nucleus enclosed in the nuclear envelope has no discernable nucleolus. One mitochondrion and one plastid lie beneath the subpellicular microtubules and plenty of free ribosomes are located in the space between the nucleus and the apical prominence [6].

Three membrane layers are present: the outermost merozoite coat containing filaments, mostly likely merozoite surface protein-1 (MSP-1), which is required to capture an erythrocyte [7], the outer plasma membrane (PM), and the inner pellicular cisterna, which might function in RBC invasion. Actin and myosin are found in the space between the PM and the pellicular cisterna, perhaps as a part of the actin-myosin motor responsible for propelling the merozoite into RBCs during invasion [6].

The apical end of the merozoite is central to the invasion process and contains special organelles, including rhoptry, microneme and dense granules. They are all membranous vesicles and discharge their contents during invasion [6].

Figure 1.2. Three-dimensional organization of the merozoite [6]. The major organelles and cytoskeletons are labeled. The right corner shows the process of a merozoite invading a RBC. The merozoite attaches the RBC membrane, reorients the apical end towards the RBC and enters. In the RBC, the merozoite transforms into a ring.

The RBC ruptures to release mature parasites in the form of merozoite, which are ready for a new round of invasion. Several proteins have been shown to be important in merozoite egress. Subtilisin-like protease 1 (SUB1), a serine protease, is activated by cGMP-dependent protein kinase G (PKG) and discharged into the parasitophorous vacuole (PV) [8]. SUB1 activates two well-known targets by cleavage. First, a cysteine protease, serine repeat antigen-5 (SERA5) mediates egress by cleaving β -spectrin, a major component of the erythrocyte cytoskeleton [9, 10]. Second, processed MSP1 on the merozoite surface binds spectrin to facilitate egress [11]. In addition to the phosphorylation of SUB1 by PKG, calcium-dependent protein kinase-5 (CDPK5) phosphorylates SERA5, which is required for egress [12].

Merozoite invasion is achieved by a series of orchestrated and regulated events. In the first step, initial attachment allows the merozoite to bind receptors on the erythrocyte. MSPs form a structural coat on the surface of merozoite and could play a role in the first phase of invasion [13]. In the second step, the tight attachment is achieved by two major families of adhesins. One family is EBAs released from the microneme, including erythrocyte binding antigen-140 (EBA140), EBA175, EBA181 and erythrocyte binding like protein-1 (EBL1). The other family is Reticulocyte binding-like homologues (RHs) released from the rhoptry, including RH1, RH2a, RH2b, RH4 and RH5. These adhesins bind different receptors on the erythrocyte and cooperatively mediate invasion [14]. RH5 plays an essential role in invasion with its Basigin receptor [15]. In addition to these two families, TRAP ligand binds CD108 receptor to assist the movement of the parasite into the erythrocyte [16]. In the third step, the merozoite reorientates its apical end to face the erythrocyte, possibly mediated by apical membrane antigen-1 (AMA1) [17, 18]. In the fourth step, tight junction is formed irreversibly between AMA1 on the merozoite surface and rhoptry neck proteins (RONs), which are injected by the parasite onto the erythrocyte surface [19]. In the fifth step, the merozoite propels itself into the erythrocyte via glideosome, which comprises glideosome-associated proteins (GAPs). GAPs are embedded in the inner membrane complex and interact with myosin A motor protein [20]. Bridging proteins, including aldolases and GAPDH, link the cytoplasmic tail of the merozoite ligands to the actin filament [21]. The myosin A motor protein glides on the actin filament to move the merozoite into the erythrocyte [22].

Ring

After invasion, the merozoite form is flattened into a cup-shaped ring form (**Figure 1.3**). The parasite begins to feed on the surrounding erythrocytic contents through the cytostome, a small dense ring at the surface of the parasite. Small portions of RBC cytosol together with the PV membrane (PVM) and the parasite PM are pulled into the parasite through this structure, pinched off as food vacuoles and then digested [23, 24]. The vacuoles either originated from the cytostome or endocytosis, are fused to form a pigment vacuole throughout the IDC. During the feeding, the heme derived from haemoglobin catabolism is converted into inert brown haemozoin crystals and accumulates in the pigment vacuole [25, 26].

Figure 1.3. Three-dimensional organization of the ring [6]. The cup-shaped structure confines an elongated nucleus, close to which the clustered vesicles and the ER suggest a small Golgi complex. Exocytic vesicles are found near the ER and the mitochondrion always attaches to the plastid.

Trophozoite

The trophozoite has similar internal organizations with the ring and differs mainly by the size and the shape (**Figure 1.4**). Along with parasite growth, the ingestion of RBC cytosol continues with multiple cytostomes until late schizont stage, resulting in a larger pigment vacuole with more haemozoin crystals. The number of free ribosomes increases greatly and the endoplasmic reticulum (ER) enlarges, indicating intense protein synthesis [6]. The Golgi complex increases in complexity, suggesting a possible mechanism of protein export into the PV space and the RBC [27].

The surface of the trophozoite possesses irregular bulges and deep tubular invaginations, which extend out into the RBC cytosol. The parasite PM is closer to the PVM with minimal space in between [6]. These invaginations are presumably involved in trafficking of proteins from parasite to the RBC [28] and perhaps vice versa [29].

Figure 1.4. Three-dimensional organization of the trophozoite [6]. The size increases and the shape becomes rounder with an irregular outline. The nucleus becomes round and a large pigment vacuole is formed. The mitochondrion and the plastid lengthen significantly.

Schizont

DNA synthesis starts from the trophozoite stage, followed by multiple rounds of mitotic duplication at schizont stage (**Figure 1.5**), producing up to 32 daughter merozoites [30].

In the schizont, the nuclear envelope remains intact and the chromosomes do not condense throughout. A pair of mitotic spindle poles is embedded in the nucleus. The microtubules capture the kinetochores on the chromosome to separate the sister chromatids, followed by nuclear membrane division [31]. In the last division, individual nucleus moves to the surface of the parasite with their spindles facing the surface. After division, the daughter nuclei space at regular intervals and bud around the circumference of the parasite [32].

Extensive cytoplasm reorganization accompanies nuclear division in preparation for merozoite formation. Multiple elements assemble in the merozoite-forming foci in a highly ordered sequence, starting with the apical organelles [6]. During the last division, the cytoskeleton components assemble beneath the merozoite surface and the

cleavage furrows form to define the merozoite shape [6]. Eventually, the PV and the RBC membranes are breached to release the mature merozoites in an event possibly dependent on SUB1 protease secreted by the parasite [33].

Figure 1.5. Three-dimensional organization of the schizont [6]. The forming merozoites are budding from the schizont surface with the developing rhoptries. The food vacuole is a dense mass with compacted haemozoin crystals. The surface knobs and the Maurer's clefts (long cleft and short cleft) are visible in the RBC.

1.2 Gene regulation in *P. falciparum*

The 23-megabase nuclear genome of *P. falciparum* consists of 14 chromosomes. They encode approximately 5,300 genes, 60% of which do not have homologues in other organisms. The overall AT composition is 80.6% and rises to 90% in the introns and the intergenic regions, making it the most AT-rich genome sequenced to date [34].

P. falciparum possesses a unique pattern of gene expression, as shown in **Figure 1.6** (adapted from Bozdech et al., 2003 [35] with open access for reuse). During one course of the IDC, the expression of most genes occurs in a continuous cascade, in which the level of mRNAs encoding functionally related proteins rises and falls together at different stages. The transcript level of most genes peaks only once at the point that

their protein products are needed, with genes for general cellular functions being expressed early in the IDC and genes for parasite specialized functions late in the IDC. Such results lead to the “just in time” model of plasmodial gene expression in which mRNAs accumulate just as their protein products are required [35].

Figure 1.6. The phaseogram of transcriptional profiling in *P. falciparum* [35]. Microarray hybridization measures mRNA abundance levels across one IDC. A total of 2,712 genes with a single peak of expression were filtered and ordered based on the phase of expression. The green-red color scale represents the mean centered $\log_2$ -transformed ratio of microarray/input values. The function annotations of the peak transcripts at different stages are aligned on the right panel.

Many factors affect the accumulation of mRNA, including epigenetic regulation, transcription and post-transcriptional regulation. In the following section, the gene regulation in *P. falciparum* is summarized in these three levels.

1.2.1 Epigenetic regulation

Genome organization

Recently, three-dimensional genome architecture was determined by chromosome conformation capture technique (Hi-C), which captures intra- and inter-chromosomal interactions [36]. The centromere and the telomere of chromosomes localize in distal regions of the nucleus, with two arms of chromosome folding in a parallel orientation across IDC. All the centromeric compartments fall in two main clusters in trophozoite [36] and these clusters were also observed at other stages [37] by immunofluorescence microscopy. The Hi-C study also found that inter-chromosomal gene pairs that are in close proximity showed more correlated expression profiles than gene pairs that are far apart, suggestive of the role of genome structure in gene expression [36].

P. falciparum maintains an unusual high proportion of its genome as relaxed, transcriptionally competent euchromatin, given that the euchromatin histone modifications, H3K9ac and H3K4me3, are enriched in the intergenic region of 91% of the genome [38, 39]. On the other hand, some regions are maintained as compact and transcriptionally silent heterochromatin, identified by heterochromatin markers H3K9me3, H3K36me3 and heterochromatin protein 1 (HP1) [40-42]. The heterochromatin regions include telomere, subtelomere and gene families encoding variant antigens (*var*, *stevor*, *rifin* and *pfmc2tm*), invasion-related genes (EBAs and CLAGs) and a gametocyte-specific transcriptional factor ApiAP2-G [43-47]. Furthermore, trophozoites show a more open chromatin structure than rings and schizonts, which may enable more interchromosomal contacts and higher levels of transcriptional activity [36].

Nucleosome occupancy is correlated with transcription. Depletion of nucleosome is observed in the promoter region, correlating with active transcription during IDC [48, 49]. The nucleosome depletion is also observed in the intergenic regions compared to the coding regions [50, 51]. The nucleosome occupancy fluctuates across parasite life

cycle. Nucleosomes are depleted at trophozoite and gametocyte stages, consistent with high levels of transcriptional activities at these stages [49, 51]. Nucleosomes are enriched in subtelomeric heterochromatin regions throughout IDC [48] and across genome at schizont stage [50].

Histone modifications

P. falciparum encodes four canonical histones (H2A, H2B, H3 and H4) and four alternative histones (H2A.Z, H2Bv, H3.3 and CenH3) but lacks the linker histone H1 [52]. H2A.Z, dimerizing with H2Bv, is enriched in the 5' UTR (untranslated region) of euchromatic genes, independent of transcriptional status [53, 54] and H2A.Z/H2Bv double-variant nucleosomes are enriched in the intergenic regions [55, 56]. H2A.Z and H2Bv carry unique acetylation modifications, suggestive their role in maintaining transcriptionally-active euchromatin [57].

A total of 232 histone post-transcriptional modifications were identified by mass spectrometry during IDC, including acetylations, methylations, phosphorylations, ubiquitylations and sumoylations [58].

Histones modifications correlate with transcription. The euchromatin markers, H3K9ac and H3K4me3 correlate well with transcriptional activities [53]. The localizations of these two modification are not limited to active promoters as reported in mammalian genomes [53, 59, 60], but extend to 5' coding regions of transcribed genes [61], intergenic regions and silenced promoters [39, 57]. The enrichment of H3K14ac, H4K5ac, H4K8ac, H4K12ac and H4K16ac at the promoter region is correlated with expression of the corresponding genes [62]. On the other hand, the enrichment of heterochromatin marker H3K9me3 shows negative correlation to gene expression [61]. Another study identified eight histone markers to be positively correlated with mRNA abundance of the corresponding genes and one histone marker, H4K5ac showing negative correlations [63].

Epigenetically regulated genes

The *var* gene family encodes erythrocyte membrane protein 1 (EMP1), which contributes to cytoadherence and antigenic variation at the surface of erythrocytes [64]. Each parasite carries around 60 *var* genes, expresses one in a mutually-exclusive

manner and switches between the *var* genes to evade immunity [65]. The single expressed *var* gene is enriched with H3K9ac and H3K4me3 near the transcription start site (TSS) [42] and the silent *var* genes are enriched with H3K9me3 in the 5' UTR and the coding region [44, 66]. The *var* gene is expressed in the early ring and repressed as the parasite matures. The same *var* gene is turned on again in the majority of progeny, possibly due to a memory mark H3K4me2 in the TSS, methylated by PfSET10 [66, 67]. On the other hand, H2A.Z alternative histone is enriched in the 5' UTR of active *var* gene and depleted at mature stages as well as in the silent *var* genes [53, 54]. A histone deacetylase PfSir2A contributes to *var* gene silencing by removing histone acetylation and triggering the depletion of H2A.Z [54, 68].

Long non-coding RNA (lncRNA) adds complexity of epigenetic regulation of *var* gene expression. Two lncRNAs, transcribed from the bidirectional promoter of *var* genes, incorporate into chromatin after being capped but not polyadenylated [69]. The antisense lncRNA activates the expression of a silent *var* gene in a sequence-specific manner [70].

Another example of epigenetic regulation is the expression of some invasion-related genes. PfBDP1 is a parasite-specific bromodomain protein and the bromodomain recognizes acetylated lysine residues [71]. PfBDP1 is enriched in the upstream of the coding regions of 50 invasion-related genes and directly controls their expression, possibly by associating with acetylated histone 3, for example H3K9ac [72]. Several studies also showed that the expression of some invasion-related genes is correlated with histone modifications, including RH4 and EBA165 [73] as well as CLAG 3.1, CLAG 3.2 and EBA140 [45, 74].

Epigenetic regulation also contributes to gametocytogenesis. The ApiAP2 transcription factor, ApiAP2-G is a key regulator of gametocyte differentiation [75]. It is localized in the heterochromatin region and silenced by H3K9me3 and HP1 in asexual stage [46].

1.2.2 Transcription

The basic mechanism of transcription in *P. falciparum* is similar to that in other eukaryotic organisms. First, the gene structure of *P. falciparum* consists of the open reading frames (ORFs), the introns in some ORFs, the intergenic regions of the 5' UTR

and 3' UTR, and the promoter regions [76]. Second, work in other model eukaryotic organisms shows that transcription of protein-coding genes requires assembly of the RNA polymerase II (RNAPII) and the general transcription factors (GTFs) to form the pre-initiation complex (PIC) at the promoter region. *P. falciparum* encodes most of these elements including all twelve subunits of the RNAPII and some GTFs, including TFIIA (Transcription Factor II A), TFIIB, TFIIE, TFIIH, TATA-binding protein (TBP) component of TFIID [77].

On the other hand, there is a clear reduction of transcriptional components in *P. falciparum* relative to other eukaryotes. First, TBP-associated factors (TAFs) are components of the TFIID complex and required for transcription [78]. The number of the TAFs reduces from 14 in the crown eukaryote to 4 without classical histone fold domain (HFD) in *P. falciparum*, suggesting an unusual architecture of the TFIID complex [77]. Second, transcription activating proteins (TAPs) bind the enhancer and the upstream element and then facilitate recruitment of the PIC to the promoter region [78]. Only one-third of the TAPs usually found in eukaryote are identified in *P. falciparum* [79]. Third, the Mediator is a large complex acting as a bridge between the TAPs at the enhancer and the PIC at the promoter [80]. A dramatically decreased number of the Mediator subunits was predicted in *P. falciparum*, though this has been contested [81].

The similar gene structure and the presence of the transcription factors listed above suggest that *P. falciparum* does control gene expression at the level of transcription. The reduced number of many transcription cofactors leads to the speculation that *P. falciparum* must compensate for their absence. Such speculation is supported by the presence of 27 *Apicomplexan*-specific transcription factors ApiAP2, which are structurally related to the plant *Apetala-2* transcription factors and contain AP2 DNA-binding domains. The 27 ApiAP2 proteins display distinct expression profiles covering the entire IDC and possess sequence-specific DNA binding activities [82]. One study focusing on two ApiAP2 proteins shows that their recognition motifs are located in the promoter region of predicted target genes, whose expression profiles are positively correlated to expression of the corresponding *ApiAP2* gene [83].

Some studies have assessed the recruitment of transcriptional components across the

Plasmodium genome by chromatin-immunoprecipitation during the intraerythrocytic IDC. In most eukaryotes, the largest subunit of RNAPII, RPB1, contains a unique C-terminal domain (CTD), made up of the conserved heptad repeats, Tyr1-Ser2-Pro3-Thr4-Ser5-Pro6-Ser/Lys7 (YSPTSPS). The transcriptional status of RNAPII is associated with sequential changes of phosphorylation in CTD, with phosphorylation of Ser5 alone during transcriptional initiation, phosphorylation of both Ser2 and Ser5 when the polymerase is engaged in elongation, and phosphorylation of Ser2 alone at termination [84]. One study showed that RPB1 occupancy regardless of the phosphorylated status display a bi-phasic binding profile across IDC, with enriched RPB1 occupancy observed in a fraction of genes at an early phase (ring and trophozoite) and switched to a mutually exclusive set of genes at a late phase (schizont). The recruitment of elongating RNAPII (RPB1 with Ser2/5-P) is positively correlated with mRNA abundance for 1,740 genes and negatively correlated with mRNA abundance for 1,200 genes [85]. Another study showed that during IDC, TBP and TFIIE components of the PIC were bound to the promoter regions of both active and inactive erythrocytic-expressed genes but not to CSP and PFS25, which are expressed in sexual stages, suggesting that the PIC might preassemble on the promoters of erythrocytic-expressed genes and poise until active transcription [62].

On the other hand, transcriptional activity was assessed in *P. falciparum* by measuring newly synthesized mRNAs. One study performed nuclear run-on to measure radiolabeled RNA and found peak transcriptional activities in the late IDC (36 hour post-invasion to 42 hpi). The transcriptional activity is only positively correlated with mRNA abundance for 7 out of 18 genes in this study [86]. Another study using nuclear run-on followed by RNA-seq analysis found low transcriptional activities at early ring stage and peak transcriptional activities at trophozoite stages. This study also showed that RPB1 with phosphorylated Ser2 is engaged in the promoter regions prior to transcriptional burst of the corresponding genes [87]. Moreover, incorporation of 4-thiouracil into newly-synthesized mRNAs via pyrimidine salvage has been employed to measure transcriptional dynamics in *P. falciparum* [88]. Genome-wide measurement of labeled RNAs by microarray showed that transcription correlates well with mRNA abundance and that the peak transcription precedes the peak mRNA abundance by 80

min [89].

A parasite-specific bromodomain protein, PfBDP1 has been shown to regulate gene expression transcriptionally in *P. falciparum*. PfBDP1 binds to acetylated histone H3 at the TSS of invasion-related genes and directly controls their expression. The loss of *Bdp1* resulted in decreased mRNA abundance of these genes at schizont stage and disrupted the invasion process [72].

1.2.3 Post-transcriptional regulation

From the perspective of transcribing DNA to mRNA, discrepancy between transcriptional activities and mRNA abundance has been reported [62, 85, 86]. Recently, an increasing number of studies have shown that post-transcriptional regulation plays important roles in gene expression in *P. falciparum*, ranging from RNA splicing, mRNA stability, mRNA degradation and translational control.

RNA splicing

Although *Plasmodium* genes have fewer and shorter introns, all the basic spliceosome components and 9 SR proteins, involved in RNA splicing, were predicted by surveying the *Plasmodium* database [90]. Alternative splicing is also found in *P. falciparum* as 84 alternative splicing events were identified by genome-wide RNA-Seq [91].

Some proteins have been shown to function in RNA splicing. Pre-mRNA processing proteins (PRPs) are present in *P. falciparum* genome, in which PRP16 is shown to function in splicing given that the mutation of *Prp16* blocked an *in-vitro* splicing reaction [92]. PfU52, an RNA helicase, is involved in RNA splicing in an *in-vitro* splicing assay [93]. SR1 is a one of the bona fide Ser/Arg-rich proteins (SRs), key regulators of alternative splicing in eukaryotes. *P. falciparum* SR1 is found to be enriched in 64 transcripts, in which two common sequence motifs were shown to be preferentially bound by SR1 [94]. Another example is CELF1, belonging to RNA-binding CUG-BP, Elav-like family (CELF), which functions in pre-mRNA splicing [95]. *P. falciparum* CELF1 binds 1,376 RNA targets and the sequence motif identified in the introns of these targets is also recognized by CELF homologs in other species, suggesting a conserved role of CELF1 in pre-mRNA splicing [96].

RNA stability and degradation

The contribution of transcriptional activity and mRNA mechanism on total mRNA abundance has been assessed recently [89]. Genome-wide microarray measurements of 4-thiouracil-labeled newly-synthesized RNAs and unlabeled RNAs at one-hour intervals reflect transcriptional activity and mRNA accumulation resulted from mRNA stability/decay, respectively. Transcription predominantly influences mRNA abundance of 2,767 genes and mRNA dynamics predominantly influences 2,660 genes. The predominant contribution of transcription was observed on genes that have peak mRNA abundance at ring and trophozoite stages, while mRNA dynamics dominated those at schizont stage [89].

On the other hand, mRNA degradation functions to regulate mRNA abundance during IDC. Orthologs encoding components of RNA degradation complexes have been predicted in *P. falciparum* using the BLASTP program, including those of the CCR4-NOT complex for mediating deadenylation, the exosome and its associated factors, and the decapping complex [90, 97, 98]. The importance of mRNA degradation might be further supported by the abundant CCCH-type zinc fingers found in *P. falciparum*, which are commonly involved in mRNA decay and translation rate [79]. Genome-wide mRNA decay rate was investigated in *P. falciparum* by blocking transcription with actinomycin D, followed by measuring transcript abundance by microarray across 240 min. The half-lives of global mRNA was shown to increase dramatically during the entire IDC, from an average 9.5 min at ring stage to an average 65.4 min at schizont stage [99]. These findings are contradictory to another study in which the slower rate of decay was only observed in a small fraction of genes that have peak mRNA abundance at late stages [89].

The importance of mRNA metabolism is further supported by the existence of abundant RNA-binding proteins in *P. falciparum*. A bioinformatics study identified 189 putative RNA-binding proteins, which comprise 3.5% of all annotated genes, by searching RNA-binding protein families in the PlasmoDB database [97]. Almost 90% of these genes belong to 6 RNA-binding protein classes, including 72 with the RNA-recognition motif, 48 putative RNA helicase, 11 with the KH domain, 2 with the PUF domain and 11 with the Acetylation Lowers Binding Affinity (ALBA) domain. Transcript

abundance of most RNA-binding proteins peaks at early stages of IDC and decreases at late stages [97]. Another study searched parasite proteome using 793 domains known to interact with RNA and predicted 924 *P. falciparum* proteins containing RNA-binding domains and 64 proteins that have annotated RNA-binding activities [98].

The functions of some RNA-binding proteins in gene regulation have been studied in *Plasmodium* parasite during asexual IDC. The ALBA family includes six DNA/RNA-binding proteins, ALBA1-6 [100]. In *P. falciparum*, ALBA1 directly binds 1,193 transcripts in trophozoites, 105 of which showed differential mRNA abundance in the parasite overexpressing ALBA1 [101]. ALBA4 in rodent-infectious *P. yoelii* interacts with proteins involved in translation machinery, mRNA export and RNA-binding proteins. The loss of *Alba4* exerted pronounced effect to increase mRNA abundance of 251 genes [102]. CAF1, a major mRNA decay factor, has been shown to regulate the abundance of 1,043 transcripts in asexual IDC, including many egress-invasion related transcripts, leading to the release of premature merozoites [103].

Translational regulation

From the perspective of translating mRNA to protein, there is a good correlation between mRNA abundance and protein level for some genes [35, 104], although a delay has been reported for many other genes, suggestive of the contribution of translational regulation [104, 105]. The translation machinery is highly conserved in *Plasmodium* parasites with regards to ribosomal components, translation factors, tRNA molecules and aminoacyl-tRNA synthetases [106].

Several studies have focused on translational regulation in *P. falciparum*. The interaction of eIF4E (eukaryotic initiation factor 4E), eIF4F and polyadenylate-binding protein (PABP) has been found in *P. falciparum*, suggesting that the formation of a closed-loop mRNA structure to promote translation and deter mRNA decay in eukaryotes might be conserved in *P. falciparum* [107-109]. PfDZ50 (DDX6 in human and DOZI in *Plasmodium* gametocyte) possesses DNA/RNA-binding activities and inhibits translation in a reticulocyte-lysate system, possibly by interacting and sequestering eIF4E from the translation machinery [110]. Translational control can also be achieved by structural features of mRNA, such as an upstream open reading frame

(uORF), which is out of frame with the coding sequence, thereby disrupting efficient translation of the downstream coding sequence [111]. Such an uORF is found in the 5' UTR of the VAR2SCA gene and suppresses its translation in *P. falciparum* [112].

Translational repression has also been reported in rodent-infectious *P. berghei*. DOZI (DDX6 RNA helicase in human), and its interacting partner CITH (CAR-I/Trailer Hitch Homolog) are central regulators of translational repression in female gametocytes. DOZI and CITH are core components of the P-granule in the cytoplasm, which also contains RNA-binding proteins like BRUNO/CELFL, ALBAs, eIF4E, PABP, but not RNases [113]. DOZI and CITH are proposed to prevent mRNA from translation and degradation in the P-granule, given that these two proteins bind 731 translationally-repressed transcripts, 211 of which were downregulated at the transcript level in the absence of DOZI/CITH [114, 115]. Some regulators of translational repression have been reported in *P. berghei* sporozoite, the invasive form that is quiescent in mosquito and activated in the liver cells. IK2, as one of the three eIF2 α kinases identified in *Plasmodium* parasites, phosphorylates eIF2 α in mosquito salivary gland, leading to reduced protein synthesis and accumulation of stalled mRNAs in granules and thus the blockage of sporozoite activation [116]. PUF2, a member of Pumilio and fem-3 binding factor (PUF) family, binds PUF-binding elements in the 5' and 3' UTRs of pfs25 and pfs28 mRNAs and represses their translation [117]. PUF2 also binds and inhibits the translation of UIS2 and UIS4 (Upregulated in Infectious Sporozoites) mRNAs, with the proteins being translated only after liver cell invasion [118, 119]. The eIF2 α is a possible target of UIS2 phosphatase, given their *in-vitro* interaction [119]. Moreover, the mutation of *Puf2* resulted in a reduced mRNA abundance of IK2 kinase [120]. These studies suggest intertwined roles of PUF2, IK2, UIS2 and eIF2 α in regulating sporozoite activation post-translationally.

Furthermore, several types of post-translational modifications on proteins have been identified, such as acetylation [40], methylation [121], phosphorylation [122], ubiquitination [123] and SUMOylation [124]. Most of them have been shown to play critical roles in gene regulation as described above, for example, acetylation and methylation of histone markers in epigenetic regulation and phosphorylation of RPB1 subunit of RNAPII in transcription.

1.3 The CCR4-NOT complex in eukaryotes

1.3.1 Components of the CCR4-NOT complex

CCR4-NOT is a highly conserved, multifunctional complex found across the eukaryotic kingdom [125]. In yeast, the complex is present in both cytoplasm and nucleus [126, 127]. In the cytoplasm, the CCR4-NOT subunits are found in the P-bodies [128] containing translationally repressed mRNAs as well as proteins for mRNA degradation and also in the polysomes [129], where mRNA is being translated.

Two forms of complex, 1.2 MDa and 1.9 MDa, have been identified in yeast [130]. The smaller complex consists of nine core subunits, NOT1-5, CCR4, CAF1, CAF40 and yeast-specific CAF130 [130-132]. The larger complex may contain the core subunits and other interacting proteins or complexes [125]. The two forms are found at similar sizes in human [133].

In yeast, NOT3 and NOT5 are orthologs of NOT3 in other eukaryotes, possibly resulting from gene duplication [134]. Human complexes contain additional NOT10 and NOT11 subunits [135], which are conserved in *Drosophila* [136]. There are two paralogs of CAF1 (CNOT7 and CNOT8) and two paralogs of CCR4 (CNOT6 and CNOT6L) in mammals. The paralogs show similar functions [137-139]. On the other hand, CNOT6L but not CNOT6 is involved in cell proliferation in mouse 3T3 fibroblasts [138]. The CAF1 paralogs are recruited to the CCR4-NOT complex in a mutually exclusive manner, so are the CCR4 paralogs as reported in human cells [140].

1.3.2 Structure of the CCR4-NOT complex in eukaryotes

The structure of the CCR4-NOT complex has been characterized in yeast [141-143], *Drosophila* [144] and human [145, 146]. **Figure 1.7** is adapted from Collart, 2016 (with open access for reuse) [147], showing the domains of NOT1 known to interact with other subunits (**Figure 1.7a**) and the structure of the complex in yeast (**Figure 1.7b**).

NOT1 is the largest subunit of the CCR4-NOT complex and acts as a scaffold to recruit other subunits. In yeast, electron microscopy studies and X-ray crystal structure show that NOT1 forms an L-shape with two arms of similar length [148].

Figure 1.7 Architecture of the CCR4-NOT complex in eukaryotes, adapted from [147]. (A) The conserved domains of NOT1. NOT1 contains the N-terminal domain (Not10/11 BD), the MIF4G domain, the DUF3819 domain and the NOT1 domain (Not1). The regions of the CCR4-NOT subunits responsible for interaction are aligned. **(B) Schematic structure of the CCR4-NOT complex in yeast.** NOT1 forms an L-shape and serves as the scaffold of the complex. NOT1 interacts with other CCR4-NOT subunits, which are depicted in different colors via different conserved domains.

The N-terminal domain of NOT1 interacts with NOT11, which does not have orthologs in yeast. The conserved DUF2363 domain (domain of unknown function) in NOT11 is not sufficient for its interaction with NOT1. In addition, human NOT11 is found to interact with NOT10 directly [135].

CAF1 associates with the central MIF4G domain of NOT1. MIF4G (Middle domain of

EIF4G) is a conserved domain known to bind eIF4A, eIF3, RNA and DNA in human cells [149]. The NOT1-CAF1 interaction is required for the efficient deadenylase activity of CAF1 in yeast, *Drosophila* and human [142, 144, 146]. CAF1 connects CCR4 to the complex by interacting with the leucine-rich repeat of CCR4. The recruitment of CCR4 to the CCR4-NOT complex is required for its function in mRNA degradation in yeast [142] and *Drosophila* [144]. These observations suggest that NOT1-CAF1-CCR4 complex is needed for the deadenylase functions of CAF1 and CCR4. Structurally, the central region of NOT1 contains an accessible cavity that could probably provide a platform to interact with mRNA that can be deadenylated by the neighboring CAF1 and CCR4 for degradation [141, 146].

The DUF3819 domain of NOT1 interacts with the Armadillo motif of CAF40 [150].

The NOT1 domain in the C-terminus of NOT1 interacts with several NOT subunits, including NOT2, NOT3, NOT4 and NOT5. NOT2 together with NOT3 or NOT5 forms a symmetric heterodimer using their C-terminal NOT box motifs. Their extended N-termini wrap around the NOT1 domain. In *Drosophila*, The C-terminal NOT box motif of NOT3 and the upstream residues 402–436 aa are required for its interaction with NOT1 [143].

In yeast, the C-terminus of NOT4 interacts with the C-terminus of NOT1. NOT4 binds the NOT1 domain independently from the NOT2/NOT5 heterodimers. NOT4 is not a stable subunit of the complex in mammals and *Drosophila* [135, 140, 151]. However, human NOT4 could interact with NOT1 in the yeast two-hybrid assay [140]. Yeast NOT4 and human NOT4 share a similar N-terminal zinc-finger RING domain but a distinct C-terminus, which might explain their conserved functions but different binding properties [152].

1.3.3 Functions of the CCR4-NOT complex

Two enzymatic activities have been attributed to this complex, with ubiquitination provided by NOT4 [153] and deadenylation by CCR4 and CAF1 [127]. Multiple functions have been attributed to the CCR4-NOT complex as well, including chromatin

modification, transcription initiation, transcription elongation, RNA nuclear export, translation and mRNA degradation. These functions are illustrated as follows.

Chromatin modification

The CCR4-NOT complex interacts with proteins involved in histone modifications and the loss of some subunits affects the status of histone markers. Two histone modifications are found to be regulated by the CCR4-NOT complex in yeast. First, the loss of NOT4 or NOT5 leads to a global reduction of acetylation on histones H3 and H4, possibly by affecting a histone acetyltransferase GCN5 [154]. Second, the level of H3K4 tri-methylation is reduced in yeast cells lacking NOT4 or NOT5 [155], possibly because NOT4 ubiquitinates Jhd2 demethylase and targets it for degradation [156].

Transcription initiation

The CCR4-NOT subunits can regulate gene expression directly. Yeast NOT proteins are transcriptional repressors of the *HIS3* promoter [134]. On the other hand, the CCR4-NOT complex regulates proteins that in turn affect transcription. For example, NOT4 targets YAP1, a transcription factor in yeast [157] and RNAPII-associated factor 1 (PAF1) in human cells for degradation [158].

Extensive interactions of the CCR4-NOT complex with transcription cofactors have been reported. NOT1 binds the promoter of retinoid acid X gene and CAF40 binds transcriptional activators, as reported in fibroblast-like monkey cell lines [159]. As for RNAPII in yeast, SRB9, SRB10 and SRB11 components interact with multiple subunits of the CCR4-NOT complex including NOT1, NOT2 and CCR4, and are functionally related to NOT4 and NOT5 in regulating gene expression [160].

Interactions of the CCR4-NOT complex with the TFIID complex and the SAGA complex have been extensively reported in yeast. NOT1 interacts with TBP [161] and TAF1 [162] with TAF1 functionally redundant to NOT5. NOT5 interacts with TBP and TAF components of the TFIID complex in a manner dependent on the integrity of the TFIID complex [163]. NOT proteins have also been reported to bind TAFII19P [164]. As for the SAGA complex, the interaction of NOT2 with ADA2 subunit of the SAGA complex is suggested to regulate expression of the *HIS3* promoter [165]. One study shows that the complex preferentially represses the SAGA-controlled genes given that these genes

are abundantly upregulated following the mutation of *Not5*, *Not4*, *Not3*, *Ccr4* or *Caf1* [166].

Transcription elongation

The role of the CCR4-NOT complex as a positive elongation factor has been reported in yeast. The complex directly interacts with the RPB4 and RPB7 components of the RNAPII elongation complex, which is required for CCR4-NOT to rescue arrested RNAPII [167, 168]. NOT3 and NOT5 might be the binding proteins of the RNAPII elongation complex because RPB4 is shown to interact with both NOT proteins [169]. Moreover, TFIIS is a transcription factor that helps RNAPII bypass the blockage during elongation [170]. The CCR4-NOT complex is suggested to recruit TFIIS to the RNAPII elongation complex and thus rescues backtracked RNAPII [171].

RNA export and nuclear RNA quality control

The CCR4-NOT subunits have been reported to interact with several proteins involved in RNA export and nuclear RNA quality control in yeast. The newly synthesized transcripts are assembled into the heterogeneous nuclear ribonucleoprotein complexes (hnRNPs) for export [172]. The Hrp1 and Nab2 components of hnRNPs interact with multiple subunits of the CCR4-NOT complex in an RNA-independent manner [173]. The CCR4-NOT complex interacts with Mlp1, Mlp2 and Nup60 in the nuclear basket of the nuclear pore complex (NPC) [173, 174]. The complex also interacts with Dbp5, Mft1 and Yra1 components of the THO complex, which is required for mRNA processing and export to the cytoplasm [173, 175, 176].

Translation

The CCR4-NOT complex is detected in the polysomes. In yeast, the mutation of *Not2*, *Not4* or *Not5* led to defective polysome profiles and aggregation of newly synthesized proteins [129].

The recruitment of the CCR4-NOT complex to mRNA lacking poly(A) tail induces translational arrest in *Xenopus* [177] and *Drosophila* [178]. In *Xenopus*, CAF1 is proposed to recruit NOT1, which in turn recruits 4E-T (eIF4E transporter protein) to mRNA, resulting in translation repression [179]. In yeast, NOT1 itself interacts with

DDX6, an activator of the decapping complex [180]. DDX6 represses translation and leads to accumulation of ribosomes on the transcripts [180, 181].

The CCR4-NOT complex is involved in co-translational quality control in yeast. NOT5 is needed for the association of RPB1 with HSP90 (heat shock protein 90) and R2TP, which are required to keep RPB1 in a soluble form [169, 182]. Yeast NOT4 is a potential translation regulator by ubiquitinating ribosomal protein RPS7A [129] and the nascent polypeptide-associated complex (NAC), which is a chaperone associated with ribosomes [183]. NOT4 also plays a role in the quality control of translational arrest products [184, 185] and has direct impacts on translation repression [186].

RNA degradation

The rate-limiting step of mRNA degradation is deadenylation, mainly contributed by CAF1 and CCR4 in eukaryotes [147]. Many RNA-binding proteins recruit the CCR4-NOT deadenylase complex to mRNA for subsequent degradation, for example, the interactions of Tristetraprolin (TTP) [187] and Nanos with NOT1 [188] as well as the interaction of TOB protein with CAF1 [189] in human cells. Moreover, a study in yeast proposed that translation release factors, eRF1 and eRF3, recruit the minor deadenylases PAN1/PAN2 and the major deadenylases CAF1/CCR4 to mRNA. The CCR4-NOT complex itself interacts with PABP via TOB protein [190].

The CCR4-NOT complex also functions in mRNA degradation mediated by microRNA (miRNA), reported in human cells. A miRNA targeting the sequence-specific mRNA is bound by Argonaute (AGO) protein, which associates with GW182 to form the minimal miRISC [191]. GW182 contains the Trp-motifs to bind CAF40, which recruits the CCR4-NOT complex to the miRNA-targeted mRNA for the downstream deadenylation and decapping [191-193]. In addition, the CCR4-NOT complex is involved in nonsense mediated mRNA decay, which is resulted from a nonsense stop codon. The decapping DCP2 protein interacts with UPF1. The phosphorylated UPF1 recruits SMG5-7, which in turn recruit the CCR4-NOT complex via CAF1 to conduct mRNA degradation in human cells [194].

1.3.4 Functions of NOT1 in eukaryotes

NOT1 is the largest subunit of the complex and the only essential subunit in yeast [147].

Structurally, NOT1 is the conserved scaffold of the complex and acts to recruit other subunits. Functions of NOT1 in transcription initiation, elongation and post-transcriptional regulation are summarized.

As mentioned above, yeast NOT1 directly interacts with TBP and TAFs specifically via its C-terminal NOT1 domain [161, 162]. The interaction of other CCR4-NOT subunits with the RNAPII elongation complex strongly suggests that yeast NOT1 associates with RNAPII as well, although it is not directly tested [167-169].

The function of NOT1 in regulating gene expression has been investigated in yeast. A *Not1* mutation caused cell cycle arrest at G1 during meiosis and constitutive activation of the mating pheromone response pathway [195-197]. Another study demonstrates that NOT1 is a nuclear protein and negatively regulates the transcription of the *HIS3* promoter [126]. In the same study the *Not1* mutation increased the transcription of genes encoding TBP, TRP3, HIS4 and BIK1 (involved in nuclear fusion). However, this negative regulation is not universal given that the transcriptions of DED1, ribosomal RNA and tryptophan tRNA were not affected in the *Not1* mutant [126, 198]. On the other hand, NOT1 is involved in the activation of transcription [131], leaving an open question as to how NOT1 regulates transcription differentially. Regarding the negative regulation, one study conducted in 2007 speculated that SPT3 may mediate the recruitment of the SAGA complex and the TFIID complex to the promoters of upregulated genes in the context of a *Not1* mutation [199].

Studies in other organisms highlight the involvement of NOT1 in translational repression and mRNA decay. NOT1 has been shown to interact with RNA-binding proteins in mammalian cells, including TTP, Xp54, eIF4E, YTHDF and Nanos2. First, TTP recognizes the A/U-rich elements in the 3' UTR of mRNA and mediates mRNA decay by recruiting human CAF1 [200]. Second, 4E-T is an eIF4E binding protein and a component of the decapping complex. NOT1 associates with 4E-T via eIF4E to stimulate mRNA decay by CCR4 and CAF1 deadenylases in HeLa cells [201]. NOT1 also interacts with 4E-T via Xp54, an RNA helicase and a component of the decapping machinery, to elicit translational repression in *Xenopus* embryos [179]. Third, methylation at the N6 of adenosine (m6A) is an abundant RNA modification in eukaryotes. A m6A reader protein, YTHDF interacts with human NOT1 to recruit the

complex for the degradation of m6A-containing RNA [202]. Fourth, NOT1 in mouse germ cells interacts with Nanos2, an RNA-binding protein with a zinc finger domain, which is speculated to promote the localization of CCR4-NOT complex to the P-bodies, the center of RNA degradation [203].

NOT1 is involved in miRNA-mediated silencing as well. In brief, a miRNA targets the sequence-specific mRNA to recruit AGO, GW182, CAF40 and then the CCR4-NOT complex as well as DDX6 for deadenylation and decapping in human cells [191-193]. NOT1 interacts with DDX6 via its MIF4G domain, leading it to mRNA for decapping [192, 193].

All these studies shed light on the possible non-exclusive roles of NOT1 in mRNA metabolism. NOT1 might serve as a platform to interact with RNA-binding proteins and recruit CCR4 and CAF1 deadenylases as well as DDX6 decapping activator to mRNA. Meanwhile, it is possible that Nanos2 directs the CCR4-NOT complex to the P-body for mRNA degradation. Another possible role of NOT1 is bridging between the 5' capping and the 3' poly(A) tail of mRNA because of the interactions of NOT1 with eIF4G, 4E-T and DDX6 at the 5' end as well as with TTP at the 3' end of mRNA. Furthermore, NOT1 could interact with these RNA-binding proteins via distinct conserved domains. TTP, 4E-T, Xp54 and DDX6 interact with the central MIF4G domain [179, 200, 204] and Nanos proteins interact with the C-terminal NOT1 domain [188].

1.3.5 Functions of other subunits in eukaryotes

CAF1 and CCR4 deadenylases

One enzymatic activity of the CCR4-NOT complex is deadenylation, contributed by CAF1 and CCR4 subunits [125]. CAF1 consists of a DEDD nuclease domain in eukaryotes [205, 206]. CCR4 contains a leucine-rich repeat in the middle and an exonuclease-endonuclease-phosphatase (EEP) nuclease domain in the C-terminus [148]. The CAF1 and CCR4 deadenylases shorten mRNA 3'-poly(A) tails, leading to decapping and exonucleolytic degradation [125, 207]. CAF1 and CCR4 deadenylases are recruited to mRNA via the interactions of CAF1 with RNA-binding proteins, for example, TOB [189], SMG5-7 [194] and BTG1 [208] in mammals. Furthermore, CCR4 failed to trigger mRNA

degradation in the context of a disruption of CAF1-CCR4 interaction in *Drosophila* [144].

NOT4

The other enzymatic activity of the CCR4-NOT complex is ubiquitination, contributed by NOT4 subunit. NOT4 contains a conserved RING/Ubox-like zinc-binding domain and an RNA recognition motif (RRM) in the N-terminus [209]. NOT4 is found to be an E3 ubiquitin ligase and active in ubiquitin-dependent degradation pathway in human [153]. It has been shown to interact with UBC4/5-type E2 ubiquitin-conjugating enzyme in both human and yeast [152, 183]. The RING/Ubox-like zinc-binding domain and the RRM motif have redundant roles in regulating the degradation of NOT4 substrates in yeast [210].

Several targets of NOT4 E3 ligase have been found in yeast. Rps7A is a small ribosomal protein [129]. Egd1 and Egd2 are components of the nascent associated polypeptide-associated complex (NAC), which assists maturation of newly synthesized peptide and protein folding [183, 211]. Jhd2 histone demethylase removes the methylation from H3K4 histone markers [156, 212]. YAP1 transcription factor is regulated by nuclear localization and NOT4 [157]. CDC17 is a catalytic subunit of DNA polymerase- α [213]. Cyclin c is a component of cyclin c-CDK8 kinase, which is a conserved negative transcriptional regulator of genes involved in stress response [214]. In addition, one target of NOT4 is found in human cells. RNAPII-associated factor 1 (PAF1) is a subunit of the PAF complex, which functions in histone modification and transcription [158, 212].

In addition, NOT4 contributes to proteasome assembly and integrity in yeast [215].

NOT2

NOT2 contains a NOT box motif for the association with NOT1 and an N-terminal domain to interact with ADA2 in yeast, a component of GCN5 histone acetyltransferase [165]. NOT2 has a potential role in translation given that a mutation of yeast *Not2* led to a reduced amount of polysomes [129]. Human NOT2 represses transcription via its NOT box motif [216].

CNOT3 or NOT3/NOT5 in yeast

Yeast encodes two paralogs of NOT3/NOT5. The names in higher eukaryotes follow NOT3, whereas NOT5 is more functionally relevant in yeast [125, 147]. Both NOT3 and NOT5 in yeast contain a C-terminal NOT box [148].

NOT5 has functions in multiple levels of gene expression in yeast. NOT5 is an important player in histone modification [154, 155]. NOT3 and NOT5 interact with RPB4, a component of RNAPII with dual functions in transcription and translation [169]. NOT5 assists RPB1 subunit of RNAPII to express in a soluble complex by interacting with R2TP [141]. The disruption of the NOT1-NOT2-NOT5 interactions caused severe effects on growth and this tertiary structure could bind poly(U) of RNA via the NOT box motif of NOT5 [143].

Human ortholog NOT3 interacts with the promoter region of *PRPF31* gene and triggers transcriptional repression in retinitis pigmentosa cell [217]. Another study in human cells shows that the intact structure of NOT1-NOT2-NOT3 complex is important for mRNA degradation [145].

CAF40

CAF40 in human cells has little affinity for poly(A) sequences and its R227 residue is important for nucleotide binding with unknown significance [218]. CAF40 does not have catalytic activities, but has been reported to interact with RNA-binding proteins. In *Drosophila* and human cells, Roquins interact with CAF40 to recruit the CCR4-NOT complex to their targets for mRNA degradation [150, 219]. The function of Bag-of-marbles (Bam), a key germline stem cells differentiation factor in *Drosophila*, in repressing the expression of *Nanos* and *E-cadherin* depends on its interaction with CAF40 [220]. Human CAF40 also assists miRNA-mediated degradation by anchoring GW182 to the CCR4-NOT complex [193]. In addition, CAF40 ortholog in mice represses transcription via c-Myb and AP-1 transcription factor [221].

CAF130

Very little is known about the yeast-specific CAF130 subunit. CAF130 interacts with BTT, a member of the NAC complex. CAF130 and BTT are functionally redundant in

control of gene expression [166].

Possible components of the 1.9 MDa complex

Some proteins identified in the 1.9 MDa complex are potential components of the larger CCR4-NOT complex. First, the RNA helicase DHH1, which is DDX6 ortholog in yeast, interacts with NOT1 and CAF1 [222]. DHH1 was found to interact with mRNA decapping proteins and stimulate mRNA decapping in yeast [223]. Furthermore, DHH1 is required for P-body assembly [224] and associates with RNAPII within the context of the CCR4-NOT complex [167].

Second, CAF4 is a WD-containing protein and CAF16 is an ATPase. Both proteins are found in the yeast 1.9 MDa complex in a CCR4-dependent manner but not in the 1.0 MDa complex. CAF4 and CAF16 also interact with SRB9-11, as do NOT1, NOT2 and CCR4 [160].

Third, DBF2 is a protein kinase and interacts with CCR4 and CAF1. The gene disruption of yeast *Dbf* leads to transcript defects as does the mutation of *Ccr4* or *Caf1*, suggesting that these three proteins are functionally correlated [225].

Fourth, MOB1 binds and facilitates DBF2 in regulating mitotic progression in the budding yeast. MOB1 interacts with CAF1, DBF2 and CAF16 [226].

1.3.6 Proposed mechanisms of the action of CCR4-NOT

Several mechanisms have been proposed to explain how CCR4-NOT can regulate gene expression at multiple levels. A review in 2012 suggests that this complex acts as a “chaperone platform” to interact with or assemble other cellular complexes [125]. However, how the complex interacts with different proteins or helps assemble them has not yet been investigated.

Another review in 2012 proposed a new model in which the complex marks mRNA for degradation during transcription [227]. The complex is loaded on mRNA in the nucleus, escorts it from the gene, exports it into the cytoplasm and then influences whether the mRNA is degraded or translated. If the complex is indeed bound to mRNA throughout transcription instead of binding mRNA only when they need to be degraded, it raises more questions: how the complex selects the mRNA marked for degradation, how the

complex binds the mRNA, why the mRNA is not immediately degraded in the cytoplasm and what triggers degradation. More studies are required to reveal the mechanisms underlying the activities of this complex.

1.4 The CCR4-NOT complex in *P. falciparum*

1.4.1 Orthologs of the CCR4-NOT subunits in *P. falciparum*

P. falciparum is predicted to encode nine orthologs of the core subunits of the CCR4-NOT complex based on the PlamoDB database. These orthologs include two paralogs of NOT1, NOT2, NOT4, NOT5, CAF1, two paralogs of CCR4, and CAF40. These proteins and their counterparts in yeast, human and *Drosophila* are listed in **Table 1.1**. In the following section, these predicted subunits are described. The core name for each subunit in *P. falciparum* is used in this thesis.

Core name	<i>P.falciparum</i>	yeast	human	<i>Drosophila</i>
NOT1	PF3D7_1103800 PF3D7_1417200	CDC39	CNOT1	Not1
NOT2	PF3D7_1128600	CDC36	CNOT2	Regena (Rga)
NOT4	PF3D7_1235300	MOT42/SIG1	CNOT4	Not4
NOT3		NOT3		
NOT5	PF3D7_1006100	NOT5	CNOT3	Not3
CAF1	PF3D7_0811300	POP2	CNOT7 CNOT8	Pop2
CCR4	PF3D7_0519500 (CCR4) PF3D7_0107200 (CCR4.1)	CCR4	CNOT6L CNOT6	twin
CAF40	PF3D7_0507600	CAF40	CNOT9	Rcd1
CAF130		CAF130		
NOT10			CNOT10	Not10
NOT11			CNOT11	Not11

Table 1.1 Predicted CCR4-NOT subunits in *P. falciparum*. The CCR4-NOT core subunits identified in yeast, human and *Drosophila* as well as their counterparts in *P. falciparum* are given. Some genes are annotated with different names according to the studies in different organisms. The gene identity of the predicted subunits in *P. falciparum* is listed and the core name for these subunits is used.

NOT1

There are two paralogs of NOT1 identified in *Plasmodium* parasites. In *P. falciparum*, PF3D7_1103800 with 3,371 aa is designated as NOT1.1 and PF3D7_1417200 with 4,466 aa is designated as NOT1.2 in this thesis. Both proteins are much larger compared to their orthologs in the other three organisms, including yeast NOT1 with 2,108 aa, human NOT1 with 2,371 aa and *Drosophila* NOT1 with 2,511 aa.

The transcript levels and protein levels of *Not1.1* and *Not1.2* at ring, trophozoite and schizont stages of one IDC are presented in **Figure 1.8**. The transcript levels of both genes peak once across one IDC: in trophozoites for *Not1.1* and in rings for *Not1.2*, whereas the proteins levels are comparable. The distinct transcript levels suggest that NOT1.1 and NOT1.2 may exert their functions in different stages across IDC. It is also possible that these two proteins play roles in different subcellular compartments or different hosts.

Figure 1.8. Expression of *Not1.1* and *Not1.2* during one IDC. The transcript abundance was analyzed by RNA-seq, done in Prof. Zbynek Bozdech's lab (data not published). The x-axis represents developmental age of parasites, harvested at 2-hour intervals. The y-axis shows the abundance of transcript in the unit of log₂-transformed FPKM (fragments per kilobase million). For protein abundance, the peptide was analyzed by LC-MS/MS, identified by matching MS/MS spectra with peptide sequence and quantified [228]. The x-axis represents three stages, ring, trophozoite and schizont and the y-axis shows the abundance at the protein level.

Other NOT proteins

The size of *P. falciparum* NOT2 is 657 aa and is larger than its orthologs in yeast (191 aa), human (540 aa) and *Drosophila* (585 aa). The conserved NOT box motif is identified in the C-terminus of NOT2 in the NCBI database. Different from NOT2 in the other three organisms, yeast NOT2 does not harbor an extended C-terminal region.

The size of NOT3/5 orthologs is 836 aa for yeast NOT3, 560 aa for yeast NOT5, 753 aa for human NOT3, 844 aa for *Drosophila* NOT3 and 178 aa for *P. falciparum* NOT5. NOT5 contains the C-terminal NOT box motif known to interact with NOT1. Another conserved domain with unknown functions at the N-terminus of NOT3/5 in other organisms is not found in *P. falciparum* NOT5.

The size of NOT4 is 587 aa in yeast, 713 aa in human, 1,051 aa in *Drosophila* and 1,662 aa in *P. falciparum*. The alignment of protein sequences shows that all proteins contain the highly conserved zinc-finger domain and the RNA recognition motif at the N-terminus although the NCBI database failed to identify the zinc-finger domain in human and *P. falciparum* as well as the RNA recognition motif in human. *P. falciparum* NOT4 and *Drosophila* NOT4 have an extended C-terminal region sharing little similarity.

CAF1 and CCR4 deadenylases

The CCR4-NOT complex includes two deadenylase enzymes, CAF1 and CCR4. CAF1 with 1,774 aa is much larger than its counterparts in yeast (433 aa), human (CNOT7 at 285 aa and CNOT8 at 292 aa) and *Drosophila* (357 aa). It contains the N-terminal ribonuclease domain, showing high similarity to its counterparts. No conserved domains are identified at its long C-terminus.

P. falciparum encodes two CCR4 paralogs. I designate the larger PF3D7_0519500 with 2,488 aa as CCR4 and the smaller PF3D7_0107200 with 337 aa as CCR4.1. The size of CCR4 is 837 aa in yeast, 557 aa in human CNOT6, 555 aa in human CNO6L and 552 aa in *Drosophila*. *P. falciparum* CCR4 contains a long N-terminal sequence followed by the highly conserved EEP domain. *P. falciparum* CCR4.1 consists of the EEP domain and does not contain the leucine-rich repeat. The LRR motif, which is known to interact with CAF1 is not identified in CCR4 and CCR4.1, raising the question as to whether CCR4 is able to interact with CAF1 in *P. falciparum*.

CAF40 is less studied compared to other CCR4-NOT subunits. CAF40 is 652 aa in *P. falciparum*, 373 aa in yeast, 331 aa in human and 304 aa in *Drosophila*. CAF40 is highly conserved throughout its protein sequences in these organisms. This conserved region in *P. falciparum* CAF40 is flanked by a divergent C-terminal region and a divergent N-terminal region.

Species-specific subunits

NOT10 and NOT11 are identified as novel subunits specific to the CCR4-NOT complex in human and *Drosophila* [135, 144]. They are not found in yeast. The sequence alignment of NOT10 to the *P. falciparum* proteome using the BLASTP program failed to identify orthologs of NOT10. As for NOT11, the 54-153aa region of PF3D7_0411100 shares 49% similarity to the 138-242aa region of *Drosophila* NOT11 but not human NOT11. This region in *Drosophila* NOT11 is inside the DUF2363 domain responsible for the NOT1-NOT11 interaction. PF3D7_0411100, a conserved protein with unknown function, is a potential NOT11 ortholog in *P. falciparum*.

CAF130 with 1,122 aa is a yeast-specific subunit of the CCR4-NOT complex. The sequence alignment of CAF130 to the *P. falciparum* proteome using the BLASTP program identified PF3D7_1037200, a conserved protein with unknown function. A 117-aa portion of the C-terminus shares 44% similarity, making PF3D7_1037200 a potential CAF130 ortholog in *P. falciparum*.

1.4.2 Conserved domains in NOT1.1 and NOT1.2

In order to characterize NOT1.1 and NOT1.2 proteins, I aligned the NOT1 protein sequences to their counterpart in yeast, *Drosophila* and human. The domains responsible for the interaction with other subunits are highly conserved, which are demonstrated as follows.

N-terminus

The N-terminal domain at 1-302 aa and 1-412 aa was found to interact with NOT11 in human NOT1 and *Drosophila* NOT1, respectively [135, 144]. The domain sequences in these two organisms were aligned to the protein sequences of NOT1.1 and NOT1.2 using the NCBI BLASTP program and no similar regions were found. The N-terminal domain is not found in yeast NOT1 neither, suggesting that NOT1 may gain the ability

to associate with NOT11 in invertebrates and mammals. Although PF3D7_0411100 might encode PfNOT11 as described previously, it is not likely to interact with NOT1.1 and NOT1.2 given the missing N-terminal domain.

MIF4G domain

The central MIF4G domain interacts with CAF1, which in turn recruits CCR4 to the complex [148]. The region of NOT1 that interacts CAF1 has been reported to be 754-1,000 aa in yeast [142], 1,093-1317 aa in human [135] and 1,152-1376 aa in *Drosophila* [144]. Both NOT1.1 and NOT1.2 contain the MIF4G domain at 623-847 aa and 1,399-1,562 aa, respectively, according to the NCBI database. NOT1.2 seems to contain a shorter MIF4G domain.

To investigate whether NOT1.1 and NOT1.2 contain sequence features that could interact with CAF1, the NOT1 protein sequences in yeast, human, *Drosophila* and *P. falciparum* were aligned using the MAFFT online service [229] and displayed in SnapGene Viewer software. The region containing the MIF4G domain is shown in **Figure 1.9**.

The MIF4G domain of NOT1.1 and NOT1.2 is conserved. In the MIF4G domain of five proteins, 74 conserved residues are found (highlighted in yellow in **Figure 1.9**). NOT1.1 shares 67 residues and NOT1.2 shares 56 residues, occupying 90.5% and 75.7% of the total conserved residues, respectively.

Furthermore, several hydrophobic residues in NOT1 and CAF1 are critical for the NOT1-CAF1 interaction in yeast, including P897, F938, P943 and W944 in NOT1 and M290 and M296 in CAF1 [142]. In human, P1257 (corresponding to yeast P943), F1252 (yeast F938) and W1258 (yeast W944) on NOT1 and A137, M141 and N171 on CAF1 mediate the NOT1-CAF1 interaction [146]. Four residues identified in the yeast study are boxed in red, while three residues identified in the human study are boxed in green in **Figure 1.9**. These residues are identical in all the five proteins, suggesting their significance in the NOT1-CAF1 interaction. Taken together, the high ratios of the conserved residues and the conserved interacting residues provide both NOT1.1 and NOT1.2 with a structural basis for the interaction with CAF1 via their MIFG4G domain.

Consensus	-----N-----	
Yeast_NOT1	-----LKFFA-----	768
Human_NOT1	-----TVAKTVT-----VTRPTGVSF-----KKDVPSPINTT	1073
Drosophila_NOT1	-----MKSIA-----	1133
PFNOT1_1	-----FYDQI-----	601
PFNOT1_2	NEGNEKNSILSQTVSLNGDAEQNINDINKNAPQKNSLQSDIYDEFSHQKISNDICINEDDVKLNNEKQKGSSKKKN	1212
Consensus	NI-----NN-----N--K-	
Yeast_NOT1	-V-----DEVSCQINQEG-----APKDVVEKVLFLNNVTLANLNNKV	805
Human_NOT1	NI-----DTLLVATDQTE-----RIVEPPENIQEKIAFIFNNLSQSNMTQKV	1115
Drosophila_NOT1	NI-----DTLLVA-NQEE-----KVTVPPEPVQDKTAFIFNNLSQLNIPOKC	1174
PFNOT1_1	NINGFGLGQIECLMDATELTN-----NIILPSSIIIGEVFSIFNTLCLFNIDEKI	651
PFNOT1_2	NIINN-KEGS-STFLDEKKKSQLTENEENNMSVDDNNNNNNKNDNIVKGGNNVTLKKNKKKGSVTVDLKNNVEIGRYGKI	1290
Consensus	-E-K-	
Yeast_NOT1	DELKKS LTPN-----	815
Human_NOT1	EELKETVKEE-----	1125
Drosophila_NOT1	DEIKETIMTKE-----	1184
PFNOT1_1	KILKDFMQPD-----	661
PFNOT1_2	QENNETKTDDQPIGENSLFDNNFEDYTYTYDYYDDPIMSDDISDDYEDGEDDEEESDDRNDNKNYITGINDRYKEIKIE	1370
Consensus	-----W--Y-V--R--E-N-H--Y--F-----T-----S--R--LKN	
Yeast_NOT1	-----YFSWFSTYLTQRAKTEPNYHDLVSKVIVAMGSGLLHQMNVNLTQLFLVLLSTKDEQAI--DKKKLKN	882
Human_NOT1	-----FMPWVSQYLVMKRVSIENFHSLSYSLDTLKNPEFNKMVLNETYRNKVLVLTSDKAAANFS--DRSLLKN	1194
Drosophila_NOT1	-----YFPLAQYLVLRASMEFNHFTLYNFDALKNGEINRFVTKETLRNIKVLVLRSDKGVINFS--DRSLLKN	1253
PFNOT1_1	-----YFNWLAFYIVKSRASKEVNLHNVFLEIDKLSYPLFIDTIISMVYDCILILFKYINELKEVSFAFRTVLKN	731
PFNOT1_2	DDVQIRRDVFFFNLLKS--VIKKKTSEDML--IYMNFLNKKIMEKIFDNIILQSCEIFKTLLNFIEITKDSSTYRLLKN	1446
Consensus	LG-WLG-ITL--N--P-----K--L-EAY--G--L--P-V-KILES--S--F--PPNPWT--I--L-ELH----	
Yeast_NOT1	LASWLGCIITLALNPKPKHKNIAREMLIEAYKENR--LEIVVPFVTIKLQRASESKIF--PPNPWTGVLKLLIELNEKAN	960
Human_NOT1	LGHWLGMITLAKNPKILHTLDLVKSLLEAYVKGQQUELLVVPFVAKVLESSIRSVVFP--PPNPWTMAIMNVLAELHHEQHD	1274
Drosophila_NOT1	LGHWLGMMITLGRNPKILQLDLKSLLAAYVKGQQUELLFVVPFVAKILESSAKSRIF--PPNPWTMGIMYVLAELHHEQPD	1333
PFNOT1_1	LGSWLGOFITLGRNPKPKSKILDLVVLEAYDKGC--LVCILPMVCKILESISKSKNEF--PPNPWTATILCLLTHEIHELPN	809
PFNOT1_2	LGAWIGIITIGRNPKPKMSKYHNIKQLILYAYDNGY--LIVTFPALCKILESISKSKNEF--PPNPWTATILCLLTHEIHEAQS	1524
Consensus	-K--L-FE-EVL-K--L-----P-N-----	
Yeast_NOT1	WKLSLTFEVEVLLKSNITTKSLK-----PSNFINT	991
Human_NOT1	LKLNLFKEIEVLCKNLALDINELK-----PQNLLKD	1305
Drosophila_NOT1	LKLNLFKEIEVLCKTLNLELAKLR-----QVIYIKD	1364
PFNOT1_1	VKTYIIFEVEVLKKNLSNIQEFQNK-TILLSKRM-----PQNKKSD	851
PFNOT1_2	LKTLIFEIEILFNIFYKINVFDDYNNKCNIIKSRNLPVNSNDLFFRNSFNEHVSVNINSINENNNMSMISVNEITPTNYTKS	1604

Figure 1.9 Alignment of protein sequences of the MIF4G domain in NOT1. The full-length NOT1 proteins from four species were aligned using the MAFFT online service [229] and the region harboring the MIF4G domain was extracted. The region of the MIF4G domain is 754-1000 aa in yeast, 1,093-1,317 aa in human, 1,152-1,376 aa in *Drosophila*, 623-847 aa in NOT1.1 and 1,399-1,562 aa in NOT1.2, deduced from published studies and the NCBI database. Residues found in 4 proteins or more are highlighted in yellow and the corresponding residue is labeled at the top. The residues, identified in yeast study and human study to mediate the NOT1-CAF1 interaction, are denoted by red boxes and green boxes, respectively.

Despite the obvious conservation of the MIF4G domain in both proteins, NOT1.1 more resembles NOT1 of the other three organisms than NOT1.2. There is a notable 80-aa insertion between 1,301-1,380 aa in the N-terminal region of NOT1.2 MIF4G domain, which is absent in NOT1.1 and NOT1 in other organisms. When aligning the protein sequence of the yeast MIF4G domain to the full-length NOT1.2 using the NCBI BALSTP program, a shorter conserved region is identified, which is consistent with the insertion. The similarities between each conserved domain in the five NOT1 proteins are shown in **Supplementary Table 1**. The extra 80-aa region inside NOT1.2 MIF4G domain may lead to missing activities or provide potential novel activities.

DUF3819 domain

The DUF3819 domain in NOT1 is responsible for the interaction with CAF40 [144, 148]. The domain is conserved to the other three organisms as shown in **Figure 1.10** with protein alignment of the regions harboring the DUF3819 domain.

Figure 1.10 Alignment of protein sequences of the DUF3819 domain in NOT1. The full-length NOT1 proteins from four species were aligned using the MAFFT online service [229] and the region harboring the DUF3819 domain was extracted. The region of NOT1 domain is 1,088-1,228 aa in yeast, 1,386-1,535 aa in human, 1,469-1,644 aa in *Drosophila*, 1,083-1,227 aa in NOT1.1 and 2,180-2,324 aa in NOT1.2. The regions are deduced from published studies and the NCBI database. Residues found in 4 proteins or more are highlighted in yellow and the corresponding residue is labeled at the top.

The sequence alignment to the DUF3819 domain in other organisms using the NCBI BLASTP program shows a similarity at around 57.7% for NOT1.1 DUF3819 domain and 52.7% for NOT1.2 DUF3819 domain (**Supplementary Table 1**). Although the critical

interacting residues for NOT1-CAF40 interaction have not been reported yet, the conserved DUF3819 domain in NOT1.1 and NOT1.2 very likely provides a platform to interact CAF40.

NOT1 domain

The C-terminal NOT box motif of NOT2 and NOT3/5 mediates the assembly of the NOT2/NOT3 heterodimer, which docks into the C-terminal NOT1 domain of NOT1 in yeast [143]. The conserved NOT box motif of NOT2 and NOT5 in *P. falciparum*, identified in the NCBI database suggests their interaction with NOT1 proteins. The full-length NOT1 proteins in the four organisms are aligned using the MAFFT online service and the region containing the NOT1 domain is shown in **Figure 1.11**.

Figure 1.11 Alignment of protein sequences of the NOT1 domain in NOT1. The full-length NOT1 proteins from four species were aligned using the MAFFT online service [229] and the region harboring the NOT1 domains was extracted. The insertion in NOT1.2 is boxed in blue. The region of NOT1 domain is 1,711-2,075 aa in yeast, 1,996-2,359 aa in human, 2,114-2,478 aa in *Drosophila*, 2,702-3,054 aa in NOT1.1 and 3,988-4,455 aa in NOT1.2, deduced from published studies and the NCBI database. Residues found in 4 proteins or more are highlighted in yellow and the corresponding residue is labeled at the top.

The abundant conserved residues reveal the highly conserved NOT1 domain in NOT1.1 and NOT1.2, which provides a platform to associate with NOT2, NOT5 and NOT4. On the other hand, an insertion in the NOT1 domain of NOT1.2 was observed, which is absent in *P. falciparum* NOT1.1 and NOT1 in other organisms. The sequences flanking the insertion are highly conserved. The NOT1 domain of NOT1.2 is thus split into the regions of 3,988-4,213 aa and 4,331-4,455 aa. The insertion might contribute new functions to NOT1.2 or lead to loss of functions, like the interaction with the other NOT proteins.

Protein architecture of NOT1.1 and NOT1.2

In summary, NOT1.1 and NOT1.2 contain the MIF4G domain, the DUF3819 domain and the NOT1 domain with high similarities to their counterparts in yeast, human and *Drosophila*. These conserved domains in yeast NOT1, NOT1.1 and NOT1.2 are depicted to scale in **Figure 1.12**.

Figure 1.12. Conserved domains of yeast NOT1, NOT1.1 and NOT1.2. Three conserved domains include MIF4G in brown, DUF3819 in orange and NOT1 in yellow. The length and location of each domain are proportional to the length of protein sequence. Yeast NOT1 contains the MIF4G domain at 754-1,000 aa, the DUF3819 domain at 1,088-1,228 aa and the NOT1 domain at 1,711-2,075 aa. NOT1.1 contains the MIF4G domain at 623-847 aa, the DUF3819 domain at 1,083-1,227 aa and the NOT1 domain at 2,702-3,054 aa. NOT1.2 contains the MIF4G domain at 1,399-1,562 aa, the DUF3819 domain at 2,180-2,324 aa and the NOT1 domain at 3,988-4,213 aa and 4,331-4,455 aa.

The conserved domains are similar between NOT1.1 and NOT1.2 given the similarity of 65%, 61% and 73.5% for the MIF4G domain, the DUF3819 domain and the NOT1 domain (average of two parts of NOT1.2 NOT1 domain), respectively. The values of similarity are shown in **Supplementary Table 1**. Comparison of protein sequences of NOT1.1, NOT1.2 and other eukaryotic NOT1 suggests that NOT1.1 is likely the canonical

NOT1 homolog, given the special characteristics of NOT1.2 including the insertion in the MIF4G domain and the insertion in the NOT1 domain.

1.4.3 Studies of the CCR4-NOT complex in *Plasmodium* parasites

Some studies on the CCR4-NOT complex in *P. falciparum* have been published. One study in 2011 focused on the function of CAF1 in gene expression in *P. falciparum* [103]. The endogenous *Caf1* was disrupted by the insertion of a *Piggybac* transposon at 1,029 bp downstream of the start codon, corresponding to the downstream of the conserved ribonuclease domain (7-260 aa) [230]. The mutation of *Caf1* resulted in mistimed mRNA accumulations of 1,031 genes during IDC. The mistimed mRNA abundance and aberrant protein accumulation of many genes involved in parasite egress and invasion possibly led to the release of premature merozoites and the reduced growth rate of the parasite [103].

Another study in 2019 reported the important roles of CCR4 and CAF1 in gametocyte development in *P. yoelii* [231]. CCR4 is shown to co-immunoprecipitate two NOT1 homologs, CAF1, CAF40, NOT2 and NOT5, but not NOT4. The mutation of *Ccr4* affected mRNA abundance in gametocytes, which led to a loss of gametocyte maturation and a reduction of gametocyte transmission into mosquito. This study failed to completely delete the coding region of *Caf1*. The mutation of *Caf1*, which mimics *Caf1* mutation in *P. falciparum* [103, 230], led to comparable but more pronounced effects to those seen in the *Ccr4* mutant, suggesting that CAF1 and CCR4 play critical and possibly intertwined roles in gametocyte maturation and transmission [231].

The CCR4-NOT complex has been predicted to be decay components in *P. falciparum* by blasting protein sequences of human and yeast decay factors in *P. falciparum* genome [99]. The protein interaction network for the CCR4-NOT complex was predicted based on the available protein network of human CCR4-NOT complex and the corresponding orthologs in *P. falciparum* [97]. In a total of 1,090 interactions, 774 proteins have orthologs in *P. falciparum*, involved in multiple metabolic processes [97].

Moreover, dispensable and essential genes have been identified in *Plasmodium* parasites. *PiggyBac*, a 'cut-and-paste' transposon that inserts into TTAA target

sequences in the presence of a *piggyBac* transposase was employed to obtain the cloned mutants and 38,000 independent insertions were achieved to cover the compact genome of *P. falciparum* [232, 233]. Eventually, 2,042 genes were disrupted by at least one insertion, including NOT1.1, NOT1.2, NOT2, CAF1, CCR4 and CCR4.1. A total of 3,357 genes had no insertions, including NOT4, NOT5 and CAF40. These genes were considered to be essential for parasite viability, because parasites with potential *piggyBac* insertions in these genes under saturated mutagenesis cannot survive. However, the essentiality of genes is presumptive in this study, because the failure to insert *piggyBac* transposons in one gene cannot be excluded [233]. Another study identified essential genes in *P. berghei*. The success to disrupt NOT4 (PBANKA_0806100) and CCR4 (PBANKA_1234300) as well as the failures to disrupt NOT1.2 (PBANKA_1025500), CAF1 (PBANKA_1426200) and CAF40 (PBANKA_1107200) were reported [234]. The essentiality of NOT1.2, NOT4 and CAF1 is contradictory in two studies, possibly due to species variations or different insertion positions in the corresponding genes.

Overall, the composition of the CCR4-NOT complex is still unknown in *P. falciparum*. The functions of the CCR4-NOT subunits and the whole complex remain unclear in the asexual intraerythrocytic stage of *P. falciparum*.

1.5 Research rationale and objectives

As described above, both transcriptional and post-transcriptional regulations have been shown to play critical roles in gene expression in *Plasmodium* parasites. RNA degradation, as an important step in post-transcriptional regulation, is largely initiated through the removal of the poly(A) tail by CAF1/CCR4 deadenylases in eukaryotes [147]. CAF1 and CCR4 are two core subunits of the conserved CCR4-NOT complex and their association with the complex is crucial for their functions [142, 144, 146]. *Plasmodium* parasites encode most CCR4-NOT subunits, suggesting that the CCR4-NOT complex is conserved in the parasite. The important roles of CAF1 and CCR4 in *P. falciparum* [103] and *P. yoelii* [231] on gene expression have been reported, suggesting that the CCR4-NOT complex might contribute significantly in post-transcriptional gene regulation in the *Plasmodium* parasites. On the other hand, the CCR4-NOT complex may also

contribute to transcriptional regulation given its interaction with many transcriptional components in eukaryotes [125, 147], which has not been studied in the *Plasmodium* parasites. Therefore, the CCR4-NOT complex is a good access to investigate gene regulation in *P. falciparum*.

NOT1 serves as the scaffold of the CCR4-NOT complex. Transcriptionally, NOT1 is shown to interact with several transcriptional components and the loss of *Not1* led to differential gene expression in yeast. Post-transcriptionally, NOT1 is proposed to be a bridge between RNA-binding proteins and mRNA decay factors for mRNA degradation, and to direct the CCR4-NOT complex to the RNA degradation center. Two NOT1 homologs in *P. falciparum* contain all the conserved domains identified in eukaryotic NOT1, suggesting that the roles of NOT1 in transcriptional and post-transcriptional regulation may be conserved in *Plasmodium* parasites.

Although the structural role of NOT1 has been extensively studied in eukaryotes, the function of NOT1 is still unclear. For example, how the mutation of *Not1* in yeast affects gene expression differentially has not been fully understood. The association of NOT1 with transcription cofactors and proteins involved in mRNA metabolism has been reported, but the direct functions of NOT1 on transcriptional and post-transcriptional regulation are unknown. Two NOT1 homologs in *Plasmodium* parasites raise more questions as to the reason to encode two NOT1 proteins rather than the unique NOT1 in other organisms, the compositions of the CCR4-NOT complex or complexes, and the correlation of two NOT1 proteins in gene regulation. This study on *P. falciparum* NOT1 proteins will reveal the mechanism of gene regulation in *P. falciparum* from a new perspective and extend the knowledge of the CCR4-NOT complex in other organisms.

This project **aims** to investigate the potential contribution of NOT1 proteins to gene expression in *P. falciparum*. To achieve this research aim, the following objectives are defined. The hypotheses and experimental designs are also elucidated.

1.5.1 To identify potential interacting partners of NOT1.1 and NOT1.2

Hypothesis: The CCR4-NOT complex may exist in more than two forms. Two differently sized complexes have been found in yeast and two NOT1 paralogs in *P. falciparum* complicate the situation. One possibility is that two NOT1 proteins maintain the

complex integrity at different stages within the IDC or in different hosts. It is also possible that these two proteins function in the same stage but form distinct complexes and/or assume distinct roles.

To determine whether *P. falciparum* NOT1 proteins provide scaffolds for a parasite CCR4-NOT complex and whether different complexes are assembled at each of the two NOT1 paralogs, NOT1.1 and NOT1.2 together with two other CCR4-NOT subunits, CAF1 and NOT4 were tagged. The potential interacting partners of NOT1 proteins were identified by co-immunoprecipitation (co-IP) and mass spectrometry (MS). These experiments may reveal novel interacting partners which will shed new light on CCR4-NOT function in *P. falciparum*.

1.5.2 To investigate the effect of *Not1* mutations on recruitment of RNAPII

Hypothesis: Because of the evidence of structural and functional interaction between the CCR4-NOT complex and RNAPII in other organisms, the distribution of RNAPII may be altered in *Not1* mutants. NOT1 was also shown to interact with TBP and TAFs, suggesting that NOT1 might functionally associate with the TFIID complex and perhaps the SAGA complex as well, which would affect transcription of some genes both positively and negatively as seen in yeast [199]. If so, the affected TFIID and SAGA complexes are very likely to influence recruitment of RNAPII. However, the influence may not be universal as the expression of only some genes was affected in yeast by a *Not1* mutation. With regards to transcriptional elongation, the ability of the CCR4-NOT complex to rescue arrested RNAPII elongation complex [235] gives rise to the possibility that NOT1 may affect RNAPII during elongation as well.

In order to investigate the function of NOT1 in transcription, recruitment of the elongating form of RNAPII across the genome was assessed by chromatin-immunoprecipitation and microarray (ChIP-on-chip). The analysis of differential RNAPII recruitment in *Not1* mutants, if any, may provide clues as to the functions of NOT1 in transcriptional regulation during the IDC.

1.5.3 To determine the effect of *Not1* mutations on transcript abundance

Hypothesis: NOT1 is presumed to affect transcript abundance by influencing both transcription initiation/elongation and post-transcriptional regulation. As elucidated

above, NOT1 may associate with key players in the general transcriptional apparatus as well as the RNAPII elongation complex as found in yeast. On the other hand, the interaction of NOT1 with RNA-binding proteins may be required for mRNA decay and translational repression, as illustrated in mammalian cells. As a result, transcript abundance may be affected by all these possible functions of NOT1 in *P. falciparum*.

In order to determine whether and how NOT1 can influence gene expression, microarray was performed to measure mRNA abundance. Differential mRNA abundance in parasites bearing a *Not1.1* mutation or a *Not1.2* mutation will reveal the effects of NOT1.1 and NOT1.2 in gene expression. Then, comparison between change in RNAPII recruitment and change in transcript abundance will reveal at which stage NOT1.1 and NOT1.2 regulate gene expression, transcriptionally or post-transcriptionally or both. Moreover, comparison between differential mRNA abundance in a *Not1.1* mutant and a *Not1.2* mutant will reveal how these two genes exert their functions in the asexual stages.

1.5.4 To outline the function of NOT1 as the scaffold for its interacting partners.

Hypothesis: NOT1 might modulate the functions of its interacting proteins given that the NOT1-CAF1 interaction is required for the deadenylase function of CAF1 in eukaryotes. If NOT1.1 and NOT1.2 are the scaffolds of the complex, this modulation might be conserved in *P. falciparum*.

To test this hypothesis, the change of transcript abundance in *Not1.1* mutants was compared to that in the mutation of its interacting partners, which were determined in Objective 1.5.1. CAF1 is a candidate given its highly conserved interaction with NOT1 in eukaryotes and the effect of *Caf1* mutation on mRNA abundance has been reported in *P. falciparum*. In addition to CAF1, the effects of potential CCR4-NOT interacting proteins in gene expression, if available, were compared to NOT1.1 and NOT1.2 as well.

Chapter 2 Materials and methods

2.1 Cloning and plasmid constructs

2.1.1 Restriction digestion and ligation-independent cloning

All restriction enzymes are from NEB. The digestion was conducted by 5 units of enzyme per microgram of plasmid in 37°C water bath for 1 hour. Digested plasmid was separated in 0.8% agarose (Vivantis) gel, cut with a clean blade and purified with Gel Extraction Kit (Qiagen).

Instead of the conventional T4 DNA ligation, ligation-independent cloning (LIC) was employed unless otherwise noted. The protocol was optimized in 2012 [236]. The method utilizes elongated overhangs on the insert as well as the vector in order to improve the efficiency of ligation. Briefly, the vector was linearized by corresponding restriction sites and the insert with 20-basepair homology to the digested end of the vector was PCR-amplified. The vector and the insert were mixed and treated with T4 DNA polymerase (Thermo Scientific) at room temperature for 2.5 min to elongate the 3' overhangs. Then, the mixture was incubated on ice for 10 min to anneal the overhangs. The reaction consists of 100 ng vector, 25 ng PCR insert, 2 µl of 10x NEB 2.1 buffer (NEB), 0.2 µl of 100x BSA (NEB), 0.5 µl of 3U/µl T4 DNA polymerase and water topping up to 20 µl. Eventually, 4 µl of the mixture was transformed into 50 µl TOP10 competent *Escherichia coli* cells.

2.1.2 Transformation of competent cells

The transformation of plasmid into the competent cell was described elsewhere [237]. The mixture of the LIC reaction and the TOP10 cells were kept on ice for 30 min, heat shocked at 42°C for 50 seconds and kept on ice for 5 min. Then, 500 µl of pre-warmed LB Broth was added to the mixture followed by shaking at 37°C at a speed of 220 rpm for 60 min. The cells were centrifuged at 4,000 rpm for 1min. The pellet was re-suspended in 200 µl LB Broth and inoculated on a LB agar plate supplemented with 100 µg/ml Ampicillin (Sigma). The agar plate was maintained in an incubator at 30°C for 20 hours to prevent potential recombination in the competent cells at 37°C.

Before scaling up colonies, colony PCR was performed on selected colonies to verify the insertion. The colonies observed on the agar plate were suspended in 10 µl DEPC-

treated water and 2 µl was used as a template for the colony PCR with Taq polymerase (Thermo Scientific). During the initial 5-min denaturation at 95°C, the cells were ruptured and plasmids were released into the reaction. The designed primers will only amplify the plasmid with the insert, usually one primer on the vector and one primer on the insert. Then, 29 cycles of amplification (95°C for 30 sec, 50°C for 30 sec and 65°C for 2 min/kb) was done followed by a final extension at 65°C for 10 min.

The colonies with PCR product at correct sizes were inoculated into 4 ml LB Broth supplemented with 100 µg/ml Ampicillin and shaken at 30°C for 20 hours. The plasmid was extracted with Plasmid Mini Kit (Qiagen) and sequenced (1st Base). The colonies with correct sequence were further cultured in 1 liter LB Broth. The plasmid was extracted with Plasmid Maxi Kit (Qiagen). The resulting plasmid was diluted at a final concentration of 1 µg/µl and stored at -20°C.

2.1.3 PCR amplification

The program of PCR amplification contains an initial denaturation at 95°C for 5 min; 34 cycles of amplification: 95°C for 30 seconds, 55°C for 45 seconds and 65°C for 2 min/kb; and a final amplification at 65°C for 5 min. The amount of template used in one reaction is 50 ng of genomic DNA, 10 ng of plasmid DNA and 50 ng of cDNA mixture.

2.1.4 pLNOT1.2-3HA construct

The plasmid pLNOT1.2-HA-ecDDD (described in Section 2.1.10) was digested with XhoI and EagI restriction enzymes to delete the homologous arm, the ecDDD cassette and three HA epitopes. A portion of the 3' genomic locus of *Not1.2* up to but not including the stop codon was PCR-amplified from T996 genomic DNA with the primers, pLNOT1.2_HA_F and pLNOT1.2_HA_R. The amplicon was inserted into the linearized vector as a homologous arm to construct the plasmid, pLNOT1.2, which was sequenced by the primer pLNOT1.2_seqR.

The primers 3HA_F and 3HA_R at a concentration of 10 µM were annealed by heating at 95°C for 5 min and cooling down at room temperature for one hour. The annealed primer pair was inserted into the plasmid pLNOT1.2 between restriction sites, XmaI and BglII. The resulting plasmid pLNOT1.2-3HA was sequenced with the primer pLNOT1.2_seqR.

2.1.5 pTNOT1.1-FLAG construct

The plasmid pTCherryHA was a gift from Dr. Brendan S. Crabb (Burnet Institute, Australia) [238]. The plasmid pTCherryHA was digested with XmaI and SpeI restriction enzymes to delete the homologous arm, the mCherry coding region and three HA epitopes. A portion of the 3' genomic locus of *Not1.1* up to but not including the stop codon was PCR-amplified from T996 genomic DNA with the primers pTNOT1.1_F and pTNOT1.1_R. The amplicon was inserted into the plasmid pTCherryHA between restriction sites, XmaI and SpeI, to construct the plasmid pTNOT1.1, which was sequenced with the primer pTNOT1.1_seqR. Then, the primers FLAG_F and FLAG_R at a concentration of 10 µM were annealed by heating at 95°C for 5 min and cooling down at room temperature for one hour. The annealed primer pair was inserted into the plasmid pTNOT1.1 between restriction sites, NcoI and SpeI. The resulting plasmid pTNOT1.1-FLAG was sequenced with the primer pTNOT1.1_seqR.

2.1.6 pTNOT1.1-3HA construct

A portion of 3' genomic locus of *Not1.1* up to but not including the stop codon was PCR-amplified with the primers pTNOT1.1_3HA _F and pTNOT1.1_3HA _R. The amplicon was inserted into the plasmid pTCherryHA between restriction sites, XmaI and BsrGI, to construct the plasmid, pTNOT1.1-3HA, which was sequenced with the primer pTNOT1.1_3HA_seqR.

2.1.7 pTNOT4-3HA construct

A portion of 3' genomic locus of *Not4* up to but not including the stop codon was PCR-amplified with the primers pTNOT4_3HA_F and pTNOT4_3HA_R. The amplicon was inserted into the plasmid pTCherryHA between restriction sites, XmaI and BsrGI to construct the plasmid, pTNOT4-3HA, which was sequenced with the primer pTNOT4_3HA_seqR.

2.1.8 pCC1-NOT1.1-KO construct

The pCC1 plasmid was a kind gift from Dr. Alan F. Cowman [239]. Two homologous arms were inserted into the plasmid pCC1. Homologous arm 1 was amplified from T996 genomic DNA with the primers, pCC1_NOT1.1_HR1_F and pCC1_NOT1.1_HR1_R. The amplicon was inserted into the plasmid pCC1 between restriction sites, SacII and SpeI,

to construct the plasmid pCC1-NOT1.1-HR1. The insertion was confirmed by sequencing (1st Base) with the primer pCC1_NOT1.1_HR1_seqR. Homologous arm 2 was amplified from T996 genomic DNA with the primer pCC1_NOT1.1_HR2_F and pCC1_NOT1.1_HR2_R. The amplicon was inserted into the plasmid pCC1-NOT1.1-HR1 between restriction sites, EcoRI and NcoI, to construct the plasmid pCC1-NOT1.1-KO. The insertion was confirmed by sequencing with the primer pCC1_NOT1.1_HR2_seqR.

2.1.9 pCC1-NOT1.2-KO construct

The primers, pCC1_NOT1.2_HR1_F and pCC1_NOT1.2_HR1_R, were used to amplify homologous arm 1 (HR1). The primers, pCC1_NOT1.2_HR2_F and pCC1_NOT1.2_HR2_R, were used to amplify homologous arm 2 (HR2). The HR1 amplicon was inserted into the plasmid pCC1 between restriction sites, SacII and SpeI, to construct the plasmid pCC1-NOT1.2-HR1. The insertion was sequenced by the primer pCC1_NOT1.2_HR1_seqR. The HR2 amplicon was then inserted into the plasmid pCC1-NOT1.2-HR1 between restriction sites, EcoRI and NcoI, to construct the plasmid pCC1-NOT1.2-KO, which was sequenced by the primer pCC1_NOT1.1_HR2_seqR.

2.1.10 pLNOT1.2-HA-ecDDD construct

The plasmid HSP101-HADB was a kind gift from Dr. Daniel E. Goldberg (Washington University, USA) [240]. The vector was digested with XhoI and AvrII restriction enzymes to delete the HSP101 homologous arm. A portion of the 3' genomic locus of *Not1.2* up to but not including the stop codon was PCR-amplified from T996 genomic DNA with the primers pLNOT1.2_HA_ecDDD_F and pLNOT1.2_HA_ecDDD_R, serving as a homologous arm for recombination. The amplicon was inserted into the plasmid HSP101-HADB between restriction sites, XhoI and AvrII. The insertion was confirmed by sequencing with the primer pLNOT1.2_HA_ecDDD_seqR.

2.1.11 pThDHFR construct

The plasmid pTCherryHA was digested with XmaI and BsrGI restriction enzymes to delete the inducible promoter part. The 5' overhang end was filled in with DNA polymerase I Large (Klenow) fragment (NEB). The reaction consists of 13.3 µl water, 5 µl of 10x NEBuffer, 30 µl of the linearized vector (4 µg), 0.16 µl of 10 mM dNTP and 1.5

μl of 5U/μl Klenow enzyme. The reaction mixture was incubated at 25°C for 15 min. The resulting vector with blunt end was purified by gel extraction and re-ligated with T4 DNA ligase (Thermo Scientific). The ligation reaction includes 10.5 μl water, 4 μl of 10x DNA ligase buffer, 4 μl of 50% 4000 PEG (Sigma), 100 ng vector and 1.5 μl T4 DNA ligase. The resulting plasmid, pThDHFR, was sequenced with the primer pThDHFR_seqR.

2.2 Culture, transfection and limiting dilution

2.2.1 Malaria culture, blood smear and parasite counting

T996 parasite strain was cultured in RPMI-HEPES medium (Gibco), supplemented with 0.25% Albumax (Gibco), 0.2% sodium bicarbonate (Sigma), 100 μM hypoxanthine (Sigma) and 10 μg/ml Gentamycin (Gibco) [241]. The medium was filtered with Rapid-Flow Filter, 0.2 μm aPES membrane (Thermo Scientific). The parasites were maintained at 2% hematocrit and at approximately 5% parasitemia.

Parasitemia was counted by blood smear. A drop of the culture containing infected RBCs (iRBCs) and medium at a ratio of 1:4 was smeared into a thin layer on a microscope slide and air-dried. The smear was fixed by 100% methanol (Merk) and stained with 10% Giemsa Staining Solution (Sigma). The slide was washed under tap water, air-dried and observed under the 100x objective lens with immersion oil.

The number of iRBC as well as the total number of RBC was counted. The parasitemia was calculated as the ratio of the number of iRBC to the total number of RBC on the slides.

2.2.2 Synchronization, cryopreserve and lysis

To ensure the population egress simultaneously, the rings were synchronized with 5% sorbitol (Sigma) for 5 min and washed with RPMI medium [242]. The synchronization was conducted at 4 hour post-invasion (hpi) and 20 hpi for two continuous IDCs. The parasites were cultured to the next cycle for experiments.

To preserve parasites for a long time, they were cryopreserved at ring stage with 1 volume of medim and 2 volumes of Cryopreserve Solution. Cryopreserve Solution

consists of 28% glycerol (Promega), 3% sorbitol (Sigma) and 0.65% NaCl (Sigma). The vials were labeled and kept in liquid nitrogen [243].

To revive parasites from liquid nitrogen, the frozen parasites were thawed in a series of solutions: 12% NaCl for 3 min, 1.6% NaCl for 5 min and 0.9% NaCl/0.2% glucose for 5min [243]. Fresh blood was added on the next morning.

To harvest parasites for genomic DNA and protein lysate, iRBCs were lysed with 0.1% saponin on ice for 5 min [244]. Parasite pellets were washed with 1x PBS until the redness cannot be seen. As for RNA extraction, iRBCs were stored at -80°C or lysed directly as described in Section 2.4.1.

2.2.3 Stable transfection

The electroporation protocol employed in this study was established elsewhere [245]. In brief, 100 µg plasmids were mixed with 300 µl Cytomix Solution and incubated at room temperature for 5 min. The composition of Cytomix Solution is 120 mM KCl, 0.15 mM CaCl₂, 10 mM K₂HPO₄/KH₂PO₄, 25 mM HEPES/2 mM EGTA (pH7.6 adjusted by 10 M KOH) and 5 mM MgCl₂ (all reagents from Sigma).

iRBCs with 20% parasitemia at early ring stage were pelleted and re-suspended in the plasmid/Cytomix solution. The mixture was then transferred to a pre-chilled Gene Pulser Cuvette (Biorad). The electroporation was accomplished with Gene Pulser Xcell Electroporation Systems (Bio-rad) using the setup: 320V voltage, 950 µF capacitance, infinite resistance and 2 mm cuvette.

One hour before the electroporation, 200 µl fresh RBC and 50 ml medium were added to a T175 flask (Corning) and incubated at 37°C. The electroporated mixture was transferred immediately into the flask. RPMI medium was changed 3 hours after the electroporation and corresponding selection drugs were added 10 hours afterward.

2.2.4 Drug cycling

After the transfection and the addition of selection drugs, parasites were expected to be cleared within 3 days and reappeared after 2 weeks. Once observing the parasites, the drug on/off cyclings were performed to select the parasites with successful integration. The drugs were removed from the culture for 10 days and added back for

10 days. The cycling was repeated until integrant locus was detected by PCR amplification. During the cycling, the stabilizing ligands were maintained in the culture if the transfected plasmid contains a destabilizing domain.

The antifolate WR99210 (Chemopharma) was prepared in RPMI medium to a stock concentration of 25 μ M. The Blastcidin (Sigma) was diluted in water to a stock concentration of 25 mg/ml. Trimethoprim (TMP) was diluted in dimethyl sulfoxide (DMSO, Sigma) to a stock concentration of 100 mM. The 5-fluorocytosine (5-FC, Valeant Pharmaceuticals) was diluted in DMSO to a stock of 10 mM. All the solutions were filtered with 0.2 μ m Acrodisc Syringe Filter (Life Sciences) and stored at -20°C. Immediately prior to use, the stock was diluted with RPMI medium to achieve the desired concentration.

The final concentration employed in the culture is 2.5 nM for WR99210, 2.5 μ g/ml for Blastcidin and 10 μ M for TMP.

2.2.5 Limiting dilution

The single clone of parasites was obtained by limiting dilution [246]. The culture was shaken at 37°C overnight to ensure that one RBC was invaded by only one merozoite. On the next day, the culture should contain approximately 4% parasitemia at ring stage. The exact parasitemia was counted by the blood smear and the culture was diluted to a concentration of 0.3 parasites per well in 96-well plates (Nunc). For each transgenic line, 60 wells were seeded for the tagging lines and 120 wells were seeded for the mutant lines.

The blood smear was used to screen for parasites from Day 14 and the wells containing parasites were scaled up. Genomic DNA was extracted for detection of integrant loci and parental loci by PCR amplification.

2.3 DNA work

2.3.1 Genomic DNA extraction

Genomic DNA was extracted from parasite pellets using DNeasy Blood & Tissue Kit (Qiagen). Briefly, saponin-lysed parasite pellet was suspended in 200 μ l of 1x PBS,

treated with 20 µl of 20 mg/ml proteinase K (Thermo Scientific) and further lysed in 200 µl Lysis Buffer (Qiagen) at 56°C for 10 min. Into the lysate, 200 µl of 100% ethanol (Merck) was added. Genomic DNA in the mixture was bound by the Qiagen column, cleaned with 700 µl Wash Buffer W1 and 700 µl Wash Buffer W2. The genomic DNA was eluted with 50 µl nuclease-free water.

2.3.2 Southern blot

Probe synthesis

To detect a specific DNA fragment, the Southern blot was conducted with a DIG-labeled DNA probe. The probe was amplified by PCR with PCR DIG Labeling Mix (Roche) instead of 10 mM dNTP mix. Out of 100 µl PCR reaction, 5 µl was tested on a 1% agarose gel to verify the size of the final product. Prior to use, the probe mixture was denatured by boiling at 100°C for 5 min and cooling on ice.

In this study, the primer pair, SouBlot_F and SouBlot_R, was used to amplify the probe targeting *Not1.1* genomic loci in wildtype and the *Not1.1* mutant.

DNA digestion and gel preparation

In total, 10 µg genomic DNA was treated with 50 units of restriction enzymes overnight in a 37°C water bath. The digested DNA was loaded with 6x Loading Buffer (Thermo Scientific) into a 1% agarose gel which was cast in a 25 cm tray. For one blot, 5 µl DIG-labeled DNA marker II (Roche) was loaded as a marker. The gel was run overnight at 30 V for 12 hours and then visualized under UV light after being stained with 0.5 µg/ml ethidium bromide for 15 min.

The following incubations were done by shaking at room temperature. The gel was depurinated in 250 mM HCl for 10 min and rinsed in distilled water. The gel was then denatured twice in Denaturation Buffer (0.5 M NaOH and 1.5 M NaCl) for 15 min and rinsed in distilled water. Next, the gel was neutralized twice in Neutralization Buffer (0.5 M Tris-HCl, pH7.5 and 1.5 M NaCl) for 15 min and rinsed in distilled water. The gel was equilibrated in 20x SSC for 10 min.

Capillary transfer and detection

The DNA fragments in the gel was transferred to a positive-charged nylon membrane (GE Health) by capillary transfer for 20 hours. The order of materials in the setup was a container with 20x SSC, a bridge glass, 4 filter papers, the agarose gel, the membrane, 5 filter papers, a stack of tissue paper, a piece of glass and a 500 g weight. On the next day, the position of the loading wells was marked with a pencil on the membrane. The membrane was then crosslinked at $1200 \times 100 \mu\text{J}/\text{cm}^2$.

DIG-High Primer DNA Labeling and Detection Starter Kit II (Roche) was employed for detection of the probe. The membrane was pre-hybridized by shaking in DIG Easy Hyb buffer (5ml/100 cm² membrane) at 37°C for 45 min and then hybridized with the DIG-labeled DNA probe (12.5 ng/ml in the pre-hybridization buffer) at 37°C for 20 hours. The membrane was washed twice with Low Stringent Wash Buffer (2x SSC, 0.1% SDS) for 5 min at room temperature, washed twice with High Stringent Wash Buffer (0.1x SSC, 0.1% SDS) at 65°C for 15 min and then incubated in Wash Buffer (Maleic acid buffer, pH7.5 supplemented with 0.3% Tween-20) for 2 min. Maleic Acid Buffer contains 0.1 M Maleic acid and 0.15 M NaCl.

After being blocked for 30 min in Blocking Buffer, the membrane was blotted with 1:10,000 anti-DIG AP-conjugated antibody in the Blocking Buffer for 30 min. Blocking Buffer was prepared by diluting 10x Blocking Solution (Roche) with Maleic Acid Buffer. The membrane was then washed twice with Wash Buffer and eventually covered with CPD Detection Buffer. The signals were exposed to an X-ray film (Canon).

2.4 RNA work

2.4.1 Total RNA extraction

Frozen or fresh iRBCs were lysed with 10 volumes of Trizol (Ambion) and then 2 volumes of chloroform (Sigma) were added. An aqueous layer was separated from the RBC residues by centrifuging at 4,500 rpm for 10 min. RNAs in this layer were precipitated with ice-cold isopropanol (Merck) at -20°C overnight. For a better visualization of RNA pellet, 1.5 μl Glycobule (Thermo Scientific) was added.

The RNA pellet was obtained by centrifuging at 9,000 rpm for 1 hour at 4°C and then washed with ice-cold 70% ethanol (Merck). The RNA pellet was air-dried for 2 hours on ice and then dissolved in DEPC-treated water (Ambion). The concentration was measured by nanodrop and 1.5 µl of the RNA solution was run on a 1% agarose gel to test the integrity of RNA samples.

2.4.2 Reverse transcription

The extracted mRNA was converted into cDNA by reverse transcription which was then used as a template for PCR analysis in Section 2.1.3 and the following real-time PCR. SuperScript III first-strand synthesis system for RT-PCR (Invitrogen) was employed for reverse transcription. In brief, 1 µg total RNA was incubated with 5 µM oligo (dT)₂₀ and 1 mM dNTP mix in the volume of 10 µl at 65°C for 5 min. The cDNA Synthesis Mix was added, consisting of 2 µl of 10x RT buffer, 4 µl of 25 mM MgCl₂, 2 µl of 0.1 M DTT, 1 µl of 40 U/µl RNaseOUT and 1 µl of 200 U/µl SuperScript III Reverse Transcriptase. The mixture was incubated at 50°C for 50 min for reverse transcription and then at 85°C for 5 min for reaction termination. Eventually, 1 µl RNase H (Invitrogen) was added to remove the RNA template. The resulting cDNA mixture at an approximate concentration of 50 ng/ µl was stored at -20°C.

2.4.3 Real-time PCR

To validate the expression of a specific gene at the transcript level, real-time PCR was performed. The cDNA mixture was diluted in a final concentration of 5 ng/µl and 20ng cDNA was used as a template for one real-time PCR reaction.

The SYBR Select Master Mix (Thermo Scientific) was employed. Technical triplicates were included for each cDNA template with a selected target and a control target. Reaction mixtures were prepared in a MicroAmp Fast Optical 96-Well Reaction Plate with Barcode (Applied Biosystems) in dark. Each reaction consists of 0.15 µl of 10 µM Forward and Reverse Primer mix, 4 µl of 5 ng/µl cDNA mixture, 5 µl of 2x SYBR Select Master Mix and water topping up to 10 µl. The plate was then sealed with a MicroAmp Clear Adhesive film (Applied Biosystems), centrifuged at 800 rpm for 3 min and scanned by StepOnePlus Real-Time PCR System (Applied Biosystems).

The thermal cycling steps include 50°C for 2 min for UDG activation, 95°C for 2 min for AmpliTaq DNA polymerase activation and 40 cycles of 95°C for 15 sec-55°C for 15 sec-65°C for 30 sec-signal scan.

2.4.4 Northern blot

DIG-labeled RNA probe synthesis

RNA probes were utilized in the Northern blot instead of DNA probes for a better binding affinity to target RNA. Probe 1 was PCR-amplified with the primers, NorBlot1_F and NorBlot1_R, with Taq polymerase. The reverse primer contains the promoter sequence of the T7 RNA polymerase. Probe 2 was amplified with the primers, NorBlot2_F and NorBlot2_R. Control probe was PCR amplified by the primers, NorBlot_control_F and NorBlot_control_R. The PCR products were purified with PCR Purification Kit (Qiagen) and served as templates for T7 in vitro transcription (Thermo Scientific). The reaction consists of 200 ng of the purified PCR product, 2 µl of 10x Transcription buffer, 2 µl DIG-labeling RNA mix (Roche), 2 µl of 10 U/µl T7 polymerase and water topping up to 20 µl. The reaction mixture was incubated in a 37°C water bath for 4 hours to synthesize the DIG-labeled RNA probe and then stored at -20°C.

The labeling efficiency of the probes was measured by spot test. Serial dilution was performed on the DIG-labeled RNA products with unknown concentration and DIG-labeled control DNA (Roche) ranging from 0.01 pg/µl to 1 ng/µl. For each diluting solution, 1 µl was dropped onto a positive-charged nylon membrane, followed by crosslinking and detection as described in Section 2.3.2. The signal intensity of the DIG-labeled RNA was compared to the signal intensity of the control DNA of a known concentration, thereby estimating the concentration of the RNA probe.

Before being added to the pre-hybridization buffer, the RNA probe was boiled at 100°C for 5 min and cooled on ice for at least 2 min.

Sample preparation

Total RNA was harvested at ring, trophozoite and schizont stages at approximately 8 hpi, 24 hpi, 40 hpi, respectively. For each sample, 30 µg was loaded for Probe 1 and Probe 2, whereas 10 µg was loaded for the control probe. The RNA samples were denatured by Loading Buffer at a ratio of 1:1. Loading Buffer contains 250 µl of 100%

formamide, 83 µl of 37% formaldehyde, 50 µl of 10x MOPS (pH7.0), 50 µl of 100% glycerol, 10 µl of 2.5% bromophenol blue, 0.5 µl of 10 mg/ml ethidium bromide and 56.5 µl DEPC-treated distilled water. The RNA/loading buffer mixture was denatured in a 65°C water bath for 10 min and chilled on ice for 1 min.

For each blot, 10 µl ssRNA ladder (NEB) and 10 µl DIG-labeled DNA marker II (Roche) were loaded as size markers.

Gel preparation, transfer and detection

The RNA samples and the ladders were loaded in a 0.8% denaturing gel (15 cm length, 15 wells) and separated in 1x MOPS buffer (pH7.0) under 60 V. A total of 2 g agarose powder was dissolved in 213 ml DEPC-treated distilled water and cooled to 65°C. Then, 23.65 ml of 10x MOPS (pH 7.0) and 13.5 ml of 37% formaldehyde (Sigma) were added. 10x MOPS Buffer contains 200 mM MOPS, 50 mM sodium acetate, 20 mM EDTA and the pH value was adjusted with sodium hydroxide and is stored in dark at room temperature.

The gel was run at 30 V for over 6 hours until bromophenol blue hit 2/3 of the gel. The gel was visualized under UV light and the distance between each marker and the loading well was recorded. The procedures of capillary transfer, crosslinking and detection were the same with those in the Southern blot in Section 2.3.2, except that nuclease-free water was used and hybridization was performed at 60°C for 20 hours.

Membrane stripping

The membrane was stripped with Stripping Buffer at 80°C for 1 hour and washed twice in 2x SSC at room temperature for 5 min. Stripping Buffer contains 50% Formamide (Sigma), 5% SDS and 50 mM Tris-HCl pH7.5. The stripped membrane can be blotted with another probe.

2.5 Microarray

2.5.1 cDNA preparation by SMART amplification

The first strand cDNA synthesis and the subsequent SMART amplification of double-stranded cDNA were established elsewhere [247].

Briefly, 500 ng total RNA was prepared in 6 µl DEPC-treated water. First, 2 µl RT primer mix was added to each sample. The stock of the RT primer mix consists of 12.5 µl SMART-Oligo dT, 12.5 µl SMARTrandom and 25 µl TS-Oligo. The concentration of each primer is 100 µM. The mixture was incubated at 65°C for 5 min and chilled on ice for 10 min to anneal the primers to the RNA template. Second, the RT Master Mix was added to each sample, containing 4 µl of 5x first-strand Buffer (Thermo Scientific), 2 µl of 100mM DTT (Thermo Scientific), 2 µl of 3x dNTP (3 mM dNTP) and 1 µl of 40 U/µl RNase OUT (Thermo Scientific). The mixture was incubated at room temperature for 2 min. Third, 1 µl of 200 U/µl Superscript II (Thermo Scientific) was added to the mixture and incubated at 42°C for 50 min to carry out reverse transcription. The resulting cDNA mixture was stored at -20°C.

The cDNA mixture was used as a template for SMART PCR in order to obtain a large quantity of cDNA. The SMART PCR reaction consists of 2.5 µl cDNA mixture and 97.5 µl SMART Master Mix. The SMART Master Mix includes 78 µl water, 10 µl of 10x Taq buffer, 1.5 µl of 30x aa-dNTPs mix, 6 µl of 100 µM Primer IIa and 2 µl of 5U/µl Taq polymerase (NEB). The 30x aa-dNTPs was prepared by adding 30 µl dATP, 15 µl dCTP, 15 µl dGTP, 15 µl dTTP and 17 µl aa-dUTP (Biotium). The original concentration of the dNTPs (Thermo Scientific) is 100 mM. The SMART program contains denaturing at 95°C for 5 min, annealing at 60°C for 1 min, initial elongating at 68°C for 10 min, 19 cycles of 95°C for 30 seconds-60°C for 30 seconds-68°C for 5min and final extending at 72°C for 5 min.

Subsequently, the amplified cDNA in the PCR reaction was bound to Minielute cleanup column (Qiagen) with 500 µl PB buffer (Qiagen) and washed with 700 µl PE buffer (Qiagen). The cDNA products were eluted with 16.5 µl Elution Buffer (Qiagen), of which 1.5 µl was measured by nanodrop and then run in a 1% gel to verify the quality of cDNA.

2.5.2 Cy-dye labeling

Reference RNA is a pooling of 3D7 parasite RNAs, harvested at 8 timepoints across one IDC. Reference cDNA was synthesized simultaneously with the test samples.

For each sample, 4 µg reference cDNA and 4 µg test cDNA were prepared in an identical volume and 0.25 volume of 0.5M sodium bicarbonate (pH9.5) was added. DMSO (17

μl) was added to each vial of Amersham Cy-dye Post-Labeling Reactive Dye (GE Health) and each sample was stained with 2 μl dye in dark for 2 hours. The reference cDNA was coupled to Cy3-dye and the test sample was coupled to Cy5 dye. The unbound Cy-dye in the reaction was removed by Minielute PCR Purification Kit (Qiagen). The labeled cDNA was eluted in 16.5 μl Elution Buffer (Qiagen) and 1.5 μl was used for concentration measurement by nanodrop.

As the same reference cDNA was used for all the test samples, the reference cDNA was labeled in the same tube. The eluted solutions of the Cy3-coupled reference cDNA were combined and 15 μl was added to each of the Cy5-coupled test cDNA. This step ensures the reference is identical in each hybridization, thereby allowing the normalization between all the test samples.

Subsequently, the cDNA mixture was denatured at 100°C for 3 min and cooled down at room temperature for 5 min. To each of the cDNA mixture, 95 μl Elution Buffer (Qiagen) and 125 μl of 2x Hybridization Buffer (Axil Technology) were added. After mixing well, 245 μl of the mixture was loaded to an array in a pre-blocked hybridization slide and sealed in a hybridization chamber.

2.5.3 Hybridization

Each hybridization slide contains two arrays and each array contains 11,664 oligos targeting 5,061 genes. The rehydration is done for the hybridization slides older than 2 weeks. The slide was rehydrated on a container filled with 0.5x SSC for 50 seconds in Slide Warmer (Cole-Parmer) at 42°C. The slide was dried on a 100°C Digital Dry Bath (Labnet) for 3 seconds. Oligos on the slide were fixed by UV crosslinking at 800 x 100 μJ/cm².

The slides were blocked in Blocking Buffer at 42°C for 1 hour. Blocking Buffer contains 12.5 ml of 20x SSC, 0.5 ml of 10% SDS, 0.5 g BSA and 37 ml water. Then, the slides were washed in MilliQ water for 5 times by dipping for 2 mins. The slides were dried by spinning down at 600 rpm for 6 min.

The hybridization chambers were rotated in a 65°C incubator for 20 hours. Then, the slides were washed twice in Solution I (1 ml of 10% SDS and 10 ml of 20x SSC in 339 ml water) and twice in Solution II (1 ml of 20x SSC in 349 ml water). Each wash was done

by dipping the slides for 4 min and letting them sit in the wash buffers for 4 min. The slides were dried by spinning down at 600 rpm for 6 min.

2.5.4 Microarray scanning and processing

The hybridized slides were scanned by PowerScanner (Tecan) and signals were captured in PowerScanner V1.2 software. The Cy5-dye was detected with the Red laser at a wavelength of 635 nm and the Cy3-dye was detected with the Green laser at a wavelength of 532 nm.

The signals captured in Cy5 channel and Cy3 channel for each array were exported in a format of TIFF picture, which were merged in Genepix 6.0 software. The oligos with bad signals were flagged for deletion. Background subtraction, normalization and filtering of raw signal intensities were accomplished by the “limma” package in R. Expression of each gene was averaged from the log₂-transformed expression of oligos for the specific gene.

2.6 Protein work

2.6.1 Protein extraction

For total protein lysate, saponin-lysed parasite pellet was treated with 1x protease inhibitor and 4 pellet volumes of Total Lysis Buffer, consisting of 4% SDS, 0.5% Triton X-100 and 50% 1xPBS. The mixture was boiled at 100°C for 10 min and centrifuged at 10,000 rpm for 10min. The protein lysate in the supernatant was stored at -20°C.

For the cytoplasmic lysate and the nuclear lysate, the lysis buffers used in co-immunoprecipitation were employed. The saponin-lysed parasite pellet was lysed with 5 pellet volumes of Lysis Buffer to obtain the cytoplasmic lysate. The resulting nuclear pellet was washed twice in Lysis Buffer, lysed with 2.5 pellet volumes of Extraction buffer and diluted with 2.5 pellet volumes of Dilution Buffer to obtain the nuclear lysate.

The recipe of buffers are listed here. Lysis Buffer contains 20 mM HEPES (pH7.8), 10 mM KCl, 1mM EDTA (pH8.0), 0.65% Nonidet P-40, fresh 1mM DTT, fresh 1 mM PMSF and fresh 1x protease inhibitor. Extraction Buffer contains 20 mM HEPES (pH7.8), 400

mM KCl, 1mM EGTA (pH8.0), 1mM EDTA (pH8.0), fresh 1mM DTT, fresh 1 mM PMSF and fresh 1x protease inhibitor. Dilution Buffer contains 20 mM HEPES (pH7.8), 1mM EDTA (pH8.0), 0.2% Triton X-100, 30% glycerol, fresh 2mM DTT, fresh 1mM PMSF and fresh 1x protease inhibitor.

2.6.2 Western blot

Sample preparation

The recipe of 5x Laemmli Buffer is 5 g SDS, 1.82 g Tris, 25 ml Glycine, 0.05% bromophenol blue and topping up with distilled water into 50 ml. Immediately prior to use, β -mercaptoethanol (β ME) was added to 5x Laemmli Buffer at a volume ratio of 1:3. The protein lysates were denatured with the 5x Laemmli Buffer/ β ME solution at a volume ratio of 4:1 at 100°C for 10 min.

The ladder used for proteins larger than 240 kDa was 12 μ l HiMark Pre-Stained Protein Standard (Thermo Scientific). The ladder used for proteins smaller than 240 kDa was 2.5 μ l BLUltra Prestained Protein Ladder (GeneDireX).

Gel preparation and transfer

A 6% SDS-PAGE gel was cast in a 1.5 mm plate for proteins larger than 100 kDa and a 12% SDS-PAGE gel was cast in a 1.0 mm plate for smaller proteins. The recipe for the 6% gel is 5.3 ml water, 2.5 ml of 1.5 M Tris-HCl (pH8.8), 100 μ l of 10% SDS, 2 ml Acrylamide/bis (30%, 37:5:1, Biorad), 100 μ l of 10% Ammonium persulfate (Biorad) and 4 μ l TEMED (Biorad). The recipe for the 12% gel is 3.3 ml water, 2.5 ml of 1.5M Tris-HCl (pH8.8), 100 μ l of 10% SDS, 4 ml Acrylamide/bis (30%, 37:5:1), 100 μ l of 10% Ammonium persulfate and 4 μ l TEMED.

A 5% stacking layer was prepared on the top for loading wells. The recipe for the 5% gel is 1.4 ml water, 250 μ l of 1M Tris-HCl (pH6.8), 20 μ l of 10% SDS, 330 μ l Acrylamide/bis (30%, 37:5:1), 20 μ l of 10% Ammonium persulfate and 2 μ l TEMED.

The gel was run in 1x Running Buffer at 70V until proteins left the stacking layer and at 120V afterward. 10x Running Buffer consists of 30.3 g Tris base, 144 g glycine and 10 g SDS dissolving in 1 L water.

Then, the transfer apparatus was assembled with PVDF membrane, which is pre-wet with methanol. Transfer Buffer contains 65 ml of 10x Transfer Buffer (30.3 g Tris base, 144 g glycine in 1 L water), 65 ml methanol and 520 ml water. An ice bucket was put in the transfer tank to prevent the heat produced from transfer to damage the membrane. Proteins larger than 100 kDa were transferred at 90 mA for 16 hours at 4°C and smaller proteins were transferred at 100V for 2 hours at room temperature.

Blocking, blotting, and detection

The PVDF membrane was blocked by 5% Blocking Buffer for 1 hour at room temperature, consisting of 5% Blocking-Grade Blocker, nonfat dry milk (Biorad) in 1x TBS. 10x TBS buffer contains 24.2 g Tris base and 80 g NaCl dissolving in 1 L water and adjusting to pH7.4 with HCl. The washes were done for three times with 1x TBS supplemented with 0.1% Tween (1x TBST) for 10 min.

The membrane was blotted with primary antibody by rotating for 16 hours in a 4°C cold room. The primary antibody was diluted in 5% Blocking Buffer. The ratios utilized were 1:10,000 for anti- β -actin antibody (Sigma, a1798, mouse), 1:1,000 for anti-HA antibody (Roche, 3F10, rat), 1:2,000 for anti-FLAG M2 antibody (Sigma, mouse), 1:2,000 for anti-aldolase antibody (Abcam, ab38905 rabbit), 1: 1,000 for anti-GAPDH antibody (Abcam, ab8245, mouse), 1:1,000 anti-histone H3 antibody (Abcam, ab18521, rabbit) and 1:1,000 for anti-Ser2/5-P antibody ([85], mouse). The membrane was washed for three times with 1x TBST.

The secondary antibody conjugated with Horseradish Peroxidase (HRP) was diluted in 2.5% Blocking Buffer and incubated for 1 hour at room temperature. The ratios were 1:5,000 for anti-rat antibody (GE Health), 1:2,000 for anti-mouse antibody (Cell Signaling) and 1:2,000 for anti-rabbit antibody (GE Health). The membrane was washed for three times with 1x TBST.

ECL Select Western blotting Detection Reagent (GE Health) was used for HRP detection. A mixture of 200 μ l of each Detection Reagent was applied to 1 membrane with 10 lanes. The membrane was kept in dark for 2 min. The signals were exposed to an X-ray film (Canon) or captured in ImageQuant LAS4000 Chemiluminescent and Fluorescence Imaging System (Fujitsu Life Sciences).

2.6.3 Co-immunoprecipitation

Three methods of co-immunoprecipitation were employed in this thesis, including HA magnetic bead (Pierce), A/G magnetic bead (Pierce) and anti-FLAG M2 affinity gel (Sigma). The proteins tagged with HA epitopes were immunoprecipitated with the HA magnetic beads. The proteins tagged with FLAG epitopes were either directly immunoprecipitated with the anti-FLAG M2 affinity gels or recognized by anti-FLAG M2 antibody followed by immunoprecipitation with the A/G magnetic beads.

The procedures of immunoprecipitation are as follows. As described in Section 2.6.1, 50 μ l parasite pellet was lysed to obtain 500 μ l protein lysate, in which 25 μ l was stored at -20°C as input and the rest was used in one immunoprecipitation reaction. For each reaction, 25 μ l HA magnetic beads or 25 μ l M2 affinity gels were washed in 175 μ l of 1x TBST. The wash buffer was removed with the help of a magnetic stand or centrifugation. The beads were washed again in 1 ml 1x TBST. The protein lysate was added to the Eppendorf tube with the beads or the gels. The protein/beads mixture and the protein/gels mixture were rotated for 16 hours in a cold room at 4°C.

For the usage of the A/G magnetic beads, the protein lysate was incubated with anti-FLAG M2 antibody for 16 hours in a cold room at 4°C. The protein/antibody mixture was added to the beads and rotated for 2 hours at 4°C.

All the following procedures were done in the cold room at 4°C. Proteins that are not bound by the beads or the gels were removed and stored at -20°C as unbound supernatant. The beads or the gels with bound proteins were washed three times with 1 ml ice-cold 1x TBST and one time with 1 ml nuclease-free water. The wash buffers were supplemented with 1x protease inhibitor to prevent protein degradation. Proteins on the beads or the gels were eluted in 40 μ l 1x Laemmli Buffer supplemented with 0.1 M DTT by boiling at 100°C for 10 min. The elution step was repeated to enhance the elution efficiency. The resulting 80 μ l eluate with the immunoprecipitated proteins was stored at -20°C.

To test whether the magnetic beads or the affinity gels successfully immunoprecipitate tagged proteins, 20 μ l of the input, 20 μ l of the unbound supernatant and 20 μ l of the eluate were analyzed by Western blot.

2.6.4 Mass spectrometry

Once the magnetic beads were confirmed to immunoprecipitate tagged protein, samples were prepared for mass spectrometry. Schizonts were harvested from 6 ml iRBCs with 15% parasitemia. The resulting 400 µl saponin-lysed parasite pellet was lysed into the cytoplasmic extract and the nucleic extract. Two lysates were combined and then split into 8 Eppendorf tubes. Immunoprecipitation was carried out as described above. The beads in the 8 Eppendorf tubes were combined into 2 tubes in the last wash. Eventually, 60 µl eluate was obtained.

Before sending the protein samples for mass spectrometry, the quality of the immunoprecipitation was tested by Western blot. The eluted proteins were prepared for mass spectrometry as follows. All the equipment was cleaned by 70% ethanol to reduce contamination. A gradient gel was cast in a 1.5 mm plate. The order of the SDS-PAGE gels is 2 cm of 20% (pH8.8), 2 cm of 8% (pH8.8) and 2 cm of 4% stacking layer (pH6.8) from bottom to top. One well was left empty between samples or ladders. Half of the eluate was loaded and run into the stacking gel for 5 min, after which the other 30 µl was loaded. For each gel, 15 µl HiMark protein ladder was loaded to estimate protein positions. The gel was run at 100 V until all the markers entered the middle 8% layer.

The Coomassie Brilliant Blue staining was employed to distinguish the proteins in the 8% layer. Briefly, the gel with three layers was transferred into a clean container and stained with Coomassie Staining Buffer for 15 min. Coomassie Staining Buffer was prepared by dissolving 0.25 g Coomassie Brilliant Blue R-250 (Biorad) in 45 ml methanol, 10 ml glacial acetic acid and 45 ml water. The stained gel was then destained with Destaining Solution until the blue color in the background disappeared. Destaining Solution contains water, methanol, and acetic acid in a volume ratio of 50/40/10.

The 8% layer with clearly stained proteins was cut with a plastic blade and transferred into a 1.5 ml Eppendorf tube. The tube was sent to Dr. Sze Siu Kwan, Newman laboratory (Nanyang Technological University, Singapore) for mass spectrometry sequencing. Briefly, peptide mass fingerprint analyses were performed using LC-MS/MS and the peptides were matched to the PlasmoDB database using Mascot search software [248].

2.7 ChIP-on-chip

2.7.1 Crosslinking

Parasites were harvested at 6 timepoints across one IDC, specifically at 4 hpi, 12 hpi, 20 hpi, 28 hpi, 36 hpi, and 44 hpi. At each timepoint, 1.5 ml iRBCs with 5% parasitemia were harvested for one ChIP reaction. The saponin-lysed parasite pellet was washed twice with 37°C 1x PBS. The pellet was then re-suspended into 3.65 ml ice-cold 1x PBS and crosslinked with 50 µl of 37% formaldehyde for 10min. The crosslinking was terminated by 300 µl of 1.67M glycine for 1min. The mixture was topped up to 40 ml with ice-cold 1x PBS and the resulted pellet was washed with ice-cold 1x PBS. The crosslinked parasites were frozen with liquid nitrogen and stored at -80°C for chromatin-immunoprecipitation. All the centrifugations were done at a speed of 4,000 rpm for 5min.

2.7.2 Chromatin-immunoprecipitation

Protein/DNA extraction

Samples were kept on ice all the time and centrifuging was done at 4°C. The parasite pellet containing crosslinked protein/DNA was lysed with 1.5 ml Buffer A (1 mM EDTA pH8.0, 0.25% NP-40, fresh 1x protease inhibitor and fresh 2 mM PMSF) by pipetting and incubating on ice for 60 min. The pellet was then thoroughly lysed in a 5ml Tissue Grinder (Wheaton) by 200 strokes. The lysate was centrifuged at 5,000 rpm for 5 min and nuclei were pelleted. The nuclei pellet was lysed with 160 µl SDS Lysis Buffer (1% SDS, 50 mM Tris pH8.1, 10 mM EDTA, fresh 1x protease inhibitor and fresh 2mM PMSF) by pipetting and incubating on ice for 15 min. The lysate was transferred to a 1.5 ml Eppendorf tube, sonicated at 25% amplitude for 8 cycles of 10-second pulse and 50-second off in 130-watt ultrasonic processor VCX 130 (SONICS) and then centrifuged at 13,000 rpm for 10 min. The supernatant containing soluble proteins and DNA fragments was diluted with 1.45ml ChIP Dilution Buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7 mM Tris pH8.1, 167 mM NaCl, fresh 1x protease inhibitor and fresh 2 mM PMSF).

Before immunoprecipitation, 100 µl of the diluted supernatant was taken for input DNA and stored at 4°C.

Immunoprecipitation

The following procedures were done in a cold room at 4°C. For each reaction, 80 µl protein A agarose /salmon DNA beads (Millipore) were added to the lysate containing crosslinked protein/DNA and rotated for 2 hours. This step aimed to reduce unspecific binding of proteins to the beads. The mixture was centrifuged at 1,000 rpm for 5 min. The supernatant was transferred to a new Eppendorf tube and 15 µl of 0.2 mg/ml anti-Ser2/5-P antibody [85] was added. The protein/antibody mixture was rotated for 16 hours.

On the next day, 80 µl protein A agarose/salmon DNA beads were blocked with 3% BSA in ChIP Dilution Buffer for 30 min at room temperature and then rinsed with ChIP Dilution Buffer. The pre-blocked beads were added to the protein/antibody mixture and rotated for 2 hours. The mixture was centrifuged at 1,000 rpm for 5min. The unbound supernatant was stored at -20°C.

The beads were washed twice with Low Salt Wash Buffer (0.1% SDS, 2 mM EDTA pH8.0, 20 mM Tris pH8.1, 150 mM NaCl and 1% Triton X-100), thrice with High Salt Wash Buffer (0.1% SDS, 2 mM EDTA pH8.0, 20 mM Tris pH8.1, 500 mM NaCl and 1% Triton X-100), once with LiCl Wash Buffer (0.25 M LiCl, 1% NP-40, 1% sodium deoxycholate, 1 mM EDTA and 10 mM Tris pH8.1) and twice with TE Buffer (10 mM Tris pH8.1 and 1 mM EDTA pH8.0). To remove supernatant, the mixture was centrifuged at 2,500 rpm for 1 min. All the wash buffers were prepared before use and filtered with Acrodisc 25 mm Syringe Filter, 0.45 µm Supor Membrane (Life Sciences).

Before performing DNA extraction, Western blot was performed to verify that anti-Ser2/5-P antibody and protein A agarose were able to immunoprecipitate the RPB1 subunit of RNAPII specifically. The proteins bound on the beads were eluted with 50 µl 1x Laemmli buffer supplemented with 0.1 M DTT. For one blot, 25 µl of the input, 25 µl of the unbound supernatant and 25 µl of the eluate was loaded in a 6% SDS-PAGE gel and blotted by anti-Ser2/5-P antibody.

Reverse crosslinking

For the following microarray, the protein bound on the beads were eluted twice with 250 µl Elution Buffer by rotating for 20 min at room temperature. Elution Buffer,

containing 1% SDS, 50 mM Tris pH8.0 and 10 mM EDTA pH8.0, was prepared freshly. The elution solutions were combined into 500 µl eluate.

To reverse crosslink protein and DNA, 20 µl of 5 M NaCl was added to the 500 µl eluate and 5 µl was added to 100 µl parasite lysate stored before ChIP. The protein/DNA mixture was heated overnight at 65°C. On the next day, 2 µl of 20 mg/ml proteinase K was added to each tube to remove proteins at 37°C for 2 hours. ChIP DNA and input DNA were cleaned by Minielute PCR Cleanup Kit. The DNAs were eluted twice with 17.5 µl Elution Buffer (Qiagen) and stored at -20°C.

2.7.3 DNA extraction and random amplification

RandomPrimer 1 harboring 9 random nucleotides was annealed to ChIP DNA and input DNA. The reaction for ChIP DNA contains 2.9 µl water, 0.6 µl DMSO, 1 µl NEB buffer 2, 0.5 µl of 100 µM RandomPrimer 1 and 5 µl ChIP DNA. The reaction for input DNA contains 6.9 µl water, 0.6 µl DMSO, 1 µl NEB buffer 2, 0.5 µl of 100 µM RandomPrimer 1 and 5 µl input DNA. The reaction was incubated at 70°C for 5 min, 94°C for 5 min and 5°C for 10 min.

The mixture was treated with 2.5 µl water, 0.5 µl NEB buffer 2 and 1 µl of 10 mM dNTP. The reaction was incubated at 37°C for 30 min, 94°C for 5 min and 5°C for 10 min. Then, 1 µl Klenow Fragment exo- (NEB) was added to each reaction and incubated at 37°C for 30 min. Into the input DNA reaction, 30 µl water was added.

The DNA was amplified by PCR reactions using RandomPrimer 2, which is identical to RandomPrimer 1 excluding the 9 random nucleotides. One PCR reaction contains 70.5 µl water, 10 µl of 10x Standard *Taq* Reaction Buffer (NEB), 2.5 µl of 30x aa-dUTP mix, 1 µl of 100 µM RandomPrimer 2, 1 µl of *Taq* polymerase (NEB) and 15 µl ChIP DNA reaction or 15 µl input DNA reaction. The PCR program is 94°C for 5 min, 30 cycles of 94°C for 30 seconds-40°C for 30 seconds-50°C for 30 seconds and 60°C for 2 min.

The amplified DNA was cleaned up with Minielute PCR Cleanup Kit and eluted with 16.5 µl Elution Buffer (Qiagen). For each sample, 4 µg ChIP DNA coupled with Cy5-dye was hybridized with 4 µg input DNA coupled with Cy3 dye.

2.8 Primer list

The primers below are in the 5'- to 3'- orientation. The 20-bp extension for ligation-independent cloning is shown in the small case. The HA and FLAG tags are underlined.

pLNOT1.2_HA_F: gacactatagaactcgagTTCATGGGTTGAACTTATATCATC
pLNOT1.2_HA_R: ttatataactcgacgcggccgagatctcagtcgccggAACGCGCCCTAGGCATGCTTG
pLNOT1.2_seqR: GTATATTGGGGTGATGATAAAATGAAAG
3HA_F:
ccgggTACCCATACGATGTTCCAGATTACGCTTACCCATACGATGTTCCAGATTACGCTTACCCATACGA
TGTTCCAGATTACGCTTGAa
3HA_R:
gatctTCAAGCGTAATCTGGAACATCGTATGGGTAAGCGTAATCTGGAACATCGTATGGGTAAGCGTA
ATCTGGAACATCGTATGGGTAc
pLNOT1.2_3HA_seqR: GTATATTGGGGTGATGATAAAATGAAAG
pTNOT1.1_F: aaatatagatctcccggtTCATAACTCTTCACAATCACATC
pTNOT1.1_R: tctgcatatttaaaaactagtTCTTTAAATGTTAAATTTTTTTTTTCTGG
pTNOT1.1_seqR: AATGCATATATAAAATATACGTCTC
FLAG_F: CATGGGACTACAAAGACGATGACGACAAGTAA
FLAG_R: CTAGTTTACTTGTCGTCATCGTCTTTGTAGTC
pTNOT1.1_3HA_F: aacaaatatagatctcccggtTCATAACTCTTCACAATCACATC
pTNOT1.1_3HA_R: gtacgggtaaagcttgtagcTCTTTAAATGTTAAATTTTTTTTTTCTGG
pTNOT1.1_3HA_seqR: AATGCATATATAAAATATACGTCTC
pTNOT4_3HA_F: aacaaatatagatctcccggtAGATGATGAAGATGAAGAAGATGAC
pTNOT4_3HA_R: cgtacgggtaaagcttgtagcTGCACTTAAATTTAGCTTCTCATTC
pTNOT4_3HA_seqR: AATGCATATATAAAATATACGTCTC
pCC1__NOT1.1_HR1_F: tccaatggccccttccgcggGATCGAAAAATAATGATGTG
pCC1__NOT1.1_HR1_R: aagacagatcttcgtactagtTTTGGTACTTGTTGTTGATA
pCC1__NOT1.1_HR1_seqR: ATAAACCAATAGATAAAATTTGTAG
pCC1__NOT1.1_HR2_F: tatttattaaatctagaattcCTCATCCATAATTATAGGAGAAG
pCC1__NOT1.1_HR2_R: cggtcggatctcaaggtagcAGTTCCTCTAATTTCTGACTATC
pCC1__NOT1.1_HR2_seqR: GCGGTGTGAAATACCGCAC
pCC1__NOT1.2_HR1_F: tccaatggccccttccgcggGAGGATGGATTACCATGA
pCC1__NOT1.2_HR1_R: aagacagatcttcgtactagtGAATGTGTTCAATTTCACTT
pCC1__NOT1.2_HR1_seqR: ATAAACCAATAGATAAAATTTGTAG
pCC1__NOT1.2_HR2_F: tatttattaaatctagaattcGATATGGTAGATACAAATGGAG
pCC1__NOT1.2_HR2_R: gccagcctaggagttccatggCTTATGTGATCTTGATTCTTC
pCC1__NOT1.2_HR2_seqR: GCGGTGTGAAATACCGCAC
pLNOT1.2_HA_ecDDD_F: gacactatagaactcgagTTCATGGGTTGAACTTATATCATC
pLNOT1.2_HA_ecDDD_R: aacatcgtatgggtacctaggAACGCGCCCTAGGCATGCTTG
pLNOT1.2_HA_ecDDD_seqR: TTTTGTAGACCCCATTTGTGAGTAC
pThDHFR_seqR: AGGTGACACTATAGAATACTC
SouBlot_F: ACCAGTAGCTGTCGATAGAGC
SouBlot_R: ACGTAATGGTTCTTTGCATGTAGC
NorBlot1_F: ACCAGTAGCTGTCGATAGAGC

NorBlot1_R: ACTGACTGTAATACGACTCACTATAGGGACGTAATGGTTCTTGCATGTAGC
 NorBlot2_F: CATCAGAAACAAATATCATCGATCC
 NorBlot2_R: ACTGACTGTAATACGACTCACTATAGGGGTGTATATATAAACACAAGATGTG
 NorBlot_control_F: TGGAAGAATGTGCCGAACCTTG
 NorBlot_control_R: actgactgtaatacgactcactatagggGCACATAATGGTTGTTCTGAGGT
 TS-Oligo: AAGCAGTGGTATCAACGCAGAGTACGCrGrGrG
 SMART-Oligo dT: AAGCAGTGGTATCAACGCAGAGTAC(T)₃₀VN (V=A,G,C; N=A,C,G,T)
 SMARTrandom: AAGCAGTGGTATCAACGCAGAGT(N)₉
 Primer IIa: AAGCAGTGGTATCAACGCAGAGT
 RandomPrimer 1: GTTCCCAGTCACGATCNNNNNNNNN
 RandomPrimer 2: GTTCCCAGTCACGATC
 1. CTTTATTATCACCTTTAAGAATTCCTG
 2. GTTAAATATTTAATTAGGAAAGGATAACATG
 3. GTATATTGGGGTGATGATAAAATGAAAG
 4. AGTGCCACCTGACGTCTAAGAAAC
 5. GAATTAGATATGGAAGGTCG
 6. ATATATAGGAATGTCATTACCATC
 7. GTGTATATATAAACACAAGATGTG
 8. TAATATAGAAAAAGTTACATGCACG
 9. CCTAAGTCATTAAGGTACTAACTTAG
 10. GAGGACCAACAAGAAGTAGATC
 11. CATATACATTATTATGTGTGAGTTC
 12. ATGATAATGAAGAAGATCAAGAG
 13. ATATGTCACTTTGTTTTGTTTC
 14. ATGTACAAATGGATGAAGAATCTG
 15. ATATTCTTGACATGTTCTTTATGC
 16. TTACCAAATAATCTGAACGCAAAC
 17. AGATTCTATAGTATTTAAGGATGGTTG
 18. TTTTTCTTTTATCATGCACATTGG
 19. GAAGTATATGAGAAGAATGATTAAGC
 20. TATATCCAATGGCCCCTTC
 21. CAAAATGCTTAAGACAGATCTTC
 22. GAAGAAAAGAAAAAGCAAGGAG
 23. TTTTATTCGTCACACTATCATTTG
 24. ATATGAACAGTATCAACAATATGAAC
 25. TGATTCATATGTTTTTCATCTTCTAC
 26. TATATCCAATGGCCCCTTC
 27. CAAAATGCTTAAGACAGATCTTC
 28. CTTTACAATATGAACATAAAGTACAAC
 29. GCGGTGTGAAATACCGCAC
 30. TTTCCATGCCGATAACGTGATCTAC
 31. ATGCATGGTTCGCTAAACTG
 32. ATCATTCTTCTCATATACTTC
 33. TGTCTTTTGTACATGAAGCAGCAG

Chapter 3 Co-immunoprecipitation of CCR4-NOT subunits

As described in Introduction, most CCR4-NOT subunits identified in eukaryotes are encoded in *P. falciparum*, including two NOT1 paralogs, NOT2, NOT4, NOT5, CAF1, two CCR4 paralogs and CAF40. These proteins share high homologies with their counterparts in other eukaryotes. The similarities suggest the existence of the CCR4-NOT complex or complexes in *P. falciparum*. Furthermore, the highly conserved domains in NOT1.1 and NOT1.2 raises the possibility that both serve as the scaffolds of alternative complexes.

In order to test this proposal, the endogenous *Not1.1* and *Not1.2* were tagged and then their potential binding partners were identified by co-immunoprecipitation and mass spectrometry. Furthermore, two other CCR4-NOT subunits, CAF1 and NOT4 were tagged and analyzed as well in order to extend the knowledge of the CCR4-NOT complex in *P. falciparum*.

3.1 NOT1.2 might be a scaffold of the CCR4-NOT complex

3.1.1 HA-tagging of endogenous *Not1.2*

Three HA epitopes were designed to tag the C-terminus of NOT1.2 protein. The integration of plasmid, pLNOT1.2-HA into the endogenous gene was accomplished by single-crossover recombination, as shown in **Figure 3.1A**.

The T996 parasite strain was transfected with plasmid pLNOT1.2-HA by electroporation. Most parasites died within 6 days, either by the stress of electroporation or by the following drug selection. The parasites that appeared after 2 weeks are referred to as transfectants. The transfectants can be integrants, in which the plasmid has successfully integrated into the endogenous genome. However, the majority of transfectants are parasites containing episomal plasmids because of the low efficiency of integration.

To increase the proportion of integrants in the culture, repeated cycles of drug addition and removal (drug on/off) were performed. The selective drug was removed from the culture to allow non-integrand parasites, which have lost their episomes to proliferate in the absence of selection pressure. Then, the drug was added back to eliminate these wildtype parasites now lacking episomes. The repeated cycling was performed until

the integrant locus was detected in the transfectants. The population of transfectants was then subcloned as described in Section 2.2.5 in order to obtain single clones of integrants.

Figure 3.1. HA-tagging of endogenous *Not1.2*. (A) Integration of plasmid, pLNOT1.2-HA into endogenous *Not1.2*. The plasmid was integrated into the genome of wildtype T996, resulting in the integrant locus of the NOT1.2-HA line. The *BSD* selectable marker is driven by the *PcDHFR-TS* promoter and has the histidine-rich protein terminator, not shown. The homologous arm (HR) of 1,516 bp is a portion of the 3' coding region of endogenous *Not1.2* excluding the stop codon. The tag is the coding sequence of 3 HA epitopes (3xHA) followed by a stop codon. The primers used to detect different loci by PCR amplification are indicated as arrows and sequences are provided in Section 2.8. **(B) Genotyping PCR of the NOT1.2-HA line.** The templates for PCR amplification are gDNA of wildtype (WT) and gDNA of the NOT1.2-HA line. The parental locus (par) was detected by the primer pair 1/2 at 1,782 bp, the integrant locus (int) was detected by the primer pair 1/3 at 1,793 bp and the episomal locus (epi) would have been detected by the primer pair 4/3 at 2,111 bp. The sizes of the ladder marker are given at the left.

3.1.2 Expression of tagged NOT1.2

The subcloned integrant is referred to as the NOT1.2-HA line. The genomic DNA of one clone was analyzed by PCR amplification with primers targeting the parental locus, the integrant locus and the episomal locus specifically. **Figure 3.1B** shows the PCR products amplified from wildtype and the NOT1.2-HA line.

In the NOT1.2-HA line, the detection of the integrant locus indicates the successful integration of the plasmid. The missing parental locus suggests that this subcloned line is not contaminated by wildtype parasites. The episomal locus on the plasmid was not detected by PCR amplification, suggesting that this parasite line may lose the episomes.

Expression of the tagged *Not1.2* at the transcript level was assessed. The mRNA of wildtype and the NOT1.2-HA line was converted into cDNA by reverse transcription, which was used as the template in PCR amplification for indirect visualization of the *Not1.2* transcripts. The PCR products are shown in **Figure 3.2**.

Figure 3.2 Expression of tagged *Not1.2* at the transcript level. PCR amplifications were performed on cDNA of wildtype (WT) and cDNA of the NOT1.2-HA line (1.2HA). The cDNA was prepared by reverse transcription with (+) or without (-) SuperScript III reverse transcriptase. The parental locus was detected by the primer pair 1/2 at 1,782 bp and the integrant locus was detected by the primer pair 5/3 at 755 bp. The primer locations are indicated in Figure 3.1A. The sizes of the ladder marker are given at the left.

The transcript specific to the integrant locus was detected in the NOT1.2-HA line, demonstrating the expression of the tagged *Not1.2* gene. Several controls were included in this experiment. First, the template with cDNA of wildtype serves as a

reference and ensures that the missing parental locus in the NOT1.2-HA line was not due to a failure of PCR reactions. Second, negative control samples were not treated with reverse transcriptase (RT- in **Figure 3.2**) and therefore the mRNA cannot be converted into cDNA. The absence of PCR products in RT- samples confirms that the PCR products were not amplified from gDNA contaminations.

The expression of HA-tagged NOT1.2 at the protein level in schizont was inspected by Western blot as shown in **Figure 3.3**. Two lysis methods in Section 2.6.1 were employed to harvest the total protein lysate, the cytoplasmic lysate and the nuclear lysate in schizonts. All three lysates, obtained from the same amount of parasites, were loaded in SDS-PAGE gel. The lysates were detected by anti-HA antibody, anti-aldolase antibody and anti-histone H3 antibody.

Figure 3.3 Expression of HA-tagged NOT1.2

protein. From lane 1-4, samples are the total lysate of wildtype (WT) and the NOT1.2-HA line (1.2HA), the cytoplasmic lysate of the NOT1.2-HA line (Cyto) and the nuclear lysate of the NOT1.2-HA line (Nuc) at schizont stage. Briefly, lysates of 4 μ l parasite pellet for the HA blot, 1 μ l for the aldolase blot and 1 μ l for the histone H3 blot were loaded. The full-length NOT1.2-HA is highlighted with an arrow. The lane number is given on the top and the sizes of the ladder marker are given at the left.

A faint full-length HA-tagged NOT1.2 at 524 kDa was seen in the total lysate of the NOT1.2-HA line (lane 2), demonstrating the expression of NOT1.2 at the protein level. NOT1.2-HA was detected in the cytoplasmic lysate (lane 3) but not the nuclear lysate (lane 4), implying that NOT1.2 is predominantly located in the cytoplasm at schizont stage. No proteins were detected in the total lysate of wildtype (lane 1), confirming the specificity of anti-HA antibody in the recognition of HA epitopes.

Multiple bands were detected in addition to the full-length protein, possibly due to the processing of the proteins in the parasites, which leads to the truncated proteins containing the C-terminal HA tags. Moreover, some bands were detected above the full-length HA-tagged NOT1.2 in the cytoplasmic lysate. These proteins migrate slower in the SDS-PAGE gel, possibly due to post-translational modifications. This speculation is supported by 17 phosphorylated residues detected in NOT1.2 at schizont stage [249]. The marker for the cytoplasmic compartment is aldolase (PF3D7_1444800) at 40.1 kDa and the marker for the nuclear compartment is histone H3 (PF3D7_0610400) at 15.4 kDa. There are no contaminations in the cytoplasmic lysate and the nuclear lysates based on the missing aldolase in the nuclear lysate and the missing histone H3 in the cytoplasmic lysate.

3.1.3 Co-immunoprecipitation of HA-tagged NOT1.2

Co-immunoprecipitation and mass spectrometry were performed to identify potential interacting partners of NOT1.2. Before preparing a large amount of parasites for mass spectrometric analysis, the specificity and efficiency of immunoprecipitation with the Pierce HA magnetic beads were assessed.

Parasites at schizont stage were harvested for co-immunoprecipitation. The HA magnetic beads (25 µl) were used to immunoprecipitate NOT1.2-HA in the protein lysates, obtained from 50 µl schizont pellet of wildtype and the NOT1.2-HA line. Proteins not bound to the beads were collected as the unbound supernatant (S/N) and the proteins bound to the beads were eluted into the eluate. In **Figure 3.4**, NOT1.2-HA in the unbound supernatant and the eluate were detected by Western blot with anti-HA antibody.

The HA magnetic beads successfully immunoprecipitated the full-length NOT1.2-HA at the correct size in the NOT1.2-HA line (lane 4). The efficiency is high given that NOT1.2 was not detected in the unbound lysate (lane 3). The wildtype lysate immunoprecipitated by the HA magnetic beads serves as a negative control. No proteins were detected in the unbound supernatant (lane 1) and the eluate (lane 2) in wildtype line, suggesting that the HA magnetic beads immunoprecipitates the HA tags specifically.

Figure 3.4 Immunoprecipitation of HA-tagged NOT1.2. The protein lysates of wildtype (WT) and the NOT1.2-HA line (1.2HA) were immunoprecipitated by the HA magnetic beads as described in Section 2.6.3. Briefly, 4% of the input, 4% of the unbound supernatant (S/N), and 50% of the eluate were loaded for Western blot analysis. NOT1.2-HA protein in each lysate was detected by anti-HA antibody. The lane number is given on the top and the sizes of the ladder marker are given at the left.

Since the HA magnetic beads have been verified to be a specific and efficient tool to immunoprecipitate the HA-tagged NOT1.2 protein, independent biological duplicates of co-immunoprecipitation were done in wildtype and the NOT1.2-HA line at schizont stage. All the proteins immunoprecipitated by the HA magnetic beads were sequenced by mass spectrometry. The immunoprecipitation of wildtype lysate was included in order to rule out false positive proteins that were non-specifically immunoprecipitated by the HA magnetic beads.

The mass spectrometry data was matched to the PlasmoDB database using Mascot search software and the sequenced peptides were identified based on Mascot probability scoring [248]. The ions score was calculated for each peptide to measure how well the peptide sequence matches the database sequences. The protein score, as the summed ions score of all peptides identified in one protein, reflects the confidence of this protein detected in the eluate.

Two criteria were employed to identify presumptive interacting partners of NOT1.2. In the first criterion, the proteins detected only in the NOT1.2-HA line were selected as positive hits. The proteins detected in both lines with a $\log_2$ -transformed ratio of protein score in the NOT1.2-HA line versus that in wildtype more than 1.5, were selected as positive hits as well. In the second criterion, the positive hits that were detected in both biological duplicates were identified as potential partners of NOT1.2.

There are 76 positive hits in the first replicate and 151 positive hits in the second replicate. In total, 20 potential partners are identified for NOT1.2, which are listed in **Supplementary Table 2**. These proteins were characterized by Gene Ontology and MPM functional analysis. The enriched functional pathways are listed in **Table 3.1**.

Pathway	Gene	Description
CCR4 - NOT complex	PF3D7_1417200	NOT family protein
	PF3D7_0507600	cell differentiation protein (CAF40)
	PF3D7_0811300	CCR4-associated factor 1 (CAF1)
mRNA-binding proteins	PF3D7_1224300	polyadenylate-binding protein (PABP)
	PF3D7_1417200	NOT family protein
	PF3D7_0814200	DNA or RNA-binding protein Alba 1 (ALBA1)
	PF3D7_0507600	cell differentiation protein (CAF40)
	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1 (PAIP1)
	PF3D7_0811300	CCR4-associated factor 1 (CAF1)
Ribosome	PF3D7_1447000	40S ribosomal protein S5
	PF3D7_1424400	60S ribosomal protein L7-3
	PF3D7_1142500	60S ribosomal protein L28 (RPL28)
	PF3D7_1019400	60S ribosomal protein L30e
Translation	PF3D7_1447000	40S ribosomal protein S5
	PF3D7_1224300	polyadenylate-binding protein (PABP)
	PF3D7_1424400	60S ribosomal protein L7-3
	PF3D7_0815600	eukaryotic translation initiation factor 3 subunit G (EIF3G)
	PF3D7_1142500	60S ribosomal protein L28 (RPL28)
	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1 (PAIP1)
	PF3D7_1019400	60S ribosomal protein L30e
Histone	PF3D7_1105100	histone H2B (H2B)
	PF3D7_0610400	histone H3 (H3)
Uncharacterized proteins	PF3D7_0813300	conserved Plasmodium protein, unknown function
	PF3D7_1471100	exported protein 2 (EXP2)
	PF3D7_1110400	asparagine-rich antigen
	PF3D7_0402000	Plasmodium exported protein (PHISTa), unknown function
	PF3D7_1015900	enolase (ENO)
	PF3D7_1342400	casein kinase II beta chain (CK2beta2)
	PF3D7_0831400	Plasmodium exported protein, unknown function

Table 3.1 Enriched functional pathways of NOT1.2 potential interacting partners. The enriched pathways were identified by Gene Ontology and MPM functional analysis. Those with a binomial $p < 0.05$ are listed. The gene identity and the description of protein product are provided. The proteins in each pathway are organized based on the average peptides detected in biological duplicates of mass spectrometry. The peptide numbers detected for each protein are provided in **Supplementary Table 2**.

The CCR4-NOT complex

NOT1.2 was detected together with two CCR4-NOT subunits, CAF40 and CAF1. Both proteins have been reported to contribute to mRNA metabolism. CAF1 is a key player in mRNA degradation [125, 207] and CAF40 interacts with RNA-binding proteins that function in mRNA degradation and translation repression [150, 250]. CAF40 is expected to interact with NOT1.2 because the DUF3819 conserved domain, known to interact with CAF40, was identified in NOT1.2. On the other hand, NOT1.2 is able to interact with CAF1, although it contains a partial MIF4G domain as described in Section 1.4.2.

NOT2, NOT4 and NOT5 subunits of the CCR4-NOT complex in eukaryotes were not detected. The interaction of these NOT proteins with NOT1.2 might not be as strong as CAF1 and CAF40. Another possibility is that these NOT proteins do not interact with NOT1.2 due to the insertion in the NOT1 domain.

RNA-binding proteins

The results provide evidence for the interaction of NOT1.2 with RNA-binding proteins. The ALBA family includes six DNA/RNA-binding proteins, ALBA1-6, in which ALBA1-4 bind RNA with differential specificities [100, 251]. ALBA1 directly binds 1,193 transcripts in trophozoites, maintains mRNA homeostasis as well as represses translation of four invasion mRNAs, including RAP1, RHOPH3, CDPK1 and AMA1 in *P. falciparum* [101].

PABP contains four RNA recognition motifs and binds the poly(A) tail of mRNA. PABP exerts multiple functions in mRNA metabolism. PABP protects polyadenylated mRNA from degradation [252]. PABP stimulates translation by forming a closed-loop mRNA structure between eIF4G at the 5' cap and PABP at the 3' poly(A) tail. [253]. In *P. falciparum*, PABP displays high specificity for the polyA sequence and interacts with eIF4G, both of which contain characteristic conserved structure features [109], implying their similar functions in the parasite. Translation initiation can also be achieved by the interaction of PABP with initiation factor eIF4B. Their interaction stimulates translation by enhancing PABP binding affinity to poly(A) tail and eIF4A helicase activity [254]. On the other hand, PABP stimulates the efficiency of release

factor eRF1 to recognize the stop codon followed by eRF3 to hydrolyze GTP and thus translation termination [255].

PABP-interacting protein 1 (PAIP1) interacts with PABP via two motifs and acts as a translation enhancer [253]. PAIP1 interacts with eIF3G, which is proposed to stabilize the closed-loop structure [256].

Histones

NOT1.2 is predominantly detected in the cytoplasm at schizont stage as shown in **Figure 3.3**, but histone H2B and H3 were found to be immunoprecipitated with NOT1.2. NOT1.2 might be located in the nucleus at a very low level such that it cannot be detected in Western blot of nuclear lysates. Alternatively, NOT1.2 may interact with newly synthesized histones in the cytoplasm. As no components of transcription machinery were immunoprecipitated by NOT1.2, we cannot conclude whether NOT1.2 plays roles in epigenetic regulation or transcription in *P. falciparum*.

Uncharacterized proteins

Seven proteins interacting with NOT1.2 have not been characterized in the enriched functional pathways.

Three exported proteins interact with NOT1.2. First, EXP2 (PF3D7_1471100) together with PTEX150 and HSP101 comprises the *Plasmodium* translocon of exported proteins (PTEX) channel across the parasitophorous vacuole membrane (PVM). PTEX plays critical roles in exporting proteins across the PVM [257, 258]. EXP2 exports proteins across the PVM in the context of the PTEX channel and also facilitates nutrient passage independent of HSP101 [259], possibly by forming PVM pores [260, 261]. Second, PHISTa (PF3D7_0402000) interacts with cytoskeleton component band 4.1R, which facilitates the interaction between spectrin and actin. PHISTa also interacts with erythrocyte ankyrin ANK1, which links spectrin to the erythrocyte membrane and reduces the deformability of iRBCs [262, 263]. The function of the third exported protein, PF3D7_0831400 is unknown.

PF3D7_0813300, a conserved Plasmodium protein of unknown function has been found in the food vacuole [264].

Asparagine-rich antigen (PF3D7_1110400) has been identified as an RNA-binding protein by bio-informatics [97].

Enolase (PF3D7_1015900) is a glycolytic enzyme shown to play multiple roles. Enolase together with RNase E (an endonuclease), polynucleotide phosphorylase (PNPase) and RhlB helicase forms the RNA degradosome in *E.coli*. The complex induces rapid mRNA decay of glucose transporter PTSG [265], suggesting that enolase is involved in RNA degradation. In *P. falciparum*, plasminogen is the precursor of plasmin, a serine protease that cleaves the extracellular matrix around the targeted host cell. Enolase, as a plasminogen receptor, assists in the parasite invasion [266]. Enolase also interacts with actin in the nucleus [267] and associates with the food vacuole [268].

PF3D7_1342400 is Casein kinase II beta chain (CK2 β 2), a regulatory component of serine/threonine protein kinase CK2 holoenzyme. It interacts with the catalytic subunit CK2 α which exhibits kinase activity in *P. falciparum* [269].

Possible functions of NOT1.2

In summary, the co-immunoprecipitation of PABP, PAIP1 and ALBA1 with NOT1.2, which is consistent with the association of NOT1 with RNA-binding proteins in other eukaryotes, highlights a close connection of NOT1.2 with mRNA. The co-immunoprecipitation of several ribosomal proteins and a translation initiation factor with NOT1.2 reinforces this connection. Asparagine-rich antigen (PF3D7_1110400) might be a new candidate of RNA-binding proteins to interact with NOT1.2. Enolase could be another candidate given its interaction with RNA degradosome machinery in *E.coli* [265]. The interaction of NOT1.2 with these RNA-binding proteins highlights the possible roles of NOT1.2 as a platform to recruit CAF1 deadenylase to mRNA targets for mRNA degradation, as proposed in other eukaryotes.

For the other uncharacterized proteins, their functions and locations are diverse. One possibility is that NOT1.2 plays diverse and distinct roles at PVM, erythrocyte membrane, nucleus and food vacuole specifically in *P. falciparum*. It is also possible that NOT1.2 is in close proximity to these proteins before they are transported to distinct locations.

3.2 Current progress of tagging endogenous *Not1.1*

3.2.1 Failure to detect FLAG-tagged NOT1.1

The tagging of a FLAG epitope to the C-terminus of endogenous *Not1.1* was accomplished by the integration of plasmid, pTNOT1.1-FLAG into the parasite genome, as shown in **Figure 3.5A**. The FLAG tag was chosen because it enables the creation of a parasite line with NOT1.1 tagged with FLAG epitopes and NOT1.2 tagged with HA epitopes.

Figure 3.5. FLAG-tagging of endogenous *Not1.1*. (A) Integration of plasmid, pTNOT1.1-FLAG into endogenous *Not1.1*. The *hDHFR* gene is driven by the calmodulin promoter and has the histidine-rich protein terminator, not shown. The homologous arm (HR) of 1,298 bp is a portion of the 3' coding region of endogenous *Not1.1* excluding the stop codon. The tag is the coding sequence of a FLAG epitope (FLAG) followed by a stop codon. The primers used to detect different loci by PCR amplification are indicated as arrows and sequences are provided in Section 2.8. (B) Genotyping PCR of the NOT1.1-FLAG line. The templates for PCR amplification are gDNA of wildtype (WT) and gDNA of the NOT1.1-FLAG line. The parental locus (par) was detected by the primer pair 6/7 at 1,569 bp, the integrant locus (int) was detected by the primer pair 6/8 at 1,523 bp and the episomal locus (epi) was detected by the primer pair 9/8 at 1,743 bp. The sizes of the ladder markers are given at the left.

The parasite with successful integration of the plasmid was subcloned and referred to as NOT1.1-FLAG lines. The genomic locus of one clone was examined by PCR amplification in **Figure 3.5B**. In the NOT1.1-FLAG line, the detection of the integrant locus indicates the successful integration of the plasmid and the missing parental locus indicates that this clone is not contaminated by wildtype parasites. The episomal locus was detected at a very low level, suggesting a low amount of episomes in this clone.

The expression of FLAG-tagged NOT1.1 at the protein level was examined by Western blot in **Figure 3.6**. The bands detected in the lysate of the NOT1.1-FLAG line were observed in the lysates of wildtype. Different amount of lysates were loaded, ranging from 10 μ g to 40 μ g. For 10 μ g lysate, the band between 171 kDa and 268 kDa can still be seen in both parasite lines, suggesting that the unspecific recognition is not caused by the overload of samples. The anti-FLAG antibody cannot specifically recognize the FLAG-tagged proteins in the T996 strain.

Figure 3.6 Unspecific detection of the FLAG tag in the T996 strain. The samples include the total lysate of wildtype (WT) and the total lysate of the NOT1.1-FLAG line at schizont stage. The lysate amount ranging from 10 μ g to 40 μ g were loaded and detected by anti-FLAG M2 antibody. The sizes of the ladders are given at the left.

To test whether the immunoprecipitation is specific for the FLAG tag, the total lysate of wildtype and the total lysate of the NOT1.1-FLAG line were immunoprecipitated by two methods. The total lysates were obtained from 50 μ l schizont pellet. In the first method, the lysates were incubated with the anti-FLAG M2 affinity gel, which binds

FLAG-tagged proteins directly. In the second method, the lysates were incubated with anti-FLAG M2 antibody to bind FLAG-tagged proteins, which were then bound by the Pierce A/G magnetic beads. The total protein lysate (Input), the proteins unbound to the gel or the beads (S/N) and the proteins bound to the gel or beads (Eluate) were analyzed by Western blot in **Figure 3.7**.

Figure 3.7 Unspecific immunoprecipitation of the FLAG tag. The protein lysates of wildtype (A) and the NOT1.1-FLAG line (B) were immunoprecipitated as described in Section 2.6.3. Briefly, 4% of the input, 4% of the unbound supernatant (S/N), and 50% of the eluate were loaded in Western blot and blotted with anti-FLAG M2 antibody. The sizes of the ladder markers are given at the left.

Some proteins were recognized by anti-FLAG M2 antibody in the eluates of both methods using the NOT1.1-FLAG line. However, the same bands were observed in the lysates of wildtype, suggesting the recognition of some untagged proteins by the anti-FLAG M2 antibody and the FLAG antibody on the affinity gel. The unspecific recognition by different FLAG antibodies made the FLAG tag an unsuitable choice in the T996 parasite strain.

3.2.2 HA-tagging of endogenous *Not1.1*

Due to the failure to detect NOT1.1 with a FLAG epitope, the tag was switched to HA epitopes. The integration of plasmid, pTNOT1.1-HA was accomplished by single-crossover recombination in Figure 3.8A. The HA epitopes were fused to the C-terminus of endogenous Not1.1, similar to the HA-tagged NOT1.2.

Figure 3.8. HA-tagging of endogenous *Not1.1*. (A) Integration of plasmid, pTNOT1.1-HA into endogenous *Not1.1*. The *hDHFR* gene is driven by the calmodulin promoter and has the histidine-rich protein terminator, not shown. The homologous arm (HR) of 1,298 bp is a portion of the 3' coding region of endogenous *Not1.1* excluding the stop codon. The tag is the coding sequence of 3 HA epitopes (3xHA) followed by a stop codon. The primers used to detect different loci by PCR amplification are indicated as arrows and sequences are provided in Section 2.8. (B) Genotyping PCR of the NOT1.1-HA line. The templates for PCR amplification are gDNA of wildtype (WT) and gDNA of the NOT1.1-HA line. The parental locus (par) was detected by the primer pair 6/7 at 1,569 bp, the integrant locus (int) was detected by the primer pair 6/8 at 1,589 bp and the episomal locus (epi) was detected by the primer pair 9/8 at 1,799 bp. The sizes of the ladders are given at the left.

The parasite with successful integration of the plasmid was subcloned and referred to as NOT1.1-HA lines. The integration in one clone was validated by PCR amplification on the gDNA template in **Figure 3.8B**. In the NOT1.1-HA line, the detection of the integrant

locus indicates the successful integration of the plasmid and the missing parental locus indicates that this line is a single clone without contaminations of wildtypes. The episomal plasmids exist in the parasite as the episomal locus was detected.

The immunoprecipitation of NOT1.1 has not been completed yet. In order to understand the compositions of the complex, the co-immunoprecipitation of other subunits, CAF1 and NOT4 were performed as followed.

3.3 CAF1 is a binding partner of NOT1.1 and NOT1.2

3.3.1 HA-tagging of endogenous *Caf1*

The tagging of CAF1 to generate the integrant parasite line CAF1-HA is shown in **Figure 3.9A**.

Figure 3.9. HA-tagging of endogenous *Caf1*. (A) Integration of plasmid, pTCAF1-HA into endogenous *Caf1*. The *hDHFR* gene is driven by the calmodulin promoter and has the histidine-rich protein terminator, not shown. The tag is the mCherry fluorescent protein and 3 HA epitopes (3xHA). The homologous arm (HR) is a portion of the 3' coding region of endogenous *Caf1* excluding the stop codon. The primers used to detect different loci are indicated as arrows and sequences are provided in Section 2.8. (B) Genotyping PCR of the CAF1-HA line. The templates for PCR amplification are gDNA of wildtype (WT) and gDNA of the CAF1-HA line. The parental locus (par) was detected by the primer pair 10/11 at 1,168 bp, the integrant locus (int) was detected by the primer pair 10/8 at 2,047bp and the episomal locus (epi) was detected by the primer pair 9/8 at 2,058 bp. The sizes of the ladder markers are given at the left.

The tag consists of three HA epitopes and the red fluorescent protein, mCherry fused to the C-terminus of endogenous *Caf1*. The transfection and subcloning were done by a post-doctoral fellow in the Featherstone and Bozdech laboratories, Dr. Ragini Rai (not published).

The integration of the plasmid pTCAF1-HA in one clone was validated by PCR amplification on gDNA as shown in **Figure 3.9B**. In the CAF1-HA line, the detection of the integrant locus indicates the successful integration of the plasmid. This clone is not contaminated by wildtype parasites given the missing parental locus. Episomal locus was detected, suggesting the existence of the episomes in this clone.

3.3.2 Co-immunoprecipitation of HA-tagged CAF1

Immunoprecipitation with the HA magnetic beads was performed on protein lysates of wildtype and the CAF1-HA line, obtained from 50 µl schizont pellet. The procedures are the same as the immunoprecipitation of NOT1.2 in Section 3.1.3. The wildtype lysate serves as a negative control. In **Figure 3.10**, the HA-tagged CAF1 in the total lysate (Input), the unbound protein lysate (S/N) and the bound proteins (Eluate) were detected by anti-HA antibody in Western blot.

No proteins were detected by anti-HA antibody in the wildtype lysates (lane 1-3), suggesting that the magnetic beads specifically recognize HA epitopes in the T996 strain. The full-length CAF1 was detected in the input (lane 4) and the eluate (lane 6) in the CAF1-HA line, confirming that the HA-tagged CAF1 was indeed immunoprecipitated by the HA magnetic beads. As CAF1 was not detected in the unbound lysate in the CAF1-HA line (lane 5), the immunoprecipitation with HA magnetic beads was thorough.

Actin, serving as the loading control, was detected by anti-actin antibody in **Figure 3.10**. The amount of actin is comparable between the input of wildtype and the input of the CAF1-HA line, confirming the comparable loading of these two parasite lines. Actin at 42 kDa was detected in the unbound supernatant at a similar amount to that in the input and actin was not detected in the eluate, suggesting that the majority of actin was not co-immunoprecipitated by HA-tagged CAF1 or HA magnetic beads. The band

at approximately 50 kDa is the heavy chain of HA antibody, which was incorporated in the magnetic beads and eluted together with the bound proteins. The heavy chain was detected by secondary anti-mouse antibody as both HA antibody and anti-actin antibody are produced in mouse.

Figure 3.10 Immunoprecipitation of HA-tagged CAF1. The protein lysates of wildtype (WT) and the CAF1-HA line were immunoprecipitated by the Pierce HA magnetic beads. Briefly, 4% of the input, 4% of the unbound supernatant (S/N), and 50% of the eluate were loaded and blotted with anti-HA antibody and anti-actin antibody. The lane number is labeled on the top and the sizes of the ladders are given at the left.

Biological duplicates of co-immunoprecipitation and mass spectrometry of wildtype and CAF1-HA lysates in schizonts were performed. The criteria used to identify the presumptive interacting partners of NOT1.2 in Section 3.1.3 were employed. In total, 391 proteins were selected in the first replicate and 501 proteins were selected in the second replicate as positive hits. Eventually, 244 proteins found in both replicates are identified as potential CAF1 partners and listed in **Supplementary Table 3**. The functional pathways in which these genes are implicated are shown in **Supplementary Table 4** and summarized as followed.

CAF1 interacts with both NOT1.1 and NOT1.2. The interaction with NOT1.2, which is consistent with the previous reciprocal immunoprecipitation of NOT1.2, reinforces the scaffold function of NOT1.2. In addition, CAF1 was found to interact with NOT1.1, suggesting that NOT1.1 is likewise a scaffold of the CCR4-NOT complex as well.

The interaction of CAF1 with CCR4 is expected because CAF1 is known to recruit CCR4 to the complex [136]. CAF40 was co-immunoprecipitated with CAF1. CAF40 is reported

to bind NOT1.1 via the DUF3819 domain, which is close to the MIF4G domain for CAF1 binding [148]. The interaction of CAF40 and CAF1 might be explained by proximity.

NOT4 was detected in the eluates of both wildtype and the CAF1-HA line with log₂-transformed protein score ratios well below 1.5. These observations suggest that NOT4 might interact with CAF1, but not as stably as NOT1.1, NOT1.2, CCR4 and CAF40. This is similar but not identical to the status of NOT4 in human and *Drosophila*, in which NOT4 is not immunoprecipitated with any of other CCR4-NOT subunits *in vivo* [135, 140]. However, NOT4 is able to interact with human NOT1 in the yeast two-hybrid assay [140].

The proteins co-immunoprecipitated by CAF1 are enriched in many functional pathways. First, many RNA-binding proteins and proteins involved in RNA degradation were detected, consistent with the role of CAF1 in deadenylation of mRNA. One interesting binding partner is mRNA-decapping enzyme 2 (DCP2), which triggers translation repression and mRNA degradation in mouse embryonic fibroblasts [270]. The interactions of CAF1 with DCP2 and PABP link decapping at the 5' end of mRNA and deadenylation at the 3' poly(A) tail, two crucial processes in mRNA degradation. Another binding partner is DNA/RNA-binding protein ALBA4, which interacts with proteins involved in translation machinery, mRNA export and RNA-binding proteins, including CAF1-interacting partner PABP in *P. yoelii* at schizont stage [102]. At both gametocyte and schizont stage, ALBA4 interacts with many proteins associated with DOZI/CITH in the stress granule in gametocyte, supporting the role of ALBA4 in mRNA metabolism [102]. In addition, multiple proteins involved in ribosome assembly and translation initiation were detected, indicating that CAF1 is in proximity to ribosomes.

Second, CAF1 might be associated with mRNA surveillance. One interacting protein is cleavage and polyadenylation specificity factor 3, CPSF3 (PF3D7_1438500). It is a component of the CPSF complex, which is essential for processing of pre-mRNA by addition of the poly(A) tail at the 3' end [271].

Third, CAF1 might function in histone modification and transcription. Histone H3 and H3 variant are specifically detected as well as multiple histone modification proteins, including histone methyltransferase, histone acetyltransferase and 2 histone

deacetylases. For transcription, CAF1 interacts with the RNAPII machinery, including RNAPII transcription factor TFB2, TFB4, transcription elongation factor SPT5 and one ApiAP2 transcription factor, PF3D7_1007700. CAF1 also interacts with DNA-directed RNA polymerase I and III subunit RPACI.

Surprisingly, CAF1 associates with DNA replication factors. The alpha subunit and the epsilon catalytic subunit of DNA polymerase, DNA primase large subunit, replication factors and putative DNA topoisomerase 3 (TOP3) were detected as potential CAF1 interacting partners. CAF1 also interacts with abundant transport proteins and vesicle membrane proteins. Thus, CAF1 might exert novel functions in DNA replication and protein transport in *P. falciparum*, which have not been reported in other eukaryotes.

Seven 26S proteasome regulatory subunits and an ubiquitin-protein ligase (PF3D7_1119400) were found to interact with CAF1. NOT4 is an ubiquitin E3 ligase, required for the integrity of proteasome in yeast [215]. CAF1 interacts with the proteasome regulatory complex possibly through NOT4, though it is unlikely given that NOT4 is not interacting with CAF1 in a stable manner as described above. The association of the proteasome with CAF1 is more likely via CAF1 directly instead of being mediated by NOT4.

Many more potential partners were co-immunoprecipitated by CAF1 than by NOT1.2. NOT1.2 as a scaffold of the CCR4-NOT complex presents in the complex with a very high molecular weight. The known subunits of the complex include 200 kDa CAF1, 73 kDa CAF40 and 289 kDa CCR4. Association with multiple partners might render limited access of the HA magnetic beads to the C-terminal tags. Another challenge is the enormous size of NOT1.2-HA at 524 kDa, making the co-immunoprecipitation more challenging than for CAF1-HA.

3.4 NOT4 might be a subunit of the CCR4-NOT complex

3.4.1 HA-tagging of the endogenous *Not4*

To elucidate whether NOT4 is a subunit of the complex, the C-terminus of endogenous *Not4* was tagged with HA epitopes, by the strategy shown in **Figure 3.11A**.

The integrant was subcloned and the genomic loci of one NOT4-HA line were tested by PCR amplification, as shown in **Figure 3.11B**. The plasmid has been successfully inserted into the genome given the detection of the integrant locus. This NOT4-HA clone is not contaminated by wildtype parasites given the missing parental locus. The episomal plasmids existed in the NOT4-HA line given the detection of the episomal locus.

Figure 3.11. HA-tagging of endogenous *Not4*. (A) Integration of plasmid, pTNOT4-HA into endogenous *Not4*. The *hDHFR* gene is driven by the calmodulin promoter and has the histidine-rich protein terminator, not shown. The tag is the coding sequence of 3 HA epitopes (3xHA) followed by a stop codon. The homologous arm (HR) of 1,045 bp is a portion of the 3' coding region of endogenous *Not1.1* excluding the stop codon. The primers used to detect different loci by PCR amplification are indicated as arrows and primer sequences are provided in Section 2.8. (B) Genotyping PCR of the NOT4-HA line. The templates for PCR amplification are gDNA of wildtype (WT) and gDNA of the NOT4-HA line. The parental locus (par) was detected by the primer pair 12/13 at 1,277 bp, the integrant locus (int) was detected by the primer pair 12/8 at 1,443 bp and the episomal locus (epi) was detected by the primer pair 8/9 at 1,546 bp. The sizes of the ladder marker are given at the left.

The expression of NOT4 protein was tested by Western blot as shown in **Figure 3.12**. The total lysate, the cytoplasmic lysate and the nuclear lysate were obtained from the same amount of parasites and loaded in SDS-PAGE gel. NOT4 was detected in both cytoplasm and nucleus at schizont stage. Markers of the cytoplasmic and nuclear compartments are GAPDH and histone H3, respectively.

Figure 3.12 Expression of HA-tagged NOT4.

The samples include the cytoplasmic lysate (Cyto) and the nuclear lysate (Nuc) of wildtype (WT) and the NOT4-HA line at schizont stage. Briefly, lysates of 4 μ l parasite pellet for the HA blot, 1 μ l for the aldolase blot and 1 μ l for the histone H3 blot were loaded. The full-length NOT4 is highlighted with the arrow. The sizes of the ladder markers are given at the left.

3.4.2 Co-immunoprecipitation of HA-tagged NOT4

To search for evidence of a physical association between NOT4 and other components of the CCR4-NOT complex, co-immunoprecipitation of NOT4-HA by the HA magnetic beads was performed on lysates of wildtype and the NOT4-HA line obtained from 50 μ l schizont pellet.

Before harvesting a large number of parasites for mass spectrometry, the efficiency of the HA magnetic beads in immunoprecipitating NOT4-HA was tested, following the procedures described in Section 3.1.3. The co-immunoprecipitation of wildtype lysates provides a negative control. In **Figure 3.13**, the HA-tagged NOT4 in the total lysate (Input), the unbound proteins (S/N) and bound proteins (Eluate) was detected by anti-HA antibody in Western blot.

The full-length NOT4-HA at 198 kDa was detected in the total lysate (lane 3) and the eluate (lane 9) of the NOT4-HA line, indicating that the immunoprecipitation was

successful. As NOT4-HA was still detected in the unbound lysate in the NOT4-HA line (lane 8), the immunoprecipitation of NOT4-HA was not complete.

Figure 3.13 Immunoprecipitation of HA-tagged NOT4. The protein lysates of wildtype (WT), the NOT1.2-HA (1.2HA) line and the NOT4-HA (4HA) line were immunoprecipitated by the Pierce HA magnetic beads as described in Section 2.6.3. Briefly, 4% of the total lysate (Input), 4% of the unbound supernatant (S/N) and 50% of the eluate (E) were loaded and probed with anti-HA antibody and anti-actin antibody. The lane number is given on the top and the sizes of the ladder markers are given at the left.

The immunoprecipitation of NOT1.2 was included in the blot as well. Compared to the intensity of NOT4-HA in lane 3, NOT1.2-HA was hardly detected in the same amount of total lysate in lane 2. NOT1.2-HA at 524 kDa is much larger than NOT4-HA at 198 kDa, challenging the detection of protein by Western blot. On the other hands, NOT1.2 as a scaffold of the complex may have been buried by other CCR4-NOT subunits, sterically obstructing contact with the HA magnetic beads.

Actin was detected by anti-actin antibody in **Figure 3.13**, serving as the loading control. The amount of actin in the input of wildtype, the NOT1.2-HA line and the NOT4-HA line is comparable, suggestive of equal loadings. Actin was detected in the unbound supernatant but not the eluate in all the three lines. The heavy chain of HA antibody was detected by the secondary anti-mouse antibody.

Biological duplicates of co-immunoprecipitation and mass spectrometry of wildtype and the NOT4-HA line at schizont stage were performed. The criteria used to identify the potential binding partners of NOT1.2 in Section 3.1.3 were employed once again. In total, 89 proteins were selected in the first replicate and 386 proteins were selected in the second replicate. Eventually, 45 proteins are filtered as potential NOT4 interacting partners and listed in **Supplementary Table 5**. Functional pathways which are implicated by potential NOT4 interacting partners are shown in **Table 3.2**.

NOT1.1 and NOT1.2 were not detected among the presumptive binding partners of NOT4. On the other hand, CAF40 subunit was detected, which was also found to bind NOT1.2 and CAF1 in this study. CAF1 was detected but only in one replicate, suggesting that NOT4 might interact with CAF1. These observations support the argument that NOT4 might not directly interact with NOT1.1 and NOT1.2, but instead with CAF40 and possibly CAF1.

NOT4 associates with abundant ribosomal proteins and two components of the translation initiation complex, suggesting that NOT4 is in proximity to the ribosome, similar to NOT1.2 and CAF1. NOT4 interacts with five RNA-binding proteins, all of which interact with NOT1.2 or CAF1 or both. The co-immunoprecipitation of ribosomal proteins and RNA-binding proteins by NOT1.2, CAF1 and NOT4 raises the possibility that NOT4 is recruited to ribosome as a subunit of the complex.

In addition to cytoplasmic proteins, co-immunoprecipitated proteins with a nuclear location were also detected, consistent with the dual locations of NOT4 in both compartments in Figure 3.12. The general transcription factor TFIIH plays major roles in transcription initiation and nucleotide excision repair. NOT4 interacting protein P44 is a component of TFIIH and assists in both roles [272, 273]. The detection of a TFIIH component, an ApiAP2 transcription factor (PF3D7_1239200) and histone H3 suggests possible roles of NOT4 in transcription.

NOT4 also interacts with three proteins involved in protein folding, HSP101, HSP90 and protein disulfide isomerase, PDI8, which is found to facilitate the conformational folding of EBA175 [274]. Moreover, NOT4 interacts with RNA helicase UAP56, which plays roles in spliceosome assembly [275, 276] and mRNA export [277, 278].

Pathways	Gene	Description
Ribosome	PF3D7_1130200	60S ribosomal protein P0 (PfP0)
	PF3D7_1447000	40S ribosomal protein S5
	PF3D7_1105400	40S ribosomal protein S4
	PF3D7_1424400	60S ribosomal protein L7-3
	PF3D7_1351400	60S ribosomal protein L17
	PF3D7_0814000	60S ribosomal protein L13-2
	PF3D7_1323400	60S ribosomal protein L23 (RPL23)
	PF3D7_1465900	40S ribosomal protein S3
	PF3D7_0309600	60S acidic ribosomal protein P2 (PfP2)
	PF3D7_1242700	40S ribosomal protein S17
	PF3D7_0520000	40S ribosomal protein S9
Translation initiation	PF3D7_1224300	polyadenylate-binding protein (PABP)
	PF3D7_1206200	eukaryotic translation initiation factor 3 subunit C (EIF3C)
	PF3D7_0815600	eukaryotic translation initiation factor 3 subunit G (EIF3G)
	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1 (PAIP1)
mRNA-binding proteins	PF3D7_1224300	polyadenylate-binding protein (PABP)
	PF3D7_1006200	DNA or RNA-binding protein Alba 3 (ALBA3)
	PF3D7_0814200	DNA or RNA-binding protein Alba 1 (ALBA1)
	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1 (PAIP1)
	PF3D7_0507600	cell differentiation protein (CAF40)
Transcription related proteins	PF3D7_1235300	CCR4-NOT transcription complex subunit 4 (NOT4)
	PF3D7_1239200	transcription factor with AP2 domain(s) (ApiAP2)
	PF3D7_1314900	general transcription factor IIH subunit 2 (P44)
	PF3D7_0610400	histone H3 (H3)
DNA binding	PF3D7_1006800	single-strand telomeric DNA-binding protein GBP2
	PF3D7_1006200	DNA or RNA-binding protein Alba 3 (ALBA3)
	PF3D7_0814200	DNA or RNA-binding protein Alba 1 (ALBA1)
	PF3D7_1239200	transcription factor with AP2 domain(s) (ApiAP2)
	PF3D7_0610400	histone H3 (H3)
Chaperones and protein folding	PF3D7_1116800	heat shock protein 101 (HSP101)
	PF3D7_0708400	heat shock protein 90 (HSP90)
	PF3D7_0827900	protein disulfide isomerase (PDI8)

Table 3.2 Enriched functional pathways of NOT4 potential interacting partners. The enriched pathways were identified by Gene Ontology and MPM functional analysis. Those with a binomial $p < 0.05$ are listed. The gene identity, gene description and their protein scores measured in the replicates of Mass Spectrometry are provided. The genes in each list are organized based on the average value of the score. The peptide numbers detected for each protein are provided in **Supplementary Table 5**.

NOT4 functions in ubiquitin-dependent protein degradation as an E3 ligase as found in eukaryotes [153]. However, the proteins involved in ubiquitination are only identified in the second biological replicate, including ubiquitin-like protein (PF3D7_0922100), polyubiquitin (PF3D7_1211800), E3 ubiquitin-protein ligase (PF3D7_0707700) and ubiquitin carboxyl-terminal hydrolase (PF3D7_1211800). One potential NOT4 interacting partner is an erythrocyte membrane-associated antigen (PF3D7_0703500), which interacts with PF3D7_0726500, a ubiquitin carboxyl-terminal hydrolase and interacts with PF3D7_0628100, a ubiquitin transferase [279]. Together, these findings are consistent with a role for NOT4 in ubiquitination and degradation by the proteasome, as for its ortholog in other species.

The failure to detect these proteins in the first replicate may be due to experimental variations. The conditions may have been more stringent for the first replicate compared to the second and only the binding partners with very strong affinity were, therefore, co-immunoprecipitated in the first replicate. In the first replicate, no proteasome regulatory subunits were detected, supporting the previous findings in this study that CAF1 might interact with proteasome directly.

Chapter 4 Functional analysis of NOT1.1 and NOT1.2

4.1 Creation of the *Not1.1* mutant

4.1.1 Disruption of endogenous *Not1.1* by a selectable marker

The pCC1 vector was initially designed by Dr. Alan F. Cowman laboratory [239]. The pCC1-NOT1.1-KO vector was constructed as described in Section 2.1.8. The vector contains a positive selectable marker, the human dihydrofolate reductase gene (*hDHFR*), conferring resistance to the selective drug, dihydrotriazine WR99210 [280]. The *hDHFR* cassette flanked by two homologous arms is designed to integrate into the endogenous *Not1.1* by double-crossover homologous recombination and thus disrupt the endogenous *Not1.1* gene, as shown in **Figure 4.1A**.

Two homologous arms include HR1, 158 bp–1,052 bp and HR2, 1,881 bp–2,822 bp downstream of the endogenous *Not1.1* start codon, corresponding to the region of 53–351 aa and 627–941 aa on NOT1.1 protein. NOT1.1 contains the MIF4G domain at 623–847 aa, the DUF3819 domain at 1,083–1,227 aa and the NOT1 domain at 2,702–3,054 aa. The insertion of the *hDHFR* cassette is expected to disrupt the expression of all the conserved domains.

The vector contains the *Saccharomyces cerevisiae* cytosine deaminase/uracil phosphoribosyltransferase chimeric gene (*ScFCU*) as a negative selectable marker. The *ScFCU* chimeric gene converts innocuous 5-fluorocytosine (5-FC) into toxic 5-fluorouracil, which inhibits RNA synthesis and thymidylate synthase [239]. The two selectable markers were included in the vector in order to increase the efficiency of the selection of integration events and thus the possibility of obtaining single clones of integrants.

The T996 parasites was transfected with the plasmid by electroporation. Drug cycling of the *hDHFR* selection was carried out for three times to increase the proportion of integrants. Then, 5-FC at a concentration of 100 μ M, which is 100 times higher than the suggested concentration [281], was added to the culture aiming to kill the parasites containing episomal plasmids. Genomic DNA of the transfectants was harvested after 7 days of 5-FC treatment and **Figure 4.1B** shows the detection of genomic loci by PCR amplification.

Figure 4.1. Disruption of *endogenous* *Not1.1*. (A) Integration of plasmid, pCC1-NOT1.1-KO into *endogenous* *Not1.1*. The *hDHFR* gene is driven by the calmodulin promoter and has the histidine-rich protein terminator. The *FCU* gene is driven by the HSP86 promoter and has the *DHFR* terminator, not shown. Two homologous arms (HR1 and HR2) flank the *hDHFR* cassette. The primers used for PCR analysis are indicated in the diagram and primer sequences are provided in Section 2.8. (B) Genotyping PCR of the *Not1.1* transfectants after 5-FC treatment. PCR amplification was conducted on genomic DNA of wildtype and the *Not1.1* transfectants. Parasites were treated with 100 μ M 5-FC from two companies, Sigma and Invivos. The size of the PCR product is 1,456 bp for primers 14/15 (par1), 1,146 bp for primers 16/17 (par2), 1,796 bp for primers 14/18 (int1), 1,646 bp for primers 19/17 (int2), 951 bp for primer 20/21 (epi1) 1,165 bp for primer 22/23 (epi2). The size markers are given at the left. (C) Genotyping PCR of the presumptive *Not1.1* mutant. PCR amplification was conducted on genomic DNA of wildtype and the cloned *Not1.1* mutant. The sizes of the ladder are given at the left.

The episomal loci can still be detected in the transfectants treated with 5-FC from two companies, suggesting that 5-FC failed to kill the parasites bearing episomal plasmids. It is possible that T996 parasite strain processes a novel 5-FC resistance mechanism, which is not found in the 3D7 strain.

Limiting dilution was then performed to obtain single clones of parasites having undergone integration. The genomic DNA of each clone was tested by PCR amplification and only one clone was confirmed to be the integrant lacking the parental locus, as shown in **Figure 4.1C**.

The detection of two parental loci in genomic DNA of wildtype ensures that the designed primers perform well in PCR amplification and lend confidence to the conclusion that the missing bands in the presumptive *Not1.1* mutant are due to the predicted double-crossover recombination event instead of an amplification failure. The detection of integrant loci in the presumptive *Not1.1* mutant confirmed the insertion of the *hDHFR* cassette. In the presumptive *Not1.1* mutant, episomal plasmids could still be detected.

4.1.2 Validation of the knockout of *Not1.1*

PCR analysis on genomic DNA suggests that double-crossover homologous recombination had occurred and that the *hDHFR* cassette had disrupted the parental loci in the clone of *Not1.1* mutant. To ensure that the parental genome is not simply rearranged, Southern blot was performed to ensure that there is only one disrupted copy of *Not1.1* gene in this presumptive *Not1.1* mutant.

The Southern blotting procedures are described in Section 2.3.2 and the design is depicted in **Figure 4.2A**. The probe is located downstream of homologous arm 2, present in both wildtype and the *Not1.1* mutant. The probe was fused with a DIG tag and detected by anti-DIG antibody followed by chemiluminescence.

Genomic DNA of wildtype and the *Not1.1* mutant was digested by BamHI and PstI restriction enzymes. The expected fragment containing the probe in wildtype is 8,216 bp. An extra BamHI site in the coding region of the *hDHFR* cassette will result in a

shorter fragment at 4,102 bp in the *Not1.1* mutant. **Figure 4.2B** shows the genomic fragments detected by Southern blot.

Figure 4.2 Southern blot on the *Not1.1* mutant. (A) Design of the Southern blot. Two homologous arms are drawn in dark grey, flanking the orange *hDHFR* cassette. The positions of restriction enzymes are labeled. The expected length of the digested fragment is labeled. The position of the probe inside the fragment to be detected is marked as a red square. The lengths of *Not1.1* gene, the homologous arms and the *hDHFR* cassette are not drawn to scale. **(B) Detection of genomic fragments in wildtype (WT) and the *Not1.1* mutant (MT).** Digested genomic DNA was separated in an agarose gel, transferred to a Nylon membrane, blotted with DIG-labeled probe and detected by the anti-DIG antibody. The signals were exposed to X-ray film. The sizes of DIG-labeled DNA ladder is shown at the left with the sizes in kb.

In **Figure 4.2B**, a single expected fragment at 8.2kb in wildtype confirms the complete digestion of genomic DNA and the reliability of the DIG-labeled system. The expected fragment at 4.1 kb in the *Not1.1* mutant confirms the insertion of the *hDHFR* cassette. The presence of the single fragment illustrates that there is only one copy of mutated *Not1.1* in the genome. The mutant does not contain another copy of wildtype *Not1.1* to support its viability.

In addition to Southern blot, Northern blot was performed and provides additional proof of the knockout of *Not1.1* at the transcript level. The *Not1.1* transcripts in wildtype and in the *Not1.1* mutant line are depicted in **Figure 4.3A** with corresponding

sizes. Two probes were employed. Probe 1 used in the Northern blot is the same with that of the Southern blot and Probe 2 is located at the end of the coding region of endogenous *Not1.1*. The DNA probes were reverse transcribed into RNA to achieve more efficient hybridization.

The length of wildtype *Not1.1* transcript is 10,100 nucleotides (nt) excluding the 5' UTR and the 3' UTR. The lengths of these two regions are difficult to determine precisely, therefore a transcript greater than 10,100 nt is expected.

Two types of potential mutated *Not1.1* transcripts can be predicted taking into account that the *hDHFR* gene is approximately 500 nt in length and harbors a 5' and 3' processing site as shown in **Figure 4.3A**. One mutated *Not1.1* transcript could extend from the start of the *Not1.1* 5' UTR through the end of the 3' UTR of the *hDHFR* cassette, encoding a transcript of more than 2,600 nt. Such a transcript would not be expected to provide NOT1.1 function, because the majority of the endogenous *Not1.1* gene (9,064 bp out of 10,116 bp) is not transcribed. This transcript would not be detected in Northern blot because the probes are located outside the transcribed region. In a second scenario, the transcript would again start from the endogenous *Not1.1* 5' UTR but extend beyond the 3' UTR of the *hDHFR* gene, ending at the endogenous 3' UTR of *Not1.1*. The transcript would contain all the remaining sequence of *Not1.1* gene and would be detected in Northern blots at a length of more than 11,400 nt.

Figure 4.3B shows the *Not1.1* transcripts at ring, trophozoite and schizont stages detected by the DIG-labeled Probe 1. Clear bands running above the 10,000 nt ladder were observed at trophozoite and schizont stages of wildtype total RNA, corresponding to the expected length of more than 10,100 nt. The transcript was barely detected in rings, most abundant in trophozoites and decreased in schizonts, matching its expression profiling in the PlasmoDB database (**Figure 1.8**). With RNA from the presumptive *Not1.1* knockout, no bands at this same size were detected, suggesting that there are no transcripts spanning Probe 1 and excluding the possibility of long transcripts containing full sequence of *Not1.1* gene.

Figure 4.3 Northern blot on the *Not1.1* mutant. (a) Design of the Northern blot. The lengths of *Not1.1* transcripts in wildtype and the presumptive *Not1.1* mutant are labeled. The locations of probes targeting *Not1.1* transcript are indicated as a red square (Probe 1) and an orange square (Probe 2). **Transcripts detected by DIG-labeled RNA Probes 1 (B) and Probe 2 (C).** The total RNA was harvested AT three stages, ring (R), trophozoite (T) and schizont (S) in wildtype and the presumptive *Not1.1* mutant. Briefly, 30 μ g of total RNA in each stage was loaded and blotted as described in Section 2.4.4. The arrow points the location of wildtype *Not1.1* transcript. The sizes of denatured Generuler 1kb ladder are labeled at the left.

Probe 2 was used to blot the membrane in **Figure 4.3B**, which was stripped after the blotting of Probe 1. The *Not1.1* transcripts detected in wildtype RNA but not mutant RNA are consistent with those detected by Probe 1. The second probe verifies that any

Not1.1 transcript, which may be transcribed do not harbor sequences downstream of the *hDHFR* insertion or are present at undetectable levels.

Serine tRNA-ligase (PF3D7_0717700) was used as the loading control for the Northern blots. The coding region of Serine tRNA-ligase is 1,620 bp, thus a transcript more than 1,620 nt is expected. The control probe is located in the coding region of Serine tRNA-ligase. The expression profile of Serine tRNA-ligase is shown in **Figure 4.4A** (adapted from the PlasmoDB database) and the Northern blot with the control probe on wildtype and the *Not1.1* mutant is shown in **Figure 4.4B**.

Figure 4.4. Northern blot with the control probe on the *Not1.1* mutant. (A) Expression of Serine tRNA-ligase across asexual and sexual stages, adapted from the PlasmoDB database. The x-axis indicates different stages, ring, trophozoite and schizont in human host as well as gametocytes and ookinete sexual stages in mosquito. The y-axis shows the level of transcripts measured by RNA-Seq. **(B) Transcripts of Serine rRNA-ligase detected by the control probe.** Briefly, 10 µg of total RNA in each stage was loaded and blotted with the control probe. The arrow points to the location of the transcript of Serine tRNA ligase. The sizes of denatured Generuler 1 kb ladder is shown at the left. The locations of abundant 28S and 18S are labeled as well.

In the blot in **Figure 4.4B**, 10 µg total RNA was used to detect the Serine tRNA-ligase transcript. The blot probed by the Probe 1 and Probe 2 was not stripped and reblotted by the control probe because 30 µg total RNA gave high levels of background signals for the control probe (data not shown).

The abundant 18S ribosomal RNA at approximately 2,100 nt and 28S ribosomal RNA at approximately 3,800 nt were visualized under UV light immediately after capillary transfer. The position of the detected band is located between these two ribosomal RNA, which is consistent with the expected size of the control transcript. Serine rRNA-ligase is more abundant in rings versus in trophozoites and schizonts. The intensities of the bands are consistent with this expression pattern (**Figure 4.4A**). The band intensity is similar at each stage between wildtype and mutant RNA sample, indicating comparable loading. Moreover, the loading control ensures that the missing *Not1.1* transcripts in the *Not1.1* mutant is not caused by the lack of sample.

In conclusion, the results of the Southern blot and the Northern blots support the notion that the selected clone is the *Not1.1* mutant and more specifically the *Not1.1* knockout parasite line. The phenotypes, transcriptional profile as well as RNAPII recruitment of this clone were analyzed as follows.

4.2 Phenotypes of the *Not1.1* mutant

Proliferation

To evaluate whether the mutation of *Not1.1* leads to any defective phenotypes, proliferation of parasite was assessed by measuring the parasitemia, which is the ratio of the number of iRBCs versus the total number of RBCs. At trophozoite stage, highly synchronized wildtype and *Not1.1* mutant were subcultured at 0.1% parasitemia in biological triplicates. Fresh medium was changed every day and the parasitemia of approximately 1,000 RBCs in each sample was measured by blood smear under microscopy for three continuous IDCs. The parasitemia values are listed in **Supplementary Table 6** and depicted in **Figure 4.5**.

The growth of the *Not1.1* mutant is significantly slower than that of wildtype in all three IDCs. One-tailed t-tests were conducted between the parasitemia of two parasite lines in each IDC. The p-values are 0.0339, 0.0091 and 0.0001 for three IDCs, suggesting that the lower growth is statistically significant ($p < 0.05$). The gap between two parasite lines becomes larger and larger until the third IDC when the parasitemia of the *Not1.1* mutant is approximately half of that in wildtype. The slower growth indicates that the missing *Not1.1* affects proliferation of parasites.

Figure 4.5. Proliferation of the *Not1.1* mutant. The x-axis represents three IDCs and the y-axis represents average parasitemia starting from 0.1% parasitemia. The one-tailed t-test was conducted on parasitemia in each IDC between wildtype (blue) and the *Not1.1* mutant (orange) with asterisk symbol * ($p < 0.05$), ** ($p < 0.01$) and *** ($p < 0.001$). The error bar is the standard deviation of biological triplicates.

Invasion rate

To further characterize the effect of the *Not1.1* mutation, the invasion rate of wildtype and the *Not1.1* mutant was assessed. The invasion rate was calculated as the ratio of the daughter rings versus the starting schizonts, representing how many merozoites produced from one schizont, can successfully invade RBCs.

The highly synchronized parasites at schizont stages were subcultured at 0.5% or 1% parasitemia and maintained in 1% hematocrit. Biological triplicates were included. The parasitemia of rings in approximately 1,000 RBCs was counted 14 hours after the subculturing. The values for the invasion rate are shown in **Supplementary Table 7** and depicted in **Figure 4.6**.

The average invasion rate for 0.5% starting parasitemia is 4.66 for wildtype and 3.04 for the *Not1.1* mutant. For 1% starting parasitemia, the invasion rate is 3.06 for wildtype and 1.68 for the *Not1.1* mutant. The ratio of the invasion rate in the *Not1.1* mutant versus that in wildtype is 0.65 and 0.55 for 0.5% and 1% starting parasitemia, respectively.

In both experiment setups, the invasion rate of the *Not1.1* mutant is significantly lower than that of wildtype ($p < 0.05$ by one-tailed t-test). The value of invasion rate varies with the starting parasitemia. The invasion rate is higher in the parasite lines starting from 0.5% parasitemia compared to 1% parasitemia. The lower invasion rate in the latter is probably due to the higher competition of more parasites.

Figure 4.6 Invasion rate of the *Not1.1* mutant.

The x-axis represents the starting 0.5% and 1% parasitemia at schizont stage. The y-axis represents the average invasion rate of biological triplicates. The one-tailed t-test was conducted between wildtype (blue) and the *Not1.1* mutant (orange) with asterisk symbol ** ($p < 0.01$) and *** ($p < 0.001$). The error bar is the standard deviation.

There are three hypotheses to explain the lower invasion rate. First, the mutation of *Not1.1* might affect the merozoite production and fewer merozoites are produced from each schizont. Therefore, less invasion rate is expected. Second, the duration of one IDC might be longer in the *Not1.1* mutant compared to 48 hours in wildtype parasites. In this case, the longer IDC length will result in a slower growth as well as invasions at different timepoints in the long term. Third, the processes of merozoite egress and/or invasion might be defective. The mutation of *Not1.1* might alter the expression of egress- or invasion-related genes, which leads to production of defective merozoite and eventually impaired egress and/or invasion processes. These three hypotheses were tested as follows.

Hypothesis 1. Fewer merozoites produced from each schizont

To measure the number of merozoites produced from one schizont, highly synchronized parasites were cultured to late schizont stages. At approximately 44 hpi, when the segmentation of merozoites was clearly observed within schizont parasites, a blood smear was made in biological duplicates and the number of merozoites in each schizont was counted under microscopy. For each sample, 50 schizonts were included. The number of merozoites is listed in **Supplementary Table 8** and shown in **Figure 4.7**.

The average number of merozoite produced in wildtype is 16.42 and 16.84 in biological duplicates. The average number of merozoites in the *Not1.1* mutant is 16.94 and 17.18. In the two replicates, the number of merozoites produced from one schizont is comparable in the two parasite lines ($p = 0.0517$ for replicate 1 and $p = 0.1197$ for replicate 2 by one-tailed t-test). Therefore, the defective invasion of the *Not1.1* mutant is not due to a decreased number of merozoites produced from each schizont.

Figure 4.7. Merozoite number per schizont in the *Not1.1* mutant. The x-axis represents the biological duplicates of wildtype (blue) and the *Not1.1* mutant (orange). The y-axis represents the number of merozoites produced in one schizont, with each dot representing one schizont. One-tailed t-test was conducted between these two parasite lines with no significant differences (ns, $p > 0.05$). The average number of merozoites in 50 parasites is shown as the middle line and the standard deviation is shown as the error bars.

Hypothesis 2. Longer length of the IDC

The transcriptional profile of parasites at 6 timepoints across one IDC was assessed by microarray, which will be described in Section 4.3.1. The parasites are expected to be harvested at 4 hpi, 12 hpi, 20 hpi, 28 hpi, 36 hpi and 42 hpi. The Spearman correlation between the transcriptional profile of each sample and the periodicity of gene expression was conducted to estimate the developmental age of parasites. The periodicity of gene expression includes the transcriptional profiles of parasites harvested every two hours throughout one IDC. The age of each sample is listed in **Supplementary Table 9**. The average ages of parasites throughout one IDC are shown in **Figure 4.8** as scatter plots with trend lines.

At TP1-6 in wildtype, the average ages of parasites are 3.33 hpi, 13.33 hpi, 18.67 hpi, 28 hpi, 36 hpi and 40 hpi. At TP1-6 in the *Not1.1* mutant, the average ages of parasites are 0 hpi, 6 hpi, 18 hpi, 24 hpi, 32 hpi and 40 hpi.

Figure 4.8 Progression of parasites in the *Not1.1* mutant. The x-axis represents 6 timepoints harvested in one IDC and y-axis represents the age of parasite in the unit of hpi. The trendline is drawn in the same color with the dot, blue color for wildtype and orange color for the *Not1.1* mutant. The error bar represents the standard deviation of technical triplicates.

One-tailed t-test was conducted on the estimated parasite ages between wildtype and the *Not1.1* mutant at each timepoint. At TP1 and TP2, the ages of wildtype (3.33 hpi and 13.33 hpi) are significantly different from the *Not1.1* mutant (0 hpi and 6 hpi) with

$p < 0.05$. The progression of the *Not1.1* mutant reflects a delayed ring stage compared to wildtype. The age of parasites at TP3 is comparable between two parasite lines ($p > 0.05$) and identical in TP6, suggesting that the progression of parasites has no significant differences at late stages.

Hypothesis 3. Defective egress and/or invasion process

To test whether the invasion or egress processes are affected in the *Not1.1* mutant, the expression of egress- and invasion-related genes will be investigated in Section 4.3.4.

4.3 Differential mRNA abundance in the *Not1.1* mutant

4.3.1 Knockout of *Not1.1* affects the expression of 715 genes

In order to characterize the functions of NOT1.1 in regulating gene expression at the transcript level, the transcriptional profile of 5,061 genes in wildtype and the *Not1.1* mutant were assessed.

The parasites were harvested at 8-hour intervals for a total of 6 timepoints, specifically at 4 hpi, 12 hpi, 20 hpi, 28 hpi, 36 hpi and 44 hpi. These 6 timepoints represent early ring, middle ring, late ring/early trophozoite, late trophozoite, early schizont and late schizont, respectively. Six timepoints were selected because of the coverage of all three stages across the 48-hour IDC with two timepoints at each stage.

Total RNA was extracted and reverse transcribed into cDNA in technical triplicates. The amino-allyl cDNA was amplified by SMART PCR program. A pool of amino-allyl cDNA from 8 timepoints of 3D7 parasites across one IDC was assembled and used as the reference. In microarray hybridization, the reference pool was coupled with Cy3-dye and the cDNA from individual timepoint was coupled with Cy5-dye. The microarray data were processed as described in Section 2.5.4 and the expression data of probes for a given transcript were averaged to obtain the expression of each gene. Eventually, the average mRNA abundance was calculated from technical triplicates.

In total, 715 genes were expressed significantly differentially at the transcript level between wildtype and the *Not1.1* mutant with an expression fold-change larger than

1.5 and $p < 0.05$ in at least one timepoint. **Table 4.1** shows the number of genes with significantly differential mRNA abundance at each timepoint.

	TP1	TP2	TP3	TP4	TP5	TP6	Total
Upregulation	128	47	48	43	26	76	368
Downregulation	30	37	221	111	136	100	635
Total in TP	158	84	269	154	162	176	1003

Table 4.1 Number of differential mRNA abundance in the *Not1.1* mutant. The differential mRNA abundance between wildtype and the *Not1.1* mutant is defined as a fold-change over 1.5 and $p < 0.05$. The number of genes being upregulated or downregulated was counted at each timepoint. The total number at each timepoint as well as the total number of up- or down-regulations in one IDC was counted.

Decreased mRNA abundance is much more common in the *Not1.1* mutant, especially at trophozoite (TP3, TP4) and schizont stages (TP5, TP6), whereas increased mRNA abundance is predominant at early ring stage (TP1). The timepoint with the most abundant dysregulations is TP3, corresponding to early trophozoites at 18 hpi.

Subsequently, these 715 genes were clustered using Cluster 3.0 software according to the timepoint at which they are severely downregulated. These genes are listed in **Supplementary Table 10** and depicted in **Figure 4.9A**. In the heatmap, differential mRNA abundance is shown in the form of $\log_2$ -transformed expression ratio of the *Not1.1* mutant versus wildtype. The red color represents upregulation in the *Not1.1* mutant and the green color represents downregulation.

The overall heatmap shows that the knockout of *Not1.1* results in altered mRNA abundances across the IDC. For the majority of genes in the heatmap, a severe downregulation is observed at one timepoint, upregulation can be observed at other timepoints. For example, genes downregulated in TP3 are upregulated in TP6 and TP1. These observations suggest that NOT1.1 may regulate mRNA abundance at multiple timepoints.

Figure 4.9 Significantly differential mRNA abundance in the *Not1.1* mutant. (A) Heatmap of 715 genes in an order according to the timepoint of the greatest downregulation. Transcriptional profiles at six timepoints in one IDC were assessed by microarray, encompassing parasites at 4 hpi, 12 hpi, 20 hpi, 28 hpi, 36 hpi and 44 hpi. Differential mRNA abundance is shown as the $\log_2$ -transformed expression ratio in the *Not1.1* mutant versus wildtype. A colored scale-bar is given at the bottom with upregulation in the *Not1.1* mutant in red and downregulation in green. **(B) Functional enrichment of differentially expressed genes at different stages.** The functions of genes in each cluster were analyzed by Gene Ontology and MPM functional analysis. The enriched functional pathways with a binomial $p < 0.05$ are listed. The number of genes in each functional pathway was counted. The green blocks represent severe downregulation and the red blocks represent upregulation at the corresponding developmental stages in the *Not1.1* mutant.

4.3.2 Functional analysis of the genes influenced by NOT1.1

From the heatmap, 6 gene clusters can be defined, in which genes are downregulated at distinct timepoint (**Figure 4.9A**). To further understand the influence of NOT1.1 on gene regulation, the functions of the genes in each cluster were analyzed. Gene Ontology term analysis (GoStat) and MEMP pathway enrichment were employed. **Figure 4.9B** shows the number of genes in each enriched pathway, which is categorized into IDC developmental stages according to the timepoint of the greatest downregulation.

Cluster 1

Cluster 1 includes 50 genes downregulated at TP1 and TP2 and upregulated at late stages. In wildtype, they show peak expression at TP3 and moderately high expression at TP2 and TP4. The transcript levels at TP1, TP5 and TP6 are very low.

There are 20 genes encoding exported proteins, 10 genes involved in the Maurer's cleft, 5 genes for metabolic pathways and 3 genes for protein folding. A large number of exported proteins facilitate the remodeling of the erythrocyte to make it a suitable host, which is critical for parasite virulence [282]. The Maurer's cleft is a parasite organelle associated with protein sorting and export [283]. In addition, 5 genes are involved in transcription, GTP metabolism, proline synthesis, glutamate synthesis and mannose metabolism. Chaperones in the parasite play critical roles in maintaining the viability in heating environments when patients suffer from the fever [284].

Cluster 2

Cluster 2 includes 46 genes downregulated at TP1-3. In wildtype, the transcript levels are high at trophozoite stage (TP3 and TP4) but not as high as those in Cluster 1. The transcript levels are low at TP6.

There are 5 genes involved in metabolic pathways, including hypoxanthine-guanine phosphoribosyltransferase (HGPRT) for purine synthesis [285], deoxyuridine 5-triphosphate nucleotidohydrolase (dUTPase) for dUTP catabolism and dTTP synthesis [286], glycine cleavage system T protein (GCVT) for glycine metabolism [287], cytidine diphosphate-diacylglycerol synthase (CDS) for lipid metabolism [288] and V-type

proton ATPase subunit D for ATP metabolism and proton translocation [289]. There are 4 genes encoding mitochondrial components, including mitochondrial ribosomal protein S8 precursor, mitochondrial import inner membrane translocase subunit TIM9, glycine cleavage system T protein and GTP-binding protein.

Cluster 3

Cluster 3 includes 235 genes downregulated at TP3. In wildtype, the transcript levels are abundant at TP4 and TP5. The expression at ring stage (TP1 and TP2) are rather low.

The enriched pathways with most genes correspond to apicoplast and mitochondrial functions. The apicoplast is an essential organelle in *P. falciparum* and consists of a 35 kb circular DNA resembling plant plastids. The apicoplast plays critical roles in type II fatty acid synthesis, isoprenoid synthesis, iron-sulfur cluster synthesis and heme synthesis [290]. The mitochondrion harbors a 6 kb genome and is a key site for energy production. Mitochondria in *P. falciparum* play roles in heme metabolism, coenzyme Q biosynthesis and iron-sulfur cluster biosynthesis, which supports the electron transport chain and pyrimidine biosynthesis [291, 292]. The apicoplast and mitochondrion are physically connected in both asexual stage in human and sexual stage in mosquito, suggesting a close association of their metabolic functions [293, 294]. In addition to enzymes for metabolic pathways in the apicoplast and mitochondrion, the genes encoding apicoplast ribosomes and mitochondrial ribosomes are expressed differentially in the *Not1.1* mutant as well.

There are 27 genes involved in metabolic pathways, including ATP synthesis, acyl-CoA synthesis, cytochrome c in the mitochondrial respiratory chain, glycosyltransferases and redox metabolism with 6 genes. In addition, 13 genes encode proteins involved in DNA repair, 11 genes encode components of DNA replication machinery and 4 genes encode ApiAP2 transcription factors. Moreover, there are 12 genes involved in proteolysis. Eight genes, like tubulin, kinetochore, centrin, and dynein play important roles in mitosis, which is critical for parasite reproduction [31]. Overall, NOT1.1 has extensive roles in regulating the expression of genes involved in distinct cellular processes.

Cluster 4

Cluster 4 includes 134 genes markedly downregulated at TP4. In wildtype, these genes are predominantly upregulated in schizonts (TP5 and TP6).

Notably, genes involved in invasion-related processes are enriched, including 48 genes for invasion, 16 genes for merozoite components and 7 genes for merozoite motility. Genes which function in phosphorylation are abundant as well. Other pathways include the Maurer's cleft with 12 genes, proteolysis with 10 genes and redox metabolism with 3 genes.

Cluster 5

Cluster 5 includes 208 genes downregulated in schizonts at TP5 and TP6. The transcript levels of these genes are high at TP6 and ring stage (TP1 and TP2) in wildtype.

Two types of pathways are remarkably abundant. First, protein export is represented by 49 genes encoding exported proteins and 15 genes are involved in the Maurer's cleft. Second, genes for egress and invasion processes are abundant, including 44 genes for egress and invasion, 45 genes encoding merozoite components and 16 genes for merozoite motility. Moreover, 16 genes are involved in phosphorylation, 4 genes encode ApiAP2 transcription factors and 6 genes are involved in protein folding.

Cluster 6

Cluster 6 includes 42 genes upregulated throughout one IDC. The transcript levels in wildtype are higher at early stages (TP1-3) compared to schizont stage (TP5 and TP6). Differential mRNA abundance in this cluster is unique compared to the other five clusters in which severe downregulation is observed at one timepoint and upregulations are observed at other timepoints. Genes in Cluster 6 are predominantly upregulated throughout the IDC and downregulation can hardly be seen.

This cluster is enriched with genes encoding exported proteins and antigenic variation, both of which are critical players in cytoadherence. Multiple var, rifin and stevor genes were affected in the *Not1.1* mutant. These genes account for the virulence and cytoadherence of parasite. Some chaperone proteins are upregulated as well.

4.3.3 Validation of microarray data by RT-qPCR in the *Not1.1* knockout

Reverse Transcription followed by Quantitative PCR (RT-qPCR) was performed to validate the accuracy of the microarray data.

Several genes with significantly differential mRNA abundance in the *Not1.1* mutant were selected. Knob-associated histidine-rich protein (KAHRP, PF3D7_0202000) is upregulated in the *Not1.1* mutant at all six timepoints. KAHRP is the major component of the knobs on the surface of infected erythrocytes, contributing to cytoadherence [295]. Invasion-related genes were selected at random, including MSP6 (PF3D7_1035500), EBA181 (PF3D7_0102500), RH4 (PF3D7_0424200) and RH5 (PF3D7_0424100). These invasion genes are all upregulated at TP1 and downregulated at TP5.

The expression of genes assessed by microarray and RT-qPCR was compared, which is listed in **Supplementary Table 11** and shown in **Figure 4.10**.

Figure 4.10 Validation of differential mRNA abundance by RT-qPCR in the *Not1.1* mutant.

The x-axis represents the transcript of selected genes at different timepoint. The y-axis represents the log₂-transformed ratio of mRNA abundance in the *Not1.1* mutant versus wildtype. The transcript levels measured by microarray were presented in orange columns and the transcript levels measured by RT-qPCR were presented in blue columns. The error bars represent the standard deviation of technical triplicates.

In **Figure 4.10**, differential mRNA abundance determined by two methods is represented by the $\log_2$ -transformed ratio of mRNA abundance in the *Not1.1* mutant versus wildtype. The value of 0 in the y-axis represents the identical mRNA abundance in the two parasite lines. The positive values show upregulation in the *Not1.1* mutant and the negative values show downregulation.

KAHRP at TP1-6 and invasion-related genes at TP1 are upregulated as assessed by both microarray and RT-qPCR in the *Not1.1* mutant. Likewise, certain invasion-related genes are downregulated at TP5 as assessed by microarray and RT-qPCR as well. The consistency of the expression changes in microarray and RT-qPCR lends confidence in the microarray data.

Furthermore, the magnitude of the changes are bigger in RT-qPCR, possibly due to the higher sensitivity of real-time PCR compared to microarray. The qPCR is able to measure slight difference in cDNA abundance by quantifying PCR product dynamically. Microarray, on the other hand, measures the amount of cDNA, which is the final product of SMART PCR amplification. The slight difference might not be sufficient to produce significantly different amount of cDNA by SMART PCR amplification, especially for the naturally low abundant transcript. Also, variations of cDNA abundance might be introduced during SMART PCR amplification.

4.3.4 NOT1.1 may be required for setting the mRNA abundance of egress- and invasion-related genes

As described previously in this study, egress- and invasion-related genes are notably enriched among genes whose mRNA abundance is altered in the *Not1.1* mutant. Therefore, the effect of NOT1.1 on the expression of egress- and invasion-related genes were evaluated. In **Figure 4.11**, the $\log_2$ -transformed ratio of mRNA abundance in the *Not1.1* mutant versus wildtype is shown in the heatmap. These genes were clustered based on their subcellular localization or site within secretory organelles.

Genes critical for merozoite egress, like PKG, SUB1, SERA6, CDPK1, SERA5 and genes critical for invasion, like MSPs, RHs, EBAs, AMA1 were found to be affected by the loss of NOT1.1 function. The egress- and invasion-related genes with differential mRNA abundance in the *Not1.1* mutant are component of most organelles of merozoites.

Figure 4.11 Egress- and invasion-related genes with significantly differential mRNA abundance in the *Not1.1* knockout. The egress- and invasion-related genes, which were expressed significantly differently in the *Not1.1* mutant, were clustered according to their subcellular localizations or secretory organelles. The differential mRNA abundance is shown as the log₂-transformed ratio of the expression in the *Not1.1* mutant versus the wildtype. The name of some genes known to play important roles in egress or invasion is labeled. The colored scale-bar is given at the bottom, with upregulation in red and downregulation in green in the *Not1.1* mutant.

These organelles include rhoptry, inner membrane complex, microneme, apical complex and dense granule, all of which function in merozoite egress and invasion as described in Introduction.

In nature, these invasion or egress-related genes achieve peak expression at schizont stage. While these genes are expressed in the *Not1.1* mutant, they are downregulated

compared to wildtype at late stages (TP4 - TP6). At early stages of the IDC, these transcripts are upregulated relative to wildtype, though still less abundant than in schizonts. The dysregulation of these genes reinforces the role of NOT1.1 in regulating expression of genes, including many egress- and invasion-related genes. Overall, the mistiming expression of egress- and invasion-related genes probably leads to the production of viable but defective merozoites, which may result in impaired invasion rate.

In summary, in the *Not1.1* mutant, the mRNA abundance of a large set of egress- and invasion-related genes at late stages is reduced, whereas that at early stages is elevated. It is possible that parasites respond to the low expression of these genes at schizont stage and complement at ring stage of the next IDC. The dysregulation at the transcript level might lead to abnormal expression at the protein level, which will be investigated in the future.

4.4 Knockout of *Not1.1* affects recruitment of RNAPII

The transcript abundance is determined by the balance between transcription and mRNA metabolism. In order to determine the mechanism underlying the functions of NOT1.1 in gene regulation, the levels of transcription were examined by assessing recruitment of the elongating RNA polymerase II to the genome.

The C-terminal domain (CTD) of the large subunit of RNAPII (RPB1) consists of multiple repeats of YSPTSPS heptapeptides. The phosphorylation on Serine 5 occurs at the promoter region, mediating transcription initiation and recruiting the capping complex [296]. The phosphorylation on Serine 2 is associated with the elongating form of RNAPII [297, 298]. One antibody against phosphorylated Ser2/5 RPB1, which is the elongating form of RNAPII, in *P. falciparum* (anti-Ser2/5-P) has been raised by Dr. Ragini Rai in the Featherstone laboratory [85]. This anti-Ser2/5-P antibody is able to immunoprecipitate phosphorylated Ser2/5 RPB1 by chromatin-immunoprecipitation and the bound DNA can be analyzed by microarray. This enables me to look into the status of transcription for each gene.

Chromatin-immunoprecipitation (ChIP) was performed on the nuclear lysate, harvested from 50 µl schizont pellet, to test whether the anti-Ser2/5-P antibody is able to immunoprecipitate RPB1 protein in wildtype and the *Not1.1* mutant. Three samples are included, wildtype lysate treated with antibody, wildtype lysate without antibody and *Not1.1* mutant lysate with antibody following the ChIP procedure described in Section 2.7.2. The cytoplasmic lysate before the ChIP procedures (Cyto), the proteins unbound to the antibody (S/N) and the proteins bound to antibody (Eluate) were collected. The RPB1 protein in all lysates were detected with the anti-Ser2/5-P antibody by Western blot (**Figure 4.12**).

Figure 4.12 Chromatin-immunoprecipitation of RPB1 tested by Western blot. The nuclear lysates of wildtype (WT) and the *Not1.1* mutant (1.1KO) were immunoprecipitated by anti-Ser2/5-P antibody and protein A agarose. The negative control is the lysate of the *Not1.1* mutant immunoprecipitated by protein A agarose. Briefly, 4% of the cytoplasmic lysate (Cyto), 4% of the unbound supernatant (S/N) and 50% of the eluate were loaded and blotted with the anti-Ser2/5-P antibody. The lane number is given on the top and the sizes of the ladder marker are given at the left.

The phosphorylated RPB1 was absent in the cytoplasmic lysate of both parasite lines (lane 1, 4 and 7), consistent with the nuclear location of RNAPII. The protein A beads

were able to immunoprecipitate the anti-Ser2/5-P antibody (two bands at around 27 kDa and 50 kDa) together with RPB1 protein in both wildtype lysate (lane 3) and *Not1.1* mutant lysate (lane 9). In the eluates, the full-length RPB1 can barely be observed and the eluted proteins were truncated at 140 kDa to 240 kDa, as shown in lane 3 and lane 9. Some proteins can still be detected in the unbound supernatant (lane 2 and lane 8), suggesting that the immunoprecipitation was not thorough.

Furthermore, for the *Not1.1* mutant which was not incubated with antibody, the anti-Ser2/5-P antibody and RPB1 protein were not detected in the eluate (lane 6) and the full-length RPB1 was detected in the unbound supernatant (lane 5), indicating the specificity of the protein A beads for the anti-Ser2/5-P antibody and also the specificity of the anti-Ser2/5-P antibody for RPB1 protein.

Subsequently, the recruitment of the elongating form of RNAPII across the parasite genome was assessed by ChIP-on-chip in one IDC. The ChIP was performed at 6 timepoints for wildtype and the *Not1.1* mutant, which are identical to those in the transcriptional profile in Section 4.3.1. The DNA bound by phosphorylated RPB1, which was immunoprecipitated by the anti-Ser2/5-P antibody (ChIP DNA) as well as genomic input DNA was extracted. The ChIP DNA and the input DNA of the same sample were amplified and hybridized in the microarray.

Enrichment of Ser2/5-phosphorylated RPB1, which is the elongating form of RNAPII, on each gene was calculated as the ratio of the amount of DNA bound by RPB1 versus the amount of input DNA. The enrichment was then compared between wildtype and the *Not1.1* mutant. In total, 150 genes show differential RNAPII recruitment with a fold-change over 1.5 and $p < 0.05$. The values of differential RNAPII recruitment, represented by the $\log_2$ -transformed ratio of RNAPII enrichment in the *Not1.1* mutant versus wildtype, are listed in **Supplementary Table 12**. The number of genes with differential RNAPII recruitment at each timepoint is shown in **Table 4.2**.

The changes in RNAPII recruitment are abundant at ring (TP1 and TP2) and early trophozoite stages (TP3). The number of genes showing enriched RNAPII binding is comparable to that showing depleted binding.

	TP1	TP2	TP3	TP4	TP5	TP6	Total
Enriched	19	32	13	7	2	0	73
Depleted	15	42	22	3	0	0	82
Total in TP	34	74	35	10	2	0	155

Table 4.2 Number of differential RNAPII recruitment in the *Not1.1* mutant. The differential RNAPII recruitment between wildtype and the *Not1.1* mutant is defined as a fold-change over 1.5 and $p < 0.05$. The number of genes with enriched or depleted RNAPII binding was counted. The total number of differential recruitment at each timepoint as well as the total number of enriched or depleted recruitment was also counted.

These 150 genes with differential RNAPII recruitment were clustered using the average linkage algorithm in Cluster 3.0 software, as shown in **Figure 4.13A**. In the left panel, differential RNAPII recruitment is shown as the $\log_2$ -transformed ratio of RNAPII enrichment in the *Not1.1* mutant versus wildtype. In the right panel, differential mRNA abundance for the corresponding gene, which is shown as the $\log_2$ -transformed ratio of mRNA abundance in the *Not1.1* mutant versus wildtype, is aligned.

There are no obvious correlations in the two heatmaps. The Spearman coefficient between the changes in RNAPII recruitment and the changes in transcript abundance at the same timepoint was calculated in **Figure 4.13B**. The coefficient near a value of 0 at each timepoint suggests that there is no correlations between two changes across one IDC.

Statistically, 17 genes with differential RNAPII recruitment are expressed differentially at the same timepoint as shown in **Figure 4.13C**. A binomial test between a subgroup of 150 genes with differential RNAPII recruitment and a subgroup of 715 genes with differential mRNA abundance suggests that an overlap of 17 genes is random ($p > 0.05$). This indicates that there is no statistical correlation between the changes in RNAPII recruitment and the changes in mRNA abundance in the *Not1.1* mutant.

In conclusion, NOT1.1 indeed affects recruitment of the elongating form of RNAPII for 150 genes. However, differential mRNA abundance of these genes is not

predominantly resulted from the effect of NOT1.1 in transcription. Therefore, the influences of NOT1.1 in gene expression are speculated to be post-transcriptional.

Figure 4.13 Correlation of RNAPII recruitment and mRNA abundance in the *Not1.1* mutant. **(A) Heatmap of differential RNAPII recruitment (left panel) and corresponding differential mRNA abundance (right panel).** Six timepoints were included representing 4 hpi, 12 hpi, 20 hpi, 28 hpi, 36 hpi and 44 hpi. The differential RNAPII recruitment and the differential mRNA abundance are represented by log₂-transformed differences in the *Not1.1* mutant versus the wildtype. The red blocks represent enriched RNAPII recruitment and increased mRNA abundance in the *Not1.1* mutant. The green blocks represent less enriched RNAPII recruitment and decreased mRNA abundance in the *Not1.1* mutant. The colored scale-bar is given at the bottom. **(B) Spearman correlation between differential RNAPII recruitment and differential mRNA abundance.** The Spearman coefficient and the p-value for the confidence of coefficient were calculated at each timepoint. The asterisk symbol represents $p < 0.05$. **(C) Overlap of genes showing differential RNAPII recruitment and differential mRNA abundance.** The number of genes in each subgroup and the overlapping genes was counted. A binomial test was conducted and the p-value is 0.4109.

4.5 Creation of the *Not1.2* mutants

4.5.1 Knockout of endogenous *Not1.2* gene was not successful

The pCC1 vector, which was used to create the *Not1.1* mutant, was employed to create the *Not1.2* mutant. The integration of the selectable marker, the *hDHFR* cassette, by double-crossover recombination is shown in **Figure 4.14A**.

Figure 4.14. Failure to disrupt endogenous *Not1.2*. (A) Integration of plasmid, pCC1-NOT1.2-KO into endogenous *Not1.2* gene. The *hDHFR* gene is driven by the calmodulin promoter and has the histidine-rich protein terminator. The *FCU* gene is driven by HSP86 promoter and has the *DHFR* terminator. Two homologous arms (HR1 and HR2) flank the *hDHFR* cassette. The primers used for PCR analysis are indicated in the diagram and sequences are provided in Section 2.8. **(B) Genotyping PCR of the *Not1.2* transfectants.** PCR amplification was conducted on gDNA of wildtype (WT) and the *Not1.2* transfectants. Three rounds of drug cycling were done. The size of PCR product is expected to be 1,164 bp for the primers 22/23 (par1), 1,063 bp for the primers 24/25 (par2), 1,135 bp for the primers 22/27 (int1), 1,162 bp for the primers 28/25 (int2), 919 bp for the primers 26/27 (epi1) and 1,144 bp for the primers 28/29 (epi2). The sizes of the ladder are given at the left.

Two homologous arms include HR1, 3 bp-865 bp and HR2, 3,021 bp-3,942 bp downstream of the endogenous *Not1.2* start codon, corresponding to the region of 1-289 aa and 1,007-1,314 aa on NOT1.2 protein. NOT1.2 contains the MIF4G domain at 1,399-1,562 aa, the DUF3819 domain at 2,180-2,324 aa and the NOT1 domain at 3,988-4,213 aa as well as 4,331-4,455 aa. The insertion of the *hDHFR* cassette is expected to disrupt the expression of all the conserved domains.

The *Not1.2* transfectants were observed after the transfection of the plasmid, pCC1-NOT1.2-KO into T996 parasite strain. Three rounds of drug cycling of the *hDHFR* selection were conducted. Then, the genomic loci in the transfectants were detected by PCR amplification in **Figure 4.14B**.

The integrant loci were hardly seen, suggesting that the plasmid failed to integrate into the genome. The subcloning has been carried out for two times but no clones contained the integrant loci, suggesting that NOT1.2 might be essential in *P. falciparum*. The failure to mutate *Not1.2* was reported in another parasite species, *P. berghei* [234], although the success to mutate *Not1.2* has been reported in *P. falciparum* [233]. The essentiality of NOT1.2 needs to be further tested based on these contradictory findings.

4.5.2 Conditional knockdown system

Although the techniques to disturb DNA and mRNA are powerful strategies to study gene function, it is difficult to use these techniques to study essential genes and the biological processes that happen in a fast timescale. Conditional protein degradation provides a means to directly regulate protein level in a specific, rapid and reversible manner. Since the regulation is conditional, it is possible to knock down an essential protein.

The system of conditional protein degradation consists of two elements, the destabilizing domain which triggers protein degradation and the stabilizing ligand, which binds and stabilizes the domain. Two destabilizing domains have been developed. The FKBP destabilizing domain is derived from human FK506-binding protein (FKBP12). The engineered mutant, L106P, renders this protein constitutively degraded when expressed. The synthetic ligand, shield-1 binds FKBP destabilizing

domain with high affinity to shield the domain from degradation [299]. The second destabilizing domain, *E.coli* DHFR destabilizing domain (ecDDD) is derived from a mutated form of *E. coli* DHFR and is stabilized by trimethoprim (TMP) [300]. The two systems have successfully been applied in *P. falciparum* [301, 302].

Some problems have been reported regarding the application of destabilizing domains in *P. falciparum*. First, the fusion protein bearing the destabilizing domain may not be completely degraded in the absence of ligand [302, 303] and therefore may not produce a loss-of-function phenotype. Second, the destabilizing domain may interfere with the function of the fusion protein [303]. For example, one study found that the HSP101 fused with ecDDD could not be degraded in the absence of TMP ligand, but interestingly the function of HSP101 was somehow perturbed [304].

In this study, the ecDDD coupled to HA epitopes was fused to the endogenous *Not1.2*. The design of conditional knockdown of NOT1.2 protein is described in **Figure 4.15**.

Figure 4.15. Design of conditional knockdown of NOT1.2. An ecDDD and HA epitopes were fused to the C-terminus of NOT1.2. Functional stabilization is achieved by exposing parasite bearing the tagged protein to TMP ligand. To achieve NOT1.2 loss of function, TMP ligand is removed in order to trigger degradation of the tagged protein.

In the presence of TMP ligand, the ecDDD is bound and stabilized by TMP ligand, thus the tagged protein is expressed and functional. On the other hand, in the absence of

TMP ligand, the destabilizing domain triggers degradation of the tagged protein, resulting in a conditional knockdown.

In this study, the FKBP destabilizing domain with a FLAG epitope has been fused to the endogenous *Not1.1*. The failure to detect FLAG-tagged NOT1.1 made it impossible to know whether NOT1.1-FLAG-FKBP is expressed or degraded (data not shown). Therefore, this strain is not analyzed in this thesis.

4.5.3 Endogenous *Not1.2* was tagged with HA epitopes and destabilizing domain

Endogenous *Not1.2* was fused in frame to the coding region of three HA epitopes followed by an ecDDD, as shown in **Figure 4.16A**. The integration was achieved by a single-crossover recombination of the plasmid pLNOT1.2-HA-ecDDD with the parasite genome.

In addition, the parasite was transfected with a second plasmid, pThDHFR encoding human DHFR (**Figure 4.16B**), in order to confer resistance to the toxic effect of TMP ligand in the parasites [300]. A control parasite line was created by transfecting the plasmid, pThDHFR into T996 parasite strain.

The TMP ligand was used at a concentration of 10 μ M [300]. Drug cyclings of the *BSD* selection were performed until the integrant locus were detected. The parasite population was subcloned by limiting dilution to obtain single clones of integrants, referred to hereafter as *Not1.2* mutant. So far, one clone of integrant has been obtained and its genomic loci were examined by PCR amplification in **Figure 4.16C**.

The detection of the integrant locus (int) along with the loss of the parental locus (par) indicates the successful integration of the ecDDD. The episomal plasmids, pLNOT1.2-HA-ecDDD and pThDHFR, were apparently maintained in the cloned parasite as also revealed by PCR (epi1 and epi2). The episomal locus on the plasmid, pThDHFR were observed in the control line, as expected.

Figure 4.16. Tagging of endogenous *Not1.2* with the destabilizing domain. (A) Integration of the plasmid into the *Not1.2* locus. The *BSD* selectable marker is driven by the *PcDHFR-TS* promoter and ending with the histidine-rich protein terminator, not shown. The homologous arm (HR) of 1,516 bp is a portion of the 3' coding region of endogenous *Not1.2* excluding the stop codon. The three HA epitopes (3xHA) and the ecDDD are followed by a stop codon. **(B) Plasmid-encoded resistance to TMP.** The parasite was co-transfected with the plasmid pThDHFR and the plasmid pLNOT1.2-HA-ecDDD. The primers used to detect different loci are indicated as arrows and primer sequences are provided in Section 2.8. **(C) Genotyping PCR of the *Not1.2* mutant.** The templates for PCR amplification are gDNA of the control line and gDNA of the *Not1.2* mutant. The parental locus (par) was detected by the primer pair 1/2 at 1,782 bp, the integrant locus (int) was detected by the primer pair 1/30 at 1,690 bp and the episomal locus (epi1) was detected by the primer pair 4/30 at 2,008 bp. The episomal locus of the plasmid pThDHFR (epi2) was detected by the primer pair 31/32 at 561 bp. The sizes of the ladder markers are given at the left.

The expression of the tagged *Not1.2* gene at the transcript level was assessed by RT-PCR as shown in **Figure 4.17**. The detection of the transcript specific to the integrant locus in the *Not1.2* mutant suggests that the tagged genes were indeed transcribed into mRNA. The detection of the transcript was independent on the presence or absence of the TMP ligand (+TMP and –TMP in **Figure 4.17B**).

Figure 4.17 Expression of the tagged *Not1.2* gene at the transcript level. PCR amplifications were performed on the cDNA of the control line, the NOT1.2-HA line (1.2HA) and the NOT1.2-HA-ecDDD parasites that had been incubated with or without TMP (+TMP and -TMP, respectively). The cDNA was prepared by reverse transcription with (+) or without (-) SuperScript III reverse transcriptase. The parental locus was detected by the primer pair 1/2 at 1,782 bp. the integrant locus was detected by the primer pair 33/3 at 314 bp for the NOT1.2-HA line and 782 bp for the NOT1.2-HA-ecDDD line. The primer locations are indicated in Figure 4.16A. The sizes of the ladder markers are given at the left.

Several controls were included in this experiment. First, the control line serves as a reference for comparison to the integrant samples. It ensures that the missing parental locus in the *Not1.2* mutant was not due to a failure of the PCR reactions. Second, the detection of the fusion transcript in the NOT1.2-HA line serves as a positive control. Third, we used a mock cDNA sample in which RNAs were not treated with reverse transcriptase. This confirms that PCR products are dependent on the production of

cDNA from RNA and are not due to contaminating genomic DNA or DNA from other sources.

4.5.4 Knockdown of ecDDD-tagged NOT1.2 in the absence of stabilizing ligand

Subsequently, the expression of the tagged NOT1.2, NOT1.2-HA-ecDDD, at the protein level in schizonts was examined by Western blot.

Parasites were harvested after maintenance in TMP ligand at a concentration of 2 μ M, 10 μ M, 25 μ M and 50 μ M for two IDCs. On the other hand, 10 μ M TMP ligand was removed from the culture and the parasites were harvested in each of the three subsequent IDCs. Total proteins lysate, obtained from schizont pellet, was prepared and blotted with anti-HA antibody in biological duplicates (**Figure 4.18A and Figure 4.18B**). Actin was used as a loading control. To reinforce the statistical analysis, another biological replicate will be performed. So far, the Western blots from biological duplicates have been analyzed.

As expected, no proteins were detected in total lysate of the control line by anti-HA antibody (lane 1). The total lysate of the previously described NOT1.2-HA line serves as a positive control for the NOT1.2-HA-ecDDD line. In lane 2, the full-length NOT1.2-HA at 524 kDa can be seen running above the 460 kDa ladder and truncated proteins can be seen at lower positions as has been observed before (**Figure 3.3**), possibly due to protein processing in the parasite.

The failure to detect the full-length NOT1.2-HA-ecDDD protein at 542 kDa is possibly due to the incomplete stabilization of the fusion protein by TMP ligand. Alternatively, the missing full-length protein might be due to the challenge to detect very large protein given that the positive control, NOT1.2-HA protein was only clearly detected in the second replication.

Figure 4.18 Expression of NOT1.2-HA-ecDDD in the presence or absence of TMP. (A, B) Biological duplicates assessed by Western blot. The samples include the total lysates of the ctrl line, the NOT1.2-HA line (1.2HA) and the *Not1.2* mutant in schizonts. In lane 3-6, the *Not1.2* mutant was incubated in different concentrations of TMP ligands. In lane 7-9, parasites were harvested in each IDC following the removal of TMP ligand. Briefly, protein lysates obtained from 5 μ l schizont pellet were loaded for the HA blot and the actin blot. Full-length

NOT1.2 is highlighted with an arrow. The sizes of the ladder markers are given at the left. **(C) The ratios of band intensity of NOT1.2-HA-ecDDD to actin.** The signal intensities of NOT1.2-HA-ecDDD protein and actin were measured by ImageJ. The ratio of band intensity of NOT1.2-HA-ecDDD against actin was calculated for each sample. The samples include the *Not1.2* mutant incubated in different concentration of the TMP ligand (2 μ M-50 μ M) and the *Not1.2* mutant with the removal of the TMP ligand (1-3 IDC-).

In the *Not1.2* mutant, only truncated NOT1.2-HA-ecDDD protein can be detected at around 117 kDa and 171 kDa (lane 3-9). These two bands probably correspond to two bands in the NOT1.2-HA line, one between 71 kDa and 117 kDa and the other between 117 kDa and 171 kDa (lane 2). The positions of the two bands in the *Not1.2* mutant are slightly above the two bands in the NOT1.2-HA line, likely due to the fusion of the ecDDD at 18 kDa.

The two bands of NOT1.2-HA-ecDDD were detected in all four samples incubated in different concentrations of TMP ligand (lane 3-6) and in three samples after the removal of TMP ligand (lane 7-9). The signal intensity of the two bands was normalized to that of the actin control in each sample, which enables comparison of NOT1.2-HA-ecDDD between different samples. The ratio of the intensity of NOT1.2-HA-ecDDD versus that of actin is shown in **Figure 4.18C**.

The amount of NOT1.2-HA-ecDDD was compared between the three samples after the removal of TMP ligands. The two bands can still be observed in the first and the second IDC (lane 7 and lane 8). After removing TMP ligand for three IDCs, the two bands were barely seen (lane 9), suggesting that NOT1.2-HA-ecDDD was indeed degraded in the absence of TMP ligand. The degradation was validated by the decreased ratio of NOT1.2-HA-ecDDD to actin after removing TMP ligand for three IDCs in both biological duplicates. Again, a third biological replicate will be performed for statistical analysis.

The amount of NOT1.2-HA-ecDDD was also compared between parasites incubated at different concentrations of TMP ligand (lane 3-6). The ratio is not increased along with the higher concentration of TMP, suggesting that the detection of the truncated proteins is independent of the concentration of TMP. The ratio at 2 μ M was as high as

those in higher concentrations, suggesting that 2 μ M TMP ligand was sufficient to stabilize the truncated NOT1.2 in the *Not1.2* mutant.

In conclusion, the incubation of parasites in TMP ligand may not stabilize the full-length NOT1.2-HA-ecDDD, although the challenge to detect a large protein needs to be considered. The knockdown of NOT1.2-HA-ecDDD at the protein level after three IDCs of TMP removal suggests that the *Not1.2* mutant is indeed a *Not1.2* knockdown line.

It is interesting to observe the degradation of NOT1.2-HA-ecDDD only after three cycles of TMP removal. It has been shown that ecDDD-tagged asparagine-repeated protein degraded several hours after the removal of TMP ligand [302]. NOT1.2 has been shown to be a scaffold of the CCR4-NOT complex in *P. falciparum*. Its potential interacting partners may preserve the intact structure of NOT1.2 from degradation, and thus the difficulty to knock down NOT1.2 efficiently. On the other hand, the conditional degradation of NOT1.2-HA-ecDDD was partial given the detection of faint bands after three IDCs of TMP removal. The failure to knock out and completely knock down NOT1.2 suggest that this protein might be essential in *P. falciparum*.

4.6 Phenotype of the *Not1.2* mutant

Proliferation

To evaluate whether the knockdown of NOT1.2 leads to any defective phenotypes, parasite proliferation of the control line and the *Not1.2* mutant was measured. At trophozoite stage, the highly synchronized control line and *Not1.2* mutant were subcultured at 0.1% parasitemia in biological triplicates. The parasitemia in approximately 1,000 RBCs was assessed by blood smear for three continuous IDCs. The parasitemia values are listed in **Supplementary Table 13** and depicted in **Figure 4.19**.

The average parasitemia of the control line is 1.49%, 8.55 % and 21.44% for three IDCs. The average parasitemia of the *Not1.2* mutant is 0.45%, 3.2% and 9.82% for three IDCs. One-tailed t-test was conducted between the parasitemia of the control line and the *Not1.2* mutant in each IDC. The p-value assessed by one-tailed t-test is 0.0048, 0.0004 and 0.0079 for the three continuous IDCs, demonstrating a significantly slower growth of the *Not1.2* knockdown.

Figure 4.19. Proliferation of the *Not1.2* mutant. The x-axis represents three IDC after the subculturing and the y-axis represents average parasitemia starting from 0.1% parasitemia. The significantly different growth between the control line (blue) and the *Not1.2* mutant (orange) in each IDC was assessed by the one-tailed t-test with asterisk symbol ** ($p < 0.01$) and *** ($p < 0.001$). The error bar is the standard deviation of biological triplicates.

Invasion rate

The highly synchronized parasites at schizont stage were subcultured at 0.5% or 1% parasitemia and maintained in 1% hematocrit. Biological triplicates were included. The parasitemia of rings was counted in approximately 1,000 RBCs 14 hours after subculturing. The invasion rate was calculated as the ratio of the daughter parasites at ring stage to the starting parasites at schizont stage. The raw data is shown in **Supplementary Table 14** and the values of invasion rate are depicted in **Figure 4.20**.

The average invasion rate for 0.5% starting parasitemia is 3.74 for the control line and 2.68 for the *Not1.2* mutant. For 1% starting parasitemia, the invasion rate is 2.79 for the control line and 1.80 for the *Not1.2* mutant. Statistically, the invasion rate of the *Not1.2* knockdown is significantly lower compared to the control line, for both 0.5% starting parasitemia ($p = 0.0020$, $p < 0.01$) and 1% starting parasitemia ($p = 0.0287$, $p < 0.05$). The p-value was calculated from the one-tailed t-test analysis.

Furthermore, the ratios of the invasion rate in the *Not1.2* mutant versus that in the control line are 0.72 and 0.65 for 0.5% and 1% starting parasitemia, respectively. The ratios are higher than 0.65 and 0.55 in the *Not1.1* mutant, suggesting that the defective

invasion rate in the *Not1.2* mutant is not as severe as the *Not1.1* mutant, possibly due to the partial knockdown of NOT1.2.

Figure 4.20 Invasion rate of the *Not1.2* mutant. The x-axis represents the starting 0.5% and 1% parasitemia at schizont stage. The y-axis represents the average invasion rate of biological triplicates. The one-tailed t-test was conducted between the control line (blue) and the *Not1.2* mutant (orange) with asterisk symbol * ($p < 0.05$) and ** ($p < 0.01$). The error bar is the standard deviation.

Number of merozoites

To test whether the lower invasion rate in the *Not1.2* mutant is caused by a defective merozoite production, the number of merozoites produced from one schizont was measured. Highly synchronized parasites were cultured to late schizont stages and a blood smear was made in biological duplicates. The number of merozoites was counted in 50 schizonts under microscopy. The number of merozoites is listed in **Supplementary Table 15** and shown in **Figure 4.21**.

The average number of merozoite produced in the control line is 14.16 and 14.46 in biological duplicates. The average number of merozoites in the *Not1.2* mutant is 14.52 and 14.70. In each replicate, the number of merozoites produced from one schizont is comparable in the two parasite lines ($p = 0.1419$ for replicate 1 and $p = 0.2328$ for replicate 2 by one-tailed t-test). Therefore, the defective invasion of the *Not1.1* mutant is not due to a decreased number of merozoites produced from each schizont.

Figure 4.21. Merozoite number per schizont in the *Not1.1* mutant. The x-axis represents the biological duplicates of the control line (blue) and the *Not1.2* mutant (orange). The y-axis represents the number of merozoites produced in one schizont, with each dot representing one schizont. One-tailed t-test was conducted between these two parasite lines, showing no significant differences (ns, $p > 0.05$). The average number of merozoites in 50 parasites is shown as the middle line and the standard deviation is shown as the error bars.

Progression of parasite

The transcriptional profile of parasites at 6 timepoints was assessed by microarray, similarly to the *Not1.1* mutant in Section 4.2. The parasites are harvested at 4 hpi, 12 hpi, 20 hpi, 28 hpi, 36 hpi and 42 hpi. The Spearman correlation between the transcriptional profile of each sample and the periodicity of gene expression was conducted to estimate the developmental age of parasites. The age of each sample is listed in **Supplementary Table 16**. The average ages of parasites throughout one IDC are shown in **Figure 4.22** as scatter plots with trend lines.

At TP1-6 in the control line, the average ages of parasites are 2 hpi, 6.67 hpi, 16 hpi, 26 hpi, 32 hpi and 40 hpi. At TP1-6 in the *Not1.2* mutant, the average ages of parasites are 2.67 hpi, 7.33 hpi, 18 hpi, 28 hpi, 36 hpi and 40 hpi.

One-tailed t-test was conducted on the estimated parasite ages between the control line and the *Not1.2* mutant at each timepoint. At TP1, TP2 and TP4, the p-value between the two parasite lines is 0.325, 0.343 and 0.211, indicating comparable ages

at each timepoint. In conclusion, the progression of parasite in one IDC is comparable between the control line and the *Not1.2* mutant.

Figure 4.22 Progression of parasites in the *Not1.2* mutant. The samples include the control line (blue) and the *Not1.2* mutant (orange). The x-axis represents 6 timepoints harvested in one IDC and y-axis represents the age of parasite in a unit of hpi. The trendline is drawn in the same color with the dot, blue color for the control line and orange color for *the Not1.2* mutant. The error bar represents the standard deviation of technical triplicates.

4.7 Differentially expressed genes in *Not1.2* knockdown

4.7.1 Knockdown of NOT1.2 affects the expression of 202 genes

The transcriptional profile of the control line and the *Not1.2* mutant at 6 timepoints of one IDC was assessed by microarray and analyzed as described in Section 4.3.1. The genes with differential mRNA abundance in the control line and the *Not1.2* mutant were filtered. There are 202 genes with an expression fold-change more than 1.5 and $p < 0.05$. **Table 4.3** shows the number of genes with significant changes in mRNA abundance at each timepoint.

The overall upregulations are much more abundant in the *Not1.2* mutant compared to downregulations. The differential mRNA abundance is enriched at late stages, especially the late schizont stage at TP6. The downregulations are relatively abundant at early stages (TP1-4).

	TP1	TP2	TP3	TP4	TP5	TP6	Total
Upregulation	4	20	14	32	38	95	203
Downregulation	4	6	6	10	0	4	30
Total in TP	8	26	20	42	38	99	233

Table 4.3 Number of differential mRNA abundance in the *Not1.2* mutant. The number of genes being upregulated or downregulated at the transcript level was counted in the *Not1.2* mutant. The total number in each timepoint as well as the total number of up- or down-regulations was calculated.

Subsequently, these 202 genes were clustered using Cluster 3.0 software according to the timepoint of the greatest upregulation. The genes are listed in **Supplementary Table 17** and depicted in **Figure 4.23A**. The differential mRNA abundance is shown in the form of log₂-transformed ratio of expression in the *Not1.2* mutant versus that in the control line. The red color represents upregulation in the *Not1.2* mutant and the green color represents downregulation.

The overall heatmap shows that the knockdown of NOT1.2 results in altered mRNA abundances across the IDC. The extent of upregulation is larger compared to downregulation. When the upregulation for one gene was observed at one timepoint, the downregulations were seen at other timepoints. The pattern contrasts that of the *Not1.1* knockout, in which the extent of downregulation is larger.

There are four clusters of genes according to the timepoint of the greatest upregulation in the heatmap. The functions of genes in each cluster were analyzed by Gene Ontology term analysis (GoStat) and MEMP pathway enrichment. **Figure 4.23B** shows the number of genes in each enriched pathway, which is categorized into IDC developmental stages according to the timepoint of the greatest upregulation.

Figure 4.23 Significantly differential mRNA abundance in *Not1.2* mutant. (A) Heatmap of 202 genes in an order according to the timepoint of the greatest upregulation. The differential mRNA abundance is shown as the log₂-transformed ratio of the expression in the *Not1.2* mutant versus that in the control line. The color-scaled bar is given at the bottom, with upregulation in red and downregulation in green in the *Not1.2* mutant. **(B) Functional enrichment of differentially expressed genes at different stages.** The functions of genes in each cluster were analyzed by Gene Ontology and MPM functional analysis. The enriched functional pathways with a binomial $p < 0.05$ are listed. The number of genes in each functional pathways was counted. The green blocks represent downregulation and the red block represent the severe upregulation in the corresponding developmental stages in the *Not1.2* mutant.

Cluster 1 includes 24 genes upregulated at TP1-TP3. In the control line, the transcript abundances are high at trophozoite stage (TP3, TP4) and relatively low at ring and schizont stages. There are 19 genes encoding exported proteins and 11 genes encoding components of the Maurer's cleft, both of which are important for protein export and parasite virulence [282, 283].

Cluster 2 includes 42 genes upregulated at trophozoite stage (TP3 and TP4). In the control line, the transcript levels are abundant at TP5 and TP6 and the expression at ring (TP1 and TP2) and trophozoite stages (TP3 and TP4) are rather low. Genes involved

in egress and invasion process are enriched, including 8 genes for egress and invasion and 3 genes encoding merozoite components. Other pathways include Maurer's cleft with 4 genes and transcription with 3 ApiAP2 transcription factors.

Cluster 3 includes 117 genes upregulated at schizont stages (TP5 and TP6). The transcript levels of these genes are high at TP6 and at ring stage (TP1 and TP2) in the control line. Two types of pathways are remarkably abundant. First, there are 46 genes encoding exported proteins and 18 genes involved in the Maurer's cleft. Second, there are 12 genes involved in egress and invasion processes, 25 genes encoding merozoite components, 13 genes involved in phosphorylation and 11 genes involved in protein folding. Moreover, 3 genes are involved in metabolic pathways, including DNA-directed RNA polymerase III subunit RPC10 for tRNA synthesis [305], RNA polymerase subunit 8c (RPABC3), which is shared by RNAPI, RNAPII and RNAPIII, and asymmetrical bis(5-nucleosyl)-tetraphosphatase, an enzyme to produce GTP and GMP.

Cluster 4 includes 19 genes downregulated throughout one IDC. The expression in the control line is higher at early stages (TP1-TP3) compared to late stages (TP4-TP6). The differential mRNA abundance in this cluster is different from the other three clusters given that upregulation was not observed and downregulation was predominant. This cluster is enriched with genes encoding exported proteins and antigenic variation. Both are critical players in cytoadherence. Three var multifamily genes, PF3D7_0324400, PF3D7_0900500 and PF3D7_0617700 were affected in the *Not1.2* mutant, which may account for the virulence and cytoadherence of the parasite.

4.7.2 Validation of microarray data by RT-qPCR in *not1.2* knockdown

To validate the accuracy of the microarray data, RT-qPCR was performed on cDNA. The genes examined in the *Not1.1* mutant in Section 4.3.3 were tested in the *Not1.2* mutant. These genes include KAHRP, MSP6, EBA181, RH4 and RH5.

The differential mRNA abundance of genes at specific timepoint was assessed by microarray and RT-qPCR and then compared. The data are shown in **Supplementary Table 18** and depicted in **Figure 4.24**. Differential mRNA abundance determined by two

methods is represented by the $\log_2$ -transformed ratio of mRNA abundance in the *Not1.2* mutant versus that in the control line, which is shown in the y-axis. The value of 0 represents the identical expression in the two parasite lines, the positive values represent upregulation in the *Not1.2* mutant and the negative values represent downregulation.

Based on microarray data, KAHRP is downregulated throughout the IDC. Invasion related genes, including MSP6, EBA181, RH4 and RH5, are downregulated at TP1 and upregulated at TP5. The expression assessed by RT-qPCR (blue bar) is consistent with those by microarray (orange bar) for the same genes at the same timepoint, demonstrating the accuracy of mRNA abundance determined by microarray.

Figure 4.24 Validation of differential mRNA abundance in the *Not1.2* mutant. The y-axis value is the $\log_2$ -transformed ratio of the expression of the selected genes in the *Not1.2* mutant versus that in the control line. The transcript abundance measured by microarray was presented in orange columns. The transcript abundance measured by RT-qPCR was presented in blue columns. The error bars represent the standard deviation of technical triplicates.

Chapter 5 Discussion

5.1 Proposed architecture of the CCR4-NOT complex in *P. falciparum*

5.1.1 Components of the CCR4-NOT complex in *P. falciparum*

The co-immunoprecipitation of NOT1.2, CAF1 and NOT4 were described in Chapter 3. In order to find possible components of the complex, the potential interactions between the CCR4-NOT core subunits, identified by co-immunoprecipitation and mass spectrometry are shown in **Table 5.1**.

Gene	Description	HA-tagged proteins		
		NOT1.2	CAF1	NOT4
PF3D7_0507600	cell differentiation protein (CAF40)	Y	Y	Y
PF3D7_0811300	CCR4-associated factor 1 (CAF1)	Y	Y	*
PF3D7_1417200	CCR4-NOT transcription complex subunit 1 (NOT1.2)	Y	Y	
PF3D7_1103800	CCR4-NOT transcription complex subunit 1 (NOT1.1)		Y	
PF3D7_0519500	carbon catabolite repressor protein 4 (CCR4)		Y	
PF3D7_1235300	CCR4-NOT transcription complex subunit 4 (NOT4)		*	Y
PF3D7_1128600	CCR4-NOT transcription complex subunit 2 (NOT2)			
PF3D7_1006100	CCR4-NOT transcription complex subunit 5 (NOT5)			
PF3D7_0107200	carbon catabolite repressor protein 4, (CCR4.1)			

Table 5.1 Interactions between the core subunits of the CCR4-NOT complex. The presumptive interactions were identified by immunoprecipitation and mass spectrometry in both of the independent biological duplicates “Y” represents the detection of proteins in the eluate of the corresponding HA-tagged protein. Asterisk symbol represents the interaction between CAF1 and NOT4 as CAF1 was co-immunoprecipitated by NOT4 in only one replicate and NOT4 was co-immunoprecipitated by CAF1 with the log₂-transformed protein score ratios below 1.5.

Based on the mass spectrometry data, the associations between CAF40, CAF1, NOT1.2, CCR4, NOT1.1 and possibly NOT4 support the existence of the CCR4-NOT complex in *P. falciparum*. In the following section, the components of the complex are nominated.

NOT1.2 was co-immunoprecipitated by CAF1 and vice versa, supporting the hypothesis that NOT1.2 is the scaffold for CAF1. NOT1.1 is also the scaffold for CAF1, given that NOT1.1 was co-immunoprecipitated by CAF1. NOT1.1 and NOT1.2 are both capable to be the scaffold because both contain conserved domains to interact with all other subunits in *P. falciparum*. The current mass spectrometry data cannot tell whether two

NOT1 paralogs form the complexes independently or physically related. However, based on the structural study of the CCR4-NOT complex in other eukaryotic organisms, one complex contains one and only one scaffold. NOT1.1 and NOT1.2 are more likely to be mutually exclusive scaffolds.

CAF40 is a stable subunit of the CCR4-NOT complex as it was co-immunoprecipitated by NOT1.2, CAF1 and NOT4.

CCR4 was co-immunoprecipitated by CAF1, but not NOT1.2 nor NOT4. CAF1 is very likely to be an anchor between CCR4 and the complex, as reported in other eukaryotes [135, 142, 144, 148]. The other CCR4 paralog, CCR4.1 was not co-immunoprecipitated by CAF1, consistent with its missing leucine-rich repeat known to interact with CAF1.

NOT1.1 and NOT1.2 were not co-immunoprecipitated by NOT4, consistent with the fact that neither was NOT4 co-immunoprecipitated by NOT1.2. On the other hand, CAF40 was co-immunoprecipitated by NOT4 and CAF1 was co-immunoprecipitated by NOT4 in one replicate. Taken together, the results strongly suggest that NOT4 associates with the CCR4-NOT complex in *P. falciparum* not as stably as CAF40, CAF1, NOT1.1, NOT1.2 and CCR4.

NOT2 and NOT5 were not co-immunoprecipitated by NOT1.2, CAF1 nor NOT4 in either of the biological duplicates. This raises the question as to the function of the NOT1 domain in NOT1.1 and NOT1.2 given that this domain bears high similarity to its counterpart in other eukaryotes and is known to bind these subunits.

In conclusion, this study shows a role of NOT1.1 and NOT1.2 as the scaffold for the CCR4-NOT complex in *P. falciparum*, consistent with the findings in other organisms. CAF40 and CAF1 associate with the complex via NOT1.2 and NOT1.1. CAF1 may bring CCR4 to the complex as in other species. NOT4 may associate with the complex via CAF40 and possibly CAF1. NOT2 and NOT5 might not be the core subunits of the complex in *P. falciparum*, though other interpretations of the data should be entertained. For example, NOT2 and NOT5 may associate strongly with a minor fraction of the total complex in the cells, or under specific conditions.

5.1.2 Architecture of the CCR4-NOT complex in *P. falciparum*

NOT1.1 and NOT1.2 contain highly conserved domains, which are known to interact with CCR4-NOT subunits in other eukaryotes. Based on protein sequence analysis in Section 1.4.2 and proposed components of the CCR4-NOT complex in Section 5.1.1, the structures of the CCR4-NOT1.1 complex and the CCR4-NOT1.2 complex are proposed in **Figure 5.1**.

Figure 5.1 Proposed architecture of the CCR4-NOT complexes in *P. falciparum*. The MIF4G domain, the DUF3819 domain and the NOT1 domain of NOT1 are labeled in yeast NOT1, NOT1.1 and NOT1.2. The structure of yeast CCR4-NOT complex is depicted based on the structural studies in other eukaryotes. The interactions of the CCR4-NOT subunits with the scaffold NOT1.1 and NOT1.2 in *P. falciparum* are proposed according to the mass spectrometry data in this study. The dotted line between CAF1 and NOT4 represents the co-immunoprecipitation of CAF1 with NOT4 in only one replicate. The question mark for NOT2 and NOT5 represents their associations to the complex in other eukaryotes, which have not been detected in this study.

NOT1.1 and NOT1.2 serve as the scaffold of the complex mutually exclusively. CAF1 interacts with both proteins via the MIF4G domain. CAF40 interacts with NOT1.2 and probably NOT1.1 via the DUF3819 domain. CCR4 is recruited to the complexes by CAF1. NOT4 is likely to be recruited to the complexes via its interaction with CAF40 or probably CAF1. The association of NOT2 and NOT5 with the complexes is unknown.

5.1.3 Candidate interacting partners and potential functions of the CCR4-NOT complex in *P. falciparum*

To define a subset of proteins which are highly likely to interact with the CCR4-NOT complex in *P. falciparum*, those proteins co-immunoprecipitated by at least two of the tagged components - NOT1.2, CAF1 or NOT4 - were chosen as listed in **Table 5.2**. The core subunits of the CCR4-NOT complex are not included.

Gene	Description	HA-tagged proteins		
		NOT1.2	CAF1	NOT4
PF3D7_1224300	polyadenylate-binding protein (PABP)	Y	Y	Y
PF3D7_1107300	polyadenylate-binding protein-interacting protein 1 (PAIP1)	Y	Y	Y
PF3D7_1015900	enolase (ENO)	Y	Y	Y
PF3D7_0610400	histone H3 (H3)	Y	Y	Y
PF3D7_1447000	40S ribosomal protein S5	Y		Y
PF3D7_0813300	conserved Plasmodium protein, unknown function	Y		Y
PF3D7_1424400	60S ribosomal protein L7-3	Y		Y
PF3D7_0815600	eukaryotic translation initiation factor 3 subunit G (EIF3G)	Y		Y
PF3D7_0814200	DNA or RNA-binding protein Alba 1 (ALBA1)	Y		Y
PF3D7_1110400	asparagine-rich antigen	Y		Y
PF3D7_0511500	RNA pseudouridylate synthase		Y	Y
PF3D7_0524000	karyopherin beta (KASbeta)		Y	Y
PF3D7_0108300	conserved Plasmodium protein, unknown function		Y	Y
PF3D7_0604500	conserved Plasmodium protein, unknown function		Y	Y
PF3D7_1423700	conserved Plasmodium protein, unknown function		Y	Y
PF3D7_0703500	erythrocyte membrane-associated antigen		Y	Y

Table 5.2 Potential interacting proteins of the CCR4-NOT complex. The presumptive interactions are identified in both of the independent biological duplicates by co-immunoprecipitation and mass spectrometry. NOT1.2, CAF1 and NOT4 in the titles are the tagged-proteins, targeted by immunoprecipitation. “Y” represents the detection of the proteins in the eluate of the corresponding HA-tagged protein.

Common interacting partners of NOT1.2, CAF1, and NOT4

Four proteins are found to be immunoprecipitated by all of NOT1.2, CAF1 and NOT4, namely PABP, PAIP1, enolase and histone H3.

Although no proteins involved in transcription were co-immunoprecipitated by NOT1.2, CAF1 interacts with histone modifying proteins, transcription factors, transcription elongation factor SPT5, DNA-directed RNA polymerase I and III and multiple ApiAP2 transcription factors. NOT4 also interacts with the P44 component of TFIIH and one ApiAP2 transcription factor. Furthermore, the co-immunoprecipitation of histone H3 by all of NOT1.2, CAF1 and NOT4 suggests that the CCR4-NOT complex can be found proximal to histone H3. It is possible that the complex binds DNA where CAF1 and NOT4 function in epigenetic modification or transcription initiation or elongation.

PABP, PAIP1 and enolase were co-immunoprecipitated by all of NOT1.2, CAF1 and NOT4. The functions of these three interacting partners have been described in Section 3.1.3. The binding of PABP and PAIP1 on the poly(A) tail of mRNA suggest that CCR4-NOT complex might be recruited to the 3' end of mRNA by RNA-binding proteins as in other eukaryotes. Enolase, which was found in the RNA degradosome of *E.coli* to induce mRNA decay, may provide a new player in RNA degradation in *P. falciparum*. Of note, both of PABP and enolase, together with ALBA1, have been reported to interact with DOZI and CITH in a P-body-like granule, which stores translational silent mRNA, in *P. berghei* gametocytes [306]. These interactions might be conserved in the asexual stage of *P. falciparum*.

Common interacting partners of NOT1.2 and NOT4

Six proteins co-immunoprecipitated by both NOT1.2 and NOT4 include two ribosomal proteins, eIF3G, ALBA1, asparagine-rich antigen and the PF3D7_0813300 conserved protein.

Ribosome protein S5 or uS5 in *P. falciparum* (PF3D7_1447000) is a component of the small 40S ribosomal subunit. Ribosomal protein L7-3 or eL8 in *P. falciparum* (PF3D7_1424400) is a component of the large 60S ribosomal subunit. The 40S and 60S subunits constitute the 80S ribosome to allow for translation. Structural study of the *P. falciparum* ribosome in 2014 located S5 at the tRNA entry side and L7-3 at the tRNA

exit side [307]. PF3D7_0815600 encodes eIF3G, which is reported to associate with the 40S ribosome and stimulate mRNA recruitment to 43S complex for the subsequent AUG recognition in yeast and mammals [308, 309]. The interactions of the CCR4-NOT complex with ribosomes and translation initiation complex reinforces its close association with mRNA and potential functions in post-transcriptional regulation.

ALBA1 is an RNA-binding protein shown to bind 1,193 transcripts and thereby inhibit translation of some invasion genes [310]. Other members of the ALBA family may associate with the CCR4-NOT complex as well. ALBA4 was co-immunoprecipitated by CAF1 and ALBA3 was co-immunoprecipitated by NOT4. ALBA proteins all contain ALBA RNA-binding domains and are reported to be RNA-binding proteins in *P. falciparum* [311]. The recruitment of ALBA1 and possibly ALBA3, ALBA4 to the CCR4-NOT complex may extend the role of the complex in post-transcriptional regulation.

Asparagine-rich antigen (PF3D7_1110400) contains an RNA-binding domain and was identified as an RNA-binding protein by bio-informatics study [97]. It may, therefore, function to recruit the CCR4-NOT complex to mRNA. The sequence of its C-terminus shows homology to the termination factor Rho according to NCBI database, which is involved in termination of transcription in *E.coli* [312].

PF3D7_0813300 is detected in the food vacuole [264] and immunoprecipitated by CDPK [313] with unknown functions.

Common interacting partners of CAF1 and NOT4

Six proteins co-immunoprecipitated by both CAF1 and NOT4 include RNA pseudouridylate synthase, karyopherin beta, erythrocyte membrane-associated antigen and three conserved proteins.

PF3D7_0703500 encodes erythrocyte membrane-associated antigen. It has been reported be an RNA-binding protein at trophozoite and schizont stages [311], making it a potential RNA-binding protein that could recruit the CCR4-NOT complex.

PF3D7_0511500 encodes RNA pseudouridylate synthase (PUS). Pseudouridine is the most abundant post-transcriptional modification on noncoding RNAs, including ribosomal RNA, small nuclear RNA and tRNA. PUS recognizes RNA substrates and catalyzes the isomerization of uridine into pseudouridine [314]. Defective rRNA

pseudouridylation led to impaired translational fidelity and IRES-dependent translation initiation in yeast and human [315].

PF3D7_0524000 encodes karyopherin beta, forming a dimer with karyopherin alpha and facilitating protein transport through the nuclear pore [316]. In *P. falciparum*, the inhibition of nuclear transport by Ivermectin caused the accumulation of karyopherin alpha and beta arrest in the cytoplasm [317]. Karyopherin beta has been immunoprecipitated by PUF3, which is a nucleolar protein and also interacts with 26 60S ribosomal proteins [318].

PF3D7_1423700 contains a nuclear transport factor (NTF) domain and associates with PfSEC13, a nucleoporin. NOT1.2, PABP and asparagine-rich antigen were immunoprecipitated by PfSEC13 as well [319].

PF3D7_0108300 shares 97% identify with the protein sequence of asparagine-rich protein in RAJ116 strain. A study in 1999 has shown this protein to be located in both nucleus and cytoplasm and both asexual stage and sexual stage with unknown functions [320].

PF3D7_0604500 encodes a protein containing a GYF domain, which recognizes the proline-rich sequence. It was found to interact with PfCYP19A, a peptidyl-prolyl cis-trans isomerase in *P. falciparum* [321].

Proposed functions of the CCR4-NOT complex

CCR4-NOT complexes might play multiple roles in gene regulation in *P. falciparum*, as reported in other eukaryotes. In the nucleus, they may influence epigenetic regulation, transcription initiation and elongation given the association of the CCR4-NOT subunits with histone H3 and components of the corresponding machinery. They are also likely to associate with the nuclear transport machinery via karyopherin beta and possibly PfSEC13.

The major role of the CCR4-NOT complexes is regulation of mRNA metabolism. The complexes are in close association with ribosomes and the translation initiation complex. Their interactions with multiple RNA-binding proteins, some of which have been confirmed to contribute to mRNA degradation and translation repression, shed light on the critical roles of the complex in regulating post-transcriptional processes.

The CCR4-NOT complexes may have novel roles that have not been reported in other organism, for example, rRNA pseudouridylation.

5.2 Hypothesized functions of NOT1.1 paralogs in gene regulation in *P. falciparum*

NOT1.1

Given that only one clone of the *Not1.1* mutant has been obtained so far, the possibility of clonal variation cannot be ruled out. In the future study, other *Not1.1* knockout parasite lines will be subcloned in order to verify the findings presented in this thesis.

For 673 genes with significantly differential mRNA abundance in Clusters 1-5, downregulation was observed at specific timepoints and upregulation was observed at other timepoints. This pattern suggests that NOT1.1 is required for setting mRNA abundance for many genes. In the absence of *Not1.1*, the expression pattern of these genes shifts. Furthermore, the downregulation occurs at one timepoint before the peak expression in wildtype. For example, the expression of genes in Cluster 4 rises at TP4 and peaks at TP5 in wildtype, whereas they are greatly downregulated at TP4 in the *Not1.1* mutant. Genes in Cluster 5 express abundantly at TP6 and TP1 in wildtype, whereas they are downregulated at TP5 and TP6 in the *Not1.1* mutant. The effects of NOT1.1 in gene regulations might be explained by either of the following two hypotheses.

One possibility is NOT1.1 affects transcription. A possible delayed transcription could result in the downregulation of genes at the timepoint before the peak expression. Transcription was later switched on to compensate the downregulation, resulting in the downregulation at a lower extent at the peak timepoint and the upregulations afterward. This hypothesis has been tested in the *Not1.1* mutant in Section 4.4, in which the changes in recruitment of the elongating form of RNAPII are not significantly correlated to the changes in mRNA abundance, suggesting that transcription is not a major mechanism underlying the changes of transcript abundance in the *Not1.1* mutant.

The other possibility is that NOT1.1 affects mRNA metabolism. Some genes may be degraded or stabilized at the wrong timepoints in the parasites lacking *Not1.1*. For example, the invasion-related genes might be degraded at TP5 and the effect is gradually minimized in the following timepoints. As a result, downregulation is severe at TP5 and TP6 and upregulation occurs in the following timepoints. This hypothesis can be tested by measuring the mRNA decay rate in the future. If NOT1.1 does have an effect in mRNA decay and this effect is the major mechanism underlying the functions of NOT1.1 in gene regulation, a correlation between differential decay rate and differential mRNA abundance in the *Not1.1* mutant will be observed.

For 42 genes in Cluster 6, which were upregulated throughout the IDC, the changes in mRNA abundance are different from those of the other 5 clusters, suggesting that NOT1.1 regulates these genes via a distinct mechanism.

NOT1.2

Similarly to the *Not1.1* mutant, only one clone of the *Not1.2* mutant has been obtained so far, so that the possibility of clonal variation cannot be ruled out. In the future study, other *Not1.2* mutant parasite lines will be subcloned in order to verify the findings presented in this thesis.

The pattern of dysregulation in the *Not1.2* mutant is essentially the mirror image of that obtained with the *Not1.1* mutant, given that upregulation was observed at specific timepoint and downregulation was observed at the other timepoint for genes in Cluster 1-3. This pattern suggests that NOT1.2 is required for the accurate expression of some genes. In the absence of NOT1.2, the expression pattern of these genes shifts. Similarly to the *Not1.1* mutant, upregulation occurs at one timepoint before the peak expression in the control line. The distinct dysregulation in Cluster 4, in which upregulation was not observed, suggests that NOT1.2 regulates genes in Cluster 4 via different mechanisms.

NOT1.2 might affect the transcript abundance by influencing transcription initiation and/or elongation or post-transcriptional processes. The discordance between the differential RNAPII recruitment and the differential mRNA abundance in the *Not1.1* mutant suggests that NOT1.1 regulates gene expression predominantly post-

transcriptionally. Given the contrasting pattern of dysregulations in the *Not1.1* mutant and the *Not1.2* mutant, NOT1.1 and NOT1.2 might be functionally correlated, making NOT1.2 another potential regulator of post-transcriptional processes. Therefore, RNAPII recruitment will not be assessed in the *Not1.2* mutant. In the future, the mRNA decay rate will be assessed in two mutants, aiming to reveal the mechanism of gene regulation by NOT1 proteins.

5.3 Functional correlation of NOT1.1 and NOT1.2

5.3.1 NOT1.1 and NOT1.2 have opposite functions in regulating mRNA abundance

The differentially expressed genes are predominantly downregulated in the *Not1.1* mutant and upregulated in the *Not1.2* mutant, suggesting a probable functional correlation. Changes in the expression of KAHPR and invasion genes examined by RT-qPCR are opposed in the two mutants, raising the possibility that two NOT1 proteins may function antagonistically in regulating mRNA abundance.

To reveal the functional correlations between NOT1.1 and NOT1.2, the genes with differential mRNA abundance in the *Not1.1* mutant and the *Not1.2* mutant were compared. There are 715 genes showing significantly differential mRNA abundance in the *Not1.1* mutant and 202 genes showing significantly differential mRNA abundance in the *Not1.2* mutant. The overlap of these two datasets is shown in Figure **5.2A**. There are 135 genes found to be differentially expressed in both mutants, 580 genes specifically in the *Not1.1* mutant and 67 genes specifically in the *Not1.2* mutant. The binomial testing ($p < 0.05$) indicates that the overlap of 135 genes is statistically significant, supporting a functional correlation between NOT1.1 and NOT1.2.

Of 135 genes showing significantly differential mRNA abundance in both mutants, 114 genes display opposing transcriptional changes at the same timepoint. The heatmap portraying the differential mRNA abundance of these 114 genes in the *Not1.1* mutant (left panel) and in the *Not1.2* mutant (right panel) is shown in **Figure 5.2B** and these genes are listed in **Supplementary Table 19**. The heatmap visually conveys the opposing nature of the changes in mRNA abundance following *Not1.1* vs *Not1.2* loss of function.

Figure 5.2. Overlap of genes with differential mRNA abundance in the *Not1.1* mutant and the *Not1.2* mutant. (A) Venn diagram of the overlap. The genes with significantly differential mRNA abundance were compared in the *Not1.1* mutant and the *Not1.2* mutant. 135 genes were found in both mutants. **(B) Heatmap of differential mRNA abundance of 114 genes in the *Not1.1* knockout and the *Not1.2* knockdown.** The differential mRNA abundance is shown as the $\log_2$ -transformed ratio of expression in the mutants versus wildtype or the control line. The scale bar is given at the bottom with upregulation in red and downregulation in green in the mutants. **(C) Spearman coefficient between differential mRNA abundance of two mutants at each timepoint.** The Spearman coefficient shows the correlation of two datasets, with a value of 1 for positive correlation and value of -1 for negative correlation (green box). The asterisk symbol (*) denotes the statistical significance of the Spearman coefficient with $p < 0.05$.

To quantify the correlation between the two mutants, the Spearman coefficient was calculated for the differential mRNA abundance observed in the two mutants at each timepoint (**Figure 5.2C**). The significant negative coefficient at each timepoint indicates a negative correlation between the differential mRNA abundance observed in the two mutants. Based on the values of Spearman coefficient, a negative correlation is very strong at late trophozoite (TP4) and schizont stages (TP5 and TP6), moderate at early trophozoite stage (TP3) and relatively weak at ring stage (TP1 and TP2). These observations imply that NOT1.1 and NOT1.2 have opposite functions in regulating expression of these 114 genes, especially at trophozoite and schizont stages.

Next, I investigated whether the opposite functions of NOT1.1 and NOT1.2 in gene regulation are limited to these 114 genes or apply more globally. There are 580 differentially expressed genes specifically in the *Not1.1* mutant, 67 differentially expressed genes specifically in the *Not1.2* mutant and 21 genes differentially expressed at different timepoints in the two mutants. The heatmap in **Figure 5.3A** presents differential mRNA abundance of these 668 genes in the *Not1.1* mutant (left panel) and in the *Not1.2* mutant (right panel). The genes are organized based on the timepoint of the greatest downregulation in the *Not1.1* mutant.

Again, the heatmap portrays an opposing nature of the changes in mRNA abundance following *Not1.1* vs *Not1.2* loss of function. A similar pattern of the greatest downregulation in the *Not1.1* mutant and the greatest upregulation in the *Not1.2* mutant is observed. In the *Not1.2* mutant, upregulation is observed for genes that are downregulated at the same timepoint in the *Not1.1* mutant and vice versa. As for the last cluster, downregulation is not observed in the *Not1.1* mutant and upregulation is barely observed in the *Not1.2* mutant.

Furthermore, the Spearman coefficients for the differential mRNA abundance in the two mutants are -0.085, 0.013, -0.449, -0.753, -0.582 and -0.677 for TP1-TP6 (**Figure 5.3B**). The differential mRNA abundance is significantly negatively correlated in the two mutants at trophozoite (TP3 and TP4) and schizont stages (TP5 and TP6). This negative correlation of differential mRNA abundance of 668 genes reinforces the opposing roles of NOT1.1 and NOT1.2 in gene regulation.

Figure 5.3 Opposite functions of NOT1.1 and NOT1.2 in regulating mRNA abundance. (A) Heatmap of differential mRNA abundance of 668 genes in both mutants. The differential mRNA abundance is shown as the $\log_2$ -transformed ratio of expression in the mutants versus wildtype or the control line. The scale bar is given at the bottom with upregulation in red and downregulation in green in the mutants. **(B) Spearman coefficient between differential mRNA abundance of two mutants at each timepoint.** The Spearman coefficient shows the correlation of two datasets, with a value of 1 for positive correlation (red box) and value of -1 for negative correlation (green box). The asterisk symbol (*) denotes the statistical significance of the Spearman coefficient with $p < 0.05$.

On the other hand, the number of differentially expressed genes is less abundant and the extent of transcriptional change is less extensive in the *Not1.2* mutant compared to those in the *Not1.1* mutant. First, there are 715 genes with significantly differential mRNA abundance in the *Not1.1* mutant and only 202 genes in the *Not1.2* mutant. Second, a majority of expression changes of the *Not1.2* mutant in **Figure 5.3B** are not considered to be significant given that the fold-change is less than 1.5. These observations may be due to the partial degradation of NOT1.2 at the protein level. The reduced level of NOT1.2 protein indeed affects expression of most of these 668 genes but the influence is not as strong as the knockout of *Not1.1* at the transcript level.

In addition, those genes categorized as Cluster 2 in the *Not1.1* mutant are downregulated at ring and early trophozoite stages in the *Not1.1* mutant but not upregulated in the *Not1.2* mutant. These genes are involved in metabolic pathways and encode mitochondrial components. It is possible that regulation of these genes is a unique function of NOT1.1. An additional possibility is that the remaining NOT1.2 is sufficient to regulate these genes.

In conclusion, NOT1.1 and NOT1.2 regulate gene expression oppositely and probably antagonistically in *P. falciparum*. The opposite functions are not only limited to 114 genes with significantly differential mRNA abundance at the same timepoint in both mutants, but also apply globally. NOT1.1 affects gene expression more extensively compared to NOT1.2, possibly due to an incomplete knockdown of NOT1.2.

5.3.2 The correlation of NOT1.1 and NOT1.2 with gametocytogenesis

The Tracking Resistance to Artemisinins Collaboration (TRAC I) is a collaborative project that analyzed 1,043 clinical samples from South East Asia and Africa [322]. The transcriptional profile and the proportion of gametocytes of each sample have been reported. The correlation between the proportion of gametocytes and the expression of *Not1.1* (**Figure 5.4A**) and *Not1.2* (**Figure 5.4B**) was analyzed. The trendlines are drawn with a positive slope for NOT1.1 and a negative slope for NOT1.2.

To quantify the extent of correlation, the Spearman coefficient was calculated between the proportion of gametocyte and mRNA abundance of *Not1.1* or *Not1.2* in

these clinical samples. The value is 0.344 for NOT1.1 and -0.449 for NOT1.2, suggesting that the proportion of gametocyte is positively correlated to the transcript level of *Not1.1* and negatively correlated to the transcript level of *Not1.2*. The correlations are significant given the p-values less than 0.05.

In conclusion, NOT1.1 and NOT1.2 transcript levels are correlated to the proportion of gametocytes in clinical samples. The correlation is antagonistic for two NOT1 paralogs.

Figure 5.4 Correlation between proportion of gametocyte and expression of *Not1.1* and *Not1.2*. (A) and *Not1.2* (B). The proportion of gametocytes was counted in clinical samples and the expression of *Not1* genes was measured by microarray. Two variables are drawn in the scatter spot and the trendline is shown. The Spearman coefficient and corresponding p-value are given on the top.

5.4 NOT1.1 and NOT1.2 are functionally correlated with their interacting partners

5.4.1 Functional correlation with CAF1

The CAF1 subunit of the CCR4-NOT complex has been extensively studied in other eukaryotes. Changes in mRNA abundance resulting from a disruption of the NOT1-CAF1 interaction are similar to those caused by *Caf1* loss of function, suggesting that the NOT1-CAF1 interaction is required for the function of CAF1 [142, 144, 146]. The effect of *Caf1* mutations on mRNA abundance in *P. falciparum* has been reported in 2011 by Dr. John Adams laboratory [103]. In this study, I assessed the correlation between the functions of CAF1 and the functions of two NOT1 paralogs in regulating mRNA abundance.

There are a total of 782 genes with significantly differential mRNA abundance either in the *Not1.1* mutant or in the *Not1.2* mutant or both. Of these, the expression of 775 genes has been assessed in the *Caf1* mutant [103]. The differential mRNA abundance of these genes is displayed in a heatmap in **Figure 5.5A** in an order according to the timepoint of the greatest downregulation in the *Not1.1* mutant. The differential mRNA abundance for each gene in the *Not1.1* mutant (left panel), in the *Caf1* mutant (middle panel) and in the *Not1.2* mutant (right panel) is aligned.

From the overall heatmap, the pattern of differential mRNA abundance in the *Caf1* mutant resembles that of the *Not1.1* mutant and contrasts with that of the *Not1.2* mutant. To quantify the correlation, the Spearman coefficient was calculated between the *Caf1* mutant and the other two mutants at each timepoint (**Figure 5.5B**). The significant positive coefficients between the *Caf1* mutant and the *Not1.1* mutant indicate a positive correlation of differential mRNA abundance in these two mutants. The significant negative coefficients between the *Caf1* mutant and the *Not1.2* mutant indicate a negative correlation of differential mRNA abundance in these two mutants.

These correlations suggest redundant functions of CAF1 with NOT1.1 and NOT1.2. CAF1 exert similar functions with NOT1.1 and opposite functions with NOT1.2 in regulating gene expression. Based on the current knowledge that NOT1-CAF1 interaction is required for the function of CAF1 in mRNA decay, NOT1.1 is proposed to

be required for the function of CAF1 in *P. falciparum*. The NOT1.2-CAF1 interaction, somehow, inactivates CAF1.

Figure 5.5. Correlation of NOT1.1, CAF1 and NOT1.2 in regulating mRNA abundance. (A) Heatmap of differential mRNA abundance in the *Not1.1* mutant, in the *Caf1* mutant and in the *Not1.2* mutant. Differential mRNA abundance is shown as the log₂-transformed ratio of expression in the mutants versus wildtype or the control line. The scale bar at the bottom shows upregulation in red and downregulation in green in the mutants. **(B) Spearman coefficient between the differential mRNA abundance in the mutation of *Caf1* and *Not1.1* as well as *Not1.2*.** The Spearman coefficient shows the correlation of two datasets. The asterisk symbol (*) denotes the statistical significance of the Spearman coefficient with $p < 0.05$.

On the other hand, NOT1.1 and NOT1.2 regulate expression of some genes independent from CAF1. In **Figure 5.5A**, the dysregulations of some genes can only be seen in the *Not1.1* mutant and the *Not1.2* mutant but not the *Caf1* mutant. This suggests that NOT1.1 and NOT1.2 have additional functions other than regulating CAF1 function. Also, CAF1 regulates expression of some genes that are not affected in the *Not1.1* mutant or in the *Not1.2* mutant (data not shown), suggesting that NOT1.1 and NOT1.2 only partially affect CAF1 function.

Based on the analysis of protein sequences of NOT1.1 and NOT1.2, NOT1.1 seems to be the canonical NOT1, given the insertion in the MIF4G domain and the insertion in the NOT1 domain of NOT1.2. This hypothesis is supported by the conserved functional correlation of NOT1.1 and CAF1. NOT1 has been shown to be an essential gene in yeast [147]. However, I was able to completely knock out the more canonical *Not1.1*, but not *Not1.2* in this study. NOT1.2 might possess some important functions independently from NOT1.1, possibly provided by the unique insertions in the MIF4G domain and the NOT1 domain. One speculation is that NOT1.2 might function in translational and post-translational regulation or the downstream metabolic pathways, which are essential in the parasite but cannot be detected at the transcript level by microarray.

5.4.2 Functional correlation with ALBA1

In Section 5.1.3, the co-immunoprecipitation of ALBA1 by NOT1.2 and NOT4 was reported, making ALBA1 a potential interacting partner of the CCR4-NOT complex. The functions of ALBA1 in gene regulation and translational repression in *P. falciparum* were reported in 2015 by Dr. Artur Scherf laboratory [310]. The expression of 117 genes and 926 genes was dysregulated in the parasite overexpressing ALBA1 at 8 hpi and 28 hpi, respectively. These timepoints correspond to TP1 and TP4 in this study.

In **Figure 5.6**, differentially expression of genes in the ALBA1 overexpressing parasite at TP4 is organized from upregulation to downregulation. Differential mRNA abundance of the corresponding genes in the *Not1.1* mutant and the *Not1.2* mutant is aligned. The overall dysregulation in parasites overexpressing ALBA1 contrasts with that of the *Not1.1* mutant and resembles that of the *Not1.2* mutant.

The Spearman coefficients between parasites overexpressing ALBA1 and parasites lacking *Not1.1* is -0.338 ($p < 0.05$), as shown in **Figure 5.6B**. The negative value suggests a negative correlation of differential mRNA abundance in two mutants, which can also be observed in the heatmap. The Spearman coefficient between parasites overexpressing ALBA1 and parasites lacking NOT1.2 is 0.225, suggesting a weak positive correlation between two mutants. These correlations demonstrate that ALBA1 has redundantly similar functions with NOT1.1 and opposite functions with NOT1.2.

Similarly to CAF1, some dysregulations in the ALBA1 mutant are not seen in the *Not1.1* mutant or in the *Not1.2* mutant, demonstrating that ALBA1 has unique functions in gene regulations independent of NOT1.1 and NOT1.2.

Figure 5.6. Correlation of NOT1.1, ALBA1 and NOT1.2 in regulating mRNA abundance. (A) Heatmap of differential mRNA abundance of genes in ALBA1 overexpressing parasite (ALBA1 OE), the *Not1.1* mutant (*Not1.1* KO) and the *Not1.2* mutant (*Not1.2* KD) at TP4. The differential mRNA abundance is shown as the log₂-transformed ratio of expression in the mutants versus wildtype or the control line. The scale bar is given at the bottom with upregulation in red and downregulation in green in the mutants. **(B) Spearman coefficient between differential mRNA abundance of ALBA1 overexpressing parasite and the *Not1.1* mutant as well as the *Not1.2* mutant.** The Spearman coefficient shows the correlation of two dataset, with value of 1 for positive correlation (red box) and value of -1 for negative correlation (green box). Star symbol (*) represents the statistical significance of the spearman coefficient.

5.5 Proposed mechanism of NOT1.1- and NOT1.2-mediated gene regulation

Based on the findings presented in this study, the mechanism underlying the functions of NOT1 proteins in regulating gene expression is proposed, as shown in **Figure 5.7**.

NOT1.1 influences recruitment of the elongating form of RNAPII for 150 genes, which however is not correlated with changes in mRNA abundance of the corresponding genes. This discordance suggests a predominant role of NOT1.1 in post-transcriptional regulation.

The positive correlation of differential mRNA abundance between the *Caf1* mutant and the *Not1.1* mutant suggests that NOT1.1-CAF1 interaction is required for the partial function of CAF1. When *Not1.1* is knocked out, the function of CAF1 is affected, leading to similar transcript changes in the *Not1.1* mutant and the *Caf1* mutant. The negative correlation of differential mRNA abundance between the *Caf1* mutant and the *Not1.2* mutant suggests that NOT1.2-CAF1 interaction is likely to inactivate CAF1.

Both NOT1.1 and NOT1.2 contain a conserved MIF4G domain, providing a structural basis for CAF1 binding. The co-immunoprecipitation of CAF1 shows that it does interact with NOT1.1 and NOT1.2 in *P. falciparum* as it does in other species. The opposing roles of NOT1.1 and NOT1.2 in modulating the functions of CAF1 may be achieved by competition to be the scaffold for CAF1 binding. The NOT1.1-CAF1 complex enables CAF1 to function properly whereas the NOT1.2-CAF1 complex sequesters CAF1 from NOT1.1 and thereby inactivates its function.

Competition between NOT1.1 and NOT1.2 appears to modulate the function of ALBA1 as well, which is a potential interacting proteins of the CCR4-NOT complexes in *P. falciparum*. NOT1.1-ALBA1 interaction may be required for the function of ALBA1 and NOT1.2-ALBA1 interaction may inactivate ALBA1.

Both CAF1 and ALBA1 are critical players in regulating gene expression post-transcriptionally. The deadenylase CAF1 shortens the poly(A) tail of mRNA and triggers mRNA degradation, which directly regulates mRNA turnover [125, 207]. ALBA1 physically interacts with 1,193 transcripts and represses translation of some invasion-related transcripts [310]. Although whether ALBA1 has a direct role in mRNA turnover

is unknown, the overexpression of ALBA1 indeed affects mRNA abundance of 117 genes at ring stage and 926 genes at trophozoite stage, suggesting a role of ALBA1 in mRNA metabolism [310]. Therefore, NOT1.1 and NOT1.2 modulate the functions of CAF1 and ALBA1 probably in regulating mRNA turnover.

Overall, NOT1.1 and NOT1.2 are proposed to regulate gene expression by modulating functions of their interacting proteins CAF1 and ALBA1, both of which are key players in post-transcriptional regulation. It is possible that the parasite could favor the formation of one complex with NOT1.1 scaffold over the other with NOT1.2 scaffold, probably depending on environment and other co-factors, thereby allowing the parasite to control gene expression and subsequent outcomes.

Figure 5.7. Proposed mechanism of NOT1.1- and NOT1.2-mediated gene regulation. NOT1.1 and NOT1.2 compete to be the scaffold for CAF1 to form NOT1.1-CAF1 complex or NOT1.2-CAF1 complex, respectively. The two forms function oppositely in regulating mRNA abundance, possibly via affecting mRNA turnover. The genes being regulated are involved in merozoite egress, merozoite invasion, parasite virulence and other cellular processes.

5.6 Improvement and future perspectives

Some experiments in this thesis will be improved and more experiments will be done in the future in order to complete this project.

Additional clones of mutants

So far, I have obtained one clone of the *Not1.1* knockout line and one clone of the *Not1.2* knockdown line. To validate the findings in this thesis, additional independent mutant parasite subclones will be obtained. The transcriptional profile of these clones will be assessed to validate the opposing functions of NOT1.1 and NOT1.2 in regulating mRNA abundance. Moreover, the expression of some multifamily genes was affected in the two mutants as shown in Section 4.3.1 and 4.7.1. The additional clones will verify whether the differential mRNA abundance is resulted from clonal variations or truly affected by NOT1.1 or NOT1.2 loss of function.

The parasite harboring both mutations

Although NOT1 is an essential gene in yeast [125, 147], this study demonstrates that NOT1.1 is dispensable, possibly due to compensation by NOT1.2. NOT1.2 may be an essential gene given the inability to isolate a knockout or achieve complete knockdown. As NOT1.1 and NOT1.2 function oppositely in regulating mRNA abundance, it would be intriguing therefore to characterize parasites lacking NOT1.1 and NOT1.2 functions.

To achieve this, the *Not1.1* mutant can be transfected with the plasmids, pLNOT1.2-HA-ecDDD and pThDHFR. It is possible that parasites lacking NOT1.1 and NOT1.2 are not viable, so that TMP ligand will be maintained in the culture aiming to stabilize NOT1.2 tagged with the ecDDD. The protein level of NOT1.2-HA-ecDDD will be measured in the presence or absence of TMP ligand. If knockdown is observed, the phenotype and transcriptional profile of the parasites lacking both NOT1.1 and NOT1.2 will be analyzed.

Phenotypes of the *Not1.1* mutant and the *Not1.2* mutant

Based on the transcriptional profile of the *Not1.1* mutant and the *Not1.2* mutant, NOT1.1 and NOT1.2 affect the expression of many egress- and invasion-related genes as well as genes encoding exported proteins. The defective invasion rate of the two

mutants may be caused by the dysregulated egress or invasion-related genes. The expression of some invasion genes at the protein level will be investigated by Western blot. Total protein lysates of the mutants and wildtype will be harvested at six timepoints across one IDC, corresponding to the six timepoints in transcriptional profiles. The expression of the selected invasion protein will be detected by antibody targeting that specific protein. Any defects in the expression of these invasion proteins will help better understand the mechanism underlying the defective invasion rate observed in the *Not1.1* mutant and the *Not1.2* mutant.

Exported proteins contribute to cytoadherence of infected erythrocytes, which is critical for parasite virulence. Also, KAHRP is a major component of the knobs on the surface of infected erythrocytes contributing to cytoadherence [295]. The expression of KAHRP genes is upregulated in the *Not1.1* mutant and downregulated in the *Not1.2* mutant throughout one IDC. Therefore, any defects in cytoadherence will be investigated in the context of *Not1.1* or *Not1.2* mutation, for example, formation of the knob on erythrocyte surface and rigidity of the infected erythrocyte.

In addition, the parasite growth and invasion rate, which have been measured by blood smear and counting in this study, will be re-assessed by flow cytometry. Parasite DNA in iRBCs will be stained by Hoechst 33342 and detected by UV laser. The parasitemia will be calculated from 100,000 total RBCs in each sample using FACS, which will reinforce the reliability of data and eliminate errors of manual counting.

mRNA decay rate

Based on the proposed functions of NOT1.1 and NOT1.2 in Section 5.5, two NOT1 paralogs are likely to regulate mRNA abundance by influencing mRNA turnover. To test this hypothesis, mRNA decay rate will be examined in the *Not1.1* mutant and the *Not1.2* mutant using actinomycin D to block transcription followed by half-life determination for existing mRNA over the subsequent four-hour period [323]. Comparison between change in half-life of mRNA and change in mRNA abundance in each mutant will define the mechanism underlying the functions of NOT1.1 and NOT1.2 in gene regulation.

DNA and RNA bound by NOT1.1 and NOT1.2

Histone H3 was co-immunoprecipitated by all of the tagged NOT1.2, CAF1 and NOT4, which the latter two co-immunoprecipitates proteins involved in histone modification and transcription. The mutation of *Not1.1* affects recruitment of the elongating form of RNAPII. All these observations suggest that the CCR4-NOT complex is recruited to the genome, possibly to specific genes. The potential DNA bound by NOT1.1 or NOT1.2 can be inspected by ChIP-on-chip in the NOT1.1-HA line and the NOT1.2-HA line. Recruitment of NOT1.1 across the genome can be compared to differential recruitment of RNAPII in the context of *Not1.1* mutation, which may reveal how NOT1.1 affects RNAPII binding.

Likewise, the RNA species that are bound to NOT1.1 or NOT1.2 can be determined by immunoprecipitation followed by microarray analysis. Although NOT1.1 and NOT1.2 do not possess RNA-binding domains, the close connections of the CCR4-NOT complex to ribosomes and the translation initiation complex suggest that NOT1.1 and NOT1.2 are very likely to interact with mRNA. The comparison between potential RNA bound by NOT1 paralogs and changes of transcript abundance in the corresponding mutants will fine-tune the post-transcriptional functions of NOT1 paralogs in gene regulations.

Subcellular localization of the CCR4-NOT complex across one IDC

The subcellular localization of NOT1.2 and NOT4 have been tested at schizont stage by Western blot. NOT1.2 is located predominantly in the cytoplasm. NOT4 as a subunit of the complex is located in both compartments. CAF1 co-immunoprecipitates proteins with both localizations. These observations suggest cytoplasmic and nuclear subcellular localizations of the complex with NOT1.2 scaffold, consistent with the localization of the complex in other organisms. On the other hand, the localization of the HA-tagged NOT1.1 has not been investigated yet.

Given the opposite functions of NOT1.1 and NOT1.2 in regulating gene expression, two NOT1 proteins are hypothesized to be located in the same compartment at the same stage. To test this hypothesis, the HA-tagged NOT1.1 will be inspected in the cytoplasmic and nuclear lysates. In addition, the localization of NOT1.1, NOT1.2, CAF1

and NOT4 at ring, trophozoite and schizont stages will be examined by immunofluorescent assay.

Structural differences of NOT1.1 and NOT1.2

The highly conserved MIF4G domain, DUF2819 domain, and NOT1 domain provide a structural basis for NOT1.1 and NOT1.2 to interact with other subunits. While some of these interactions have been verified by mass spectrometry as summarized in Section 5.1.1, interactions with NOT2 and NOT5 were not detected. Both are stable subunits of the CCR4-NOT complex in other eukaryotes, raising the question as to why the conserved NOT1 domains do not interact with NOT2 and NOT5 at least not as strongly as other subunits in *P. falciparum*.

Furthermore, given the high similarity between the conserved domains in NOT1.1 and NOT1.2, the paralogs exert opposite functions in regulating gene expression, probably through modulating the function of their interacting proteins. In the future, the structure of NOT1.1 and NOT1.2 can be investigated in order to decipher as to how the two paralogs can modulate the functions of CAF1 and ALBA1 in an opposing manner with similar binding domains.

Chapter 6 Conclusion

The CCR4-NOT complex, involved in multiple levels of gene regulation in eukaryotes, is highly conserved in *P. falciparum*. Its CAF1 subunit has been shown to influence transcript abundance, highlighting the significant role of this complex in gene regulation within *P. falciparum*, which is rather mysterious. My project aims to investigate the functions of two NOT1 paralogs in gene regulation.

NOT1.1 and NOT1.2 are shown to be the scaffold of the complexes for other core subunits, including CAF1, CAF40, CCR4 and possibly NOT4. Some potential interacting proteins of the complexes were identified, for example, PABP, PAIP, ALBA1 and histone H3, suggesting the roles of the complexes in transcription and mRNA metabolism.

The transcriptional profiles of the two mutants indicate that both NOT1 paralogs regulate mRNA abundance, including 715 genes in the *Not1.1* mutant and 202 genes in the *Not1.2* mutant. The egress- and invasion-related genes were abundantly affected, which may result in defective growth and invasion rate in the two mutants.

Furthermore, I found that the functions of two NOT1 paralogs in gene regulations are opposing. The opposite functions of NOT1.1 and NOT1.2 might be achieved by modulating the activity of their interacting proteins, including CAF1, a CCR4-NOT subunit and ALBA1, a potential interacting partner of the complexes in *P. falciparum*. The interaction with NOT1.1 is proposed to be required for the function of CAF1 and ALBA1, whereas NOT1.2 may provide a competing scaffold that shields CAF1 and ALBA1 from NOT1.1.

This study reveals the striking opposing or antagonistic functions of NOT1.1 and NOT1.2 in regulating gene expression and extends the knowledge of the CCR4-NOT complex in *P. falciparum*. *Plasmodium* is unique among organisms studied to date in harboring two NOT1 paralogs, a feature that could be exploited to reveal novel targets for antimalarial drug development.

Appendix

Supplementary Table 1. Similarity between the conserved domains of NOT1 in yeast, human, *Drosophila* and *P. falciparum*. The conserved domains were pair-wise blasted using the NCBI BLASTP program. The lengths and the similarities of the identified regions are given. Due to the partial NOT1.2 MIF4G domain, the full-length NOT1.2 was blasted with the MIF4G domains of other proteins. The NOT1 domain of NOT1.2 consists of two regions, 3,988-4,213 aa (part1) and 4,331-4,455 aa (part2), and blasted separately.

MIF4G domain				
	yeast	human	<i>Drosophila</i>	NOT1.1
human	222 aa (63%)			
<i>Drosophila</i>	216 aa (64%)	247 aa (80%)		
NOT1.1	203 aa (57%)	231 aa (60%)	212 aa (58%)	
NOT1.2	97 aa (65%)	98 aa (69%)	98 aa (68%)	199 aa (65%)
DUF2819 domain				
	yeast	human	<i>Drosophila</i>	NOT1.1
human	185 aa (54%)			
<i>Drosophila</i>	178 aa (52%)	182 aa (70%)		
NOT1.1	141 aa (56%)	154 aa (59%)	148 aa (58%)	
NOT1.2	193 aa (52%)	153 aa (54%)	147 aa (52%)	197 aa (61%)
NOT1 domain				
	yeast	human	<i>Drosophila</i>	NOT1.1
human	358 aa (55%)			
<i>Drosophila</i>	347 aa (56%)	364 aa (87%)		
NOT1.1	354 aa (52%)	365 aa (67%)	239 aa (66%)	
NOT1.2 part1	218 aa (52%)	216 aa (62%)	213 aa (60%)	209 aa (69%)
NOT1.2 part2	95 aa (63%)	105 aa (80%)	106 aa (80%)	100 aa (78%)

Supplementary Table 2. 20 NOT1.2 potential interacting proteins at schizont stage. A total of 20 proteins were detected in both of the independent biological duplicates of co-immunoprecipitation. The number of distinct peptides identified by mass spectrometry and the corresponding protein score are given. The genes are ordered according to the average number of the detected peptides.

No	Gene	Description	Peptide	Protein score
1	PF3D7_0813300	conserved Plasmodium protein, unknown function	8/10	369.36/187.31
2	PF3D7_1447000	40S ribosomal protein S5	12/4	374/37.68
3	PF3D7_1224300	polyadenylate-binding protein, putative (PABP)	4/9	110.76/43.24
4	PF3D7_1424400	60S ribosomal protein L7-3, putative	1/10	42.47/34.69
5	PF3D7_1417200	NOT family protein, putative	2/9	98.21/412.16
6	PF3D7_1471100	exported protein 2 (EXP2)	4/4	111.04/216.96
7	PF3D7_0815600	eukaryotic translation initiation factor 3 subunit G, putative (EIF3G)	4/4	250.67/227.49
8	PF3D7_0814200	DNA or RNA-binding protein Alba 1 (ALBA1)	3/4	63/105.07
9	PF3D7_1110400	asparagine-rich antigen	3/4	97.13/183.27
10	PF3D7_1142500	60S ribosomal protein L28 (RPL28)	2/4	102.99/292.88
11	PF3D7_0507600	cell differentiation protein, putative (CAF40)	3/2	106.73/100.78
12	PF3D7_1105100	histone H2B (H2B)	4/1	169.61/24.17
13	PF3D7_1015900	enolase (ENO)	1/2	41.05/82.48
14	PF3D7_0610400	histone H3 (H3)	2/1	47.26/27.21
15	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1, putative (PAIP1)	1/2	52.53/61.82
16	PF3D7_0811300	CCR4-associated factor 1 (CAF1)	2/1	66.47/47.11
17	PF3D7_0402000	Plasmodium exported protein (PHISTa), unknown function	2/1	67.07/73.18
18	PF3D7_1019400	60S ribosomal protein L30e, putative	1/1	31.08/53.75
19	PF3D7_1342400	casein kinase II beta chain (CK2beta2)	1/1	47.87/96.24
20	PF3D7_0831400	Plasmodium exported protein, unknown function	1/1	60.06/56.32

Supplementary Table 3. 244 CAF1 potential interacting proteins at schizont stage. A total of 244 proteins were detected in both of the independent biological duplicates of co-immunoprecipitation. The number of distinct peptides identified by mass spectrometry and the corresponding protein score are given. The genes are ordered according to the average number of the detected peptides.

No	Gene	Description	Peptide	Protein score
1	PF3D7_1417200	NOT family protein, putative	27/41	888.46/1555.65
2	PF3D7_1359600	conserved Plasmodium protein, unknown function	21/25	1190.37/1223.63
3	PF3D7_1224300	polyadenylate-binding protein 1, putative	19/23	1092.45/1562.07
4	PF3D7_1021900	PHAX domain-containing protein, putative	20/21	600.5/729.19
5	PF3D7_0612100	eukaryotic translation initiation factor 3 subunit L, putative	19/19	1230.85/1219.84
6	PF3D7_0524000	karyopherin beta	19/19	970.26/910.81
7	PF3D7_0511500	RNA pseudouridylate synthase, putative	19/19	411.1/399.08
8	PF3D7_0305500	conserved Plasmodium protein, unknown function	19/17	489.31/379.43
9	PF3D7_1145400	dynammin-like protein	17/18	632.34/599.51
10	PF3D7_0108300	conserved Plasmodium protein, unknown function	17/17	679.75/724.4
11	PF3D7_0811300	CCR4-associated factor 1	13/15	789.53/976.89
12	PF3D7_0811600	conserved Plasmodium protein, unknown function	13/14	289.54/341.46
13	PF3D7_1036900	conserved Plasmodium protein, unknown function	10/15	720.85/1000.37
14	PF3D7_1103800	CCR4-NOT transcription complex subunit 1, putative	13/12	525.22/414.77
15	PF3D7_0811200	ER membrane protein complex subunit 1, putative	12/13	389.96/413.19
16	PF3D7_1408700	conserved protein, unknown function	13/11	422.17/354.72
17	PF3D7_1423700	conserved Plasmodium protein, unknown function	10/12	609.58/788.53
18	PF3D7_0519500	carbon catabolite repressor protein 4, putative	10/11	286.73/468.4
19	PF3D7_0510100	conserved Plasmodium protein, unknown function	9/11	308.9/636.05
20	PF3D7_0507600	cell differentiation protein, putative	8/10	794.96/885.29
21	PF3D7_0810600	ATP-dependent RNA helicase DBP1, putative	9/9	373.08/394.62
22	PF3D7_0609000	conserved Plasmodium protein, unknown function	9/8	288.85/225.09

23	PF3D7_0624600	SNF2 helicase, putative	10/7	284.02/149.87
24	PF3D7_1005500	regulator of nonsense transcripts 1, putative	6/10	238.83/369.28
25	PF3D7_0612900	nucleolar GTP-binding protein 1, putative	8/8	255.39/254.92
26	PF3D7_0706500	conserved Plasmodium protein, unknown function	8/7	443.51/475.03
27	PF3D7_0604500	conserved Plasmodium protein, unknown function	5/10	187.28/301.82
28	PF3D7_0302500	cytoadherence linked asexual protein 3.1	7/8	271.47/292.38
29	PF3D7_1352700	intron-binding protein aquarius, putative	8/7	274.84/161.24
30	PF3D7_1019000	eukaryotic translation initiation factor subunit eIF2A, putative	5/9	149.85/329.98
31	PF3D7_0112200	multidrug resistance-associated protein 1	11/3	381.22/206.8
32	PF3D7_0608700	T-complex protein 1 subunit zeta	6/7	189.38/241.69
33	PF3D7_0904900	copper-transporting ATPase	7/6	248.23/206.84
34	PF3D7_0720700	phosphoinositide-binding protein, putative	7/6	295.74/184.79
35	PF3D7_0527500	Hsc70-interacting protein	7/6	161.43/118.77
36	PF3D7_1103100	60S acidic ribosomal protein P1, putative	6/6	427.82/373.37
37	PF3D7_1347500	DNA/RNA-binding protein Alba 4	6/6	165.42/179.42
38	PF3D7_0630300	DNA polymerase epsilon catalytic subunit A, putative	7/5	156.32/136.35
39	PF3D7_1211400	heat shock protein DNAJ homologue Pfj4	6/5	249.32/271.73
40	PF3D7_1231800	asparagine-rich protein, putative	5/6	174.47/250.73
41	PF3D7_1472200	histone deacetylase, putative	4/7	142.6/219.02
42	PF3D7_1406200	conserved Plasmodium protein, unknown function	6/5	104.59/138.55
43	PF3D7_0212500	conserved Plasmodium protein, unknown function	7/4	135.35/39.8
44	PF3D7_1305300	conserved Plasmodium protein, unknown function	5/5	318.69/330.96
45	PF3D7_1419400	conserved Plasmodium membrane protein, unknown function	3/7	157.1/260.87
46	PF3D7_1318800	translocation protein SEC63, putative	4/6	157.47/234.38
47	PF3D7_0615400	ribonuclease, putative	4/6	157.27/205.53
48	PF3D7_1020700	N-acetyltransferase, GNAT family, putative	6/4	143.25/178.42
49	PF3D7_0607700	conserved Plasmodium protein, unknown function	1/8	35.56/226.75
50	PF3D7_1463300	DNA polymerase alpha subunit B, putative	5/4	150.69/150.16

51	PF3D7_1136400	signal recognition particle subunit SRP72, putative	4/5	160.82/147.24
52	PF3D7_1142600	60S ribosomal protein L35ae, putative	5/4	145.9/93.75
53	PF3D7_0212400	conserved Plasmodium membrane protein, unknown function	6/3	146.86/82.91
54	PF3D7_0528100	AP-1 complex subunit beta, putative	5/4	195.54/78.55
55	PF3D7_1007700	AP2 domain transcription factor AP2-I	5/4	179.3/62.16
56	PF3D7_1236100	clustered-asparagine-rich protein	4/4	157.78/151.86
57	PF3D7_0103900	parasite-infected erythrocyte surface protein	5/3	160.84/138.05
58	PF3D7_0628200	eukaryotic translation initiation factor 2-alpha kinase	4/4	136.66/127.06
59	PF3D7_0413500	phosphoglucomutase-2	5/3	181.4/125.9
60	PF3D7_1142100	conserved Plasmodium protein, unknown function	3/5	65.73/113.83
61	PF3D7_0910900	DNA primase large subunit, putative	6/2	241.38/99.13
62	PF3D7_1006700	conserved Plasmodium protein, unknown function	4/4	98.73/85.15
63	PF3D7_0704000	conserved Plasmodium membrane protein, unknown function	6/2	130.7/28.12
64	PF3D7_1412500	actin II	3/4	259.33/340.61
65	PF3D7_0413600	26S protease regulatory subunit 6B, putative	1/6	43.21/260.29
66	PF3D7_0310500	ATP-dependent RNA helicase DHX57, putative	1/6	45.05/251.1
67	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1, putative	3/4	78.14/170.41
68	PF3D7_1227800	elongator complex protein 3, putative	4/3	265.99/155.35
69	PF3D7_0303000	N-ethylmaleimide-sensitive fusion protein	4/3	195/142.02
70	PF3D7_1417800	DNA replication licensing factor MCM2	4/3	94.61/92.63
71	PF3D7_0804500	conserved Plasmodium membrane protein, unknown function	3/4	46.04/66.19
72	PF3D7_1308900	mRNA-decapping enzyme 2, putative	3/4	63.5/51.73
73	PF3D7_1130700	structural maintenance of chromosomes protein 1, putative	5/2	147.8/47.15
74	PF3D7_0905100	nucleoporin NUP100/NSP100, putative	6/1	65.12/37.25

75	PF3D7_1313800	conserved Plasmodium membrane protein, unknown function	4/3	63.68/34.54
76	PF3D7_0935400	gametocyte development protein 1	2/4	82.47/188.11
77	PF3D7_1238600	sphingomyelin phosphodiesterase	2/4	85.23/168.35
78	PF3D7_1111100	replication factor C subunit 5, putative	3/3	127.22/156.94
79	PF3D7_0725300	conserved Plasmodium protein, unknown function	2/4	93.27/123.01
80	PF3D7_1366900	conserved Plasmodium protein, unknown function	2/4	28.01/118.97
81	PF3D7_1248900	26S protease regulatory subunit 8, putative	2/3	111.27/169.3
82	PF3D7_1426700	phosphoenolpyruvate carboxylase	2/3	166.63/142.41
83	PF3D7_1462800	glyceraldehyde-3-phosphate dehydrogenase	2/3	60.93/136.14
84	PF3D7_0613500	AP-3 complex subunit beta, putative	2/3	91.81/129.31
85	PF3D7_1126500	conserved Plasmodium protein, unknown function	2/3	77.08/120.69
86	PF3D7_0703500	erythrocyte membrane-associated antigen	2/3	43.25/119.99
87	PF3D7_0215800	origin recognition complex subunit 5	2/3	81.62/117.43
88	PF3D7_1244600	ADP-ribosylation factor GTPase-activating protein	2/3	64.77/113.78
89	PF3D7_1014700	prohibitin, putative	3/2	100.31/105.67
90	PF3D7_0204200	conserved Plasmodium protein, unknown function	2/3	54.69/93.06
91	PF3D7_1409600	conserved Plasmodium protein, unknown function	4/1	189.46/90.51
92	PF3D7_1320000	golgi protein 1	3/2	60.27/83.45
93	PF3D7_1439100	DEAD/DEAH box helicase, putative	2/3	88.65/76.3
94	PF3D7_0935200	vacuolar protein sorting-associated protein 33, putative	2/3	28.52/72.35
95	PF3D7_1366300	conserved Plasmodium protein, unknown function	2/3	31.75/67.74
96	PF3D7_1441400	FACT complex subunit SSRP1, putative	2/3	78.96/67.57
97	PF3D7_1004200	WD repeat-containing protein, putative	3/2	111.52/64.82
98	PF3D7_1359000	conserved Plasmodium protein, unknown function	3/2	65.67/63.37
99	PF3D7_1456700	conserved Plasmodium protein, unknown function	2/3	42.33/59.34

100	PF3D7_0921800	ribosome production factor 1, putative	2/3	91.52/56.02
101	PF3D7_1139800	conserved Plasmodium protein, unknown function	3/2	74.5/47.01
102	PF3D7_1352400	conserved Plasmodium protein, unknown function	3/2	64.66/46.49
103	PF3D7_0605800	DNA repair protein RAD50, putative	3/2	131.55/45.48
104	PF3D7_1130400	26S protease regulatory subunit 6A, putative	2/2	55.06/95.65
105	PF3D7_0308200	T-complex protein 1 subunit eta	2/2	56.08/94.36
106	PF3D7_1244200	RNA polymerase II transcription factor B subunit 2, putative	1/3	32.98/92.13
107	PF3D7_0204700	hexose transporter	1/3	70.53/90.46
108	PF3D7_0205900	26S proteasome regulatory subunit RPN1, putative	2/2	60.91/86.65
109	PF3D7_0208100	conserved Plasmodium protein, unknown function	2/2	46.5/76.34
110	PF3D7_1423000	nucleolar GTP-binding protein 2, putative	1/3	21.48/69
111	PF3D7_1107700	pescadillo homolog	1/3	40.91/62.72
112	PF3D7_1237200	conserved Plasmodium protein, unknown function	3/1	98.16/60.68
113	PF3D7_0929200	RNA-binding protein, putative	2/2	52.47/60.41
114	PF3D7_1311400	AP-1 complex subunit mu-1	2/2	58.06/58.94
115	PF3D7_1352500	thioredoxin-related protein, putative	2/2	99.76/58.65
116	PF3D7_1321500	3',5'-cyclic nucleotide phosphodiesterase, putative	2/2	59.46/58.13
117	PF3D7_0412000	LITAF-like zinc finger protein, putative	3/1	152.3/55.53
118	PF3D7_1208100	conserved Plasmodium protein, unknown function	2/2	94.72/52.09
119	PF3D7_1317400	zinc finger protein, putative	2/2	45.25/44.31
120	PF3D7_1330800	RNA-binding protein, putative	3/1	103.78/43.7
121	PF3D7_1034000	Sec1 family protein, putative	2/2	41.92/31.77
122	PF3D7_0724000	Rab GTPase activator and protein kinase, putative	2/2	41.98/30.1
123	PF3D7_0610400	histone H3	1/2	190.2/255.85
124	PF3D7_1015900	enolase	1/2	134.91/194.67
125	PF3D7_0424200	reticulocyte binding protein homologue 4	1/2	68.43/117.28
126	PF3D7_1008400	26S protease regulatory subunit 4, putative	1/2	28.27/106.91
127	PF3D7_1321700	splicing factor 1	1/2	58.92/101.7
128	PF3D7_1013900	translation initiation factor eIF-2B subunit delta, putative	1/2	31.13/93.67
129	PF3D7_1143300	DNA-directed RNA polymerases I and III subunit RPAC1, putative	1/2	37.62/91.28

130	PF3D7_1332000	syntaxin, Qa-SNARE family	1/2	78.63/90.08
131	PF3D7_1315300	conserved Plasmodium protein, unknown function	1/2	46.32/85.91
132	PF3D7_0212300	peptide chain release factor subunit 1, putative	1/2	32.46/84.86
133	PF3D7_1471000	RNA 3'-terminal phosphate cyclase-like protein, putative	2/1	115.01/78.9
134	PF3D7_1414500	atypical protein kinase, ABC-1 family, putative	1/2	30.28/75.77
135	PF3D7_1347100	DNA topoisomerase 3, putative	1/2	65.16/73.72
136	PF3D7_1124100	beige/BEACH domain protein, putative	1/2	38.24/72.22
137	PF3D7_1410600	eukaryotic translation initiation factor 2 subunit gamma, putative	1/2	20.24/61.12
138	PF3D7_0210200	conserved Plasmodium protein, unknown function	1/2	49.75/59.79
139	PF3D7_1366600	signal recognition particle receptor subunit alpha, putative	2/1	96.17/57.73
140	PF3D7_0304200	EH domain-containing protein	1/2	43.94/55.61
141	PF3D7_0610900	transcription elongation factor SPT5, putative	1/2	53.43/54.8
142	PF3D7_0703000	conserved Plasmodium protein, unknown function	2/1	104.4/53.1
143	PF3D7_1230900	serine/threonine protein kinase RIO1, putative	1/2	36.04/50.75
144	PF3D7_0302800	conserved Plasmodium protein, unknown function	2/1	88.93/49.5
145	PF3D7_1437000	N-acetyltransferase, GNAT family, putative	2/1	71.73/47.8
146	PF3D7_1312900	eukaryotic translation initiation factor 4 gamma	2/1	43.54/46.39
147	PF3D7_1459600	conserved Plasmodium protein, unknown function	1/2	32.73/44.8
148	PF3D7_1134200	conserved Plasmodium protein, unknown function	1/2	31.05/43.45
149	PF3D7_1224400	conserved Plasmodium protein, unknown function	1/2	30.21/40.99
150	PF3D7_0618700	trafficking protein particle complex subunit 6A, putative	2/1	71.42/40.93
151	PF3D7_0420100	serine/threonine protein kinase RIO2	2/1	73.3/35.54
152	PF3D7_1471200	inorganic anion exchanger, inorganic anion antiporter	1/2	42.03/33.39
153	PF3D7_1005100	U3 small nucleolar RNA-associated protein 25, putative	2/1	53.48/33.27
154	PF3D7_0323800	conserved Plasmodium protein, unknown function	2/1	31.79/32.27
155	PF3D7_0629400	polyadenylate-binding protein 3, putative	2/1	49.27/31.6

156	PF3D7_1225600	conserved Plasmodium protein, unknown function	2/1	44.9/30.14
157	PF3D7_1439300	Sad1/UNC domain-containing protein, putative	2/1	42.22/29.48
158	PF3D7_1119500	AP-4 complex subunit mu, putative	1/2	32.96/26.87
159	PF3D7_0409800	zinc finger protein, putative	2/1	31.33/26.34
160	PF3D7_1025300	conserved Plasmodium protein, unknown function	2/1	51.29/24.63
161	PF3D7_1309700	vacuolar protein sorting-associated protein 18, putative	2/1	50.99/21.59
162	PF3D7_1123900	13 kDa ribonucleoprotein-associated protein, putative	1/1	176.77/148.59
163	PF3D7_1307700	conserved Plasmodium protein, unknown function	1/1	34.77/124.44
164	PF3D7_0912500	SAP domain-containing protein, putative	1/1	25.47/115.48
165	PF3D7_1469100	conserved Plasmodium protein, unknown function	1/1	54.99/93.15
166	PF3D7_1314500	transmembrane emp24 domain-containing protein, putative	1/1	58.11/87.78
167	PF3D7_0322400	regulator of initiation factor 2 (eIF2)	1/1	97.08/84.7
168	PF3D7_0203700	protein MAK16, putative	1/1	84.82/84.15
169	PF3D7_1368900	conserved Plasmodium protein, unknown function	1/1	28.9/73.74
170	PF3D7_1139100	RNA-binding protein, putative	1/1	57.5/68.85
171	PF3D7_1105700	tRNA-splicing ligase RtcB, putative	1/1	24.14/65.13
172	PF3D7_0618200	conserved Plasmodium protein, unknown function	1/1	59.39/63.57
173	PF3D7_0312400	glycogen synthase kinase 3	1/1	139.27/59.74
174	PF3D7_1448400	ubiquitin-protein ligase, putative	1/1	72.04/59.2
175	PF3D7_1034900	methionine--tRNA ligase	1/1	50.37/58.33
176	PF3D7_0611400	SWIB/MDM2 domain-containing protein	1/1	41.9/57.75
177	PF3D7_1438500	cleavage and polyadenylation specificity factor subunit 3, putative	1/1	44.97/55.32
178	PF3D7_1037100	pyruvate kinase 2	1/1	63.44/53.78
179	PF3D7_0219400	unspecified product	1/1	84.61/53.75
180	PF3D7_0214100	protein transport protein SEC31	1/1	34.65/52.94
181	PF3D7_1214100	GPI ethanolamine phosphate transferase 3, putative	1/1	50/50.98
182	PF3D7_0404600	conserved Plasmodium membrane protein, unknown function	1/1	50.55/50.56
183	PF3D7_0722100	protein pelota homolog, putative	1/1	50.35/50.44

184	PF3D7_1223200	nucleus export protein BRR6, putative	1/1	48.73/48.33
185	PF3D7_0617900	histone H3 variant	1/1	34.06/48.27
186	PF3D7_0318900	conserved protein, unknown function	1/1	48.12/45.48
187	PF3D7_0726400	conserved Plasmodium membrane protein, unknown function	1/1	48.16/45.44
188	PF3D7_1466300	26S proteasome regulatory subunit RPN2, putative	1/1	24.91/44.83
189	PF3D7_0829000	conserved Plasmodium membrane protein, unknown function	1/1	40.79/44.6
190	PF3D7_1452500	SNARE protein, putative	1/1	45.04/44.37
191	PF3D7_0705500	inositol-phosphate phosphatase, putative	1/1	48.18/44.26
192	PF3D7_1354400	V-type proton ATPase 21 kDa proteolipid subunit, putative	1/1	43.7/44.12
193	PF3D7_0405100	protein transport protein Sec24B, putative	1/1	40/42.78
194	PF3D7_0928200	conserved Plasmodium protein, unknown function	1/1	44.3/41.9
195	PF3D7_0519200	V-type proton ATPase 16 kDa proteolipid subunit	1/1	36.47/40.36
196	PF3D7_1408500	conserved Plasmodium protein, unknown function	1/1	50.11/39.44
197	PF3D7_0716300	conserved Plasmodium protein, unknown function	1/1	31.84/38.89
198	PF3D7_0213100	protein SIS1	1/1	31.7/38.82
199	PF3D7_0111300	P-loop containing nucleoside triphosphate hydrolase, putative	1/1	51.48/38.51
200	PF3D7_1353500	RNA polymerase II transcription factor B subunit 4, putative	1/1	40.04/38.43
201	PF3D7_0925700	histone deacetylase 1	1/1	40.75/38.23
202	PF3D7_0505000	conserved Plasmodium membrane protein, unknown function	1/1	42.39/37.99
203	PF3D7_0813000	conserved Plasmodium protein, unknown function	1/1	31.59/37.65
204	PF3D7_0703200	conserved Plasmodium protein, unknown function	1/1	34.54/37.42
205	PF3D7_1326400	translation initiation factor eIF-2B subunit gamma, putative	1/1	40.37/37.15
206	PF3D7_0418600	regulator of chromosome condensation, putative	1/1	30.62/36.82
207	PF3D7_0822600	protein transport protein SEC23	1/1	73/36.24
208	PF3D7_1119400	ubiquitin-protein ligase, putative	1/1	46.19/35.94
209	PF3D7_0607400	G-protein associated signal transduction protein, putative	1/1	23.84/35.38
210	PF3D7_0530100	SNARE protein, putative	1/1	51.29/35.22

211	PF3D7_0505400	conserved Plasmodium protein, unknown function	1/1	39.82/34.27
212	PF3D7_1121700	protein GCN20	1/1	25.69/33.99
213	PF3D7_0321400	protein kinase, putative	1/1	32.03/33.47
214	PF3D7_0302100	serine/threonine protein kinase	1/1	37.1/33.21
215	PF3D7_1125600	conserved Plasmodium protein, unknown function	1/1	41.9/32.07
216	PF3D7_0715300	calcium/calmodulin-dependent protein kinase, putative	1/1	38.23/31.55
217	PF3D7_1213900	conserved Plasmodium protein, unknown function	1/1	35.18/31.43
218	PF3D7_1029900	conserved Plasmodium protein, unknown function	1/1	30.45/30.53
219	PF3D7_0729300	60S ribosomal export protein NMD3, putative	1/1	33.67/30.06
220	PF3D7_0528800	nucleolar preribosomal GTPase, putative	1/1	22.14/29.91
221	PF3D7_1206300	conserved Plasmodium protein, unknown function	1/1	24.08/29.86
222	PF3D7_1243000	syntaxin, Qa-SNARE family	1/1	28.9/29.76
223	PF3D7_1338100	26S proteasome regulatory subunit RPN3, putative	1/1	27.54/29.73
224	PF3D7_1250100	osmiophilic body protein G377	1/1	32.57/28.68
225	PF3D7_1140800	conserved Plasmodium protein, unknown function	1/1	29.2/27.85
226	PF3D7_1226900	parasitophorous vacuolar protein 2	1/1	60.5/26.1
227	PF3D7_1250600	translation initiation factor eIF-2B subunit beta, putative	1/1	22.63/25.92
228	PF3D7_1412800	glycylpeptide N-tetradecanoyltransferase	1/1	27.73/25.75
229	PF3D7_1362400	calpain	1/1	20.55/25.11
230	PF3D7_0316700	protein YOP1, putative	1/1	39.23/24.72
231	PF3D7_1324700	SNARE protein, putative	1/1	23.14/24.41
232	PF3D7_1319100	ubiquitin carboxyl-terminal hydrolase MINDY, putative	1/1	79.47/24.04
233	PF3D7_0415300	cdc2-related protein kinase 3	1/1	43.77/23.86
234	PF3D7_0807200	conserved Plasmodium membrane protein, unknown function	1/1	44.32/23.29
235	PF3D7_1405600	ribonucleoside-diphosphate reductase small chain, putative	1/1	28.59/22.95
236	PF3D7_0321700	conserved Plasmodium protein, unknown function	1/1	32.39/22.88
237	PF3D7_0704300	conserved Plasmodium membrane protein, unknown function	1/1	50.6/22.4
238	PF3D7_0411900	DNA polymerase alpha catalytic subunit A	1/1	65.76/22.29

239	PF3D7_1416500	NADP-specific glutamate dehydrogenase	1/1	24.97/22.04
240	PF3D7_1205500	zinc finger protein, putative	1/1	25.9/21.7
241	PF3D7_1448300	conserved Plasmodium protein, unknown function	1/1	77.44/21.59
242	PF3D7_1215900	serpentine receptor, putative	1/1	27.67/21.17
243	PF3D7_0320700	signal peptidase complex subunit SPC2, putative	1/1	30.9/21.06
244	PF3D7_0721300	ATP-dependent RNA helicase DBP7, putative	1/1	29.47/20.09

Supplementary Table 4. Enriched functions of CAF1 potential interacting proteins.

The enriched pathways of proteins found in both biological duplicates of mass spectrometry were identified by Gene Ontology and MPM functional analysis. Those with a binomial $p < 0.05$ are listed with gene identity and description. The proteins in each pathway are ordered according to the average number of the detected peptides.

Pathway	Gene	Description
CCR4 - NOT complex	PF3D7_1417200	NOT family protein, putative
	PF3D7_0811300	CCR4-associated factor 1 (CAF1)
	PF3D7_0507600	cell differentiation protein, putative (CAF40)
	PF3D7_1103800	CCR4-NOT transcription complex subunit 1, putative (NOT1)
	PF3D7_0519500	carbon catabolite repressor protein 4, putative (CCR4)
RNA-binding proteins	PF3D7_1224300	polyadenylate-binding protein, putative (PABP)
	PF3D7_1417200	NOT family protein, putative
	PF3D7_0811300	CCR4-associated factor 1 (CAF1)
	PF3D7_0507600	cell differentiation protein, putative (CAF40)
	PF3D7_0810600	RNA helicase, putative
	PF3D7_0615400	ribonuclease, putative
	PF3D7_1347500	DNA or RNA-binding protein Alba 4 (ALBA4)
	PF3D7_1236100	clustered-asparagine-rich protein
	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1, putative (PAIP1)
	PF3D7_0703500	erythrocyte membrane-associated antigen
	PF3D7_0929200	RNA-binding protein, putative
	PF3D7_1410600	eukaryotic translation initiation factor 2 gamma subunit, putative
	PF3D7_0629400	RNA-binding protein, putative
RNA degradation	PF3D7_1224300	polyadenylate-binding protein, putative (PABP)
	PF3D7_1417200	NOT family protein, putative
	PF3D7_0811300	CCR4-associated factor 1 (CAF1)
	PF3D7_1103800	CCR4-NOT transcription complex subunit 1, putative (NOT1)
	PF3D7_0519500	carbon catabolite repressor protein 4, putative (CCR4)
	PF3D7_1005500	regulator of nonsense transcripts, putative
	PF3D7_1308900	mRNA-decapping enzyme 2, putative (DCP2)
mRNA surveillance	PF3D7_1224300	polyadenylate-binding protein, putative (PABP)
	PF3D7_1005500	regulator of nonsense transcripts, putative
	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1, putative (PAIP1)
	PF3D7_0703500	erythrocyte membrane-associated antigen

	PF3D7_0212300	peptide chain release factor subunit 1, putative
	PF3D7_0722100	PelOta protein homologue, putative
	PF3D7_1438500	cleavage and polyadenylation specificity factor, putative
Ribosome assembly	PF3D7_0612900	nucleolar GTP-binding protein 1, putative
	PF3D7_0615400	ribonuclease, putative
	PF3D7_1123900	splicing factor, putative
	PF3D7_1020700	histone acetyltransferase, putative
	PF3D7_1471000	RNA 3-terminal phosphate cyclase-like protein, putative
	PF3D7_0203700	protein MAK16, putative (MAK16)
	PF3D7_0921800	ribosome production factor 1, putative (RPF1)
	PF3D7_0420100	serine or threonine protein kinase RIO2 (RIO2)
	PF3D7_1107700	pescadillo homolog (PES)
	PF3D7_1423000	nucleolar GTP-binding protein 2, putative (NOG2)
	PF3D7_1230900	serine or threonine protein kinase RIO1, putative (RIO1)
	PF3D7_0729300	60S ribosomal export protein NMD3, putative (NMD3)
	PF3D7_1121700	protein GCN20 (GCN20)
	PF3D7_0528800	nucleolar preribosomal GTPase, putative
	PF3D7_0721300	DEAD or DEAH box ATP-dependent RNA helicase, putative
Translation initiation	PF3D7_1224300	polyadenylate-binding protein, putative (PABP)
	PF3D7_0612100	eukaryotic translation initiation factor 3 subunit L, putative (EIF3L)
	PF3D7_1019000	eukaryotic translation initiation factor subunit eIF2A, putative
	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1, putative (PAIP1)
	PF3D7_0322400	regulator of initiation factor 2 (eIF2)
	PF3D7_1013900	initiation factor 2 subunit family, putative
	PF3D7_1312900	eukaryotic translation initiation factor 4 gamma (EIF4G)
	PF3D7_1410600	eukaryotic translation initiation factor 2 gamma subunit, putative
	PF3D7_1326400	translation initiation factor EIF-2B gamma subunit, putative
	PF3D7_1250600	eukaryotic translation initiation factor 2b, subunit 2, putative
Histones	PF3D7_0610400	histone H3 (H3)
	PF3D7_0617900	histone H3 variant, putative (H3.3)
	PF3D7_0624600	SNF2 helicase, putative (ISWI)
	PF3D7_1227800	histone S-adenosyl methyltransferase, putative

Chromatin modifying proteins	PF3D7_1472200	histone deacetylase, putative (HDA1)
	PF3D7_1437000	N-acetyltransferase, putative
	PF3D7_0925700	histone deacetylase 1 (HDAC1)
Transcription-related proteins	PF3D7_1417200	NOT family protein, putative
	PF3D7_0811300	CCR4-associated factor 1 (CAF1)
	PF3D7_0507600	cell differentiation protein, putative (CAF40)
	PF3D7_1103800	CCR4-NOT transcription complex subunit 1, putative (NOT1)
	PF3D7_0519500	carbon catabolite repressor protein 4, putative (CCR4)
	PF3D7_0610400	histone H3 (H3)
	PF3D7_0624600	SNF2 helicase, putative (ISWI)
	PF3D7_1472200	histone deacetylase, putative (HDA1)
	PF3D7_1007700	transcription factor with AP2 domain(s) (ApiAP2)
	PF3D7_1441400	FACT complex subunit SSRP1, putative (FACT-S)
	PF3D7_1244200	RNA polymerase II transcription factor B subunit 2, putative (TFB2)
	PF3D7_1437000	N-acetyltransferase, putative
	PF3D7_0610900	transcription elongation factor SPT5, putative (SPT5)
	PF3D7_0617900	histone H3 variant, putative (H3.3)
	PF3D7_0925700	histone deacetylase 1 (HDAC1)
	PF3D7_1353500	RNA polymerase II transcription factor B subunit 4, putative (TFB4)
	PF3D7_0409800	zinc finger protein, putative
	PF3D7_1205500	zinc finger protein, putative
	PF3D7_1015900	enolase (ENO)
	PF3D7_1417800	DNA replication licensing factor MCM2 (MCM2)
	PF3D7_1143300	DNA-directed RNA polymerase I and III subunit RPACI, putative
DNA-directed DNA polymerase	PF3D7_1463300	DNA polymerase alpha subunit, putative
	PF3D7_0630300	DNA polymerase epsilon, catalytic subunit a, putative
	PF3D7_0411900	DNA polymerase alpha
DNA replication	PF3D7_0910900	DNA primase large subunit, putative
	PF3D7_1463300	DNA polymerase alpha subunit, putative
	PF3D7_0310500	DEAD box helicase, putative
	PF3D7_0630300	DNA polymerase epsilon, catalytic subunit a, putative
	PF3D7_1111100	replication factor C subunit 5, putative (RFC5)
	PF3D7_1417800	DNA replication licensing factor MCM2 (MCM2)
	PF3D7_0605800	DNA repair protein RAD50, putative (RAD50)
	PF3D7_1347100	DNA topoisomerase 3, putative (TOP3)
	PF3D7_0111300	replication factor c protein, putative
	PF3D7_0411900	DNA polymerase alpha
	PF3D7_1405600	ribonucleotide reductase small subunit (RNR)

Protein transport	PF3D7_1145400	dynamin-like protein (DYN1)
	PF3D7_0528100	AP-1 complex subunit beta, putative
	PF3D7_0613500	AP-3 complex subunit beta, putative
	PF3D7_1244600	ADP-ribosylation factor GTPase-activating protein, putative (ARFGAP)
	PF3D7_1332000	syntaxin, Qa-SNARE family (SYN5)
	PF3D7_1314500	cop-coated vesicle membrane protein p24 precursor, putative
	PF3D7_1311400	AP-1 complex subunit mu-1 (AP1M1)
	PF3D7_0618700	trafficking protein particle complex subunit 6A, putative
	PF3D7_0822600	protein transport protein SEC23 (SEC23)
	PF3D7_0935200	vacuolar protein sorting-associated protein 33, putative (VPS33)
	PF3D7_0304200	EH (Eps15 homology) protein (PAST1)
	PF3D7_1452500	SNARE protein, putative (BET1)
	PF3D7_0214100	protein transport protein SEC31 (SEC31)
	PF3D7_0530100	SNARE protein, putative (SYN6)
	PF3D7_0405100	protein transport protein Sec24B (SEC24B)
	PF3D7_1034000	Sec1 family protein, putative
	PF3D7_1309700	vacuolar protein sorting-associated protein 18, putative (VPS18)
	PF3D7_1119500	AP-4 complex subunit mu, putative
	PF3D7_1243000	syntaxin, Qa-SNARE family (SYN16)
	PF3D7_0320700	signal peptidase complex subunit 2, putative (SPC2)
	PF3D7_1324700	SNARE protein, putative (YKT6.2)
Proteasome and ubiquitin-dependent protein degradation	PF3D7_0413600	26S protease regulatory subunit 6B, putative (RPT3)
	PF3D7_1248900	26S protease regulatory subunit 8, putative (RPT6)
	PF3D7_1130400	26S protease regulatory subunit 6A, putative (RPT5)
	PF3D7_0205900	26S proteasome regulatory subunit RPN1, putative (RPN1)
	PF3D7_1008400	26S protease regulatory subunit 4, putative (RPT2)
	PF3D7_1119400	ubiquitin-protein ligase, putative
	PF3D7_1466300	26S proteasome regulatory subunit RPN2, putative (RPN2)
	PF3D7_1338100	26S proteasome regulatory subunit RPN3, putative (RPN3)

Supplementary Table 5. 45 NOT4 potential interacting proteins at schizont stage. A total of 45 proteins were detected in both of the independent biological duplicates of immunoprecipitation. The number of distinct peptides identified by mass spectrometry and the corresponding protein score are given. The genes are ordered according to the average number of the detected peptides.

No	Gene	Description	Peptide	Protein score
1	PF3D7_1130200	60S ribosomal protein P0 (PfP0)	8/23	546.26/2006.95
2	PF3D7_0523000	multidrug resistance protein 1 (MDR1)	1/26	31.94/1338.41
3	PF3D7_1447000	40S ribosomal protein S5	7/19	160.59/1157.84
4	PF3D7_1206200	eukaryotic translation initiation factor 3 subunit C, putative (EIF3C)	1/21	29.91/917.13
5	PF3D7_1116800	heat shock protein 101 (HSP101)	1/18	23.59/910.44
6	PF3D7_1424400	60S ribosomal protein L7-3, putative	4/14	174.77/707.03
7	PF3D7_1224300	polyadenylate-binding protein, putative (PABP)	5/13	175.98/829.71
8	PF3D7_1105400	40S ribosomal protein S4, putative	6/11	203.23/733.94
9	PF3D7_0511500	RNA pseudouridylate synthase, putative	1/14	50.14/565.61
10	PF3D7_0813300	conserved Plasmodium protein, unknown function	4/10	203.06/926.46
11	PF3D7_0524000	karyopherin beta (KASbeta)	2/10	77.53/749.04
12	PF3D7_1006800	single-strand telomeric DNA-binding protein GBP2, putative	2/10	66.5/503.2
13	PF3D7_0604500	conserved Plasmodium protein, unknown function	4/8	124.57/295.3
14	PF3D7_0818200	14-3-3 protein (14-3-3I)	2/9	63.6/720.65
15	PF3D7_1118500	box C or D snoRNP rRNA 2-O-methylation factor, putative	1/10	27.24/588.48
16	PF3D7_0815600	eukaryotic translation initiation factor 3 subunit G, putative (EIF3G)	5/6	339.97/578.72
17	PF3D7_0703500	erythrocyte membrane-associated antigen	1/9	75.22/369.14
18	PF3D7_0708400	heat shock protein 90 (HSP90)	4/6	203.2/292.9
19	PF3D7_1465900	40S ribosomal protein S3	3/6	98.73/344.94
20	PF3D7_1235300	CCR4-NOT transcription complex subunit 4, putative (NOT4)	2/7	97.68/291.54
21	PF3D7_0814200	DNA or RNA-binding protein Alba 1 (ALBA1)	2/7	84.46/246.04
22	PF3D7_0108300	conserved Plasmodium protein, unknown function	1/8	35.18/401.18
23	PF3D7_0814000	60S ribosomal protein L13-2, putative	1/7	83.07/633.05

24	PF3D7_0517000	60S ribosomal protein L12, putative	4/4	178.21/272.38
25	PF3D7_0827900	protein disulfide isomerase (PDI8)	1/6	29.47/239.4
26	PF3D7_1323400	60S ribosomal protein L23 (RPL23)	2/5	110.51/483.51
27	PF3D7_1459400	conserved Plasmodium protein, unknown function	2/4	92.44/157.98
28	PF3D7_1107300	polyadenylate-binding protein-interacting protein 1, putative (PAIP1)	1/5	52.02/271.87
29	PF3D7_1239200	transcription factor with AP2 domain(s) (ApiAP2)	1/5	43.35/274.44
30	PF3D7_1468100	conserved Plasmodium protein, unknown function	1/5	34.18/264.63
31	PF3D7_1351400	60S ribosomal protein L17, putative	1/5	25.17/805.02
32	PF3D7_0309600	60S acidic ribosomal protein P2 (PfP2)	2/3	87.57/192.63
33	PF3D7_1006200	DNA or RNA-binding protein Alba 3 (ALBA3)	1/4	76.55/254.75
34	PF3D7_1423700	conserved Plasmodium protein, unknown function	1/4	67.21/175.37
35	PF3D7_1110400	asparagine-rich antigen	1/4	27.65/203.47
36	PF3D7_1117700	GTP-binding nuclear protein RAN or TC4 (RAN)	3/1	66.62/30.81
37	PF3D7_1224000	GTP cyclohydrolase I (GCH1)	1/3	35.46/262.39
38	PF3D7_0507600	cell differentiation protein, putative (CAF40)	1/2	47.73/128.37
39	PF3D7_0209800	ATP-dependent RNA helicase UAP56 (UAP56)	1/2	43.39/130.59
40	PF3D7_0520000	40S ribosomal protein S9, putative	1/2	36.79/56.2
41	PF3D7_1242700	40S ribosomal protein S17, putative	1/2	25.18/169.21
42	PF3D7_1314900	general transcription factor IIH subunit 2 (P44)	1/1	144.9/119.65
43	PF3D7_0905400	high molecular weight rhoptry protein 3 (RhopH3)	1/1	47.78/22.33
44	PF3D7_1015900	enolase (ENO)	1/1	39.14/78.18
45	PF3D7_0610400	histone H3 (H3)	1/1	29.46/43.53

Supplementary Table 6. Proliferation of the *Not1.1* mutant. Parasites were grown in 2% hematocrit for three IDCs. Initial parasitemia was 0.1%. The number of iRBC in around 1,000 total RBCs was counted in biological triplicates (r1, r2 and r3) for each parasite line in each IDC. The data was plotted in **Figure 4.5**. The average parasitemia and the standard deviation were calculated. One-tail t-test was conducted between parasitemia of the *Not1.1* mutant and wildtype in each IDC.

IDC cycles	0	1			2			3		
Replicate		r1	r2	r3	r1	r2	r3	r1	r2	r3
Parasite line	wildtype									
iRBC		5	5	8	49	49	44	215	231	205
Total RBC		1000	1000	1000	960	1038	967	1015	1134	939
Parasitemia (%)	0.10	0.50	0.50	0.80	5.10	4.72	4.55	21.18	20.37	21.83
Average		0.60			4.79			21.13		
Standard deviation		0.17			0.28			0.73		
Parasite line	Not1.1 mutant									
iRBC		3	1	4	31	33	20	114	130	126
Total RBC		1000	1000	1000	1003	1006	941	1065	1077	1012
Parasitemia (%)	0.10	0.30	0.10	0.40	3.09	3.28	2.13	10.70	12.07	12.45
Average		0.27			2.83			11.74		
Standard deviation		0.15			0.62			0.92		
p-value		0.0339			0.0091			0.0001		

Supplementary Table 7. Invasion rate of the *Not1.1* mutant. Parasites were grown in 1% hematocrit in biological triplicates (r1, r2 and r3). Initial parasitemia was 0.5% and 1% synchronized parasites at late schizont stage. The number of iRBC in around 1,000 total RBCs was counted 14 hours afterward. Invasion rate was calculated as the ratio of the parasitemia after invasion versus the initial parasitemia. The average parasitemia and standard deviation were calculated. The data was plotted in **Figure 4.6**. One-tail t-test was conducted between the invasion rate of the *Not1.1* mutant and wildtype for each setup.

Hematocrit (%)		1	1
Parasitemia (%)		0.5	1
Parasite line	wildtype		
Parasitemia (%)	r1	2.53	3.06
	r2	2.37	3.18
	r3	2.10	2.92
Invasion rate	r1	5.06	3.06
	r2	4.73	3.18
	r3	4.20	2.92
Average		4.66	3.06
Standard deviation		0.43	0.13
Parasite line	<i>Not1.1</i> mutant		
Parasitemia (%)	r1	1.56	1.47
	r2	1.68	1.78
	r3	1.32	1.81
Invasion rate	r1	3.12	1.47
	r2	3.37	1.78
	r3	2.63	1.81
Average		3.04	1.68
Standard deviation		0.38	0.19
p-value		0.0042	0.0004

Supplementary Table 8. Merozoite number per schizont in the *Not1.1* mutant.

Parasites were grown in 2% hematocrit in biological duplicates with 5% parasitemia. The number of merozoites in each schizont was counted for 50 iRBCs once segmentation of merozoites was clearly observed within schizont parasites. The data was plotted in **Figure 4.7**. One-tail t-test was conducted between the number of merozoites in the *Not1.1* mutant and wildtype for each biological replicate.

Replicate	Replicate 1				Replicate 2			
Parasite line	wildtype		Not1.1 mutant		wildtype		Not1.1 mutant	
Number of Merozoites	17	16	22	15	19	18	18	18
	18	18	17	16	15	16	16	20
	17	16	15	19	19	19	21	19
	18	17	18	19	18	18	14	19
	18	15	17	15	17	15	16	16
	16	16	15	16	17	16	16	17
	20	16	17	17	18	16	16	18
	16	17	18	17	16	17	18	17
	15	16	16	16	16	15	16	16
	18	16	16	16	20	16	18	18
	17	16	16	15	16	16	20	19
	20	17	20	18	19	17	18	16
	14	15	16	18	16	18	15	20
	14	15	17	16	18	18	18	17
	16	15	17	16	18	18	17	17
	17	16	18	17	18	15	19	17
	15	15	21	15	16	16	17	18
	15	15	18	16	18	17	17	16
	16	17	16	17	15	18	15	16
	14	20	20	15	17	16	16	16
	15	17	15	15	15	17	15	16
	17	18	18	16	15	16	19	18
	18	16	20	18	15	17	17	16
	18	15	16	15	16	19	17	17
	15	17	18	17	15	16	18	15
Average	16.42		17.07		16.84		17.18	
Standard deviation	1.47		1.68		1.36		1.51	
p-value	0.0517				0.1197			

Supplementary Table 9. Estimated age of parasites harvested at 6 timepoints across one IDC in the *Not1.1* mutant. The Spearman correlation between the transcriptional profile of each sample and the periodicity of gene expression was conducted. The age was estimated with the highest Spearman coefficient. The average of the ages and the standard deviation were calculated. The data was plotted in **Figure 4.8**. One-tail t-test was conducted between the ages of parasites in the *Not1.1* mutant and wildtype at the same timepoint. One-tailed t-test cannot calculate the p-value for samples with identical ages, which was denoted as NA.

	TP1	TP2	TP3	TP4	TP5	TP6
wildtype						
rep1	2	12	18	28	36	40
rep2	4	14	20	28	36	40
rep3	4	14	18	28	36	40
average	3.333	13.333	18.667	28	36	40
standard deviation	1.155	1.155	1.155	0	0	0
<i>Not1.1</i> mutant						
rep1	0	6	18	24	32	40
rep2	0	6	18	24	32	40
rep3	0	6	18	24	32	40
average	0	6	18	24	32	40
standard deviation	0	0	0	0	0	0
p-value	0.0189	0.0041	0.2113	NA	NA	NA

Supplementary Table 10. 715 genes with significantly differential mRNA abundance in the *Not1.1* mutant. The differential mRNA abundance is represented by the log₂-transformed ratio of mRNA abundance in the *Not1.1* mutant versus that in wildtype. The order of the genes corresponds to the heatmap in Figure 4.9. The significant change is defined as an expression fold-change more than 1.5 (corresponding to the log₂ ratio at 0.585) and p-value less than 0.05. There are six clusters of genes according to the timepoint of the greatest downregulation and genes in each cluster are labeled.

No.	Gene	Description	TP1	TP2	TP3	TP4	TP5	TP6	Cluster
1	PF3D7_0608800	ornithine aminotransferase (OAT)	0.636	-0.510	0.046	0.299	0.582	0.549	1
2	PF3D7_1115300	cysteine proteinase falcipain 2b (FP2B)	-0.473	0.008	0.106	0.639	0.560	-0.093	1
3	PF3D7_0621200	pyridoxine biosynthesis protein PDX1 (PDX1)	-0.233	-0.235	-0.056	0.732	0.649	0.171	1
4	PF3D7_1364800	DNA-directed RNA polymerase II, putative	-0.032	-0.201	-0.129	0.270	0.626	-0.143	1
5	PF3D7_1149400	Plasmodium exported protein, unknown function	-0.209	-0.365	0.513	0.814	0.446	-0.040	1
6	PF3D7_0532300	Plasmodium exported protein (PHISTb), unknown function	-0.195	-0.209	0.237	0.789	0.452	-0.047	1
7	PF3D7_0213100	protein SIS1 (SIS1)	-0.169	-0.119	-0.217	0.295	0.417	0.616	1
8	PF3D7_1022400	serine or arginine-rich splicing factor 4 (SRSF4)	-0.252	-0.241	-0.083	0.189	0.264	0.616	1
9	PF3D7_0512000	prefoldin subunit, putative	-0.547	-0.311	-0.063	0.333	0.339	0.717	1
10	PF3D7_1149100	conserved Plasmodium protein, unknown function	-0.225	-0.478	-0.253	0.442	0.915	0.427	1
11	PF3D7_1252300	conserved Plasmodium protein, unknown function	-0.133	-0.259	-0.023	0.270	0.817	0.580	1
12	PF3D7_0831400	Plasmodium exported protein, unknown function	0.150	-0.258	-0.248	0.496	0.740	0.189	1
13	PF3D7_0212700	SRR1-like protein	-0.763	-0.077	-0.071	0.019	0.116	0.228	1
14	PF3D7_0407100	methyltransferase, putative	-0.601	-0.058	-0.082	-0.001	-0.016	0.182	1
15	PF3D7_1038700	Plasmodium exported protein, unknown function	-1.013	-0.176	0.031	0.270	0.098	0.319	1
16	PF3D7_1451000	conserved Plasmodium protein, unknown function	-0.853	-0.344	-0.116	0.069	-0.068	0.244	1
17	PF3D7_0313500	GTP-binding protein EngA, putative	-0.778	-0.312	-0.047	0.154	0.004	0.123	1
18	PF3D7_1028000	methyltransferase, putative	-0.675	-0.262	-0.085	0.214	-0.014	0.043	1
19	PF3D7_1412600	deoxyhypusine synthase (DHS)	-0.758	-0.189	-0.101	0.275	0.108	0.000	1
20	PF3D7_1314800	ubiquitin-like protein, putative	-0.711	-0.223	0.165	0.106	-0.097	0.047	1
21	PF3D7_1473000	conserved Plasmodium protein, unknown function	-0.637	-0.175	0.087	0.144	-0.118	0.064	1
22	PF3D7_0113000	glutamic acid-rich protein (GARP)	-0.807	-0.716	0.082	0.640	0.652	0.235	1
23	PF3D7_0501000	Plasmodium exported protein, unknown function	-0.707	-0.681	0.046	0.150	0.417	0.547	1
24	PF3D7_1017400	phosphomannomutase, putative (PMM)	-0.625	-0.132	0.146	0.259	0.048	-0.573	1
25	PF3D7_1353700	reactive oxygen species modulator 1, putative (ROMO1)	-0.632	-0.048	0.085	0.182	0.021	-0.191	1
26	PF3D7_1035800	probable protein, unknown function (M712)	-0.630	-0.509	0.121	0.768	0.192	-0.220	1
27	PF3D7_1301400	Plasmodium exported protein (hyp12), unknown function (HYP12)	-0.409	-0.343	0.308	0.779	-0.039	-0.171	1
28	PF3D7_0318100	stomatin-like protein	0.100	-0.809	-0.314	0.248	0.336	0.589	1
29	PF3D7_0501200	parasite-infected erythrocyte surface protein (PIESP2)	0.321	-1.084	-0.332	0.360	0.457	0.629	1
30	PF3D7_0802000	glutamate dehydrogenase, putative (GDH3)	0.150	-0.617	0.178	0.099	0.489	0.125	1
31	PF3D7_0713100	Pfmc-2TM Maurers cleft two transmembrane protein (MC-2TM)	0.414	-1.001	0.477	0.374	0.672	0.294	1

32	PF3D7_0817300	asparagine-rich antigen	0.183	-0.767	0.259	0.198	0.306	0.136	1
33	PF3D7_0532400	lysine-rich membrane-associated PHISTb protein (LyMP)	-0.141	-0.773	0.249	0.592	0.499	0.088	1
34	PF3D7_1001000	Plasmodium exported protein (hyp12), unknown function (Pf13)	0.120	-1.047	0.179	0.454	0.398	0.161	1
35	PF3D7_1149000	antigen 332, DBL-like protein (Pf332)	0.158	-0.543	0.059	0.363	0.703	0.235	1
36	PF3D7_1201000	Plasmodium exported protein (PHISTb), unknown function	-0.048	-0.928	0.064	0.185	0.855	0.349	1
37	PF3D7_0301700	Plasmodium exported protein, unknown function	-0.292	-0.731	0.030	0.047	0.188	0.113	1
38	PF3D7_0114500	Plasmodium exported protein (hyp10), unknown function	-0.455	-1.543	0.254	0.636	0.370	0.104	1
39	PF3D7_1317200	transcription factor with AP2 domain(s) (ApiAP2)	-0.109	-0.589	0.069	0.204	0.191	0.039	1
40	PF3D7_1039700	Pfmc-2TM Maurers cleft two transmembrane protein (MC-2TM)	-0.095	-0.655	0.118	-0.030	0.294	0.204	1
41	PF3D7_0402000	Plasmodium exported protein (PHISTa), unknown function	-0.066	-0.805	-0.240	0.055	0.556	0.002	1
42	PF3D7_1215300	10 kDa chaperonin (CPN10)	-0.270	-0.495	0.011	0.166	0.612	0.013	1
43	PF3D7_0201500	Plasmodium exported protein (hyp9), unknown function	0.008	-0.786	0.249	0.220	-0.033	0.497	1
44	PF3D7_0532500	Plasmodium exported protein, unknown function	0.098	-0.673	0.283	0.325	0.327	0.414	1
45	PF3D7_0424000	Plasmodium exported protein (PHISTc), unknown function	-0.018	-0.775	-0.484	0.032	-0.140	-0.137	1
46	PF3D7_1223900	50S ribosomal protein L24, putative	0.157	-1.040	-0.411	0.079	0.189	0.305	1
47	PF3D7_1224000	GTP cyclohydrolase I (GCH1)	-0.061	-0.977	-0.421	-0.061	0.071	0.093	1
48	PF3D7_0832200	conserved Plasmodium protein, unknown function	0.122	-0.897	-0.135	0.068	0.053	-0.115	1
49	PF3D7_1031300	conserved Plasmodium protein, unknown function	0.079	-0.666	-0.011	-0.067	0.090	0.018	1
50	PF3D7_1442400	conserved Plasmodium protein, unknown function	0.104	-0.726	0.090	-0.022	0.000	0.042	1
51	PF3D7_1131100	serpentine receptor, putative (SR1)	0.255	-0.831	-0.170	-0.539	-0.206	0.426	2
52	PF3D7_1317600	conserved Plasmodium protein, unknown function	-0.915	-0.906	-0.408	-0.295	-0.183	0.282	2
53	PF3D7_0322800	conserved Plasmodium protein, unknown function	-0.406	-0.621	-0.378	-0.249	-0.219	-0.021	2
54	PF3D7_0718400	mitochondrial ribosomal protein S8 precursor, putative	-0.809	-0.694	-0.525	-0.264	-0.578	0.182	2
55	PF3D7_0302300	erythrocyte membrane protein 1 (PfEMP1), pseudogene	-0.281	-0.635	-0.122	-0.072	-0.328	-0.208	2
56	PF3D7_1338400	conserved Plasmodium protein, unknown function	-0.230	-0.656	-0.215	-0.323	-0.440	-0.165	2
57	PF3D7_1245300	NEDD8-conjugating enzyme UBC12, putative (UBC12)	-0.539	-0.306	-0.625	-0.124	-0.191	-0.238	2
58	PF3D7_1230600	sun-family protein, putative	-0.977	-0.486	-0.930	-0.065	-0.145	0.107	2
59	PF3D7_0726100	Plasmodium exported protein, unknown function	-0.785	-0.621	-0.293	-0.076	-0.193	-0.673	2
60	PF3D7_0424800	Plasmodium exported protein (PHISTb), unknown function	-0.997	-0.674	-0.290	0.156	-0.027	-0.325	2
61	PF3D7_1305400	AAR2 protein, putative	-0.630	-0.238	-0.148	-0.101	0.010	0.043	2
62	PF3D7_0318000	conserved Plasmodium protein, unknown function	-0.659	-0.216	-0.292	-0.230	-0.249	0.075	2
63	PF3D7_0822400	conserved Plasmodium protein, unknown function	-0.672	-0.258	-0.171	-0.173	-0.316	0.019	2
64	PF3D7_1237800	ubiquitin-like modifier HUB1, putative (HUB1)	-0.616	-0.129	-0.269	-0.051	-0.270	0.152	2
65	PF3D7_1248100	methyltransferase, putative	-0.621	-0.295	-0.288	0.013	-0.293	0.255	2
66	PF3D7_0720200	conserved Plasmodium protein, unknown function	-0.664	-0.079	-0.387	0.006	-0.013	0.087	2
67	PF3D7_1210300	conserved Plasmodium protein, unknown function	-0.958	-0.177	-0.447	0.182	-0.150	0.234	2
68	PF3D7_0829300	U6 snRNA-associated Sm-like protein LSM8, putative (LSM8)	-1.028	-0.091	-0.543	0.023	-0.180	0.680	2
69	PF3D7_0913400	conserved Plasmodium protein, unknown function	-0.485	-0.123	-0.883	0.154	0.118	0.335	2
70	PF3D7_1313700	septum formation protein MAF homologue, putative	-0.292	-0.193	-0.641	-0.095	0.072	0.252	2
71	PF3D7_1127100	deoxyuridine 5-triphosphate nucleotidohydrolase (dUTPase)	-0.668	-0.336	-0.778	0.003	0.607	0.468	2

72	PF3D7_1234000	conserved Plasmodium protein, unknown function	-0.383	-0.626	-0.622	-0.105	-0.270	0.225	2
73	PF3D7_1368600	mitochondrial import inner membrane translocase subunit TIM9, putative (TIM9)	-0.668	-0.541	-0.214	0.261	0.171	0.233	2
74	PF3D7_0810500	protein phosphatase PPM7, putative (PPM7)	-0.664	-0.128	-0.344	0.346	0.096	0.298	2
75	PF3D7_1365500	glycine cleavage system T protein, putative (GCVT)	-0.201	-0.613	-0.575	0.051	0.024	0.498	2
76	PF3D7_1443900	heat shock protein 90, putative (HSP90)	-0.275	-0.609	-0.446	0.176	0.057	0.544	2
77	PF3D7_0810400	aquaporin, putative (AQP2)	-0.089	-0.424	-0.288	0.213	0.150	0.602	2
78	PF3D7_1409900	cytidine diphosphate-diacylglycerol synthase (CDS)	-0.258	-0.225	-0.448	0.134	0.121	0.750	2
79	PF3D7_1468200	conserved protein, unknown function	-0.159	-0.141	-0.500	0.088	0.133	0.757	2
80	PF3D7_1314700	conserved Plasmodium protein, unknown function	-0.353	-0.327	-0.692	-0.021	0.080	0.408	2
81	PF3D7_0502700	conserved Plasmodium protein, unknown function	-0.689	-0.397	-0.634	-0.067	0.073	0.385	2
82	PF3D7_0915200	ribonuclease H2 subunit C, putative	-0.682	-0.241	-0.463	-0.201	0.043	0.388	2
83	PF3D7_0810700	conserved Plasmodium protein, unknown function	-0.869	-0.638	-0.809	-0.148	-0.202	0.444	2
84	PF3D7_0513400	GTP-binding protein, putative	-0.551	-0.460	-0.654	-0.019	-0.050	0.388	2
85	PF3D7_0413900	ubiquitin carboxyl-terminal hydrolase 13, putative (USP13)	-0.606	-0.344	-0.755	-0.377	-0.331	0.621	2
86	PF3D7_1250900	conserved Plasmodium protein, unknown function	-0.917	-0.469	-0.863	-0.246	-0.128	0.615	2
87	PF3D7_1341900	V-type proton ATPase subunit D, putative	-0.487	-0.069	-0.218	-0.137	-0.011	0.627	2
88	PF3D7_1012400	hypoxanthine-guanine phosphoribosyltransferase (HGPRT)	-0.540	-0.039	-0.224	-0.064	-0.005	0.878	2
89	PF3D7_1106100	apicoplast ribosomal protein S15 precursor, putative	-0.434	-0.291	-0.536	-0.261	-0.193	0.882	2
90	PF3D7_0211800	asparagine--tRNA ligase (AsnRS)	-0.589	-0.276	-0.213	0.106	0.078	0.543	2
91	PF3D7_0511100	conserved Plasmodium protein, unknown function	-0.714	-0.233	-0.345	-0.074	-0.096	0.513	2
92	PF3D7_1233100	conserved protein, unknown function	-0.747	-0.331	-0.319	-0.074	-0.274	0.666	2
93	PF3D7_1360200	ER membrane protein complex subunit 3, putative (EMC3)	-0.620	-0.077	-0.301	0.019	-0.078	0.519	2
94	PF3D7_1209000	conserved Plasmodium protein, unknown function	-0.879	-0.187	-0.546	0.093	-0.063	0.877	2
95	PF3D7_1469100	conserved Plasmodium protein, unknown function	-0.598	-0.109	-0.349	-0.029	-0.202	0.650	2
96	PF3D7_0702900	centrin, putative	-0.481	-0.035	-0.437	0.179	-0.018	0.703	2
97	PF3D7_0721100	conserved Plasmodium protein, unknown function	0.412	-0.579	-0.634	-0.019	0.407	0.402	3
98	PF3D7_1426900	ubiquinol-cytochrome c reductase hinge protein, putative	0.101	-0.465	-0.586	-0.092	0.214	0.342	3
99	PF3D7_1303000	conserved Plasmodium protein, unknown function	0.006	-0.490	-0.725	-0.174	0.181	0.311	3
100	PF3D7_1024500	conserved Plasmodium protein, unknown function	0.082	-0.344	-0.640	-0.191	0.200	0.231	3
101	PF3D7_0510100	conserved Plasmodium protein, unknown function	0.720	0.157	-0.916	-0.053	0.268	0.351	3
102	PF3D7_1372400	acyl-CoA synthetase (ACS4)	0.382	0.084	-0.625	-0.052	0.184	0.230	3
103	PF3D7_0513600	deoxyribodipyrimidine photo-lyase, putative	0.409	0.055	-0.628	-0.041	0.162	0.361	3
104	PF3D7_1239200	transcription factor with AP2 domain(s) (ApiAP2)	0.465	0.135	-0.614	-0.151	0.154	0.392	3
105	PF3D7_1002600	conserved Plasmodium protein, unknown function	0.570	0.071	-0.811	-0.205	0.224	0.504	3
106	PF3D7_0729600	conserved Plasmodium protein, unknown function	0.480	-0.017	-0.765	-0.054	-0.006	0.304	3
107	PF3D7_0211200	ras-related protein Rab-5A (RAB5a)	0.498	0.041	-0.781	-0.087	0.043	0.413	3
108	PF3D7_1023300	conserved Plasmodium protein, unknown function	0.360	0.057	-0.805	0.040	0.323	0.294	3
109	PF3D7_0219500	pseudouridine synthase, putative	0.442	0.042	-1.045	-0.001	0.267	0.475	3
110	PF3D7_0504400	ATP-dependent helicase, putative	0.279	0.087	-0.623	-0.006	0.195	0.333	3
111	PF3D7_1246900	RAC-beta serine or threonine protein kinase (PKB)	0.454	0.109	-0.974	0.049	0.161	0.391	3
112	PF3D7_1115000	conserved Plasmodium protein, unknown function	0.237	0.077	-0.823	0.030	0.179	0.284	3

113	PF3D7_1204700	conserved Plasmodium protein, unknown function	0.360	0.166	-1.096	-0.067	0.130	0.295	3
114	PF3D7_1209900	ABC transporter B family member 7, putative (ABCB7)	0.231	0.120	-0.641	-0.047	0.134	0.249	3
115	PF3D7_1008700	tubulin beta chain	0.599	0.006	-0.901	-0.079	0.354	0.454	3
116	PF3D7_0105900	DNA binding protein, putative	0.481	-0.021	-0.637	-0.002	0.329	0.337	3
117	PF3D7_1446800	heme detoxification protein (HDP)	0.410	0.001	-0.696	-0.030	0.423	0.323	3
118	PF3D7_0611200	transcription factor with AP2 domain(s), putative (ApiAP2)	0.397	-0.240	-1.014	0.004	0.459	0.304	3
119	PF3D7_0414000	chromosome associated protein, putative	0.269	-0.133	-0.707	-0.084	0.174	0.288	3
120	PF3D7_1001200	acyl-CoA binding protein, isoform 2, ACBP2 (ACBP2)	0.331	-0.193	-0.753	-0.103	0.192	0.187	3
121	PF3D7_1021900	conserved Plasmodium protein (10b antigen), unknown function	0.196	-0.108	-0.610	-0.074	0.198	0.163	3
122	PF3D7_0320000	protein phosphatase inhibitor 2 (I2)	0.450	-0.153	-1.576	-0.225	0.024	0.846	3
123	PF3D7_1438700	DNA primase small subunit	0.338	-0.108	-1.364	-0.199	0.068	0.709	3
124	PF3D7_0608400	conserved Plasmodium protein, unknown function	0.218	-0.063	-1.339	-0.062	0.071	0.737	3
125	PF3D7_1213600	dynein light chain 1, putative (DLC8)	0.479	-0.011	-1.343	-0.075	0.123	0.613	3
126	PF3D7_0215800	origin recognition complex subunit 5 (ORC5)	0.210	0.004	-0.672	-0.056	0.134	0.335	3
127	PF3D7_1304500	small heat shock protein, putative	0.306	-0.007	-1.102	-0.012	0.204	0.578	3
128	PF3D7_0218000	replication factor C subunit 2, putative (RFC2)	0.420	0.166	-1.047	-0.123	0.175	0.787	3
129	PF3D7_0719100	conserved Plasmodium protein, unknown function	0.293	0.116	-1.056	-0.222	0.074	0.586	3
130	PF3D7_0217900	conserved Plasmodium protein, unknown function	0.214	0.173	-0.655	-0.022	0.247	0.549	3
131	PF3D7_1226900	conserved Plasmodium protein, unknown function	0.441	-0.236	-0.568	-0.057	0.237	0.657	3
132	PF3D7_1230400	ATP-dependent protease subunit ClpQ (ClpQ)	0.277	-0.429	-0.900	0.183	0.088	0.904	3
133	PF3D7_0907900	peptide deformylase (PDF)	0.347	-0.122	-0.801	-0.135	0.036	0.662	3
134	PF3D7_0715800	drug or metabolite exporter, drug or metabolite transporter	0.318	-0.150	-0.893	-0.137	-0.021	1.279	3
135	PF3D7_0526200	conserved Plasmodium protein, unknown function	0.464	0.098	-1.133	-0.343	-0.203	0.036	3
136	PF3D7_1472700	DNA-directed RNA polymerase, alpha subunit, putative	0.306	0.125	-0.668	-0.070	-0.211	0.009	3
137	PF3D7_0929700	conserved Plasmodium protein, unknown function	0.523	0.363	-1.391	-0.413	-0.339	0.474	3
138	PF3D7_0305900	conserved Plasmodium protein, unknown function	0.388	0.291	-1.385	-0.530	-0.431	0.304	3
139	PF3D7_0320700	signal peptidase complex subunit 2, putative (SPC2)	0.215	0.340	-1.328	0.002	-0.165	0.555	3
140	PF3D7_0406600	conserved Plasmodium protein, unknown function	0.193	0.295	-1.403	-0.077	-0.019	0.406	3
141	PF3D7_0520500	phosphomethylpyrimidine kinase, putative	0.325	0.287	-0.764	-0.015	-0.098	0.312	3
142	PF3D7_1024200	RNA-binding protein, putative	0.142	0.280	-1.108	0.138	-0.501	0.327	3
143	PF3D7_0409500	conserved Plasmodium protein, unknown function	0.563	-0.339	-1.628	-0.349	0.061	0.750	3
144	PF3D7_0208600	ribosome-recycling factor (RRF1)	0.198	-0.152	-0.650	-0.027	0.054	0.224	3
145	PF3D7_1405600	ribonucleotide reductase small subunit (RNR)	0.445	-0.359	-1.346	-0.062	0.243	0.488	3
146	PF3D7_1415000	uracil-DNA glycosylase (UDG)	0.531	-0.461	-1.382	-0.145	0.101	0.220	3
147	PF3D7_0531200	mitochondrial ribosomal protein S16 precursor, putative	0.157	0.024	-0.907	0.009	0.090	0.115	3
148	PF3D7_0622400	conserved Apicomplexan protein, unknown function	0.280	0.068	-0.890	-0.109	0.148	0.109	3
149	PF3D7_1428600	peptide chain release factor 1, putative	0.349	0.043	-0.960	-0.196	0.140	0.105	3
150	PF3D7_0610000	mitochondrial ribosomal protein L19 precursor, putative	0.014	-0.127	-0.726	-0.126	0.173	0.065	3
151	PF3D7_1120600	conserved Plasmodium protein, unknown function	0.004	-0.123	-0.625	-0.172	0.094	0.051	3
152	PF3D7_1117600	conserved Plasmodium protein, unknown function	0.107	-0.114	-0.816	-0.085	0.232	0.158	3
153	PF3D7_0204900	ubiE or COQ5 methyltransferase, putative	0.218	-0.133	-1.664	-0.286	-0.072	0.578	3

154	PF3D7_0505500	DNA mismatch repair protein MSH6, putative (MSH6)	0.186	-0.078	-1.119	-0.273	0.007	0.359	3
155	PF3D7_0616200	kinetochore protein NDC80, putative (NDC80)	0.180	-0.074	-1.052	-0.228	0.001	0.319	3
156	PF3D7_0503000	50S ribosomal protein L28, apicoplast, putative	0.210	-0.145	-1.034	-0.233	-0.222	0.375	3
157	PF3D7_0827500	apicoplast ribosomal protein L21 precursor, putative	0.319	-0.176	-1.669	-0.418	-0.130	0.729	3
158	PF3D7_0911600	conserved Plasmodium protein, unknown function	0.161	-0.112	-1.201	-0.349	-0.125	0.476	3
159	PF3D7_0312000	conserved Plasmodium protein, unknown function	0.170	-0.031	-2.371	-0.067	-0.091	0.913	3
160	PF3D7_0513800	ras-related protein Rab-1A (RAB1a)	0.108	-0.001	-2.111	-0.182	-0.204	0.929	3
161	PF3D7_1322700	conserved Plasmodium protein, unknown function	0.029	-0.080	-0.716	-0.133	-0.068	0.364	3
162	PF3D7_0812200	trypsin-like serine protease, putative	0.104	0.059	-0.822	0.096	0.021	0.209	3
163	PF3D7_0919600	dolichyl-diphosphooligosaccharide--protein glycosyltransferase, putative	0.243	0.127	-1.659	-0.008	0.016	0.540	3
164	PF3D7_1003200	conserved Plasmodium protein, unknown function	0.088	-0.078	-1.303	-0.024	-0.087	0.355	3
165	PF3D7_1135500	conserved Plasmodium protein, unknown function	0.193	-0.005	-0.879	0.015	-0.102	0.282	3
166	PF3D7_1120800	conserved Plasmodium protein, unknown function	0.035	-0.114	-0.643	-0.029	-0.056	0.047	3
167	PF3D7_1338500	conserved Plasmodium protein, unknown function	0.042	-0.153	-0.877	-0.100	-0.058	0.109	3
168	PF3D7_1244000	glucose inhibited division protein a homologue, putative	0.035	-0.143	-0.832	-0.092	0.010	0.177	3
169	PF3D7_1474300	DNA repair metallo-beta-lactamase protein, putative	0.058	-0.253	-0.946	0.084	-0.068	0.247	3
170	PF3D7_0422900	conserved Plasmodium protein, unknown function	-0.013	-0.173	-1.177	0.051	-0.110	0.692	3
171	PF3D7_0109400	tubulin-specific chaperone a, putative	0.110	-0.247	-1.318	0.153	0.081	0.591	3
172	PF3D7_0802200	1-cys peroxiredoxin (1-CysPxn)	0.249	-0.239	-1.473	0.162	0.115	0.943	3
173	PF3D7_0419100	conserved Plasmodium protein, unknown function	0.079	-0.218	-0.800	0.130	0.176	0.224	3
174	PF3D7_1120900	heat shock factor-binding protein 1 (HSBP)	-0.093	-0.087	-1.252	0.190	0.126	0.408	3
175	PF3D7_1207900	conserved Plasmodium protein, unknown function	-0.137	-0.173	-0.837	0.094	0.113	0.385	3
176	PF3D7_1431500	mitogen-activated protein kinase 1 (MAPK1)	-0.024	-0.106	-0.661	0.068	0.106	0.292	3
177	PF3D7_0715500	mitochondrial ATP synthase F1, epsilon subunit, putative	-0.104	0.135	-0.770	-0.119	0.220	0.226	3
178	PF3D7_0720900	conserved Plasmodium protein, unknown function	-0.036	-0.014	-1.124	-0.083	0.192	0.308	3
179	PF3D7_1309900	conserved Plasmodium protein, unknown function	-0.155	0.021	-0.811	0.044	0.194	0.171	3
180	PF3D7_0710500	conserved Plasmodium protein, unknown function	0.127	-0.006	-1.189	0.154	0.213	0.416	3
181	PF3D7_1419800	conserved Plasmodium protein, unknown function	0.050	0.044	-0.591	0.079	0.156	0.257	3
182	PF3D7_1001300	Plasmodium exported protein (PHISTa), unknown function	0.002	0.126	-1.122	-0.095	0.243	0.365	3
183	PF3D7_1330600	elongation factor Tu, putative	0.034	0.193	-1.214	-0.019	0.214	0.539	3
184	PF3D7_0417200	bifunctional dihydrofolate reductase-thymidylate synthase (DHFR-TS)	0.107	0.018	-0.626	-0.180	0.037	0.183	3
185	PF3D7_0906100	vacuolar protein sorting-associated protein 46, putative (VPS46)	0.273	-0.018	-1.242	-0.309	0.150	0.462	3
186	PF3D7_0913600	conserved Plasmodium protein, unknown function	0.119	0.120	-0.971	-0.195	0.127	0.308	3
187	PF3D7_0405800	conserved Plasmodium protein, unknown function	0.005	0.138	-0.916	-0.258	0.162	0.338	3
188	PF3D7_1245100	kinesin-13, putative (KLP8)	-0.012	0.051	-0.741	-0.283	0.090	0.327	3
189	PF3D7_1239400	conserved Plasmodium protein, unknown function	0.063	0.130	-0.875	-0.356	-0.027	0.394	3
190	PF3D7_0911500	conserved Plasmodium protein, unknown function	-0.052	-0.028	-0.862	-0.339	0.031	0.329	3
191	PF3D7_1462100	conserved Plasmodium protein, unknown function	-0.023	-0.017	-0.708	-0.371	0.068	0.306	3
192	PF3D7_0217100	ATP synthase F1, alpha subunit	-0.017	-0.122	-1.095	-0.314	-0.167	0.636	3
193	PF3D7_1362600	conserved Plasmodium protein, unknown function	0.020	-0.140	-0.685	-0.235	-0.134	0.309	3

194	PF3D7_0413700	lysine decarboxylase-like protein, putative	-0.050	-0.085	-1.041	-0.294	-0.076	0.367	3
195	PF3D7_1242400	GA58-like protein, putative	-0.029	-0.107	-0.811	-0.245	-0.054	0.274	3
196	PF3D7_1106200	conserved Plasmodium protein, unknown function	-0.026	-0.069	-0.696	-0.185	-0.113	0.258	3
197	PF3D7_1311300	ATP synthase subunit gamma, mitochondrial	-0.076	-0.055	-0.979	-0.243	-0.121	0.365	3
198	PF3D7_1303100	methyltransferase-like protein, putative	-0.054	-0.128	-0.692	-0.110	-0.025	0.316	3
199	PF3D7_1363600	conserved Plasmodium protein, unknown function	-0.182	-0.260	-1.395	-0.198	-0.154	0.545	3
200	PF3D7_0727100	conserved Plasmodium protein, unknown function	-0.155	-0.105	-0.758	-0.216	0.022	0.352	3
201	PF3D7_1130900	conserved Plasmodium protein, unknown function	-0.088	-0.174	-0.796	-0.273	0.008	0.366	3
202	PF3D7_1114500	conserved Plasmodium protein, unknown function	0.028	0.087	-0.826	-0.129	-0.054	0.490	3
203	PF3D7_1349900	conserved Plasmodium protein, unknown function	0.018	0.142	-0.792	-0.126	-0.044	0.494	3
204	PF3D7_0801600	conserved Plasmodium protein, unknown function	-0.144	-0.039	-1.161	-0.198	-0.120	0.774	3
205	PF3D7_1320200	TBC domain protein, putative	-0.099	0.001	-0.863	-0.125	-0.094	0.602	3
206	PF3D7_1121200	conserved Plasmodium protein, unknown function	0.041	-0.059	-0.845	-0.145	-0.105	0.646	3
207	PF3D7_1233800	conserved Plasmodium protein, unknown function	0.063	-0.239	-0.825	-0.264	0.164	0.686	3
208	PF3D7_1429600	conserved Plasmodium protein, unknown function (GEXP03)	0.042	-0.088	-0.758	-0.303	0.149	0.685	3
209	PF3D7_1432700	protein-L-isoaspartate(D-aspartate) O-methyltransferase, putative	-0.008	-0.019	-0.766	-0.270	0.122	0.651	3
210	PF3D7_0504900	conserved protein, unknown function	0.458	0.272	-0.565	-0.449	0.152	1.033	3
211	PF3D7_1121000	palmitoyltransferase, putative (DHHC3)	0.737	0.444	-1.117	-0.681	0.214	1.052	3
212	PF3D7_0904400	signal peptidase complex subunit 3, putative (SPC3)	-0.019	0.147	-0.355	-0.440	-0.030	0.735	3
213	PF3D7_0909600	conserved Plasmodium protein, unknown function	0.145	0.022	-0.601	-0.368	-0.112	0.727	3
214	PF3D7_0904700	bacterial histone-like protein (HU)	1.284	0.294	-1.037	-0.758	0.187	0.982	3
215	PF3D7_1325600	mitochondrial fission 1 protein, putative (FIS1)	0.946	0.473	-0.578	-0.312	0.130	0.378	3
216	PF3D7_0107000	centrin-1 (CEN1)	0.600	0.112	-0.798	-0.502	0.214	0.444	3
217	PF3D7_1137500	apicoplast ribosomal protein S14p or S29e precursor, putative	0.455	0.253	-0.696	-0.485	0.163	0.255	3
218	PF3D7_0915700	conserved Plasmodium protein, unknown function	0.512	0.363	-0.721	-0.331	0.254	0.363	3
219	PF3D7_0503400	actin-depolymerizing factor 1 (ADF1)	0.747	0.763	-1.015	-0.184	-0.021	0.027	3
220	PF3D7_1336400	RanBPM and CLTH-like protein, putative	0.149	0.584	-0.885	-0.292	-0.022	0.151	3
221	PF3D7_1008400	26S protease regulatory subunit 4, putative (RPT2)	0.372	0.328	-0.679	-0.200	0.085	-0.093	3
222	PF3D7_1251300	thymidylate kinase (TMK)	0.296	0.148	-0.698	-0.148	0.059	-0.070	3
223	PF3D7_1410700	conserved Plasmodium protein, unknown function	0.828	0.334	-1.115	-0.707	-0.080	0.293	3
224	PF3D7_1471400	diacylglycerol kinase, putative	0.486	0.006	-0.645	-0.378	-0.239	0.164	3
225	PF3D7_0106000	conserved Plasmodium protein, unknown function	0.344	-0.134	-0.588	-0.472	-0.094	0.203	3
226	PF3D7_1471500	conserved Plasmodium membrane protein, unknown function	0.300	0.284	-0.628	-0.674	-0.278	0.296	3
227	PF3D7_1012300	ubiquinol-cytochrome c reductase complex subunit, putative	0.275	0.145	-0.640	-0.533	-0.266	0.216	3
228	PF3D7_1460900	conserved Plasmodium protein, unknown function	0.430	0.241	-0.753	-0.698	-0.102	0.361	3
229	PF3D7_0419300	conserved Plasmodium protein, unknown function	0.255	0.133	-0.597	-0.327	-0.029	0.140	3
230	PF3D7_0625200	conserved Plasmodium protein, unknown function	0.222	0.114	-0.719	-0.416	-0.099	0.213	3
231	PF3D7_1028600	conserved Plasmodium protein, unknown function	0.356	0.102	-1.030	-0.777	0.026	0.294	3
232	PF3D7_0821000	conserved Plasmodium protein, unknown function	0.491	0.146	-1.113	-0.359	-0.152	0.270	3
233	PF3D7_0828500	translation initiation factor EIF-2b alpha subunit, putative	0.612	0.111	-1.285	-0.323	-0.073	0.438	3

234	PF3D7_1215200	peptidyl-prolyl cis-trans isomerase (CYP32)	0.318	0.193	-0.721	-0.275	-0.027	0.426	3
235	PF3D7_1440100	conserved Plasmodium protein, unknown function	0.397	0.303	-0.842	-0.322	-0.005	0.362	3
236	PF3D7_0522500	50S ribosomal protein L17, apicoplast, putative (RPL17)	0.416	0.173	-1.054	-0.329	0.235	0.445	3
237	PF3D7_1141400	phosphatidylinositol N-acetylglucosaminyltransferase subunit H, putative (PIGH)	0.364	-0.058	-1.043	-0.357	0.210	0.285	3
238	PF3D7_1456000	transcription factor with AP2 domain(s) (ApiAP2)	0.323	-0.058	-0.921	-0.265	0.114	0.273	3
239	PF3D7_1360900	polyadenylate-binding protein, putative	0.295	0.051	-0.770	-0.228	0.080	0.130	3
240	PF3D7_0110000	conserved Plasmodium protein, unknown function	0.205	0.048	-0.695	-0.272	-0.037	0.154	3
241	PF3D7_1304000	chromosome condensation protein, putative	0.224	-0.001	-0.850	-0.406	-0.150	0.138	3
242	PF3D7_1022800	4-hydroxy-3-methylbut-2-en-1-yl diphosphate synthase (GcpE)	0.070	0.088	-0.673	-0.240	-0.008	0.158	3
243	PF3D7_1242500	conserved protein, unknown function	0.062	0.028	-0.808	-0.259	-0.065	0.183	3
244	PF3D7_0613400	apicoplast ribosomal protein L18 precursor, putative	0.303	-0.007	-0.932	-0.459	-0.099	0.354	3
245	PF3D7_1319400	conserved Plasmodium protein, unknown function	0.164	0.051	-0.700	-0.300	-0.058	0.249	3
246	PF3D7_1117100	deubiquinating or deneddylating enzyme (UCH54)	0.340	-0.021	-1.079	-0.388	-0.061	0.559	3
247	PF3D7_0923600	lipoate-protein ligase 2 (LipL2)	0.201	0.057	-0.797	-0.227	-0.235	0.361	3
248	PF3D7_1430900	cytochrome c oxidase subunit 2, putative	0.365	0.136	-1.415	-0.437	-0.282	0.425	3
249	PF3D7_0307400	ATP-dependent Clp protease proteolytic subunit (ClpP)	0.254	0.110	-1.067	-0.295	-0.142	0.579	3
250	PF3D7_1451300	E3 SUMO-protein ligase NSE2, putative (NSE2)	0.204	0.081	-0.821	-0.201	-0.097	0.438	3
251	PF3D7_0307500	conserved protein, unknown function	0.129	0.095	-0.708	-0.192	-0.141	0.360	3
252	PF3D7_1364700	conserved Plasmodium protein, unknown function, pseudogene	0.284	0.086	-1.552	-0.381	-0.360	0.816	3
253	PF3D7_1008200	endonuclease, putative	0.239	-0.073	-0.665	-0.408	-0.221	0.359	3
254	PF3D7_1470800	conserved Plasmodium protein, unknown function	0.187	-0.110	-0.682	-0.364	-0.078	0.402	3
255	PF3D7_0316300	conserved Plasmodium protein, unknown function	0.306	-0.088	-0.701	-0.339	0.196	0.374	3
256	PF3D7_1462700	cytochrome c1 precursor, putative	0.301	-0.038	-0.657	-0.410	0.103	0.454	3
257	PF3D7_0911200	conserved Plasmodium protein, unknown function	0.311	0.027	-0.785	-0.470	-0.043	0.449	3
258	PF3D7_1036800	acetyl-CoA transporter, putative	0.295	-0.027	-0.639	-0.423	-0.006	0.442	3
259	PF3D7_1334000	conserved Plasmodium protein, unknown function	0.272	0.137	-0.768	-0.513	0.004	0.745	3
260	PF3D7_1422500	ERAD-associated E3 ubiquitin-protein ligase HRD1 (HRD1)	0.334	0.133	-0.848	-0.549	-0.047	0.630	3
261	PF3D7_1033200	early transcribed membrane protein 10.2 (ETRAMP10.2)	0.896	-0.650	-0.154	0.066	0.398	1.118	3
262	PF3D7_1207500	RNA-binding protein, putative	1.132	0.513	-0.756	-0.190	0.423	1.116	3
263	PF3D7_1415300	RNA-binding protein Nova-1, putative	1.012	0.536	-0.643	0.234	0.366	0.784	3
264	PF3D7_1476300	Plasmodium exported protein (PHISTb), unknown function	0.461	0.007	0.111	0.060	0.381	0.822	3
265	PF3D7_1302100	gamete antigen 27 or 25 (G27 or 25)	0.159	0.143	-0.540	0.653	0.991	1.004	3
266	PF3D7_0729200	1-cys peroxiredoxin (AOP)	-0.348	-0.094	-0.776	-0.142	-0.289	0.268	3
267	PF3D7_1210900	GPI mannosyltransferase I (PIGM)	-0.404	-0.044	-0.740	-0.266	-0.268	0.603	3
268	PF3D7_1404300	secreted ookinete adhesive protein, putative (SOAP)	-0.302	0.052	-0.571	-0.597	-0.010	0.205	3
269	PF3D7_0818600	BEM46-like protein, putative (PBLP)	-0.333	0.188	-0.567	-0.274	-0.786	-0.225	3
270	PF3D7_0308500	activator of Hsp90 ATPase, putative	-0.142	0.115	-0.714	0.314	0.510	-0.130	3
271	PF3D7_0110700	chromatin assembly factor 1 protein WD40 domain, putative	0.070	-0.083	-0.617	0.143	0.344	-0.057	3
272	PF3D7_0613000	conserved Plasmodium protein, unknown function	-0.011	0.062	-0.813	0.086	0.451	0.171	3
273	PF3D7_1410500	conserved Plasmodium protein, unknown function	0.005	0.094	-0.723	0.052	0.174	-0.102	3

274	PF3D7_0521200	conserved Plasmodium protein, unknown function	-0.233	-0.136	-0.739	-0.092	0.174	0.195	3
275	PF3D7_0720800	Ham1-like protein, putative	-0.224	-0.049	-0.606	-0.047	0.119	0.236	3
276	PF3D7_1334500	MSP7-like protein (MSRP6)	-0.134	-0.207	-0.676	0.094	0.306	0.210	3
277	PF3D7_1205600	conserved Plasmodium protein, unknown function	-0.157	-0.227	-0.988	0.148	0.358	0.244	3
278	PF3D7_1310100	conserved Plasmodium protein, unknown function	-0.125	-0.187	-0.667	0.020	0.266	0.040	3
279	PF3D7_0824500	conserved Plasmodium protein, unknown function	-0.243	-0.126	-0.731	0.089	0.265	0.404	3
280	PF3D7_1309800	conserved Plasmodium protein, unknown function	-0.222	-0.141	-1.080	0.111	0.308	0.534	3
281	PF3D7_1360000	conserved Plasmodium protein, unknown function	-0.099	-0.115	-1.015	0.103	0.398	0.528	3
282	PF3D7_1012100	conserved Plasmodium protein, unknown function	-0.016	-0.184	-0.599	0.011	-0.183	-0.009	3
283	PF3D7_0724800	kelch protein, putative	-0.039	-0.120	-0.750	-0.295	-0.305	-0.150	3
284	PF3D7_1454100	tRNA intron endonuclease, putative	-0.033	-0.392	-0.589	-0.277	0.073	0.006	3
285	PF3D7_1021000	conserved Plasmodium protein, unknown function	-0.135	-0.294	-1.095	-0.198	-0.045	0.281	3
286	PF3D7_0824200	conserved Plasmodium protein, unknown function	-0.083	-0.246	-0.726	-0.096	-0.024	0.164	3
287	PF3D7_1314300	conserved Plasmodium protein, unknown function	-0.218	-0.357	-0.715	-0.092	0.043	0.270	3
288	PF3D7_0503200	conserved Plasmodium protein, unknown function	-0.048	-0.399	-0.690	-0.128	-0.012	0.286	3
289	PF3D7_0111200	conserved Plasmodium protein, unknown function	-0.111	-0.371	-0.810	-0.107	-0.012	0.421	3
290	PF3D7_1411600	GTP-binding protein, putative	-0.154	-0.285	-0.708	-0.115	-0.024	0.290	3
291	PF3D7_1430600	exodeoxyribonuclease III, putative	-0.132	-0.447	-0.655	-0.282	-0.114	0.184	3
292	PF3D7_0108100	conserved Plasmodium protein, unknown function	-0.035	-0.211	-0.594	-0.176	-0.030	0.531	3
293	PF3D7_0709400	Cg7 protein	-0.011	-0.399	-0.641	-0.217	-0.012	0.550	3
294	PF3D7_0934400	transcription factor with AP2 domain(s), putative (ApiAP2)	0.227	-0.341	-0.706	-0.141	-0.133	0.533	3
295	PF3D7_1129500	A or G-specific adenine glycosylase, putative	0.021	-0.296	-0.792	-0.051	-0.040	0.525	3
296	PF3D7_0211300	ubiquinol-cytochrome-c reductase complex assembly factor 1, putative	0.165	-0.623	-1.341	-0.212	-0.158	0.640	3
297	PF3D7_0303700	lipoamide acyltransferase component of branched-chain alpha-keto acid dehydrogenase complex (BCKDH-E2)	0.079	-0.225	-0.772	-0.108	-0.057	0.384	3
298	PF3D7_0618600	rhomboid protease ROM10 (ROM10)	0.119	-0.309	-1.115	-0.154	-0.152	0.543	3
299	PF3D7_0813500	conserved Plasmodium protein, unknown function	0.097	-0.310	-1.021	-0.158	-0.095	0.434	3
300	PF3D7_1433600	microsomal signal peptidase protein, putative (SPC1)	0.151	-0.466	-1.296	-0.239	-0.214	0.655	3
301	PF3D7_1141600	dolichol-phosphate mannosyltransferase (DPM1)	0.029	-0.447	-1.450	-0.169	-0.409	0.978	3
302	PF3D7_1318100	ferredoxin, putative	-0.094	-0.357	-1.212	-0.326	-0.223	0.798	3
303	PF3D7_0515200	conserved Plasmodium protein, unknown function	0.186	-0.550	-1.392	-0.488	-0.335	0.574	3
304	PF3D7_1403100	conserved Plasmodium protein, unknown function	0.031	-0.191	-0.665	-0.202	-0.063	0.200	3
305	PF3D7_0703200	conserved Plasmodium protein, unknown function	0.153	-0.165	-0.679	-0.291	0.000	0.206	3
306	PF3D7_0914800	GIN5 complex subunit Psf3, putative	0.077	-0.226	-0.754	-0.300	-0.168	0.033	3
307	PF3D7_0111300	replication factor c protein, putative	0.098	-0.161	-0.792	-0.520	-0.260	0.370	3
308	PF3D7_0624700	N-acetylglucosaminylphosphatidylinositol deacetylase, putative	0.110	-0.236	-1.315	-0.706	-0.183	0.464	3
309	PF3D7_0212000	GDP-fructose--GMP antiporter, putative	-0.305	-0.206	-0.906	-0.619	-0.189	0.557	3
310	PF3D7_1201700	conserved Plasmodium membrane protein, unknown function	-0.030	-0.207	-0.681	-0.397	-0.081	0.442	3
311	PF3D7_0605200	conserved Plasmodium protein, unknown function	-0.108	-0.327	-0.984	-0.366	-0.139	0.429	3
312	PF3D7_1463300	DNA polymerase alpha subunit, putative	-0.154	0.082	-0.597	-0.052	-0.018	0.234	3
313	PF3D7_0204000	conserved Plasmodium protein, unknown function	-0.035	-0.031	-0.723	0.006	0.018	0.287	3

314	PF3D7_0716600	cysteine desulfurase (SufS)	-0.185	-0.055	-0.970	-0.055	-0.018	0.377	3
315	PF3D7_0715600	GTP-binding translation elongation factor tu family protein, putative	-0.112	-0.031	-1.048	-0.133	-0.126	0.420	3
316	PF3D7_1242300	conserved Plasmodium membrane protein, unknown function	-0.074	-0.050	-1.342	-0.159	-0.066	0.594	3
317	PF3D7_1234300	DNA polymerase epsilon subunit b, putative	-0.042	0.059	-1.039	-0.130	-0.026	0.355	3
318	PF3D7_1024600	RAP protein, putative	-0.106	-0.001	-0.609	-0.141	-0.017	0.097	3
319	PF3D7_1025700	conserved Plasmodium protein, unknown function	-0.016	0.002	-0.961	-0.128	-0.077	0.176	3
320	PF3D7_0107700	conserved Plasmodium membrane protein, unknown function	-0.231	-0.362	-1.439	-0.153	-0.253	0.455	3
321	PF3D7_1135600	conserved Plasmodium protein, unknown function	-0.138	-0.214	-0.908	-0.096	-0.150	0.341	3
322	PF3D7_0825400	conserved Plasmodium protein, unknown function	-0.294	-0.218	-0.960	-0.069	-0.022	0.410	3
323	PF3D7_1245400	mitochondrial ribosomal protein L3 precursor, putative	-0.261	-0.261	-0.960	-0.034	-0.028	0.465	3
324	PF3D7_0626900	mitochondrial ribosomal protein L46 precursor, putative	-0.023	0.107	-0.758	-0.113	-0.272	0.241	3
325	PF3D7_0603100	conserved Plasmodium protein, unknown function	-0.016	0.042	-0.757	-0.193	-0.217	0.475	3
326	PF3D7_0320600	conserved Plasmodium protein, unknown function	-0.026	0.098	-1.302	-0.307	-0.438	0.608	3
327	PF3D7_0306000	conserved Plasmodium protein, unknown function	-0.065	-0.152	-0.936	-0.453	-0.306	0.305	3
328	PF3D7_1249400	conserved Plasmodium protein, unknown function	-0.080	-0.194	-0.994	-0.432	-0.201	0.303	3
329	PF3D7_1461100	conserved Plasmodium protein, unknown function	0.013	-0.076	-1.006	-0.293	-0.380	0.309	3
330	PF3D7_1112200	ubiquinone biosynthesis protein COQ4, putative (COQ4)	-0.069	-0.188	-1.168	-0.228	-0.323	0.362	3
331	PF3D7_1227300	conserved Plasmodium protein, unknown function	-0.234	-0.199	-0.085	-0.586	-0.027	0.143	3
332	PF3D7_0102400	lysophospholipase, putative, pseudogene	-0.290	0.037	-0.834	-1.312	-1.505	-0.634	4
333	PF3D7_1103800	CCR4-NOT transcription complex subunit 1, putative (NOT1)	-0.223	-0.047	-0.597	-1.093	-0.922	-0.855	4
334	PF3D7_0303200	HAD superfamily protein, putative	0.610	0.268	-0.187	-0.403	-0.009	-0.174	4
335	PF3D7_1116100	serine esterase, putative	0.636	0.084	-0.055	-0.421	0.028	-0.302	4
336	PF3D7_0302500	cytoadherence linked asexual protein 3.1 (CLAG3.1)	0.680	0.380	0.284	-0.571	-0.230	-0.285	4
337	PF3D7_0501500	rho-try-associated protein 3 (RAP3)	0.680	0.353	0.356	-0.640	-0.304	-0.283	4
338	PF3D7_0905400	high molecular weight rho-try protein 3 (RhopH3)	0.805	0.350	0.342	-0.772	-0.124	-0.230	4
339	PF3D7_0528400	palmitoyltransferase (DHHC7)	0.752	0.483	0.095	-0.576	-0.377	-0.213	4
340	PF3D7_0817700	rho-try neck protein 5 (RON5)	0.761	0.469	0.174	-0.687	-0.431	-0.149	4
341	PF3D7_1007200	rhoGAP GTPase, putative	0.649	0.539	0.199	-0.739	-0.361	-0.190	4
342	PF3D7_0515700	glideosome-associated protein 40, putative (GAP40)	0.784	0.270	0.069	-0.346	-0.135	-0.173	4
343	PF3D7_1222700	glideosome-associated protein 45 (GAP45)	1.140	0.506	0.171	-0.321	-0.313	-0.358	4
344	PF3D7_0707300	rho-try-associated membrane antigen (RAMA)	0.886	0.287	0.274	-0.473	-0.355	-0.155	4
345	PF3D7_1116000	rho-try neck protein 4 (RON4)	0.676	0.254	0.157	-0.418	-0.280	-0.149	4
346	PF3D7_1035600	merozoite surface protein (H101)	0.739	0.234	0.241	-0.389	-0.154	-0.143	4
347	PF3D7_1410400	rho-try-associated protein 1 (RAP1)	1.100	0.491	0.196	-0.742	-0.365	-0.332	4
348	PF3D7_0930300	merozoite surface protein 1 (MSP1)	1.401	0.700	0.647	-0.944	-0.017	-0.244	4
349	PF3D7_1035700	duffy binding-like merozoite surface protein (DBLMSP)	1.527	0.986	0.801	-0.675	0.125	-0.008	4
350	PF3D7_0829400	prolyl 4-hydroxylase subunit alpha, putative	0.830	0.469	0.241	-0.760	0.134	0.148	4
351	PF3D7_0423500	glideosome associated protein with multiple membrane spans 2 (GAPM2)	0.872	0.509	0.248	-0.819	-0.177	0.076	4
352	PF3D7_1335100	merozoite surface protein 7 (MSP7)	1.069	0.904	0.429	-0.908	-0.062	-0.016	4
353	PF3D7_1320700	conserved Plasmodium protein, unknown function	0.495	0.149	0.191	-0.590	-0.030	0.179	4
354	PF3D7_1437300	conserved Plasmodium protein, unknown function	0.750	0.312	0.402	-0.773	-0.335	0.139	4

355	PF3D7_0932200	profilin, putative (PFN)	0.584	0.882	-0.768	-0.591	-0.348	-0.220	4
356	PF3D7_0214900	rhostry neck protein 6 (RON6)	0.623	0.304	-0.228	-0.622	-0.441	-0.074	4
357	PF3D7_0810300	protein phosphatase PPM5, putative (PPM5)	0.744	0.368	-0.075	-0.625	-0.331	-0.169	4
358	PF3D7_1409400	conserved Plasmodium membrane protein, unknown function	0.749	0.396	-0.077	-0.509	-0.353	-0.204	4
359	PF3D7_1246200	actin I (ACT1)	1.031	0.643	-0.306	-0.352	-0.265	-0.206	4
360	PF3D7_0617800	histone H2A (H2A)	1.108	0.708	-0.582	-0.306	0.261	-0.166	4
361	PF3D7_0610400	histone H3 (H3)	1.086	0.530	-0.586	-0.545	0.042	-0.044	4
362	PF3D7_1112100	conserved Plasmodium protein, unknown function	0.720	0.363	-0.417	-0.169	-0.020	-0.041	4
363	PF3D7_1247500	serine or threonine protein kinase, putative	0.619	0.185	-0.293	-0.245	-0.096	-0.025	4
364	PF3D7_1327400	conserved Plasmodium protein, unknown function	0.976	0.224	-0.287	-0.457	-0.051	0.012	4
365	PF3D7_1414400	serine or threonine protein phosphatase PP1 (PP1)	0.939	0.191	-0.201	-0.305	0.186	-0.072	4
366	PF3D7_1001600	alpha or beta hydrolase, putative	0.921	0.457	-0.255	-0.671	0.227	0.145	4
367	PF3D7_1356100	conserved Plasmodium protein, unknown function	0.731	0.330	-0.133	-0.473	0.082	0.045	4
368	PF3D7_1439600	conserved Plasmodium protein, unknown function	0.617	0.470	-0.064	-0.148	0.115	0.079	4
369	PF3D7_0404700	dipeptidyl aminopeptidase 3 (DPAP3)	0.461	0.190	0.247	-0.874	-0.703	-0.175	4
370	PF3D7_1436600	cGMP-dependent protein kinase (PKG)	0.279	0.146	0.007	-0.635	-0.440	-0.104	4
371	PF3D7_1136200	conserved Plasmodium protein, unknown function	0.193	0.230	0.161	-0.677	-0.589	-0.065	4
372	PF3D7_1232500	CG2-related protein, putative	0.329	0.240	0.142	-1.384	-0.557	-0.247	4
373	PF3D7_0501600	rhostry-associated protein 2 (RAP2)	0.633	0.592	0.400	-1.037	-0.788	-0.211	4
374	PF3D7_1140400	conserved Plasmodium protein, unknown function	0.728	0.806	0.345	-1.498	-0.681	-0.088	4
375	PF3D7_0507200	subtilisin-like protease 3 (SUB3)	0.270	0.213	0.204	-0.625	-0.279	-0.062	4
376	PF3D7_0105300	cyclase-associated protein (CAP)	0.284	0.530	0.089	-0.784	-0.618	-0.110	4
377	PF3D7_1145200	serine or threonine protein kinase, putative	0.533	0.179	0.282	-0.629	-0.569	-0.192	4
378	PF3D7_0722200	rhostry-associated leucine zipper-like protein 1 (RALP1)	0.806	0.244	0.294	-0.594	-0.618	-0.202	4
379	PF3D7_1461400	conserved Plasmodium protein, unknown function	1.025	0.625	0.485	-0.979	-0.787	-0.397	4
380	PF3D7_0810900	conserved Plasmodium protein, unknown function	0.411	0.076	-0.111	-0.389	-0.635	-0.156	4
381	PF3D7_1017100	rhostry neck protein 12 (RON12)	0.636	-0.372	0.326	-0.422	-0.279	-0.363	4
382	PF3D7_1223100	cAMP-dependent protein kinase regulatory subunit (PKAr)	0.729	0.090	0.176	-0.514	-0.183	-0.424	4
383	PF3D7_1452000	rhostry neck protein 2 (RON2)	0.697	0.123	0.101	-0.506	-0.516	-0.345	4
384	PF3D7_0518200	SWIB or MDM2 domain-containing protein, putative	0.763	0.323	-0.022	-0.925	0.039	0.235	4
385	PF3D7_1235100	conserved Plasmodium protein, unknown function	0.944	0.522	-0.060	-1.005	0.075	0.508	4
386	PF3D7_0504600	2-oxoisovalerate dehydrogenase subunit beta, mitochondrial, putative (BCKDHB)	0.587	0.246	-0.077	-1.077	-0.086	0.421	4
387	PF3D7_1315300	conserved Plasmodium protein, unknown function	0.714	0.316	0.052	-0.957	-0.072	0.367	4
388	PF3D7_0207400	serine repeat antigen 7 (SERA7)	0.680	0.187	0.027	-1.168	0.278	0.345	4
389	PF3D7_0207700	serine repeat antigen 4 (SERA4)	0.948	0.335	0.002	-1.542	0.240	0.476	4
390	PF3D7_0613800	transcription factor with AP2 domain(s) (ApiAP2)	0.566	0.189	0.021	-0.789	0.177	0.116	4
391	PF3D7_0207600	serine repeat antigen 5 (SERA5)	1.102	0.251	0.285	-1.882	0.304	0.421	4
392	PF3D7_0607700	conserved Plasmodium protein, unknown function	0.600	0.184	0.231	-1.084	0.055	0.298	4
393	PF3D7_0604100	transcription factor with AP2 domain(s) (SIP2)	0.615	0.204	0.298	-1.396	0.093	0.251	4
394	PF3D7_0611600	conserved Plasmodium protein, unknown function	0.475	0.150	0.207	-1.032	-0.028	0.133	4
395	PF3D7_1143200	DnaJ protein, putative	0.491	0.007	0.076	-1.075	0.179	0.239	4

396	PF3D7_1122700	conserved Plasmodium protein, unknown function	0.495	0.040	0.081	-1.057	0.053	0.331	4
397	PF3D7_0511200	stearoyl-CoA desaturase (SCD)	0.349	0.288	0.093	-0.621	0.194	-0.096	4
398	PF3D7_0918000	glideosome-associated protein 50 (GAP50)	0.812	0.141	0.035	-0.678	0.114	0.005	4
399	PF3D7_1228600	merozoite surface protein 9 (MSP9)	1.488	0.060	0.358	-1.555	0.148	0.199	4
400	PF3D7_1252100	rhoptry neck protein 3 (RON3)	0.997	0.251	0.243	-0.962	-0.084	-0.101	4
401	PF3D7_0207500	serine repeat antigen 6 (SERA6)	0.891	0.037	0.230	-1.005	0.593	0.617	4
402	PF3D7_0204100	conserved Plasmodium protein, unknown function	0.471	-0.090	-0.026	-0.599	0.179	0.181	4
403	PF3D7_0903300	conserved Plasmodium membrane protein, unknown function	0.487	0.049	-0.099	-0.678	0.160	0.219	4
404	PF3D7_0618500	malate dehydrogenase (MDH)	0.650	0.165	-0.301	-0.744	0.211	0.595	4
405	PF3D7_1364100	6-cysteine protein (P92)	0.893	0.223	0.102	-1.184	0.169	0.812	4
406	PF3D7_0310400	parasite-infected erythrocyte surface protein (PIESP1)	0.704	0.094	-0.026	-1.099	0.329	1.021	4
407	PF3D7_0522600	inner membrane complex protein	0.367	0.127	0.099	-0.650	-0.205	0.043	4
408	PF3D7_1128900	conserved Plasmodium protein, unknown function	0.468	0.068	0.182	-0.838	-0.255	0.017	4
409	PF3D7_1019100	conserved Plasmodium protein, unknown function	0.549	0.013	0.059	-0.863	-0.248	0.164	4
410	PF3D7_0929400	high molecular weight rhoptry protein 2 (RhopH2)	0.802	0.332	0.012	-1.410	-0.198	0.002	4
411	PF3D7_1035100	probable protein, unknown function	0.525	0.185	0.156	-1.452	-0.367	0.107	4
412	PF3D7_1133300	LEM3 or CDC50 family protein, putative	0.408	0.250	0.226	-1.261	-0.250	0.188	4
413	PF3D7_0806200	conserved Plasmodium membrane protein, unknown function	0.292	0.163	0.147	-0.812	-0.101	0.087	4
414	PF3D7_0919900	regulator of chromosome condensation, putative	0.481	0.047	0.194	-0.832	0.020	0.076	4
415	PF3D7_1018200	serine or threonine protein phosphatase 8, putative (PPP8)	0.644	0.069	0.136	-1.175	-0.110	0.025	4
416	PF3D7_0508900	conserved Plasmodium protein, unknown function	0.401	0.162	0.173	-0.725	-0.034	0.025	4
417	PF3D7_1030800	calmodulin, putative	0.439	0.158	0.288	-0.850	-0.056	-0.012	4
418	PF3D7_1025900	conserved Plasmodium protein, unknown function	0.382	0.075	0.241	-0.758	-0.152	-0.084	4
419	PF3D7_0617300	conserved Plasmodium protein, unknown function	0.338	0.024	0.093	-0.726	-0.062	-0.218	4
420	PF3D7_0704500	serine or threonine protein kinase, putative	0.415	0.156	0.278	-0.915	-0.152	0.171	4
421	PF3D7_1138000	conserved Plasmodium protein, unknown function	0.280	-0.002	0.284	-0.687	-0.070	0.115	4
422	PF3D7_0705100	conserved Plasmodium protein, unknown function	0.126	0.023	0.221	-0.712	-0.119	0.171	4
423	PF3D7_0413500	phosphoglucosyltransferase-2 (PGM2)	0.299	-0.127	0.062	-0.660	-0.167	0.270	4
424	PF3D7_0612200	leucine-rich repeat protein (LRR6)	0.313	-0.053	-0.103	-0.873	-0.098	0.256	4
425	PF3D7_1344500	U6 snRNA phosphodiesterase, putative (USB1)	0.292	-0.100	-0.158	-1.025	-0.152	0.318	4
426	PF3D7_1440400	conserved Plasmodium protein, unknown function	0.303	0.008	-0.152	-0.757	-0.266	0.260	4
427	PF3D7_1224100	conserved Plasmodium protein, unknown function	0.418	-0.129	-0.301	-0.812	-0.106	0.312	4
428	PF3D7_0417700	conserved Plasmodium protein, unknown function	0.037	0.140	-0.013	-0.616	-0.321	0.177	4
429	PF3D7_0914400	conserved Plasmodium protein, unknown function	0.353	0.037	0.045	-1.342	-0.595	0.252	4
430	PF3D7_1033900	ubiquitin-conjugating enzyme, putative	0.220	0.117	0.099	-1.555	0.003	0.008	4
431	PF3D7_1222500	ATP-binding protein, putative	0.004	0.277	0.079	-1.277	0.058	0.237	4
432	PF3D7_1306000	conserved Plasmodium protein, unknown function	0.143	0.079	-0.142	-0.826	-0.023	0.147	4
433	PF3D7_0806300	ferlin, putative	0.048	0.053	0.008	-0.721	-0.106	0.235	4
434	PF3D7_0508100	SET domain protein, putative (SET9)	0.111	-0.233	0.352	-0.603	0.055	-0.102	4
435	PF3D7_1029900	conserved Plasmodium protein, unknown function	0.234	-0.195	0.169	-0.848	0.080	0.071	4
436	PF3D7_1121300	tyrosine kinase-like protein (TKL2)	0.125	-0.099	0.164	-0.768	0.120	0.104	4

437	PF3D7_0606800	probable protein, unknown function	-0.352	-0.216	-0.071	-0.938	-0.025	0.385	4
438	PF3D7_1319600	conserved Plasmodium protein, unknown function	-0.123	-0.073	-0.194	-0.662	0.011	0.266	4
439	PF3D7_1334600	MSP7-like protein (MSRP3)	-0.350	0.091	-0.058	-1.666	-0.078	0.575	4
440	PF3D7_1227000	conserved Plasmodium protein, unknown function	0.611	0.313	-0.446	-1.145	-0.029	0.260	4
441	PF3D7_1322500	palmitoyltransferase, putative (DHHC5)	0.387	0.224	-0.224	-0.954	-0.438	0.117	4
442	PF3D7_1451700	calcineurin subunit B (CNB)	0.299	0.431	-0.339	-0.957	-0.431	0.366	4
443	PF3D7_1028200	zinc finger, C3HC4 type, putative	0.166	0.019	-0.344	-0.618	-0.172	-0.104	4
444	PF3D7_1031800	conserved Plasmodium protein, unknown function	0.193	0.169	-0.361	-0.823	-0.213	-0.144	4
445	PF3D7_0106200	conserved Plasmodium protein, unknown function	0.185	-0.168	-0.157	-0.979	-0.140	0.439	4
446	PF3D7_1245500	conserved Plasmodium protein, unknown function	-0.034	-0.077	-0.390	-1.218	-0.329	0.556	4
447	PF3D7_0504500	conserved Plasmodium protein, unknown function	0.137	0.017	-0.424	-0.959	-0.746	0.110	4
448	PF3D7_0105400	conserved Plasmodium protein, unknown function	0.228	0.231	-0.325	-0.578	-0.678	-0.051	4
449	PF3D7_0729700	conserved Plasmodium protein, unknown function	0.169	0.203	-0.154	-0.626	-0.796	0.017	4
450	PF3D7_0916300	conserved Plasmodium protein, unknown function	-0.277	0.119	-0.094	-1.292	-0.568	0.351	4
451	PF3D7_1349100	nucleoside diphosphate hydrolase, putative	-0.169	0.009	-0.045	-0.897	-0.460	0.272	4
452	PF3D7_1238000	COPI associated protein, putative	-0.532	-0.268	-0.193	-1.662	-0.602	0.266	4
453	PF3D7_0404800	conserved Plasmodium protein, unknown function	-0.161	0.147	0.145	-0.948	-0.812	0.046	4
454	PF3D7_1430200	plasmepsin IX	-0.151	0.126	-0.171	-0.927	-0.838	-0.052	4
455	PF3D7_1449600	conserved Plasmodium protein, unknown function	-0.023	-0.090	-0.160	-0.884	-0.702	0.039	4
456	PF3D7_0918500	telomerase RNA (TR)	0.587	0.049	-0.579	-0.602	-0.028	0.078	4
457	PF3D7_1301600	erythrocyte binding antigen-140 (EBA140)	1.167	0.213	0.254	0.098	-0.400	0.616	4
458	PF3D7_0917900	heat shock protein 70 (HSP70-2)	0.578	0.431	0.369	-0.152	0.720	0.712	4
459	PF3D7_0927600	conserved Plasmodium protein, unknown function	0.731	0.354	-0.540	-0.127	0.081	-0.108	4
460	PF3D7_1124600	ethanolamine kinase (EK)	0.614	0.249	-0.528	-0.209	0.096	-0.120	4
461	PF3D7_0416500	repressor of RNA polymerase III transcription MAF1, putative	0.157	-0.037	-0.482	-0.632	-0.165	0.245	4
462	PF3D7_0717600	conserved Plasmodium protein, unknown function	0.155	-0.041	-0.527	-0.620	-0.252	0.220	4
463	PF3D7_1331600	protein tyrosine phosphatase-like protein, putative (PTPLA)	0.405	0.098	-0.807	-0.877	-0.309	0.171	4
464	PF3D7_1339700	conserved Plasmodium protein, unknown function	0.032	-0.059	-0.497	-0.656	-0.060	0.274	4
465	PF3D7_1348100	GTPase, putative	0.140	0.119	-0.643	-1.065	-0.568	0.816	4
466	PF3D7_0206900	conserved Plasmodium protein, unknown function	0.627	0.411	-0.728	0.078	-0.598	-0.458	5
467	PF3D7_1247700	conserved Plasmodium protein, unknown function	0.599	0.341	-0.222	0.011	-0.571	-0.431	5
468	PF3D7_1348600	conserved Plasmodium protein, unknown function	0.621	0.543	-0.434	-0.232	-0.520	-0.265	5
469	PF3D7_0321400	protein kinase, putative	0.358	0.124	0.108	-0.984	-1.166	-0.419	5
470	PF3D7_0217500	calcium-dependent protein kinase 1 (CDPK1)	0.675	0.444	0.342	-0.878	-0.949	-0.296	5
471	PF3D7_1423100	ankyrin, putative	0.585	0.455	0.241	-0.249	-0.735	-0.036	5
472	PF3D7_0613300	rhoptry protein ROP14 (ROP14)	0.347	0.187	0.394	-0.241	-0.655	-0.385	5
473	PF3D7_1035900	probable protein, unknown function (M566)	0.666	0.411	0.383	-0.430	-1.020	-0.540	5
474	PF3D7_1036000	merozoite surface protein (MSP11)	0.482	0.367	0.376	-0.615	-0.947	-0.587	5
475	PF3D7_0423300	conserved Plasmodium protein, unknown function	0.512	0.225	0.240	-0.505	-0.765	-0.260	5
476	PF3D7_0405900	apical sushi protein (ASP)	0.455	0.115	0.219	-0.309	-0.632	-0.143	5
477	PF3D7_1460600	inner membrane complex sub-compartment protein 3 (ISP3)	0.364	0.800	-0.122	-0.482	-0.795	-0.083	5

478	PF3D7_0930200	leucine-rich repeat protein (LRR8)	0.070	0.516	-0.020	-0.553	-1.080	-0.224	5
479	PF3D7_0817600	conserved Plasmodium protein, unknown function	0.478	0.314	0.005	-0.472	-0.843	-0.302	5
480	PF3D7_1361500	conserved Plasmodium protein, unknown function	0.293	0.180	-0.345	-0.307	-0.711	-0.435	5
481	PF3D7_1031200	MORN repeat-containing protein 1 (MORN1)	0.693	0.098	0.231	-0.207	-0.677	-0.400	5
482	PF3D7_1335300	reticulocyte binding protein 2 homologue b (RH2b)	0.706	0.008	0.253	0.133	-0.605	-0.348	5
483	PF3D7_1017500	conserved Plasmodium protein, unknown function	1.353	1.145	0.811	0.139	-1.479	-0.902	5
484	PF3D7_1136900	subtilisin-like protease 2 (SUB2)	0.696	0.378	0.288	0.116	-0.615	-0.628	5
485	PF3D7_1238900	protein kinase 2 (PK2)	0.888	0.548	0.440	0.192	-0.632	-0.756	5
486	PF3D7_1133400	apical membrane antigen 1 (AMA1)	0.964	0.758	0.523	-0.108	-0.576	-0.558	5
487	PF3D7_0102500	erythrocyte binding antigen-181 (EBA181)	1.078	0.489	0.279	0.063	-1.305	-0.369	5
488	PF3D7_1035500	merozoite surface protein 6 (MSP6)	1.082	0.537	0.228	0.102	-1.504	-0.373	5
489	PF3D7_0507500	subtilisin-like protease 1 (SUB1)	0.777	0.416	0.186	0.133	-0.916	-0.407	5
490	PF3D7_0725400	conserved Plasmodium protein, unknown function	1.294	0.725	0.270	0.108	-1.393	-0.530	5
491	PF3D7_0731500	erythrocyte binding antigen-175 (EBA175)	1.172	0.716	0.506	0.145	-1.423	-0.468	5
492	PF3D7_0613900	myosin E, putative (myoE)	0.724	0.226	0.284	0.017	-1.292	-0.449	5
493	PF3D7_0822900	conserved Plasmodium protein, unknown function	0.856	0.341	0.305	-0.083	-1.704	-0.597	5
494	PF3D7_1014100	conserved Plasmodium protein, unknown function	0.833	0.385	0.201	-0.036	-1.231	-0.476	5
495	PF3D7_1310700	RNA-binding protein, putative	0.852	0.523	0.488	-0.167	-1.572	-0.612	5
496	PF3D7_0620400	merozoite surface protein 10 (MSP10)	0.477	0.491	0.345	0.008	-0.873	-0.459	5
497	PF3D7_1468400	zinc finger protein, putative (D13)	0.769	0.736	0.456	-0.013	-1.381	-0.376	5
498	PF3D7_1252400	reticulocyte binding protein homologue 3, pseudogene (RH3)	0.853	0.521	0.264	0.135	-1.290	-0.593	5
499	PF3D7_1251200	coronin	0.831	0.568	0.340	0.008	-1.390	-0.743	5
500	PF3D7_1335400	reticulocyte binding protein 2 homologue a (RH2a)	0.768	0.488	0.299	0.097	-1.339	-0.640	5
501	PF3D7_1342600	myosin A (MyoA)	0.943	0.540	0.443	0.039	-1.331	-0.662	5
502	PF3D7_1013800	conserved Plasmodium protein, unknown function	0.590	0.397	0.082	0.109	-0.555	-0.376	5
503	PF3D7_0109100	LCCL domain-containing protein (CCp5)	0.790	0.500	0.099	-0.045	-0.765	-0.509	5
504	PF3D7_0323400	Rh5 interacting protein (RIPR)	0.774	0.400	0.346	-0.070	-1.004	-0.730	5
505	PF3D7_0503600	myosin B (MyoB)	0.568	0.310	0.187	-0.028	-0.764	-0.613	5
506	PF3D7_0109000	photosensitized INA-labeled protein PHIL1, putative (PHIL1)	1.107	0.666	0.325	-0.090	-1.318	-1.024	5
507	PF3D7_1332200	conserved Plasmodium protein, unknown function	0.851	0.739	0.099	-0.424	-1.101	-0.579	5
508	PF3D7_0920500	ADP-ribosylation factor, putative	0.516	0.217	-0.037	-0.212	-0.683	-0.393	5
509	PF3D7_1003600	inner membrane complex protein 1c, putative (IMC1c)	1.102	0.646	0.246	-0.500	-1.603	-0.933	5
510	PF3D7_0212600	secreted protein with altered thrombospondin repeat domain (SPATR)	0.846	0.661	0.336	-0.257	-1.636	-0.888	5
511	PF3D7_0919200	PPPDE peptidase, putative	0.572	0.281	0.161	-0.103	-1.230	-0.586	5
512	PF3D7_1143100	transcription factor with AP2 domain(s) (AP2-O)	0.491	0.344	0.321	-0.072	-1.220	-0.569	5
513	PF3D7_0507400	conserved Plasmodium protein, unknown function	0.343	0.306	0.156	-0.258	-0.935	-0.475	5
514	PF3D7_0308700	conserved Plasmodium protein, unknown function	0.335	0.211	-0.042	-0.140	-0.766	-0.449	5
515	PF3D7_1243700	ubiquitin-conjugating enzyme E2, putative	0.426	0.296	0.054	-0.139	-0.826	-0.662	5
516	PF3D7_1361400	actin-depolymerizing factor 2 (ADF2)	0.300	0.442	0.017	-0.140	-0.737	-0.437	5
517	PF3D7_0525300	conserved Plasmodium protein, unknown function	0.485	0.317	-0.078	0.059	-1.171	-0.635	5
518	PF3D7_1127900	conserved Plasmodium protein, unknown function	0.296	0.120	-0.006	0.095	-0.800	-0.528	5
519	PF3D7_0316000	microneme associated antigen (MA)	0.469	0.341	-0.193	-0.060	-1.028	-0.775	5

520	PF3D7_0629300	phospholipase, putative (PL)	0.174	0.092	-0.128	0.019	-0.610	-0.444	5
521	PF3D7_0902100	serine or threonine protein kinase, FIKK family (FIKK9.2)	0.381	0.284	0.162	0.051	-0.678	-0.761	5
522	PF3D7_1030200	conserved Plasmodium protein, unknown function	0.450	0.257	0.233	0.213	-0.725	-0.707	5
523	PF3D7_1411000	conserved Plasmodium protein, unknown function	0.600	0.243	0.146	0.107	-0.939	-0.763	5
524	PF3D7_1435600	conserved Plasmodium protein, unknown function	0.170	0.034	0.137	0.003	-0.592	-0.548	5
525	PF3D7_1243400	conserved Plasmodium protein, unknown function	0.234	0.160	0.201	0.160	-0.764	-0.679	5
526	PF3D7_0920700	conserved Plasmodium protein, unknown function	0.293	0.201	-0.028	-0.050	-0.558	-0.639	5
527	PF3D7_0102200	ring-infected erythrocyte surface antigen (RESA)	0.595	0.980	0.947	0.272	-1.113	-0.904	5
528	PF3D7_1149200	ring-infected erythrocyte surface antigen	0.542	0.724	0.823	0.485	-0.872	-0.862	5
529	PF3D7_1201200	Plasmodium exported protein (PHISTa-like), unknown function	0.104	0.289	0.601	0.271	-0.892	-0.468	5
530	PF3D7_0605900	long chain polyunsaturated fatty acid elongation enzyme, putative	-0.013	0.908	0.496	-0.448	-0.712	-0.583	5
531	PF3D7_0918700	conserved Plasmodium protein, unknown function	-0.226	0.286	0.353	-0.032	-0.724	-0.494	5
532	PF3D7_0805200	gamete release protein, putative (GAMER)	0.315	0.748	0.553	-0.128	-0.751	-0.708	5
533	PF3D7_1240100	early transcribed membrane protein 12 (ETRAPP12)	0.356	0.523	0.666	-0.029	-0.807	-0.630	5
534	PF3D7_1035300	glutamate-rich protein (GLURP)	0.429	0.536	0.505	-0.221	-1.220	-0.526	5
535	PF3D7_1028900	inner membrane complex protein 1m, putative (IMC1m)	0.402	0.405	0.578	-0.161	-1.057	-0.676	5
536	PF3D7_1302200	early transcribed membrane protein 13 (ETRAPP13)	0.326	0.260	0.413	0.053	-0.717	-0.591	5
537	PF3D7_1246400	myosin A tail domain interacting protein (MTIP)	0.207	0.323	0.393	0.011	-0.669	-0.617	5
538	PF3D7_1423300	serine or threonine protein phosphatase 7 (PP7)	0.349	-0.076	0.664	-0.274	-1.277	-0.329	5
539	PF3D7_0418600	regulator of chromosome condensation, putative	0.734	0.112	0.341	-0.080	-1.021	-0.389	5
540	PF3D7_0402300	reticulocyte binding protein homologue 1 (RH1)	0.663	0.099	0.270	0.040	-1.019	-0.590	5
541	PF3D7_0828800	GPI-anchored micronemal antigen (GAMA)	0.838	0.314	0.294	0.050	-1.213	-0.662	5
542	PF3D7_0424200	reticulocyte binding protein homologue 4 (RH4)	0.593	0.034	0.355	0.155	-0.987	-0.456	5
543	PF3D7_0818100	zinc finger protein, putative	0.790	0.348	0.373	0.236	-1.301	-0.569	5
544	PF3D7_1327300	conserved Plasmodium protein, unknown function	0.810	0.283	0.312	0.255	-1.188	-0.510	5
545	PF3D7_1337800	calcium-dependent protein kinase 5 (CDPK5)	0.657	0.554	0.507	0.211	-1.618	-0.679	5
546	PF3D7_0402200	surface-associated interspersed protein 4.1 (SURFIN 4.1), pseudogene (SURF4.1)	0.462	0.350	0.494	0.053	-1.065	-0.599	5
547	PF3D7_0507300	conserved Plasmodium protein, unknown function	0.333	0.246	0.227	0.015	-0.684	-0.434	5
548	PF3D7_0424100	reticulocyte binding protein homologue 5 (RH5)	0.533	0.219	0.274	0.065	-1.058	-0.569	5
549	PF3D7_1036500	probable protein, unknown function	0.587	0.230	0.287	0.130	-1.099	-0.576	5
550	PF3D7_1026600	conserved Plasmodium protein, unknown function	0.544	0.163	0.295	0.145	-1.226	-0.699	5
551	PF3D7_1125800	kelch protein, putative	0.541	0.202	0.312	0.230	-1.314	-0.374	5
552	PF3D7_0911100	conserved Plasmodium protein, unknown function	0.346	0.238	0.296	0.032	-1.378	-0.548	5
553	PF3D7_0934600	conserved Plasmodium protein, unknown function	0.374	0.231	0.224	0.090	-1.190	-0.503	5
554	PF3D7_0402100	Plasmodium exported protein (PHISTb), unknown function	0.263	0.122	0.324	0.220	-1.177	-0.431	5
555	PF3D7_0914100	conserved Plasmodium protein, unknown function	0.391	0.150	0.206	-0.163	-1.254	-0.550	5
556	PF3D7_1126700	autophagy-related protein 23, putative (ATG23)	0.426	0.180	0.295	-0.070	-1.754	-0.444	5
557	PF3D7_1204900	probable protein, unknown function	0.532	0.229	0.339	-0.152	-1.610	-0.515	5
558	PF3D7_0903600	conserved Plasmodium protein, unknown function	0.417	-0.004	0.142	-0.011	-1.196	-0.387	5
559	PF3D7_0214600	serine or threonine protein kinase, putative	0.623	0.240	0.462	0.040	-1.487	-0.481	5

560	PF3D7_0614700	conserved Plasmodium protein, unknown function	0.391	0.097	0.247	-0.015	-0.777	-0.310	5
561	PF3D7_1401600	Plasmodium exported protein (PHISTb), unknown function	0.298	0.075	0.287	-0.036	-0.758	-0.269	5
562	PF3D7_0515400	conserved Plasmodium protein, unknown function	0.473	0.020	-0.007	-0.128	-0.680	-0.495	5
563	PF3D7_1229800	myosin D (MyoD)	0.370	0.027	0.241	-0.060	-0.697	-0.443	5
564	PF3D7_1356800	serine or threonine protein kinase, putative (ARK3)	0.359	0.084	0.204	-0.092	-0.724	-0.492	5
565	PF3D7_1344300	conserved protein, unknown function	0.505	0.157	0.244	-0.108	-0.642	-0.488	5
566	PF3D7_0808200	plasmepsin X	0.628	0.213	0.380	-0.124	-0.844	-0.643	5
567	PF3D7_0618000	conserved Plasmodium membrane protein, unknown function	0.507	0.129	0.249	-0.230	-0.917	-0.657	5
568	PF3D7_1007300	RING zinc finger protein, putative	0.306	0.047	0.146	-0.116	-0.628	-0.566	5
569	PF3D7_1218000	thrombospondin-related apical membrane protein (TRAMP)	0.255	0.152	0.330	-0.252	-1.122	-0.752	5
570	PF3D7_1107800	transcription factor with AP2 domain(s) (ApiAP2)	0.349	-0.405	0.155	0.234	-0.797	-0.506	5
571	PF3D7_0203100	protein kinase, putative	0.168	-0.105	-0.029	0.000	-0.717	-0.505	5
572	PF3D7_0908500	conserved Plasmodium protein, unknown function	0.137	-0.023	0.120	-0.011	-0.606	-0.331	5
573	PF3D7_1108100	conserved Plasmodium protein, unknown function	-0.012	-0.163	0.040	-0.155	-0.951	-0.550	5
574	PF3D7_0210600	conserved Plasmodium protein, unknown function	0.680	0.698	0.270	-0.041	-1.534	-0.456	5
575	PF3D7_0414900	armadillo-domain containing rhoptry protein (ARO)	0.670	0.828	0.362	-0.019	-1.805	-0.624	5
576	PF3D7_1104900	calcium or calmodulin-dependent protein kinase, putative	0.433	0.402	0.118	0.017	-1.147	-0.458	5
577	PF3D7_1125700	kelch protein, putative	0.565	0.381	0.305	0.017	-1.448	-0.419	5
578	PF3D7_1351700	inner membrane complex protein 1f, putative (IMC1f)	0.619	0.507	0.347	-0.063	-1.552	-0.457	5
579	PF3D7_1455300	conserved Plasmodium protein, unknown function	0.592	0.843	0.416	0.044	-2.056	-0.584	5
580	PF3D7_0318700	trophozoite stage antigen	0.564	0.776	0.449	0.291	-2.257	-0.613	5
581	PF3D7_0423800	cysteine-rich protective antigen (CyRPA)	0.249	0.396	0.268	-0.198	-1.432	-0.492	5
582	PF3D7_0913800	conserved Plasmodium protein, unknown function	0.264	0.446	0.363	-0.106	-1.535	-0.555	5
583	PF3D7_0525800	inner membrane complex protein 1g, putative (IMC1g)	0.224	0.749	0.238	0.001	-1.852	-0.783	5
584	PF3D7_1307500	conserved Plasmodium protein, unknown function	0.140	0.419	0.259	0.019	-1.393	-0.548	5
585	PF3D7_0408000	conserved Plasmodium protein, unknown function	0.339	0.421	0.140	0.012	-1.454	-0.840	5
586	PF3D7_0404000	conserved Plasmodium protein, unknown function	0.211	0.301	0.144	-0.107	-1.000	-0.502	5
587	PF3D7_0308300	conserved Plasmodium protein, unknown function	0.597	0.729	0.063	-0.289	-1.973	-0.577	5
588	PF3D7_0423400	asparagine-rich protein (AARP)	0.272	0.426	-0.036	-0.053	-0.728	-0.165	5
589	PF3D7_0423900	probable protein, unknown function	-0.287	0.258	0.298	-0.076	-1.846	-0.418	5
590	PF3D7_1016800	Plasmodium exported protein (PHISTc), unknown function	0.027	0.298	-0.037	0.137	-1.455	-0.317	5
591	PF3D7_1304200	conserved Plasmodium protein, unknown function	0.152	-0.004	-0.136	0.071	-0.802	-0.340	5
592	PF3D7_1316200	ADP-ribosylation factor, putative	0.133	0.138	-0.463	-0.082	-1.138	-0.337	5
593	PF3D7_1310200	conserved Plasmodium protein, unknown function	-0.155	-0.023	-0.001	-0.280	-0.797	-0.318	5
594	PF3D7_0214700	conserved Plasmodium protein, unknown function	-0.026	0.120	0.293	-0.455	-0.980	-0.414	5
595	PF3D7_0210700	syntaxin, Qa-SNARE family (SYN17)	-0.248	0.100	-0.361	-0.132	-0.872	-0.766	5
596	PF3D7_0315900	conserved Plasmodium protein, unknown function	-0.122	0.406	-0.346	-0.022	-0.887	-0.585	5
597	PF3D7_0220400	DnaJ protein, putative, pseudogene	-0.132	0.062	-0.020	-0.007	-0.670	-0.564	5
598	PF3D7_1415100	conserved Plasmodium protein, unknown function	-0.002	0.094	-0.095	0.037	-0.618	-0.390	5
599	PF3D7_0919300	thioredoxin-like protein 1, putative (TrxL1)	-0.019	-0.039	0.124	-0.369	-0.700	-1.166	5
600	PF3D7_0104200	StAR-related lipid transfer protein	0.725	0.481	-0.355	0.104	-0.289	-0.321	5

601	PF3D7_0528300	conserved Plasmodium protein, unknown function	0.587	0.037	0.339	-0.077	-0.243	-0.293	5
602	PF3D7_1028700	merozoite TRAP-like protein (MTRAP)	0.696	0.052	0.387	0.047	-0.314	-0.458	5
603	PF3D7_1469200	shewanella-like protein phosphatase 1, putative (SHLP1)	0.607	0.219	0.426	-0.076	-0.210	-0.170	5
604	PF3D7_1146600	CCAAT-box DNA binding protein subunit B (NFYB)	0.648	0.129	0.019	-0.164	-0.253	-0.470	5
605	PF3D7_1012200	conserved Plasmodium protein, unknown function	0.704	0.118	0.144	-0.043	-0.420	-0.199	5
606	PF3D7_0805300	conserved Plasmodium protein, unknown function	0.559	0.579	0.344	0.250	-0.442	-0.593	5
607	PF3D7_0628300	choline or ethanolaminephosphotransferase, putative (CEPT)	0.818	0.739	0.077	-0.007	-0.178	-0.503	5
608	PF3D7_1416500	NADP-specific glutamate dehydrogenase (GDH1)	0.804	0.619	0.142	0.029	-0.098	-0.234	5
609	PF3D7_1001800	Plasmodium exported protein (PHISTc), unknown function	0.421	0.513	0.640	0.219	0.064	-0.371	5
610	PF3D7_1135100	protein phosphatase PPM8, putative (PPM8)	0.312	0.346	0.629	0.033	-0.040	-0.337	5
611	PF3D7_1001100	conserved Plasmodium protein, unknown function	0.283	0.617	0.329	-0.005	-0.213	-0.654	5
612	PF3D7_1002100	EMP1-trafficking protein (PTP5)	0.601	0.499	0.730	0.234	-0.378	-0.907	5
613	PF3D7_0606000	conserved Plasmodium protein, unknown function	0.288	0.291	0.330	0.252	-0.134	-0.601	5
614	PF3D7_1218500	conserved Plasmodium protein, unknown function	0.273	0.299	0.406	0.200	-0.094	-0.606	5
615	PF3D7_1407800	plasmepsin IV (PM4)	0.511	0.423	0.344	0.204	0.278	-0.621	5
616	PF3D7_0407900	AAA family ATPase, putative	0.456	0.204	0.030	0.078	-0.491	-0.675	5
617	PF3D7_0830500	tryptophan or threonine-rich antigen (TryThrA)	0.186	0.085	0.626	0.242	-0.701	-0.780	5
618	PF3D7_0100100	erythrocyte membrane protein 1, PfEMP1 (VAR)	-0.460	0.579	0.748	0.081	-0.093	-0.117	5
619	PF3D7_0712800	erythrocyte membrane protein 1, PfEMP1 (VAR)	-0.036	0.079	0.680	0.149	-0.006	-0.201	5
620	PF3D7_1254000	rifin (RIF)	0.001	0.311	0.856	0.152	0.258	-0.070	5
621	PF3D7_0732900	rifin (RIF)	0.048	-0.114	0.615	0.264	-0.323	-0.212	5
622	PF3D7_1301700	Plasmodium exported protein (hyp8), unknown function (GEXP07)	0.090	0.412	0.807	0.847	0.132	-0.512	5
623	PF3D7_1100400	rifin (RIF)	0.066	0.060	0.771	0.666	0.066	-0.165	5
624	PF3D7_1477200	Plasmodium exported protein (hyp15), unknown function	0.153	0.046	0.741	0.888	0.065	-0.165	5
625	PF3D7_1221900	conserved Plasmodium membrane protein, unknown function	0.272	0.064	0.502	0.753	-0.154	-0.338	5
626	PF3D7_0424700	serine or threonine protein kinase, FIKK family (FIKK4.2)	-0.120	-0.037	0.477	0.697	0.039	-0.452	5
627	PF3D7_1002000	Plasmodium exported protein (hyp2), unknown function	0.052	0.065	0.535	0.272	-0.030	-0.671	5
628	PF3D7_1401100	DnaJ protein, putative	0.334	0.088	0.777	0.311	-0.234	-1.144	5
629	PF3D7_0424600	Plasmodium exported protein (PHISTb), unknown function	0.541	0.831	0.257	0.741	-0.235	-0.889	5
630	PF3D7_0501100	conserved Plasmodium protein, unknown function	0.412	0.447	0.499	0.262	0.069	-0.678	5
631	PF3D7_1329500	conserved Plasmodium protein, unknown function	0.375	0.279	0.369	0.243	0.123	-0.605	5
632	PF3D7_1001500	early transcribed membrane protein 10.1 (ETRAP10)	0.481	0.790	0.601	0.644	0.133	-1.049	5
633	PF3D7_0501300	skeleton-binding protein 1 (SBP1)	0.504	0.598	0.555	0.645	0.001	-1.015	5
634	PF3D7_1370300	membrane associated histidine-rich protein (MAHRP1)	0.484	0.488	0.538	0.735	-0.003	-1.026	5
635	PF3D7_0113300	Plasmodium exported protein (hyp1), unknown function	0.670	0.331	0.606	0.521	-0.080	-0.702	5
636	PF3D7_1105600	translocon component PTEX88 (PTEX88)	0.758	0.370	0.591	0.439	-0.051	-0.651	5
637	PF3D7_0731100	EMP1-trafficking protein (PTP2)	0.468	0.281	0.536	0.572	0.091	-0.633	5
638	PF3D7_0501400	interspersed repeat antigen (FIRA)	0.377	0.098	0.630	0.515	-0.269	-0.633	5
639	PF3D7_1116800	heat shock protein 101 (HSP101)	0.345	0.158	0.469	0.561	-0.462	-0.730	5
640	PF3D7_1436300	translocon component PTEX150 (PTEX150)	0.535	0.293	0.569	0.614	-0.434	-0.630	5

641	PF3D7_0623800	tyrosine kinase-like protein, putative (TKL4)	0.415	0.198	0.264	0.432	-0.238	-0.675	5
642	PF3D7_1353100	Plasmodium exported protein, unknown function	0.545	0.254	0.389	0.567	-0.096	-0.672	5
643	PF3D7_0108500	conserved Plasmodium protein, unknown function	0.354	0.237	0.419	0.204	-0.427	-0.697	5
644	PF3D7_1115800	conserved Plasmodium protein, unknown function	0.775	0.514	0.453	0.247	-0.417	-0.850	5
645	PF3D7_0730300	transcription factor with AP2 domain(s) (AP2-L)	0.442	0.254	0.204	0.154	-0.211	-0.624	5
646	PF3D7_0104300	ubiquitin carboxyl-terminal hydrolase 1, putative (UBP1)	0.431	0.201	0.228	0.277	-0.369	-0.595	5
647	PF3D7_0420300	transcription factor with AP2 domain(s) (ApiAP2)	0.490	0.135	0.144	0.285	-0.330	-0.618	5
648	PF3D7_1405700	RING zinc finger protein, putative	0.337	0.118	0.178	0.143	-0.304	-0.620	5
649	PF3D7_0414800	conserved Plasmodium protein, unknown function	0.032	-0.083	0.234	0.272	-0.171	-0.619	5
650	PF3D7_0424500	serine or threonine protein kinase, FIKK family (FIKK4.1)	-0.019	0.072	0.293	0.399	-0.280	-0.615	5
651	PF3D7_1404800	conserved Plasmodium protein, unknown function	0.068	0.235	0.163	0.475	-0.291	-0.784	5
652	PF3D7_1353200	membrane associated histidine-rich protein (MAHRP2)	-0.129	-0.051	-0.019	0.775	-0.125	-1.523	5
653	PF3D7_1237900	conserved Plasmodium protein, unknown function	0.179	0.049	0.170	0.307	-0.126	-0.648	5
654	PF3D7_1372200	histidine-rich protein III (HRPIII)	0.189	0.133	0.033	0.589	-0.027	-1.067	5
655	PF3D7_0622900	transcription factor with AP2 domain(s), putative (ApiAP2)	-0.087	-0.253	0.032	-0.085	-0.513	-0.843	5
656	PF3D7_1231000	conserved Plasmodium protein, unknown function	0.063	-0.220	0.090	-0.085	-0.298	-0.675	5
657	PF3D7_1366100	conserved Plasmodium protein, unknown function	0.072	-0.072	0.043	0.137	-0.574	-0.875	5
658	PF3D7_1114200	GTPase-activating protein, putative	-0.124	0.141	0.042	0.231	-0.358	-0.744	5
659	PF3D7_1477800	acyl-CoA binding protein (ACBP)	-0.206	0.247	0.277	0.326	-0.480	-0.901	5
660	PF3D7_0102700	merozoite-associated tryptophan-rich antigen (MaTrA)	-0.220	0.316	0.128	0.073	-0.326	-0.621	5
661	PF3D7_0219700	Plasmodium exported protein (PHISTc), unknown function (GEXP20)	-0.518	0.088	0.204	0.362	-0.797	-0.834	5
662	PF3D7_1125900	conserved Plasmodium protein, unknown function	-0.257	-0.070	0.077	0.202	-0.363	-0.663	5
663	PF3D7_1446100	conserved Plasmodium protein, unknown function	-0.385	-0.033	-0.094	0.098	-0.159	-0.608	5
664	PF3D7_1001900	Plasmodium exported protein (hyp16), unknown function (PfJ23)	-0.181	-0.294	0.266	0.038	0.097	-1.005	5
665	PF3D7_0209500	conserved Plasmodium protein, unknown function	-0.054	0.014	-0.136	0.093	0.066	-0.673	5
666	PF3D7_1011500	conserved Plasmodium membrane protein, unknown function	0.102	0.136	-0.025	-0.045	0.132	-0.632	5
667	PF3D7_1302000	EMP1-trafficking protein (PTP6)	0.139	0.279	0.089	0.477	0.217	-0.794	5
668	PF3D7_0220700	Plasmodium exported protein (hyp9), unknown function	0.066	0.139	0.296	0.214	0.103	-0.739	5
669	PF3D7_0830600	Plasmodium exported protein (PHISTc), unknown function	0.030	0.080	0.290	0.425	0.117	-0.587	5
670	PF3D7_0831700	heat shock protein 70 (HSP70-x)	0.001	0.130	0.162	0.233	0.486	-0.702	5
671	PF3D7_0730900	EMP1-trafficking protein (PTP4)	0.135	-0.590	0.248	0.758	0.394	-0.660	5
672	PF3D7_0731400	serine or threonine protein kinase, FIKK family, pseudogene (FIKK7.2)	0.213	0.060	0.511	0.703	0.390	-0.290	5
673	PF3D7_0601900	conserved Plasmodium protein, unknown function	-0.001	-0.042	0.107	0.713	0.433	-0.101	5
674	PF3D7_0112900	Plasmodium exported protein, unknown function	0.991	-0.198	0.101	0.279	0.314	0.312	6
675	PF3D7_0909000	conserved Plasmodium protein, unknown function	0.474	0.435	0.636	0.010	0.236	0.217	6
676	PF3D7_0902600	serine or threonine protein kinase, FIKK family (FIKK9.7)	0.511	0.738	0.921	0.446	0.047	-0.089	6
677	PF3D7_0816900	adenylate kinase 2 (AK2)	0.320	0.487	0.724	0.333	0.205	0.106	6
678	PF3D7_1139100	RNA-binding protein, putative	0.173	0.471	0.657	0.302	0.100	0.004	6
679	PF3D7_0902500	serine or threonine protein kinase, FIKK family (FIKK9.6)	1.124	1.111	0.691	0.304	-0.060	0.040	6
680	PF3D7_1478800	Plasmodium exported protein, unknown function	0.420	0.664	0.414	-0.076	0.010	0.076	6

681	PF3D7_1477500	Plasmodium exported protein (PHISTb), unknown function	0.085	0.358	0.429	0.757	0.101	-0.108	6
682	PF3D7_1479000	acyl-CoA synthetase (ACS1a)	0.662	0.304	0.391	0.207	0.055	-0.109	6
683	PF3D7_1240900	erythrocyte membrane protein 1, PfEMP1 (VAR)	0.362	0.182	0.586	0.456	0.097	-0.065	6
684	PF3D7_0708400	heat shock protein 90 (HSP90)	0.485	0.218	0.243	0.310	0.712	-0.218	6
685	PF3D7_0202100	liver stage associated protein 2 (LSAP2)	1.920	1.746	1.577	1.230	0.651	1.274	6
686	PF3D7_0201700	DnaJ protein, putative	1.159	1.081	0.995	0.512	0.085	0.463	6
687	PF3D7_0420700	erythrocyte membrane protein 1, PfEMP1 (VAR)	0.954	0.917	1.015	0.588	-0.005	-0.050	6
688	PF3D7_1476000	conserved Plasmodium protein, unknown function	1.225	1.775	1.899	1.003	0.349	-0.085	6
689	PF3D7_0201800	knob associated heat shock protein 40 (KAHsp40)	1.149	0.666	0.645	1.036	0.933	0.185	6
690	PF3D7_1129000	spermidine synthase (SpdSyn)	0.803	0.288	0.717	0.652	0.569	0.011	6
691	PF3D7_0202000	knob-associated histidine-rich protein (KAHRP)	1.313	1.783	1.568	1.690	1.722	1.227	6
692	PF3D7_0201600	PHISTb domain-containing RESA-like protein 1 (RLP1)	0.557	0.329	0.303	0.995	0.799	0.361	6
693	PF3D7_1148700	Plasmodium exported protein (PHISTc), unknown function (GEXP12)	0.736	1.446	1.505	2.138	0.459	0.345	6
694	PF3D7_1340900	sodium-dependent phosphate transporter (PIT)	1.101	0.930	1.014	0.887	0.710	-0.023	6
695	PF3D7_1219400	erythrocyte membrane protein 1 (PfEMP1), pseudogene (VAR)	0.768	0.647	0.863	0.750	0.174	-0.010	6
696	PF3D7_1476400	serine or threonine protein kinase, FIKK family, pseudogene (FIKK14)	0.976	0.998	0.903	0.995	0.221	-0.052	6
697	PF3D7_0324400	rifin (RIF)	0.481	0.392	1.003	0.941	0.180	-0.021	6
698	PF3D7_1254700	rifin (RIF)	-0.211	-0.243	1.127	0.865	0.126	0.084	6
699	PF3D7_0532600	Plasmodium exported protein, unknown function	-0.077	-0.130	1.746	1.101	0.374	0.158	6
700	PF3D7_0632100	rifin (RIF)	0.112	-0.080	1.130	0.882	0.146	0.042	6
701	PF3D7_1000200	rifin (RIF)	-0.045	-0.126	1.076	0.766	0.400	0.262	6
702	PF3D7_1372600	rifin (RIF)	-0.058	0.109	1.511	1.193	0.419	0.411	6
703	PF3D7_0731300	Plasmodium exported protein (PHISTb), unknown function (PfG174)	0.250	0.419	1.082	1.221	0.730	0.430	6
704	PF3D7_1372500	stevor, pseudogene	0.054	0.191	1.376	1.998	1.243	0.909	6
705	PF3D7_0300900	stevor-like	0.108	-0.189	1.770	2.378	1.390	0.615	6
706	PF3D7_1476200	Plasmodium exported protein (PHISTb), unknown function	-0.079	0.293	1.498	1.934	0.697	0.123	6
707	PF3D7_1478600	EMP1-trafficking protein (PTP3)	0.180	-0.447	0.634	1.165	0.715	0.372	6
708	PF3D7_0202300	Plasmodium exported protein (hyp11), unknown function	0.164	0.148	0.302	1.042	0.540	0.122	6
709	PF3D7_0201900	erythrocyte membrane protein 3 (EMP3)	0.653	-0.451	0.410	1.174	1.469	1.076	6
710	PF3D7_1478100	Plasmodium exported protein (hyp13), unknown function	0.226	0.347	0.849	0.181	0.401	0.731	6
711	PF3D7_0322000	peptidyl-prolyl cis-trans isomerase (CYP19A)	0.138	0.079	-0.040	0.299	0.702	0.226	6
712	PF3D7_1437900	HSP40, subfamily A, putative (ERdj3)	0.055	-0.031	0.031	0.246	0.612	0.022	6
713	PF3D7_0731600	acyl-CoA synthetase (ACS5)	0.211	-0.004	-0.155	0.222	0.428	0.590	6
714	PF3D7_1476800	lysophospholipase, putative	-0.091	0.375	0.129	0.508	0.457	1.433	6
715	PF3D7_1439900	triosephosphate isomerase (TIM)	0.161	0.391	-0.049	0.135	0.613	0.331	6

Supplementary Table 11. Differential mRNA abundance assessed by microarray and RT-qPCR in the *Not1.1* mutant. The log₂-transformed mRNA abundance in the *Not1.1* mutant versus wildtype are given. The expression was determined by microarray and RT-qPCR. The total RNA was reversed transcribed into cDNA, amplified by SMART-PCR and hybridized to microarray chips. For RT-qPCR, the total RNA was reverse transcribed into cDNA, 20 ng of which served as template in real-time PCR. The average value of expression (ave) and the standard deviation (stdev) were calculated from technical triplicates (r1, r2 and r3).

	microarray					RT-qPCR				
	r1	r2	r3	ave	stdev	r1	r2	r3	ave	stdev
KAHRP_TP1	0.25	0.49	0.20	0.32	0.15	4.35	4.83	4.59	4.59	0.24
KAHRP_TP2	2.54	0.89	2.34	1.92	0.90	2.95	2.77	3.31	3.01	0.27
KAHRP_TP3	2.10	1.54	1.55	1.73	0.33	3.50	3.52	3.51	3.51	0.01
KAHRP_TP4	0.47	0.25	0.42	0.38	0.12	4.80	4.38	4.97	4.72	0.31
KAHRP_TP5	0.85	1.05	0.97	0.96	0.10	4.53	4.38	4.19	4.37	0.17
KAHRP_TP6	1.05	1.13	1.00	1.06	0.07	3.81	4.26	3.70	3.92	0.30
EBA181_TP1	0.63	1.44	1.16	1.08	0.41	1.11	4.02	1.62	2.25	1.55
EBA181_TP5	-1.53	-1.24	-1.14	-1.31	0.20	-3.06	-3.72	-4.33	-3.70	0.64
MSP6_TP1	0.77	1.42	1.06	1.08	0.33	0.23	3.52	2.35	2.04	1.67
MSP6_TP5	-1.77	-1.46	-1.28	-1.50	0.25	-2.15	-3.70	-3.57	-3.14	0.86
RH4_TP1	0.59	0.55	0.64	0.59	0.04	0.36	3.31	1.19	1.62	1.52
RH4_TP5	-0.97	-1.17	-0.82	-0.99	0.17	-1.77	-3.54	-3.81	-3.04	1.11
RH5_TP5	-1.04	-1.16	-0.98	-1.06	0.09	-2.39	-4.57	-7.40	-4.78	2.51

Supplementary Table 12. 150 genes with significantly differential RNAPII recruitment in *Not1.1* mutant. Recruitment of the elongating form of RNAPII was assessed by chromatin-immunoprecipitation followed by microarray. The differential recruitment was calculated between the *Not1.1* mutant and wildtype at six timepoints. The significant change is defined as a fold-change more than 1.5 (corresponding to the log₂ ratio at 0.585) and p-value less than 0.05. The values are given in the form of log₂-transformed differential recruitment. The order of the genes corresponds to the heatmap in Figure 4.13A.

No.	Gene	Description	TP1	TP2	TP3	TP4	TP5	TP6
1	PF3D7_0109700	conserved protein, unknown function	0.634	0.355	-0.314	-0.400	-0.231	-0.054
2	PF3D7_1421500	mitochondrial import inner membrane translocase subunit TIM8, putative (TIM8)	0.636	0.152	-0.297	0.394	-0.128	-0.029
3	PF3D7_0320000	protein phosphatase inhibitor 2 (I2)	0.702	0.022	-0.255	0.417	0.229	-0.081
4	PF3D7_1357900	pyrroline-5-carboxylate reductase	0.791	0.480	-0.314	0.126	0.118	-0.133
5	PF3D7_1002900	conserved Plasmodium protein, unknown function	0.585	-0.118	0.070	0.099	0.093	0.141
6	PF3D7_1415900	zinc finger protein, putative	0.719	-0.016	-0.096	0.079	0.083	0.031
7	PF3D7_1310100	conserved Plasmodium protein, unknown function	0.691	0.119	0.054	0.235	0.098	0.085
8	PF3D7_0624800	conserved Plasmodium protein, unknown function	0.719	0.194	-0.091	0.148	0.291	0.081
9	PF3D7_1130900	conserved Plasmodium protein, unknown function	0.712	0.177	0.085	-0.054	0.301	0.137
10	PF3D7_0318000	conserved Plasmodium protein, unknown function	0.653	0.064	-0.068	-0.194	0.213	0.215
11	PF3D7_1214700	conserved Plasmodium protein, unknown function	0.602	0.288	0.119	-0.003	0.080	0.109
12	PF3D7_0201300	stevor	0.634	0.094	0.097	-0.120	-0.049	0.200
13	PF3D7_1005800	conserved Plasmodium membrane protein, unknown function	0.604	-0.191	0.172	-0.161	-0.051	-0.098
14	PF3D7_1239400	conserved Plasmodium protein, unknown function	0.762	-0.154	-0.078	0.337	-0.133	-0.080
15	PF3D7_1254700	rifin (RIF)	0.710	-0.346	0.283	0.387	-0.206	-0.006
16	PF3D7_1474800	proteasome subunit alpha type-1, putative	0.648	1.131	0.252	-0.177	0.049	-0.118
17	PF3D7_0825200	translation initiation factor IF-3 (IF3a)	0.529	0.892	0.214	0.092	0.063	0.129
18	PF3D7_1214600	adrenodoxin-type ferredoxin, putative	0.464	0.595	0.085	-0.136	0.242	-0.064
19	PF3D7_1230300	subpellicular microtubule protein 2, putative (SPM2)	0.277	0.598	-0.083	0.064	0.283	-0.004
20	PF3D7_1204900	probable protein, unknown function	0.465	0.684	-0.026	0.358	0.135	0.017
21	PF3D7_0529300	apicoplast TIC22 protein (TIC22)	0.141	0.602	0.345	0.154	-0.024	0.095
22	PF3D7_1367600	ribosome biogenesis protein MRT4, putative	0.188	0.806	-0.094	-0.367	0.004	0.085
23	PF3D7_0629600	conserved Plasmodium protein, unknown function	0.089	0.686	0.128	-0.278	0.343	-0.114

24	PF3D7_0720500	conserved Plasmodium protein, unknown function	0.092	0.695	0.313	-0.201	-0.015	-0.186
25	PF3D7_0301100	Plasmodium exported protein (hyp13), unknown function	-0.049	0.956	0.051	0.039	0.100	0.039
26	PF3D7_1141500	conserved Plasmodium protein, unknown function	-0.043	0.759	0.001	0.010	0.239	-0.041
27	PF3D7_0111200	conserved Plasmodium protein, unknown function	0.128	1.606	0.372	0.048	0.329	-0.052
28	PF3D7_1435100	ribonuclease H2 subunit B, putative	0.053	0.949	0.284	-0.045	0.137	0.080
29	PF3D7_1469300	small subunit rRNA processing KH domain protein, putative	0.112	1.613	0.186	0.082	-0.249	-0.178
30	PF3D7_1339900	ABC transporter B family member 5, putative (ABCB5)	0.016	0.586	0.067	0.039	0.001	-0.039
31	PF3D7_0105600	conserved Plasmodium protein, unknown function	0.048	0.802	0.064	-0.177	-0.024	-0.017
32	PF3D7_1434100	queuine tRNA-ribosyltransferase, putative	0.086	0.640	-0.012	0.147	-0.120	-0.243
33	PF3D7_0502700	conserved Plasmodium protein, unknown function	0.151	0.750	0.050	0.354	0.053	-0.178
34	PF3D7_1473900	conserved Plasmodium protein, unknown function	-0.330	0.757	-0.117	0.021	-0.042	-0.146
35	PF3D7_0724100	conserved Plasmodium protein, unknown function	-0.262	1.011	-0.074	0.019	-0.090	0.018
36	PF3D7_0306300	glutaredoxin 1 (GRX1)	-0.252	0.627	-0.127	-0.181	-0.074	-0.116
37	PF3D7_1228200	conserved Plasmodium protein, unknown function	-0.508	0.946	0.290	0.094	-0.023	-0.001
38	PF3D7_0615300	GPI-anchored wall transfer protein 1, putative (GWT1)	0.139	0.708	-0.127	0.126	-0.009	-0.047
39	PF3D7_1451600	LCCL domain-containing protein (LAP5)	0.209	0.682	-0.201	-0.038	0.103	-0.110
40	PF3D7_0916300	conserved Plasmodium protein, unknown function	-0.037	0.817	-0.598	0.079	0.057	-0.358
41	PF3D7_1120800	conserved Plasmodium protein, unknown function	0.000	0.611	-0.297	0.138	0.161	-0.125
42	PF3D7_1456900	conserved protein, unknown function	-0.027	0.670	-0.354	0.036	0.110	0.052
43	PF3D7_0922900	3-oxoacyl-(acyl-carrier-protein) reductase (FabG)	0.098	0.796	-0.440	0.041	-0.097	0.054
44	PF3D7_1248300	conserved Plasmodium membrane protein, unknown function	0.075	0.860	-0.668	-0.043	0.045	0.004
45	PF3D7_1034700	ADP-ribosylation factor, putative	-0.444	0.692	0.214	0.485	0.253	0.166
46	PF3D7_0115200	rifin (RIF)	-0.081	0.493	-0.159	0.587	-0.041	0.065
47	PF3D7_1302900	conserved Plasmodium protein, unknown function	0.037	0.179	0.093	0.641	0.081	-0.209
48	PF3D7_1364700	conserved Plasmodium protein, unknown function, pseudogene	0.152	0.547	0.053	0.654	0.378	-0.070
49	PF3D7_0602300	liver merozoite formation protein, putative (PALM)	0.621	0.052	0.604	0.112	-0.100	0.081
50	PF3D7_1363600	conserved Plasmodium protein, unknown function	0.245	-0.158	0.589	0.072	0.099	0.035
51	PF3D7_0524200	conserved Plasmodium membrane protein, unknown function	0.259	0.085	0.597	-0.001	0.131	0.064
52	PF3D7_1461100	conserved Plasmodium protein, unknown function	0.076	-0.145	0.624	-0.146	0.188	-0.163
53	PF3D7_1139900	conserved protein, unknown function	-0.176	-0.002	0.609	-0.440	-0.317	0.028
54	PF3D7_1005100	conserved protein, unknown function	-0.490	0.100	0.830	-0.354	0.002	0.002
55	PF3D7_0632000	rifin (RIF)	-0.085	0.901	1.174	-0.586	-0.261	0.670

56	PF3D7_0718800	conserved Plasmodium protein, unknown function	-0.446	0.769	0.682	-0.126	0.300	0.041
57	PF3D7_1444900	conserved Plasmodium protein, unknown function	-0.257	0.332	0.595	0.117	-0.028	0.117
58	PF3D7_1326100	WD repeat-containing protein, putative	-0.094	0.347	0.646	0.118	0.239	0.138
59	PF3D7_0810700	conserved Plasmodium protein, unknown function	0.781	-0.699	-0.075	0.371	-0.060	0.167
60	PF3D7_1006500	conserved Plasmodium protein, unknown function	0.641	-0.484	-0.137	0.325	0.180	-0.072
61	PF3D7_0728500	conserved Plasmodium protein, unknown function	0.405	-0.693	-0.146	0.191	0.225	0.200
62	PF3D7_0110400	DNA-directed RNA polymerase 2, putative	0.150	-0.666	-0.074	-0.196	-0.022	-0.137
63	PF3D7_1454100	tRNA intron endonuclease, putative	0.401	-0.648	-0.215	0.047	-0.034	-0.006
64	PF3D7_1228900	conserved Plasmodium protein, unknown function	0.440	-0.720	-0.002	-0.047	-0.144	0.111
65	PF3D7_1031500	DEAD or DEAH box helicase, putative	0.060	-0.616	0.039	0.101	0.089	0.079
66	PF3D7_1362500	exonuclease, putative	0.066	-0.758	0.084	0.310	0.278	0.015
67	PF3D7_0816700	trafficking protein particle complex subunit 2-like protein, putative (TRAPPC2L)	-0.069	-0.778	-0.001	0.276	0.133	-0.010
68	PF3D7_0318600	cleavage and polyadenylation specificity factor, putative	0.264	-0.631	0.222	0.103	0.061	0.050
69	PF3D7_1243100	zinc finger protein, putative	0.246	-0.595	0.159	0.317	0.104	-0.053
70	PF3D7_1302800	40S ribosomal protein S7, putative	0.136	-0.604	0.098	-0.216	0.170	0.026
71	PF3D7_1336400	RanBPM and CLTH-like protein, putative	0.186	-0.818	0.452	-0.426	0.121	0.040
72	PF3D7_0218200	conserved Plasmodium protein, unknown function	0.078	-0.744	0.429	-0.072	0.169	-0.119
73	PF3D7_1474700	protein kinase, putative	0.019	-0.603	0.154	-0.072	0.076	-0.204
74	PF3D7_1232400	conserved protein, unknown function	-0.073	-0.622	0.202	-0.238	-0.191	-0.183
75	PF3D7_0934600	conserved Plasmodium protein, unknown function	-0.390	-0.717	-0.322	-0.127	-0.133	-0.116
76	PF3D7_0931600	conserved protein, unknown function	-0.505	-0.726	-0.175	-0.362	-0.303	-0.010
77	PF3D7_0519800	conserved protein, unknown function	-0.413	-0.599	-0.182	-0.232	-0.061	0.054
78	PF3D7_0617800	histone H2A (H2A)	-0.122	-0.591	0.028	-0.301	-0.145	0.139
79	PF3D7_1112500	conserved protein, unknown function	0.040	-0.728	-0.006	-0.160	0.035	0.229
80	PF3D7_1130100	60S ribosomal protein L38 (RPL38)	-0.151	-0.646	0.143	-0.197	0.141	0.096
81	PF3D7_1213300	conserved Plasmodium protein, unknown function	-0.054	-1.026	0.016	-0.151	0.141	-0.077
82	PF3D7_1311400	AP-1 complex subunit mu-1 (AP1M1)	-0.295	-0.665	-0.091	-0.246	0.163	0.016
83	PF3D7_1018100	conserved Plasmodium protein, unknown function	-0.289	-0.956	0.190	0.060	0.016	0.282
84	PF3D7_1451500	pre-mRNA-splicing factor CWF18, putative	-0.429	-0.801	0.054	-0.166	-0.048	0.164
85	PF3D7_0532100	early transcribed membrane protein 5 (ETRAMP5)	-0.191	-0.588	0.174	-0.131	-0.084	0.044
86	PF3D7_0621900	signal recognition particle subunit SRP68, putative (SRP68)	-0.048	-0.605	-0.161	0.039	-0.112	0.033
87	PF3D7_0109600	cold-shock protein, putative	-0.089	-0.905	-0.023	0.099	-0.062	-0.100

88	PF3D7_0618100	conserved Plasmodium protein, unknown function	-0.030	-0.645	-0.110	0.061	-0.069	-0.073
89	PF3D7_0406800	conserved Plasmodium protein, unknown function	-0.154	-0.746	-0.146	-0.007	-0.007	-0.053
90	PF3D7_1329600	conserved Plasmodium protein, unknown function	-0.131	-0.698	-0.011	-0.061	-0.135	0.039
91	PF3D7_1030800	calmodulin, putative	-0.089	-1.067	0.123	-0.033	-0.075	-0.010
92	PF3D7_0501600	rhoptry-associated protein 2 (RAP2)	-0.031	-0.596	-0.012	-0.026	-0.213	-0.079
93	PF3D7_1335400	reticulocyte binding protein 2 homologue a (RH2a)	-0.052	-0.594	0.101	0.174	-0.083	0.028
94	PF3D7_1145800	conserved Plasmodium protein, unknown function	-0.190	-0.740	0.177	0.091	-0.005	-0.143
95	PF3D7_0830500	tryptophan or threonine-rich antigen (TryThrA)	-0.234	-0.606	-0.387	0.197	0.099	0.003
96	PF3D7_0407100	methyltransferase, putative	-0.072	-0.700	-0.465	-0.142	-0.036	0.131
97	PF3D7_1446800	heme detoxification protein (HDP)	-0.044	-0.654	-0.322	-0.112	0.185	-0.096
98	PF3D7_1326000	conserved Plasmodium protein, unknown function	0.182	-0.705	-0.342	0.213	-0.115	-0.053
99	PF3D7_1117400	conserved Plasmodium protein, unknown function	0.044	-0.745	-0.492	0.089	-0.092	0.030
100	PF3D7_0933300	conserved Plasmodium protein, unknown function	-0.702	-0.195	0.432	0.377	0.003	0.053
101	PF3D7_1437300	conserved Plasmodium protein, unknown function	-0.784	-0.509	0.629	0.359	0.349	-0.287
102	PF3D7_0531000	conserved Plasmodium protein, unknown function	-0.220	-0.768	0.081	0.566	0.061	-0.053
103	PF3D7_1253700	rifin (RIF)	-0.288	-0.265	0.109	0.616	-0.014	0.070
104	PF3D7_1336100	conserved Plasmodium protein, unknown function	-0.207	-0.611	0.567	0.072	0.033	-0.111
105	PF3D7_1317700	conserved Plasmodium protein, unknown function	-0.432	-0.657	0.651	0.306	-0.131	-0.048
106	PF3D7_0109400	tubulin-specific chaperone a, putative	0.096	-0.576	0.660	-0.166	-0.202	0.003
107	PF3D7_0502900	mitochondrial inner membrane TIM10 associated protein, putative	0.337	-0.249	-0.492	0.598	-0.037	0.022
108	PF3D7_1408300	conserved Plasmodium protein, unknown function	0.083	-0.366	-0.170	0.607	-0.204	0.037
109	PF3D7_1475900	conserved Plasmodium protein, unknown function	-0.712	-0.288	-0.397	0.000	-0.151	0.063
110	PF3D7_1446300	conserved Plasmodium membrane protein, unknown function	-0.621	0.054	0.027	0.150	-0.027	-0.002
111	PF3D7_0705700	40S ribosomal protein S29, putative	-0.610	0.007	-0.180	0.403	-0.092	0.076
112	PF3D7_1245800	epsin, putative	-0.756	-0.357	0.087	0.011	0.054	-0.113
113	PF3D7_1459500	conserved Plasmodium protein, unknown function	-0.722	-0.223	-0.095	-0.093	-0.053	-0.189
114	PF3D7_1230400	ATP-dependent protease subunit ClpQ (ClpQ)	-0.619	-0.042	0.119	-0.163	0.123	0.126
115	PF3D7_1141200	conserved Plasmodium protein, unknown function	-0.670	0.022	0.197	-0.011	0.038	0.201
116	PF3D7_1217600	anaphase promoting complex subunit 10, putative	-0.603	-0.063	0.100	-0.088	0.076	-0.082
117	PF3D7_1353600	ER lumen protein retaining receptor (ERD2)	-0.716	-0.182	0.190	-0.207	-0.095	0.048
118	PF3D7_1366200	conserved Plasmodium protein, unknown function	-0.981	0.511	-0.216	-0.419	-0.290	-0.122
119	PF3D7_1410100	alpha or beta hydrolase, putative	-0.964	0.185	-0.365	0.000	-0.075	-0.226
120	PF3D7_0811300	CCR4-associated factor 1 (CAF1)	-0.589	0.009	-0.125	-0.095	0.045	-0.051

121	PF3D7_1015800	ribonucleotide reductase small subunit, putative	-0.891	0.099	-0.120	-0.154	-0.268	-0.232
122	PF3D7_0702200	lysophospholipase, putative	-0.228	-0.699	-0.642	-0.701	0.116	0.225
123	PF3D7_1353900	proteasome subunit alpha type-7, putative	0.077	-0.258	-0.022	-0.601	-0.010	-0.001
124	PF3D7_0308300	conserved Plasmodium protein, unknown function	-0.379	-0.085	-0.119	-0.856	0.116	-0.107
125	PF3D7_0937100	Plasmodium exported protein, unknown function	-0.294	0.249	-0.458	-1.075	-0.104	-0.055
126	PF3D7_0404200	conserved Plasmodium protein, unknown function	-0.587	0.233	-0.499	0.054	0.045	0.027
127	PF3D7_0528000	proteasome maturation factor UMP1, putative (UMP1)	-0.799	0.746	-0.661	-0.044	0.067	-0.129
128	PF3D7_0302900	exportin-1, putative	-0.395	0.236	-0.589	0.292	0.113	0.112
129	PF3D7_1001500	early transcribed membrane protein 10.1 (ETRAPM10)	-0.325	-0.105	-0.715	0.152	-0.051	-0.140
130	PF3D7_1007000	conserved membrane protein, unknown function	-0.643	0.038	-1.008	0.088	0.205	-0.222
131	PF3D7_0922300	conserved Plasmodium protein, unknown function	-0.220	0.604	-0.755	-0.158	0.286	-0.027
132	PF3D7_0402900	probable protein, unknown function	-0.257	0.453	-0.681	-0.385	0.270	-0.209
133	PF3D7_1236600	p25-alpha family protein, putative	-0.314	0.064	-0.790	-0.444	-0.341	0.251
134	PF3D7_1449300	transcription factor IIIb subunit, putative	-0.451	0.170	-0.749	-0.515	-0.048	-0.101
135	PF3D7_1431700	60S ribosomal protein L14, putative	0.126	-0.335	-0.600	-0.180	0.125	-0.039
136	PF3D7_1202800	conserved Plasmodium protein, unknown function	0.433	-0.323	-0.603	-0.413	-0.073	-0.014
137	PF3D7_1311700	cytochrome c2 precursor, putative	0.159	0.058	-0.651	-0.106	0.243	-0.039
138	PF3D7_0621500	ribonuclease P or MRP protein subunit RPP1, putative (RPP1)	-0.046	0.004	-0.691	-0.078	-0.043	-0.011
139	PF3D7_1357800	T-complex protein 1 subunit delta (CCT4)	-0.071	-0.134	-0.645	-0.253	0.020	-0.237
140	PF3D7_1003200	conserved Plasmodium protein, unknown function	-0.237	-0.352	-0.630	-0.252	0.003	-0.060
141	PF3D7_0618800	conserved Plasmodium protein, unknown function	0.082	-0.295	-0.609	0.176	-0.007	-0.144
142	PF3D7_0307500	conserved protein, unknown function	0.201	0.608	-0.458	0.231	-0.429	0.077
143	PF3D7_1326800	syntaxin, Qa-SNARE family (SYN18)	0.446	0.409	-0.507	0.659	-0.347	-0.095
144	PF3D7_0209400	conserved protein, unknown function	0.285	0.259	-0.677	0.211	-0.061	0.048
145	PF3D7_0823700	conserved Plasmodium protein, unknown function	0.122	0.488	-0.688	0.422	0.050	-0.156
146	PF3D7_0207300	serine repeat antigen 8 (SERA8)	-0.181	0.446	-0.648	0.292	-0.099	-0.010
147	PF3D7_0412300	phosphopantothienoylcysteine synthetase, putative	-0.113	0.287	-0.626	0.010	-0.154	-0.129
148	PF3D7_1132700	mitochondrial ribosomal protein L2 precursor	-0.144	-0.026	-0.603	0.391	0.242	0.274
149	PF3D7_0508200	longevity-assurance (LAG1) protein, putative	-0.206	0.184	-0.254	-0.136	0.636	0.242
150	PF3D7_1319800	probable protein, unknown function	0.098	0.076	-0.185	0.249	0.595	0.179

Supplementary Table 13. Proliferation of the *Not1.2* mutant. Parasites were grown in 2% hematocrit for three IDCs. Initial parasitemia was 0.1%. The number of iRBC in around 1,000 total RBCs was counted in biological triplicates (r1, r2 and r3) for each parasite line in each IDC. The average parasitemia and the standard deviation were calculated. The data was plotted in **Figure 4.19**. One-tail t-test was conducted between the parasitemia of the *Not1.2* mutant and the control line in each IDC.

IDC cycles	0	1			2			3		
Replicate		r1	r2	r3	r1	r2	r3	r1	r2	r3
Parasite line	control									
iRBC		18	16	12	85	85	98	208	276	204
Total RBC		1035	1044	1005	1029	1049	1056	1052	1132	1012
Parasitemia (%)	0.10	1.74	1.53	1.19	8.26	8.10	9.28	19.77	24.38	20.16
Average		1.49			8.55			21.44		
Standard deviation		0.28			0.64			2.56		
Parasite line	Not1.2 mutant									
iRBC		3	6	5	32	29	38	103	107	101
Total RBC		1033	1019	1070	1018	1012	1047	1055	1090	1022
Parasitemia (%)	0.10	0.29	0.59	0.47	3.14	2.87	3.63	9.76	9.82	9.88
Average		0.45			3.21			9.82		
Standard deviation		0.15			0.39			0.06		
p-value		0.0048			0.0004			0.0079		

Supplementary Table 14. Invasion rate of the *Not1.2* mutant. Parasites were grown in 1% hematocrit in biological triplicates (r1, r2 and r3). Initial parasitemia was 0.5% and 1% synchronized parasites in the late schizont stage. The number of iRBC in around 1,000 total RBCs was counted 14 hours afterward in biological triplicates for each parasite line. Invasion rate was calculated as the ratio of the parasitemia after invasion versus the initial parasitemia. The data was plotted in **Figure 4.20**. The average parasitemia and the standard deviation were calculated. One-tail t-test was conducted between the invasion rate of the *Not1.2* mutant and the control line.

Hematocrit (%)		1	1
Parasitemia (%)		0.5	1
Parasite line	control		
Parasitemia (%)	r1	1.74	2.78
	r2	1.96	2.58
	r3	1.90	3.00
Invasion rate	r1	3.48	2.78
	r2	3.92	2.58
	r3	3.80	3.00
Average		3.74	2.79
Standard deviation		0.23	0.21
Parasite line	<i>Not1.2</i> mutant		
Parasitemia (%)	r1	1.23	1.85
	r2	1.40	1.28
	r3	1.40	2.27
Invasion rate	r1	2.46	1.85
	r2	2.80	1.28
	r3	2.79	2.27
Average		2.68	1.80
Standard deviation		0.20	0.49
p-value		0.0020	0.0287

Supplementary Table 15. Merozoite number per schizont in the *Not1.2* mutant.

Parasites were grown in 2% hematocrit in biological duplicates with 5% parasitemia. The number of merozoites in each schizont was counted for 50 iRBCs once segmentation of merozoites was clearly observed within schizont parasites. The data was plotted in **Figure 4.21**. One-tail t-test was conducted between the number of merozoites in the *Not1.2* mutant and the control line for each biological replicate.

Replicate	Replicate 1				Replicate 2			
Parasite line	wildtype		Not1.1 mutant		wildtype		Not1.1 mutant	
Number of Merozoites	16	13	12	14	13	13	14	20
	13	14	13	16	13	15	13	13
	15	14	16	14	14	13	14	16
	17	14	14	17	12	13	14	15
	15	13	15	12	12	14	15	14
	14	16	16	15	15	14	12	12
	16	14	14	14	14	16	14	14
	15	13	13	18	14	13	16	18
	14	13	16	13	15	15	18	14
	15	13	13	16	14	16	13	13
	15	12	15	16	13	14	13	17
	13	13	13	17	15	15	16	16
	13	16	12	13	12	14	18	14
	13	13	14	14	12	16	15	17
	12	16	16	14	14	15	13	14
	18	13	16	13	14	13	15	13
	15	14	13	13	16	15	15	13
	12	13	12	14	14	15	17	13
	13	14	13	13	14	16	13	14
	13	13	15	14	15	16	13	15
	14	16	14	13	15	17	15	18
	14	14	13	18	14	18	16	13
	13	15	22	16	15	19	13	16
	14	17	13	16	15	14	16	14
	15	15	16	14	15	15	15	13
Average	14.16		14.52		14.46		14.70	
Standard deviation	1.39		1.91		1.45		1.81	
p-value	0.1419				0.2328			

Supplementary Table 16. Estimated age of parasites harvested at 6 timepoints across one IDC in the *Not1.2* mutant. The Spearman correlation between the transcriptional profile of each sample and the periodicity of gene expression was conducted. The age was estimated with the highest Spearman coefficient. The average of the ages and the standard deviation were calculated. The data was plotted in **Figure 4.22**. One-tail t-test was conducted between the ages of parasites in *Not1.2* mutant and the control line at the same timepoint. One-tailed t-test cannot calculate the p-value for samples with identical ages, which is denoted as NA.

	TP1	TP2	TP3	TP4	TP5	TP6
control						
rep1	0	6	16	22	32	40
rep2	2	8	16	28	32	40
rep3	4	6	16	28	32	40
average	2	6.667	16	26	32	40
standard deviation	2	1.155	0	3.464	0	0
<i>Not1.2</i> mutant						
rep1	2	10	18	28	36	40
rep2	2	6	18	28	36	40
rep3	4	6	18	28	36	40
average	2.667	7.333	18	28	36	40
standard deviation	1.155	2.309	0	0	0	0
p-value	0.3247	0.3428	NA	0.2113	NA	NA

Supplementary Table 17. 202 genes with significantly differential mRNA abundance in the *Not1.2* mutant. The value is the log₂-transformed ratio of mRNA abundance in the *Not1.2* mutant versus that in the control line. The order of the genes corresponds to the heatmap in **Figure 4.23**. The significant change is defined as an expression fold-change more than 1.5 (corresponding to the log₂ ratio at 0.585) and p-value less than 0.05. There are four clusters of genes according to the timepoint of the greatest upregulation and genes in each cluster are labeled.

No.	Gene	Description	TP1	TP2	TP3	TP4	TP5	TP6	Cluster
1	PF3D7_0424800	Plasmodium exported protein (PHISTb), unknown function	1.978	1.406	0.712	0.132	0.421	2.312	1
2	PF3D7_1301700	Plasmodium exported protein (hyp8), unknown function (GEXP07)	1.250	0.992	0.923	0.306	0.323	1.286	1
3	PF3D7_0114500	Plasmodium exported protein (hyp10), unknown function	0.039	1.682	0.907	0.819	1.088	1.043	1
4	PF3D7_0424900	Plasmodium exported protein (PHISTa), unknown function	0.930	0.903	0.563	0.066	-0.052	0.648	1
5	PF3D7_1033200	early transcribed membrane protein 10.2 (ETRAPM10.2)	-0.190	0.604	0.229	-0.240	-0.162	-0.127	1
6	PF3D7_0832200	conserved Plasmodium protein, unknown function	0.041	0.602	0.197	-0.621	-0.160	0.128	1
7	PF3D7_0402000	Plasmodium exported protein (PHISTa), unknown function	-0.034	0.726	0.797	-0.115	-0.250	-0.085	1
8	PF3D7_1223900	50S ribosomal protein L24, putative	0.009	1.364	0.350	-0.073	-0.010	0.003	1
9	PF3D7_0532400	lysine-rich membrane-associated PHISTb protein (LyMP)	0.073	0.822	0.505	-0.131	-0.043	0.070	1
10	PF3D7_0408600	sporozoite invasion-associated protein 1 (SIAP1)	0.059	0.725	0.112	0.117	0.479	0.245	1
11	PF3D7_0713100	Pfmc-2TM Maurers cleft two transmembrane protein (MC-2TM)	-0.479	0.783	0.622	-0.066	-0.096	0.065	1
12	PF3D7_1201000	Plasmodium exported protein (PHISTb), unknown function	-0.094	0.660	0.537	-0.226	-0.254	-0.136	1
13	PF3D7_0401800	Plasmodium exported protein (PHISTb), unknown function (PfD80)	0.092	0.498	0.123	-0.693	-0.102	0.120	1
14	PF3D7_0731300	Plasmodium exported protein (PHISTb), unknown function (PfG174)	-0.095	0.704	0.528	0.062	0.157	0.159	1
15	PF3D7_1149000	antigen 332, DBL-like protein (Pf332)	-0.043	0.507	0.688	0.040	-0.204	-0.192	1
16	PF3D7_0201900	erythrocyte membrane protein 3 (EMP3)	0.305	1.351	0.119	-0.065	0.308	0.071	1
17	PF3D7_1476200	Plasmodium exported protein (PHISTb), unknown function	-0.165	0.674	0.377	-0.360	-0.007	-0.039	1
18	PF3D7_0301700	Plasmodium exported protein, unknown function	-0.109	0.592	0.521	0.005	0.001	-0.194	1
19	PF3D7_0525000	zinc finger protein, putative	0.191	0.965	0.368	-0.561	-0.112	-0.103	1
20	PF3D7_1035800	probable protein, unknown function (M712)	0.384	1.035	0.157	-0.638	-0.101	0.066	1

21	PF3D7_0702500	Plasmodium exported protein, unknown function	0.245	0.710	0.287	-0.467	-0.351	0.060	1
22	PF3D7_1478600	EMP1-trafficking protein (PTP3)	0.072	0.862	0.325	-0.182	0.084	-0.133	1
23	PF3D7_1001000	Plasmodium exported protein (hyp12), unknown function (Pfj13)	0.089	0.509	0.607	-0.051	0.029	0.101	1
24	PF3D7_1039700	Pfmc-2TM Maurers cleft two transmembrane protein (MC-2TM)	0.006	0.593	0.610	0.146	0.419	0.428	1
25	PF3D7_1405600	ribonucleotide reductase small subunit (RNR)	-0.037	0.057	0.710	0.268	-0.094	-0.446	2
26	PF3D7_1410700	conserved Plasmodium protein, unknown function	-0.033	0.129	0.876	0.556	-0.139	-0.216	2
27	PF3D7_1143000	alpha or beta hydrolase, putative	-0.107	-0.350	0.614	0.343	0.082	-0.430	2
28	PF3D7_0409500	conserved Plasmodium protein, unknown function	-0.050	0.171	1.053	0.473	-0.096	-0.346	2
29	PF3D7_1304000	chromosome condensation protein, putative	0.066	-0.056	0.607	0.319	0.020	-0.141	2
30	PF3D7_0932200	profilin, putative (PFN)	-0.132	0.087	0.550	0.633	0.254	-0.055	2
31	PF3D7_0904700	bacterial histone-like protein (HU)	0.061	0.108	0.517	0.686	0.098	-0.253	2
32	PF3D7_0828500	translation initiation factor EIF-2b alpha subunit, putative	-0.312	-0.001	0.611	0.315	-0.104	-0.487	2
33	PF3D7_1033900	ubiquitin-conjugating enzyme, putative	0.055	-0.040	-0.406	1.103	0.083	-0.110	2
34	PF3D7_1137500	apicoplast ribosomal protein S14p or S29e precursor, putative	0.103	0.255	0.642	0.400	-0.112	-0.120	2
35	PF3D7_0401700	Plasmodium exported protein, unknown function, pseudogene	-0.201	0.089	0.681	0.173	0.197	0.005	2
36	PF3D7_1331600	protein tyrosine phosphatase-like protein, putative (PTPLA)	-0.239	-0.196	0.322	0.604	0.136	-0.108	2
37	PF3D7_1228600	merozoite surface protein 9 (MSP9)	-0.029	-0.013	-0.049	0.987	0.597	-0.119	2
38	PF3D7_1327500	conserved Plasmodium protein, unknown function	0.041	-0.340	0.257	0.644	0.285	0.015	2
39	PF3D7_0501600	rhopty-associated protein 2 (RAP2)	0.090	0.003	-0.073	0.930	0.530	-0.152	2
40	PF3D7_1035100	probable protein, unknown function	0.026	-0.400	-0.118	0.975	0.245	-0.272	2
41	PF3D7_0929400	high molecular weight rhopty protein 2 (RhopH2)	0.037	-0.159	-0.059	0.923	0.304	-0.226	2
42	PF3D7_1028600	conserved Plasmodium protein, unknown function	-0.130	-0.269	0.726	0.627	0.061	-0.263	2
43	PF3D7_0604100	transcription factor with AP2 domain(s) (SIP2)	-0.208	-0.171	-0.024	0.853	0.312	-0.305	2
44	PF3D7_1235100	conserved Plasmodium protein, unknown function	0.288	-0.020	0.399	0.816	0.014	-0.119	2
45	PF3D7_1001600	alpha or beta hydrolase, putative	0.070	0.136	0.096	0.809	0.392	0.278	2
46	PF3D7_1344500	U6 snRNA phosphodiesterase, putative (USB1)	-0.071	-0.168	0.005	0.598	0.025	-0.348	2
47	PF3D7_1335100	merozoite surface protein 7 (MSP7)	0.037	0.109	0.086	0.627	0.515	0.065	2
48	PF3D7_1141800	phd finger protein, putative	0.100	-0.040	0.358	0.915	0.358	-0.129	2
49	PF3D7_1334600	MSP7-like protein (MSRP3)	0.122	-0.215	0.048	0.973	0.208	-0.203	2
50	PF3D7_0107000	centrin-1 (CEN1)	-0.037	-0.029	0.469	0.590	0.045	-0.159	2
51	PF3D7_0704500	serine or threonine protein kinase, putative	0.070	-0.004	0.040	0.748	0.160	-0.295	2

52	PF3D7_0613800	transcription factor with AP2 domain(s) (ApiAP2)	-0.042	0.073	-0.020	0.702	0.244	-0.218	2
53	PF3D7_0207400	serine repeat antigen 7 (SERA7)	-0.133	0.132	-0.136	0.605	0.061	-0.230	2
54	PF3D7_1245500	conserved Plasmodium protein, unknown function	0.001	-0.111	0.200	0.690	0.213	-0.395	2
55	PF3D7_0106200	conserved Plasmodium protein, unknown function	-0.151	-0.199	0.198	0.754	0.228	-0.377	2
56	PF3D7_0504900	conserved protein, unknown function	-0.050	0.041	0.399	0.639	-0.043	-0.560	2
57	PF3D7_0930300	merozoite surface protein 1 (MSP1)	-0.106	-0.071	0.010	0.623	0.221	0.280	2
58	PF3D7_1143200	DnaJ protein, putative	0.090	-0.014	-0.025	0.718	0.121	-0.101	2
59	PF3D7_1417300	cysteine protease ATG4, putative (ATG4)	-0.085	-0.181	0.115	0.945	0.538	0.017	2
60	PF3D7_0423500	glideosome associated protein with multiple membrane spans 2 (GAPM2)	0.063	0.060	0.134	0.844	0.266	-0.358	2
61	PF3D7_1002900	conserved Plasmodium protein, unknown function	0.140	-0.047	-0.332	0.698	0.098	-0.014	2
62	PF3D7_1364100	6-cysteine protein (P92)	-0.083	-0.027	0.184	0.853	0.035	-0.354	2
63	PF3D7_0114400	erythrocyte membrane protein 1 (PfEMP1), exon 2, pseudogene (VAR)	0.107	0.191	0.911	0.663	0.706	0.378	2
64	PF3D7_0501500	rhoptry-associated protein 3 (RAP3)	0.258	0.060	0.055	0.636	0.744	-0.110	2
65	PF3D7_1432700	protein-L-isoaspartate(D-aspartate) O-methyltransferase, putative	0.249	-0.797	0.175	-0.008	0.614	-0.718	2
66	PF3D7_1439800	vesicle-associated membrane protein, putative	0.216	-0.005	-0.044	-0.001	0.056	-0.602	2
67	PF3D7_0903600	conserved Plasmodium protein, unknown function	-0.365	-0.116	-0.006	-0.023	0.608	-0.137	3
68	PF3D7_1461400	conserved Plasmodium protein, unknown function	0.103	0.175	-0.104	0.511	0.740	0.035	3
69	PF3D7_1468400	zinc finger protein, putative (D13)	-0.170	-0.031	-0.098	0.045	0.586	0.001	3
70	PF3D7_0217500	calcium-dependent protein kinase 1 (CDPK1)	0.041	-0.114	-0.123	0.487	0.603	0.132	3
71	PF3D7_0321400	protein kinase, putative	0.056	-0.046	-0.413	0.484	0.590	-0.034	3
72	PF3D7_1126700	autophagy-related protein 23, putative (ATG23)	-0.093	-0.126	-0.090	0.204	0.658	-0.067	3
73	PF3D7_1140400	conserved Plasmodium protein, unknown function	0.305	0.239	-0.155	0.925	0.733	-0.362	3
74	PF3D7_0830500	tryptophan or threonine-rich antigen (TryThrA)	-0.299	-0.219	-0.338	-0.099	0.707	0.708	3
75	PF3D7_0731500	erythrocyte binding antigen-175 (EBA175)	-0.424	0.003	-0.011	-0.050	0.655	0.605	3
76	PF3D7_1016800	Plasmodium exported protein (PHISTc), unknown function	-0.171	-0.198	-0.413	-0.105	1.001	0.888	3
77	PF3D7_1337800	calcium-dependent protein kinase 5 (CDPK5)	-0.311	-0.200	-0.218	-0.173	0.809	0.422	3
78	PF3D7_1201200	Plasmodium exported protein (PHISTa-like), unknown function	-0.398	-0.222	-0.404	-0.358	1.053	0.353	3
79	PF3D7_0402100	Plasmodium exported protein (PHISTb), unknown function	-0.350	-0.094	-0.269	-0.416	0.566	0.838	3
80	PF3D7_1204900	probable protein, unknown function	-0.147	-0.205	-0.066	-0.040	0.630	0.428	3

81	PF3D7_0414900	armadillo-domain containing rhopty protein (ARO)	-0.004	-0.070	-0.100	0.023	0.637	0.388	3
82	PF3D7_1036500	probable protein, unknown function	-0.304	0.028	-0.106	-0.075	0.837	0.435	3
83	PF3D7_0102500	erythrocyte binding antigen-181 (EBA181)	-0.307	0.051	-0.146	0.018	0.731	0.347	3
84	PF3D7_1125800	kelch protein, putative	-0.244	-0.053	-0.163	-0.258	0.589	0.486	3
85	PF3D7_1026600	conserved Plasmodium protein, unknown function	-0.257	-0.038	0.057	-0.055	0.682	0.426	3
86	PF3D7_0911100	conserved Plasmodium protein, unknown function	-0.121	0.195	0.113	-0.018	0.597	0.365	3
87	PF3D7_0308300	conserved Plasmodium protein, unknown function	-0.044	0.041	-0.187	0.062	1.018	0.496	3
88	PF3D7_1401600	Plasmodium exported protein (PHISTb), unknown function	-0.016	0.117	-0.104	0.005	0.635	0.236	3
89	PF3D7_0822900	conserved Plasmodium protein, unknown function	-0.081	-0.017	-0.052	0.027	0.714	0.184	3
90	PF3D7_0102200	ring-infected erythrocyte surface antigen (RESA)	-0.173	0.037	0.056	-0.140	0.700	0.928	3
91	PF3D7_0501400	interspersed repeat antigen (FIRA)	-0.261	-0.310	-0.342	-0.308	0.838	0.764	3
92	PF3D7_0214600	serine or threonine protein kinase, putative	-0.232	0.014	-0.156	-0.049	0.665	0.237	3
93	PF3D7_1351700	inner membrane complex protein 1f, putative (IMC1f)	-0.184	-0.105	0.081	-0.050	0.738	0.238	3
94	PF3D7_0913800	conserved Plasmodium protein, unknown function	-0.056	-0.072	-0.274	-0.068	0.643	0.436	3
95	PF3D7_0318700	trophozoite stage antigen	-0.188	-0.112	-0.145	-0.104	0.762	0.727	3
96	PF3D7_1125700	kelch protein, putative	-0.245	0.000	-0.040	-0.047	0.658	0.376	3
97	PF3D7_0805200	gamete release protein, putative (GAMER)	-0.286	-0.167	-0.136	0.032	0.487	0.739	3
98	PF3D7_1251200	coronin	-0.203	-0.105	-0.228	-0.063	0.640	0.216	3
99	PF3D7_1149200	ring-infected erythrocyte surface antigen	-0.219	-0.066	0.078	-0.042	0.632	0.729	3
100	PF3D7_0934600	conserved Plasmodium protein, unknown function	0.038	0.054	-0.099	-0.099	0.731	0.333	3
101	PF3D7_1017500	conserved Plasmodium protein, unknown function	0.190	0.178	0.084	-0.081	0.535	0.624	3
102	PF3D7_0112800	Plasmodium exported protein (hyp11), unknown function	0.087	-0.197	0.230	0.293	0.773	0.922	3
103	PF3D7_1370300	membrane associated histidine-rich protein (MAHRP1)	0.603	0.468	0.024	-0.454	-0.124	1.101	3
104	PF3D7_1002100	EMP1-trafficking protein (PTP5)	0.143	0.189	0.017	-0.098	0.491	1.402	3
105	PF3D7_1149600	DnaJ protein, putative	-0.279	-0.047	-0.226	-0.297	0.283	0.952	3
106	PF3D7_0219900	Plasmodium exported protein, unknown function	0.114	0.212	-0.154	-0.079	0.142	0.657	3
107	PF3D7_0424600	Plasmodium exported protein (PHISTb), unknown function	-0.235	0.156	0.310	-0.077	0.338	1.198	3
108	PF3D7_0726100	Plasmodium exported protein, unknown function	0.239	-0.008	-0.096	-0.185	0.303	0.999	3
109	PF3D7_1456800	V-type H(-)-translocating pyrophosphatase, putative (VP1)	-0.057	-0.032	-0.048	-0.249	-0.145	0.666	3

110	PF3D7_1353200	membrane associated histidine-rich protein (MAHRP2)	-0.004	-0.086	-0.084	-0.762	0.032	1.943	3
111	PF3D7_0206600	DNA-directed RNA polymerase III subunit RPC10, putative	-0.001	-0.001	-0.140	-0.340	-0.145	0.634	3
112	PF3D7_0220100	DnaJ protein, putative	0.059	-0.119	-0.097	-0.395	0.020	0.607	3
113	PF3D7_1436300	translocon component PTEX150 (PTEX150)	-0.056	0.108	0.191	-0.418	0.098	0.806	3
114	PF3D7_1116800	heat shock protein 101 (HSP101)	-0.099	0.269	-0.020	-0.435	0.262	0.909	3
115	PF3D7_0610700	conserved Plasmodium protein, unknown function	0.154	0.097	-0.018	-0.197	-0.314	0.823	3
116	PF3D7_0219700	Plasmodium exported protein (PHISTc), unknown function (GEXP20)	-0.129	-0.343	-0.561	-0.285	0.822	1.453	3
117	PF3D7_1401100	DnaJ protein, putative	-0.137	0.085	-0.074	-0.155	0.318	1.371	3
118	PF3D7_1301600	erythrocyte binding antigen-140 (EBA140)	0.163	0.027	-0.159	-0.010	0.352	1.172	3
119	PF3D7_1001900	Plasmodium exported protein (hyp16), unknown function (PfJ23)	0.137	0.207	0.026	0.309	0.589	1.731	3
120	PF3D7_1404900	conserved Plasmodium protein, unknown function	-0.129	0.092	0.416	-0.222	-0.137	0.660	3
121	PF3D7_1411900	p1 or s1 nuclease, putative	0.037	-0.361	-0.187	-0.570	0.003	0.882	3
122	PF3D7_1200800	serine or threonine protein kinase, FIKK family (FIKK12)	-0.130	0.010	-0.068	-0.115	0.112	1.426	3
123	PF3D7_1372200	histidine-rich protein III (HRPIII)	0.500	-0.014	-0.186	-0.664	-0.302	1.349	3
124	PF3D7_0906500	arginase	-0.058	-0.080	0.126	-0.283	-0.074	0.654	3
125	PF3D7_0301600	Plasmodium exported protein (hyp1), unknown function (GEXP21)	0.028	0.149	-0.110	-0.162	-0.037	0.939	3
126	PF3D7_0501100	conserved Plasmodium protein, unknown function	-0.029	0.016	-0.195	-0.353	-0.108	0.697	3
127	PF3D7_0108500	conserved Plasmodium protein, unknown function	-0.083	-0.007	-0.153	-0.011	0.395	0.852	3
128	PF3D7_1006200	DNA or RNA-binding protein Alba 3 (ALBA3)	0.358	0.313	0.219	0.315	-0.131	0.812	3
129	PF3D7_0202500	early transcribed membrane protein 2 (ETRAPM2)	0.103	0.242	-0.127	-0.241	0.421	1.286	3
130	PF3D7_1302200	early transcribed membrane protein 13 (ETRAPM13)	0.027	-0.141	-0.306	0.129	0.221	0.623	3
131	PF3D7_1404800	conserved Plasmodium protein, unknown function	-0.072	-0.160	-0.287	-0.640	0.167	0.944	3
132	PF3D7_0629600	conserved Plasmodium protein, unknown function	0.011	0.000	0.059	-0.389	-0.116	0.865	3
133	PF3D7_0106400	pre-rRNA-processing protein TSR2, putative	0.000	0.031	-0.081	-0.231	-0.539	0.703	3
134	PF3D7_0402900	probable protein, unknown function	-0.232	-0.044	0.155	-0.086	0.449	0.859	3
135	PF3D7_0315100	eukaryotic translation initiation factor 4E (eIF4E)	-0.015	-0.200	-0.212	-0.193	-0.261	0.603	3
136	PF3D7_1407800	plasmepsin IV (PM4)	0.065	0.251	-0.133	-0.060	-0.198	0.720	3
137	PF3D7_1401400	early transcribed membrane protein 14.1 (ETRAPM14)	-0.041	-0.042	-0.179	-0.599	0.592	0.988	3
138	PF3D7_0605900	long chain polyunsaturated fatty acid elongation enzyme, putative	0.133	0.029	0.888	-0.559	0.282	0.960	3

139	PF3D7_1354300	large subunit rRNA methyltransferase, putative	-0.031	0.018	-0.002	-0.264	-0.034	0.595	3
140	PF3D7_0301300	alpha or beta hydrolase, putative	0.049	0.138	-0.124	-0.186	0.055	0.626	3
141	PF3D7_1471100	exported protein 2 (EXP2)	-0.181	0.174	0.198	-0.204	-0.130	0.603	3
142	PF3D7_1351800	conserved Plasmodium protein, unknown function	0.138	0.155	-0.242	-0.296	-0.268	0.888	3
143	PF3D7_1001800	Plasmodium exported protein (PHISTc), unknown function	0.022	0.120	-0.160	0.109	0.101	0.911	3
144	PF3D7_1420700	surface protein P113 (P113)	-0.095	0.062	0.175	-0.136	-0.122	0.723	3
145	PF3D7_0502800	DnaJ protein, putative	0.204	-0.028	0.084	-0.132	-0.099	0.628	3
146	PF3D7_1451200	conserved Plasmodium protein, unknown function	-0.160	0.097	0.244	0.038	0.078	0.677	3
147	PF3D7_0424500	serine or threonine protein kinase, FIKK family (FIKK4.1)	-0.232	-0.008	-0.173	-0.438	0.146	0.835	3
148	PF3D7_0424300	erythrocyte binding antigen-165, pseudogene (EBA165)	-0.131	-0.026	-0.069	-0.165	-0.228	0.593	3
149	PF3D7_0517300	serine or arginine-rich splicing factor 1 (SR1)	0.010	0.044	0.016	-0.221	0.064	0.820	3
150	PF3D7_1002000	Plasmodium exported protein (hyp2), unknown function	-0.026	-0.002	-0.069	0.081	0.079	1.420	3
151	PF3D7_0913000	conserved protein, unknown function	0.098	0.025	-0.066	-0.200	-0.112	0.721	3
152	PF3D7_1476400	serine or threonine protein kinase, FIKK family, pseudogene (FIKK14)	-0.186	-0.072	0.085	-0.157	0.281	0.784	3
153	PF3D7_0502400	merozoite surface protein 8 (MSP8)	0.009	-0.004	-0.056	-0.279	0.203	0.673	3
154	PF3D7_1016700	Plasmodium exported protein (PHISTc), unknown function	-0.109	-0.175	-0.095	-0.419	0.314	0.653	3
155	PF3D7_0906600	zinc finger protein, putative	-0.021	0.056	0.148	-0.174	-0.170	0.773	3
156	PF3D7_0301200	serine or threonine protein kinase, FIKK family (FIKK3)	-0.271	-0.178	-0.047	0.159	0.121	0.823	3
157	PF3D7_1416500	NADP-specific glutamate dehydrogenase (GDH1)	-0.125	0.006	-0.193	-0.042	0.092	0.724	3
158	PF3D7_0623800	tyrosine kinase-like protein, putative (TKL4)	-0.176	0.086	-0.062	-0.218	0.220	0.926	3
159	PF3D7_0114600	stevor, pseudogene	-0.049	0.601	0.027	-0.315	0.588	1.159	3
160	PF3D7_0830600	Plasmodium exported protein (PHISTc), unknown function	-0.213	-0.082	-0.219	-0.219	0.085	0.967	3
161	PF3D7_1114200	GTPase-activating protein, putative	-0.053	-0.108	-0.136	-0.195	0.011	0.696	3
162	PF3D7_0308500	activator of Hsp90 ATPase, putative	0.368	0.111	-0.077	-0.146	-0.388	0.725	3
163	PF3D7_0220700	Plasmodium exported protein (hyp9), unknown function	-0.096	0.181	-0.069	-0.169	-0.018	0.594	3
164	PF3D7_1302000	EMP1-trafficking protein (PTP6)	0.083	0.046	0.118	-0.294	-0.132	1.253	3
165	PF3D7_0520600	bis(5-nucleosyl)-tetraphosphatase, asymmetrical	0.193	0.121	0.001	0.002	-0.134	0.827	3
166	PF3D7_1225400	conserved Plasmodium protein, unknown function	-0.111	-0.321	-0.127	-0.254	-0.332	0.841	3
167	PF3D7_1465400	conserved Plasmodium protein, unknown function	0.010	0.035	-0.002	-0.121	0.209	0.889	3
168	PF3D7_0501300	skeleton-binding protein 1 (SBP1)	0.139	0.306	-0.127	-0.258	-0.006	1.226	3
169	PF3D7_1001100	conserved Plasmodium protein, unknown function	-0.069	-0.099	-0.009	0.046	0.211	1.059	3

170	PF3D7_0113300	Plasmodium exported protein (hyp1), unknown function	-0.089	-0.137	-0.487	-0.265	-0.004	1.080	3
171	PF3D7_1115800	conserved Plasmodium protein, unknown function	-0.402	-0.010	-0.127	-0.059	0.426	1.090	3
172	PF3D7_0301100	Plasmodium exported protein (hyp13), unknown function	-0.003	0.052	-0.041	-0.238	-0.068	0.856	3
173	PF3D7_1329500	conserved Plasmodium protein, unknown function	0.048	0.161	-0.105	-0.060	0.079	0.845	3
174	PF3D7_0902300	serine or threonine protein kinase, FIKK family (FIKK9.4)	-0.107	0.095	0.015	-0.312	0.133	0.585	3
175	PF3D7_0731200	Plasmodium exported protein, unknown function	-0.128	-0.054	-0.114	0.043	0.516	0.897	3
176	PF3D7_1127000	protein phosphatase, putative	0.104	0.327	-0.008	0.005	-0.185	0.602	3
177	PF3D7_1324400	MSF1-like protein, putative	-0.025	0.061	-0.086	0.095	0.068	0.877	3
178	PF3D7_0731100	EMP1-trafficking protein (PTP2)	0.125	-0.044	-0.096	-0.167	0.076	0.918	3
179	PF3D7_1102900	Plasmodium exported protein (hyp11), unknown function	-0.137	-0.107	-0.213	-0.245	0.320	1.073	3
180	PF3D7_0113900	Plasmodium exported protein (hyp8), unknown function (GEXP10)	0.062	0.184	-0.069	-0.140	0.167	0.789	3
181	PF3D7_0102700	merozoite-associated tryptophan-rich antigen (MaTrA)	-0.143	0.005	-0.111	-0.227	0.179	0.950	3
182	PF3D7_0202200	EMP1-trafficking protein (PTP1)	-0.043	0.165	-0.096	-0.179	-0.251	0.626	3
183	PF3D7_1213700	RNA polymerase subunit 8c, putative	-0.147	-0.144	-0.005	-0.259	-0.130	0.679	3
184	PF3D7_0113400	Plasmodium exported protein, unknown function	0.107	0.246	-0.687	-0.275	0.384	-0.048	4
185	PF3D7_1455300	conserved Plasmodium protein, unknown function	-0.181	-0.142	-0.189	-0.472	0.805	-0.122	4
186	PF3D7_0202400	gamete antigen 27 or 25, putative	-0.037	0.311	-0.065	-0.618	0.080	-0.212	4
187	PF3D7_1477800	acyl-CoA binding protein (ACBP)	-1.330	-0.938	-0.522	-0.427	-0.001	-0.844	4
188	PF3D7_0202100	liver stage associated protein 2 (LSAP2)	-0.614	-0.582	-0.331	-0.235	-0.064	0.155	4
189	PF3D7_1001500	early transcribed membrane protein 10.1 (ETRAP10)	-1.679	-1.256	-1.578	-1.428	-0.222	0.039	4
190	PF3D7_1410100	alpha or beta hydrolase, putative	-0.619	-1.039	-0.153	-0.074	-0.094	-0.484	4
191	PF3D7_1139100	RNA-binding protein, putative	-0.388	-0.564	-0.756	-0.333	-0.088	0.012	4
192	PF3D7_0201700	DnaJ protein, putative	-0.989	-1.051	-1.351	-0.780	-0.011	-0.356	4
193	PF3D7_0115000	surface-associated interspersed protein 1.3 (SURFIN 1.3) (SURF1.3)	-0.612	-0.731	0.241	-0.087	-0.139	-0.397	4
194	PF3D7_1413200	conserved Plasmodium protein, unknown function	-0.276	-0.854	-0.090	0.054	-0.096	-0.518	4
195	PF3D7_0900500	rifin (RIF)	-0.057	-0.744	0.103	-0.081	0.187	0.013	4
196	PF3D7_1467600	conserved Plasmodium protein, unknown function	-0.072	-0.260	-0.610	-0.163	0.018	0.033	4
197	PF3D7_0201800	knob associated heat shock protein 40 (KAHsp40)	-0.686	-0.602	-0.698	-1.088	-0.324	-0.002	4
198	PF3D7_1148700	Plasmodium exported protein (PHISTc), unknown function (GEXP12)	-0.173	-0.179	-0.540	-0.970	-0.258	-0.012	4
199	PF3D7_0617700	rifin (RIF)	0.218	-0.124	-0.074	-0.809	0.053	0.036	4
200	PF3D7_0324400	rifin (RIF)	0.283	0.324	-0.467	-1.018	-0.415	0.080	4

201	PF3D7_0424000	Plasmodium exported protein (PHISTc), unknown function	-0.207	0.132	-0.050	-0.888	-0.086	0.149	4
202	PF3D7_1476300	Plasmodium exported protein (PHISTb), unknown function	-0.196	-0.158	-0.161	-0.113	-0.347	-0.778	4

Supplementary Table 18. Differential mRNA abundance assessed by microarray and RT-qPCR in the *Not1.2* mutant. The differential mRNA abundance in the form of log₂-transformed mRNA abundance in the *Not1.2* mutant versus the control line is given. The differential mRNA abundance was determined by microarray and RT-qPCR. The total RNA was reversed transcribed into cDNA, amplified by SMART-PCR and hybridized to microarray chips. For RT-qPCR, the total RNA was reverse transcribed into cDNA, 20 ng of which served as the template in real-time PCR. The average value of the expression (ave) and the standard deviation (stdev) were calculated from technical triplicates (r1, r2 and r3).

	microarray					RT-qPCR				
	r1	r2	r3	ave	stdev	r1	r2	r3	ave	stdev
KAHRP_TP1	-0.17	-0.25	-0.25	-0.22	0.05	-1.43	-2.16	-1.86	-1.82	0.37
KAHRP_TP2	0.02	-0.22	-0.10	-0.10	0.12	-1.10	-0.90	0.20	-0.60	0.70
KAHRP_TP3	-0.48	-0.36	-0.51	-0.45	0.08	-1.06	-0.41	-0.73	-0.73	0.32
KAHRP_TP4	-0.57	-0.36	-0.53	-0.49	0.11	-0.94	-1.31	0.03	-0.74	0.69
KAHRP_TP5	-0.38	-0.47	-0.56	-0.47	0.09	-2.42	-1.87	-1.65	-1.98	0.40
KAHRP_TP6	-0.06	-0.04	-0.34	-0.15	0.17	-0.95	-0.41	-2.13	-1.17	0.88
EBA181_TP1	-0.34	-0.23	-0.35	-0.31	0.07	-1.18	-2.03	-1.65	-1.62	0.43
EBA181_TP5	0.77	0.62	0.81	0.73	0.10	2.42	1.87	1.65	1.98	0.40
MSP6_TP5	0.60	0.76	0.28	0.54	0.24	0.15	0.79	1.20	0.71	0.53
RH4_TP1	-0.32	-0.23	-0.32	-0.29	0.05	-0.69	-1.75	-1.48	-1.31	0.55
RH4_TP5	0.31	0.68	0.43	0.47	0.19	1.47	2.11	1.79	1.79	0.32
RH5_TP1	-0.20	-0.16	-0.21	-0.19	0.03	0.08	-3.88	-4.25	-2.69	2.40
RH5_TP5	0.36	0.42	0.50	0.43	0.07	2.06	2.33	1.54	1.97	0.40

Supplementary Table 19. 114 genes with significantly differential mRNA abundance at the same timepoint in the *Not1.1* mutant and the *Not1.2* mutant. The order of the genes corresponds to the heatmap in Figure 5.2. The order of the genes in the heatmap of the *Not1.1* mutant (**Figure 4.9**) is given in the lane “1.1KO”. The order of the genes in the heatmap of the *Not1.2* mutant (**Figure 4.23**) is given in the lane “1.2KD”.

No	Gene	Description	1.1KO	1.2KD
1	PF3D7_1035800	probable protein, unknown function (M712)	26	20
2	PF3D7_0713100	Pfmc-2TM Maurer's cleft two transmembrane protein (MC-2TM)	31	11
3	PF3D7_0532400	lysine-rich membrane-associated PHISTb protein (LyMP)	33	9
4	PF3D7_1201000	Plasmodium exported protein (PHISTb), unknown function	36	12
5	PF3D7_0301700	Plasmodium exported protein, unknown function	37	18
6	PF3D7_0114500	Plasmodium exported protein (hyp10), unknown function	38	3
7	PF3D7_0402000	Plasmodium exported protein (PHISTa), unknown function	41	7
8	PF3D7_1223900	50S ribosomal protein L24, putative	46	8
9	PF3D7_0832200	conserved Plasmodium protein, unknown function	48	6
10	PF3D7_0726100	Plasmodium exported protein, unknown function	59	108
11	PF3D7_0424800	Plasmodium exported protein (PHISTb), unknown function	60	1
12	PF3D7_0409500	conserved Plasmodium protein, unknown function	143	28
13	PF3D7_1405600	ribonucleotide reductase small subunit (RNR)	145	25
14	PF3D7_1432700	protein-L-isoaspartate(D-aspartate) O-methyltransferase, putative	209	65
15	PF3D7_0904700	bacterial histone-like protein (HU)	214	31
16	PF3D7_1137500	apicoplast ribosomal protein S14p or S29e precursor, putative	217	34
17	PF3D7_1410700	conserved Plasmodium protein, unknown function	223	26
18	PF3D7_1028600	conserved Plasmodium protein, unknown function	231	42
19	PF3D7_0828500	translation initiation factor EIF-2b alpha subunit, putative	233	32
20	PF3D7_1304000	chromosome condensation protein, putative	241	29
21	PF3D7_1033200	early transcribed membrane protein 10.2 (ETRAMP10.2)	261	5
22	PF3D7_1476300	Plasmodium exported protein (PHISTb), unknown function	264	202
23	PF3D7_0930300	merozoite surface protein 1 (MSP1)	348	57
24	PF3D7_0423500	glideosome associated protein with multiple membrane spans 2 (GAPM2)	351	60
25	PF3D7_0932200	profilin, putative (PFN)	355	30
26	PF3D7_1001600	alpha or beta hydrolase, putative	366	45
27	PF3D7_0501600	rhoptry-associated protein 2 (RAP2)	373	39
28	PF3D7_1140400	conserved Plasmodium protein, unknown function	374	73
29	PF3D7_1461400	conserved Plasmodium protein, unknown function	379	68
30	PF3D7_1235100	conserved Plasmodium protein, unknown function	385	44
31	PF3D7_0207400	serine repeat antigen 7 (SERA7)	388	53
32	PF3D7_0613800	transcription factor with AP2 domain(s) (ApiAP2)	390	52
33	PF3D7_0604100	transcription factor with AP2 domain(s) (SIP2)	393	43
34	PF3D7_1143200	DnaJ protein, putative	395	58

35	PF3D7_1228600	merozoite surface protein 9 (MSP9)	399	37
36	PF3D7_1364100	6-cysteine protein (P92)	405	62
37	PF3D7_0929400	high molecular weight rhoptry protein 2 (RhopH2)	410	41
38	PF3D7_1035100	probable protein, unknown function	411	40
39	PF3D7_0704500	serine or threonine protein kinase, putative	420	51
40	PF3D7_1344500	U6 snRNA phosphodiesterase, putative (USB1)	425	46
41	PF3D7_1033900	ubiquitin-conjugating enzyme, putative	430	33
42	PF3D7_1334600	MSP7-like protein (MSRP3)	439	49
43	PF3D7_0106200	conserved Plasmodium protein, unknown function	445	55
44	PF3D7_1245500	conserved Plasmodium protein, unknown function	446	54
45	PF3D7_1331600	protein tyrosine phosphatase-like protein, putative (PTPLA)	463	36
46	PF3D7_0321400	protein kinase, putative	469	71
47	PF3D7_0217500	calcium-dependent protein kinase 1 (CDPK1)	470	70
48	PF3D7_0102500	erythrocyte binding antigen-181 (EBA181)	487	83
49	PF3D7_0731500	erythrocyte binding antigen-175 (EBA175)	491	75
50	PF3D7_0822900	conserved Plasmodium protein, unknown function	493	89
51	PF3D7_1468400	zinc finger protein, putative (D13)	497	69
52	PF3D7_1251200	coronin	499	98
53	PF3D7_0102200	ring-infected erythrocyte surface antigen (RESA)	527	90
54	PF3D7_1149200	ring-infected erythrocyte surface antigen	528	99
55	PF3D7_1201200	Plasmodium exported protein (PHISTa-like), unknown function	529	78
56	PF3D7_0805200	gamete release protein, putative (GAMER)	532	97
57	PF3D7_1302200	early transcribed membrane protein 13 (ETRAPM13)	536	130
58	PF3D7_1337800	calcium-dependent protein kinase 5 (CDPK5)	545	77
59	PF3D7_1036500	probable protein, unknown function	549	82
60	PF3D7_1026600	conserved Plasmodium protein, unknown function	550	85
61	PF3D7_1125800	kelch protein, putative	551	84
62	PF3D7_0911100	conserved Plasmodium protein, unknown function	552	86
63	PF3D7_0934600	conserved Plasmodium protein, unknown function	553	100
64	PF3D7_0402100	Plasmodium exported protein (PHISTb), unknown function	554	79
65	PF3D7_1126700	autophagy-related protein 23, putative (ATG23)	556	72
66	PF3D7_1204900	probable protein, unknown function	557	80
67	PF3D7_0903600	conserved Plasmodium protein, unknown function	558	67
68	PF3D7_0214600	serine or threonine protein kinase, putative	559	92
69	PF3D7_1401600	Plasmodium exported protein (PHISTb), unknown function	561	88
70	PF3D7_0414900	armadillo-domain containing rhoptry protein (ARO)	575	81
71	PF3D7_1125700	kelch protein, putative	577	96
72	PF3D7_1351700	inner membrane complex protein 1f, putative (IMC1f)	578	93
73	PF3D7_1455300	conserved Plasmodium protein, unknown function	579	185
74	PF3D7_0318700	trophozoite stage antigen	580	95
75	PF3D7_0913800	conserved Plasmodium protein, unknown function	582	94
76	PF3D7_0308300	conserved Plasmodium protein, unknown function	587	87
77	PF3D7_1016800	Plasmodium exported protein (PHISTc), unknown function	590	76

78	PF3D7_1001100	conserved Plasmodium protein, unknown function	611	169
79	PF3D7_1002100	EMP1-trafficking protein (PTP5)	612	104
80	PF3D7_1407800	plasmepsin IV (PM4)	615	136
81	PF3D7_0830500	tryptophan or threonine-rich antigen (TryThrA)	617	74
82	PF3D7_1002000	Plasmodium exported protein (hyp2), unknown function	627	150
83	PF3D7_1401100	DnaJ protein, putative	628	117
84	PF3D7_0424600	Plasmodium exported protein (PHISTb), unknown function	629	107
85	PF3D7_0501100	conserved Plasmodium protein, unknown function	630	126
86	PF3D7_1329500	conserved Plasmodium protein, unknown function	631	173
87	PF3D7_1001500	early transcribed membrane protein 10.1 (ETRAMP10)	632	189
88	PF3D7_0501300	skeleton-binding protein 1 (SBP1)	633	168
89	PF3D7_1370300	membrane associated histidine-rich protein (MAHRP1)	634	103
90	PF3D7_0113300	Plasmodium exported protein (hyp1), unknown function	635	170
91	PF3D7_0731100	EMP1-trafficking protein (PTP2)	637	178
92	PF3D7_0501400	interspersed repeat antigen (FIRA)	638	91
93	PF3D7_1116800	heat shock protein 101 (HSP101)	639	114
94	PF3D7_1436300	translocon component PTEX150 (PTEX150)	640	113
95	PF3D7_0623800	tyrosine kinase-like protein, putative (TKL4)	641	158
96	PF3D7_0108500	conserved Plasmodium protein, unknown function	643	127
97	PF3D7_1115800	conserved Plasmodium protein, unknown function	644	171
98	PF3D7_0424500	serine or threonine protein kinase, FIKK family (FIKK4.1)	650	147
99	PF3D7_1404800	conserved Plasmodium protein, unknown function	651	131
100	PF3D7_1353200	membrane associated histidine-rich protein (MAHRP2)	652	110
101	PF3D7_1372200	histidine-rich protein III (HRPIII)	654	123
102	PF3D7_1114200	GTPase-activating protein, putative	658	161
103	PF3D7_0102700	merozoite-associated tryptophan-rich antigen (MaTrA)	660	181
104	PF3D7_0219700	Plasmodium exported protein (PHISTc), unknown function (GEXP20)	661	116
105	PF3D7_1001900	Plasmodium exported protein (hyp16), unknown function (PfJ23)	664	119
106	PF3D7_1302000	EMP1-trafficking protein (PTP6)	667	164
107	PF3D7_0220700	Plasmodium exported protein (hyp9), unknown function	668	163
108	PF3D7_0830600	Plasmodium exported protein (PHISTc), unknown function	669	160
109	PF3D7_1139100	RNA-binding protein, putative	678	191
110	PF3D7_0202100	liver stage associated protein 2 (LSAP2)	685	188
111	PF3D7_0201700	DnaJ protein, putative	686	192
112	PF3D7_0201800	knob associated heat shock protein 40 (KAHsp40)	689	197
113	PF3D7_1148700	Plasmodium exported protein (PHISTc), unknown function (GEXP12)	693	198
114	PF3D7_0324400	rifin (RIF)	697	200

References

1. WHO, *World malaria report*. World health organization, Geneva, Switzerland, 2017.
2. Mutabingwa, T.K., *Artemisinin-based combination therapies (ACTs): Best hope for malaria treatment but inaccessible to the needy!* Acta Tropica, 2005. **95**(3): p. 305-315.
3. Dondorp, A.M., et al., *Artemisinin resistance in Plasmodium falciparum malaria*. The New England journal of medicine, 2009. **361**(5): p. 455-467.
4. Pasvol, G., *Protective hemoglobinopathies and Plasmodium falciparum transmission*. Nature Genetics, 2010. **42**: p. 284.
5. Cowman, A.F. and B.S. Crabb, *Invasion of Red Blood Cells by Malaria Parasites*. Cell, 2006. **124**(4): p. 755-766.
6. Bannister, L.H., et al., *A Brief Illustrated Guide to the Ultrastructure of Plasmodium falciparum Asexual Blood Stages*. Parasitology Today, 2000. **16**(10): p. 427-433.
7. Baldwin, M.R., et al., *Merozoite surface protein 1 recognition of host glycophorin A mediates malaria parasite invasion of red blood cells*. Blood, 2015. **125**(17): p. 2704-2711.
8. Collins, C.R., et al., *Malaria parasite cGMP-dependent protein kinase regulates blood stage merozoite secretory organelle discharge and egress*. PLoS pathogens, 2013. **9**(5): p. e1003344-e1003344.
9. Ruecker, A., et al., *Proteolytic activation of the essential parasitophorous vacuole cysteine protease SERA6 accompanies malaria parasite egress from its host erythrocyte*. The Journal of biological chemistry, 2012. **287**(45): p. 37949-37963.
10. Thomas, J.A., et al., *A protease cascade regulates release of the human malaria parasite Plasmodium falciparum from host red blood cells*. Nature microbiology, 2018. **3**(4): p. 447-455.
11. Das, S., et al., *Processing of Plasmodium falciparum Merozoite Surface Protein MSP1 Activates a Spectrin-Binding Function Enabling Parasite Egress from RBCs*. Cell Host & Microbe, 2015. **18**(4): p. 433-444.
12. Iyer, G.R., et al., *Calcium-dependent phosphorylation of Plasmodium falciparum serine repeat antigen-5 triggers merozoite egress*. Journal of Biological Chemistry, 2018.
13. Cowman, A.F., et al., *The Molecular Basis of Erythrocyte Invasion by Malaria Parasites*. Cell Host & Microbe, 2017. **22**(2): p. 232-245.
14. Lopaticki, S., et al., *Reticulocyte and erythrocyte binding-like proteins function cooperatively in invasion of human erythrocytes by malaria parasites*. Infection and immunity, 2011. **79**(3): p. 1107-1117.
15. Volz, Jennifer C., et al., *Essential Role of the PfRh5/PfRipr/CyRPA Complex during Plasmodium falciparum Invasion of Erythrocytes*. Cell Host & Microbe, 2016. **20**(1): p. 60-71.
16. Bartholdson, S.J., et al., *Semaphorin-7A is an erythrocyte receptor for P. falciparum merozoite-specific TRAP homolog, MTRAP*. PLoS pathogens, 2012. **8**(11): p. e1003031-e1003031.
17. Gilson, P.R. and B.S. Crabb, *Morphology and kinetics of the three distinct phases of red blood cell invasion by Plasmodium falciparum merozoites*. International Journal for Parasitology, 2009. **39**(1): p. 91-96.
18. Mitchell, G.H., et al., *Apical membrane antigen 1, a major malaria vaccine candidate, mediates the close attachment of invasive merozoites to host red blood cells*. Infection and immunity, 2004. **72**(1): p. 154-158.
19. Riglar, D.T., et al., *Super-Resolution Dissection of Coordinated Events during Malaria Parasite Invasion of the Human Erythrocyte*. Cell Host & Microbe, 2011. **9**(1): p. 9-20.

20. Frénal, K., et al., *Functional Dissection of the Apicomplexan Glideosome Molecular Architecture*. Cell Host & Microbe, 2010. **8**(4): p. 343-357.
21. Pal-Bhowmick, I., et al., *Binding of aldolase and glyceraldehyde-3-phosphate dehydrogenase to the cytoplasmic tails of Plasmodium falciparum merozoite duffy binding-like and reticulocyte homology ligands*. mBio, 2012. **3**(5): p. e00292-12.
22. Boucher, L.E. and J. Bosch, *The apicomplexan glideosome and adhesins - Structures and function*. Journal of structural biology, 2015. **190**(2): p. 93-114.
23. Klemmba, M., et al., *Trafficking of plasmepsin II to the food vacuole of the malaria parasite Plasmodium falciparum*. The Journal of cell biology, 2004. **164**(1): p. 47-56.
24. Dluzewski, A.R., et al., *Formation of the food vacuole in Plasmodium falciparum: a potential role for the 19 kDa fragment of merozoite surface protein 1 (MSP1(19))*. PloS one, 2008. **3**(8): p. e3085-e3085.
25. Milani, K.J., T.G. Schneider, and T.F. Taraschi, *Defining the Morphology and Mechanism of the Hemoglobin Transport Pathway in Plasmodium falciparum-Infected Erythrocytes*. Eukaryotic Cell, 2015. **14**(4): p. 415-426.
26. Chugh, M., et al., *Protein complex directs hemoglobin-to-hemozoin formation in Plasmodium falciparum*. Proceedings of the National Academy of Sciences, 2013. **110**(14): p. 5392-5397.
27. Römisch, K., *Diversion at the ER: How Plasmodium falciparum exports proteins into host erythrocytes*. F1000Research, 2012. **1**: p. 12-12.
28. Trelka, D.P., et al., *Evidence for vesicle-mediated trafficking of parasite proteins to the host cell cytosol and erythrocyte surface membrane in Plasmodium falciparum infected erythrocytes*. Molecular and Biochemical Parasitology, 2000. **106**(1): p. 131-145.
29. Lauer, S.A., et al., *A Membrane Network for Nutrient Import in Red Cells Infected with the Malaria Parasite*. Science, 1997. **276**(5315): p. 1122-1125.
30. Mitra, P., A.S. Deshmukh, and S.K. Dhar, *DNA replication during intra-erythrocytic stages of human malarial parasite Plasmodium falciparum*. Current Science (00113891), 2012. **102**(5): p. 725-740.
31. Gerald, N., B. Mahajan, and S. Kumar, *Mitosis in the Human Malaria Parasite Plasmodium falciparum*. Eukaryotic Cell, 2011. **10**(4): p. 474-482.
32. Francia, M.E. and B. Striepen, *Cell division in apicomplexan parasites*. Nat Rev Micro, 2014. **12**(2): p. 125-136.
33. Blackman, M.J., *Malarial proteases and host cell egress: an 'emerging' cascade*. Cellular Microbiology, 2008. **10**(10): p. 1925-1934.
34. Gardner, M.J., et al., *Genome sequence of the human malaria parasite Plasmodium falciparum*. Nature, 2002. **419**(6906): p. 498-511.
35. Bozdech, Z., et al., *The Transcriptome of the Intraerythrocytic Developmental Cycle of Plasmodium falciparum*. PLOS Biology, 2003. **1**(1): p. e5.
36. Ay, F., et al., *Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression*. Genome Research, 2014.
37. Hoeijmakers, W.A.M., et al., *Plasmodium falciparum centromeres display a unique epigenetic makeup and cluster prior to and during schizogony*. Cellular Microbiology, 2012. **14**(9): p. 1391-1401.
38. Duffy, M.F., et al., *The role of chromatin in Plasmodium gene expression*. Cellular Microbiology, 2012. **14**(6): p. 819-828.
39. Salcedo-Amaya, A.M., et al., *Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of Plasmodium falciparum*. Proceedings of the National Academy of Sciences, 2009. **106**(24): p. 9655-9660.

40. Salcedo-Amaya, A.M., et al., *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, 2009. **106**(24): p. 9655-9660.
41. Flueck, C., et al., *Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors*. PLoS pathogens, 2009. **5**(9): p. e1000569-e1000569.
42. Lopez-Rubio, J.-J., L. Mancio-Silva, and A. Scherf, *Genome-wide Analysis of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear Repressive Centers in Malaria Parasites*. Cell Host & Microbe, 2009. **5**(2): p. 179-190.
43. Batugedara, G. and K.G. Le Roch, *Unraveling the 3D genome of human malaria parasites*. Seminars in Cell & Developmental Biology, 2019. **90**: p. 144-153.
44. Chookajorn, T., et al., *Epigenetic memory at malaria virulence genes*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(3): p. 899-902.
45. Crowley, V.M., et al., *Heterochromatin formation in bistable chromatin domains controls the epigenetic repression of clonally variant Plasmodium falciparum genes linked to erythrocyte invasion*. Molecular Microbiology, 2011. **80**(2): p. 391-406.
46. Brancucci, Nicolas M.B., et al., *Heterochromatin Protein 1 Secures Survival and Transmission of Malaria Parasites*. Cell Host & Microbe, 2014. **16**(2): p. 165-176.
47. Bunnik, E.M., et al., *Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages*. Nature Communications, 2018. **9**(1): p. 1910.
48. Ponts, N., et al., *Nucleosome occupancy at transcription start sites in the human malaria parasite: a hard-wired evolution of virulence?* Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 2011. **11**(4): p. 716-724.
49. Bunnik, E.M., et al., *DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite Plasmodium falciparum*. BMC Genomics, 2014. **15**(1): p. 347.
50. Westenberger, S.J., et al., *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, 2009. **10**: p. 610-610.
51. Ponts, N., et al., *Nucleosome landscape and control of transcription in the human malaria parasite*. Genome research, 2010. **20**(2): p. 228-238.
52. Miao, J., et al., *The malaria parasite Plasmodium falciparum histones: Organization, expression, and acetylation*. Gene, 2006. **369**: p. 53-65.
53. Bártfai, R., et al., *H2A.Z Demarcates Intergenic Regions of the Plasmodium falciparum Epigenome That Are Dynamically Marked by H3K9ac and H3K4me3*. PLOS Pathogens, 2010. **6**(12): p. e1001223.
54. Petter, M., et al., *Expression of P. falciparum var Genes Involves Exchange of the Histone Variant H2A.Z at the Promoter*. PLOS Pathogens, 2011. **7**(2): p. e1001292.
55. Hoeijmakers, W.A.M., et al., *H2A.Z/H2B.Z double-variant nucleosomes inhabit the AT-rich promoter regions of the Plasmodium falciparum genome*. Molecular microbiology, 2013. **87**(5): p. 1061-1073.
56. Petter, M., et al., *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, 2013. **87**(6): p. 1167-1182.
57. Trelle, M.B., et al., *Global Histone Analysis by Mass Spectrometry Reveals a High Content of Acetylated Lysine Residues in the Malaria Parasite Plasmodium falciparum*. Journal of Proteome Research, 2009. **8**(7): p. 3439-3450.

58. Saraf, A., et al., *Dynamic and Combinatorial Landscape of Histone Modifications during the Intraerythrocytic Developmental Cycle of the Malaria Parasite*. Journal of proteome research, 2016. **15**(8): p. 2787-2801.
59. Bernstein, B.E., et al., *Genomic Maps and Comparative Analysis of Histone Modifications in Human and Mouse*. Cell, 2005. **120**(2): p. 169-181.
60. Hiromi Nishida, T.S., Shinji Kondo, Hisashi Miura, Yu-ichi Fujimura, Yoshihide Hayashizaki, *Histone H3 acetylated at lysine 9 in promoter is associated with low nucleosome density in the vicinity of transcription start site in human cell*. Chromosome Research, 2006. **14**(2): p. 203-211.
61. Cui, L., et al., *PfGCN5-mediated histone H3 acetylation plays a key role in gene expression in Plasmodium falciparum*. Eukaryotic cell, 2007. **6**(7): p. 1219-1227.
62. Gopalakrishnan, A.M., et al., *Plasmodium falciparum: Preinitiation complex occupancy of active and inactive promoters during erythrocytic stage*. Experimental Parasitology, 2009. **121**(1): p. 46-54.
63. Gupta, A.P., et al., *Dynamic Epigenetic Regulation of Gene Expression during the Life Cycle of Malaria Parasite Plasmodium falciparum*. PLoS Pathogens, 2013. **9**(2): p. e1003170.
64. Batugedara, G., et al., *The Role of Chromatin Structure in Gene Regulation of the Human Malaria Parasite*. Trends in parasitology, 2017. **33**(5): p. 364-377.
65. Scherf, A., J.J. Lopez-Rubio, and L. Riviere, *Antigenic Variation in Plasmodium falciparum*. Annual Review of Microbiology, 2008. **62**(1): p. 445-470.
66. Lopez-Rubio, J.J., et al., *5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites*. Molecular microbiology, 2007. **66**(6): p. 1296-1305.
67. Volz, Jennifer C., et al., *PfSET10, a Plasmodium falciparum Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division*. Cell Host & Microbe, 2012. **11**(1): p. 7-18.
68. French, J.B., Y. Cen, and A.A. Sauve, *Plasmodium falciparum Sir2 is an NAD⁺-dependent deacetylase and an acetyllysine-dependent and acetyllysine-independent NAD⁺ glycohydrolase*. Biochemistry, 2008. **47**(38): p. 10227-10239.
69. Epp, C., et al., *Chromatin associated sense and antisense noncoding RNAs are transcribed from the var gene family of virulence genes of the malaria parasite Plasmodium falciparum*. RNA (New York, N.Y.), 2009. **15**(1): p. 116-127.
70. Amit-Avraham, I., et al., *Antisense long noncoding RNAs regulate var gene activation in the malaria parasite Plasmodium falciparum*. Proceedings of the National Academy of Sciences of the United States of America, 2015. **112**(9): p. E982-E991.
71. Josling, G.A., et al., *The role of bromodomain proteins in regulating gene expression*. Genes, 2012. **3**(2): p. 320-343.
72. Josling, Gabrielle A., et al., *A Plasmodium Falciparum Bromodomain Protein Regulates Invasion Gene Expression*. Cell Host & Microbe, 2015. **17**(6): p. 741-751.
73. Jiang, L., et al., *Epigenetic control of the variable expression of a Plasmodium falciparum receptor protein for erythrocyte invasion*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(5): p. 2224-2229.
74. Comeaux, C.A., et al., *Functional analysis of epigenetic regulation of tandem RhopH1/clag genes reveals a role in Plasmodium falciparum growth*. Molecular microbiology, 2011. **80**(2): p. 378-390.
75. Kafsack, B.F.C., et al., *A transcriptional switch underlies commitment to sexual development in malaria parasites*. Nature, 2014. **507**(7491): p. 248-252.
76. Lanzer, M., et al., *Plasmodium: Control of Gene Expression in Malaria Parasites*. Experimental Parasitology, 1993. **77**(1): p. 121-128.

77. Callebaut, I., et al., *Prediction of the general transcription factors associated with RNA polymerase II in Plasmodium falciparum: conserved features and differences relative to other eukaryotes*. BMC Genomics, 2005. **6**: p. 100-100.
78. Lee, T.I. and R.A. Young, *Transcription of Eukaryotic Protein-Coding Genes*. Annual Review of Genetics, 2000. **34**(1): p. 77-137.
79. Coulson, R.M.R., N. Hall, and C.A. Ouzounis, *Comparative Genomics of Transcriptional Control in the Human Malaria Parasite Plasmodium falciparum*. Genome Research, 2004. **14**(8): p. 1548-1554.
80. Kornberg, R.D., *Mediator and the mechanism of transcriptional activation*. Trends in Biochemical Sciences. **30**(5): p. 235-239.
81. Bourbon, H.-M., *Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex*. Nucleic Acids Research, 2008. **36**(12): p. 3993-4008.
82. Campbell, T.L., et al., *Identification and Genome-Wide Prediction of DNA Binding Specificities for the ApiAP2 Family of Regulators from the Malaria Parasite*. PLoS Pathogens, 2010. **6**(10): p. e1001165.
83. De Silva, E.K., et al., *Specific DNA-binding by Apicomplexan AP2 transcription factors*. Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(24): p. 8393-8398.
84. Cho, E.-J., *RNA polymerase II carboxy-terminal domain with multiple connections*. Experimental & Molecular Medicine, 2007. **39**: p. 247.
85. Rai, R., et al., *Genome-wide analysis in Plasmodium falciparum reveals early and late phases of RNA polymerase II occupancy during the infectious cycle*. BMC genomics, 2014. **15**(1): p. 959-959.
86. Sims, J.S., et al., *Patterns of gene-specific and total transcriptional activity during the Plasmodium falciparum intraerythrocytic developmental cycle*. Eukaryotic cell, 2009. **8**(3): p. 327-338.
87. Lu, X.M., et al., *Nascent RNA sequencing reveals mechanisms of gene regulation in the human malaria parasite Plasmodium falciparum*. Nucleic acids research, 2017. **45**(13): p. 7825-7840.
88. Painter, H.J., M. Carrasquilla, and M. Llinás, *Capturing in vivo RNA transcriptional dynamics from the malaria parasite Plasmodium falciparum*. Genome Research, 2017.
89. Painter, H.J., et al., *Genome-wide real-time in vivo transcriptional dynamics during Plasmodium falciparum blood-stage development*. Nature Communications, 2018. **9**(1): p. 2656.
90. Hughes, K.R., et al., *From cradle to grave: RNA biology in malaria parasites*. Wiley Interdisciplinary Reviews - RNA, 2010. **1**(2): p. 287-303.
91. Otto, T.D., et al., *New insights into the blood-stage transcriptome of Plasmodium falciparum using RNA-Seq*. Molecular Microbiology, 2010. **76**(1): p. 12-24.
92. Singh, P.K., et al., *Plasmodium falciparum Prp16 homologue and its role in splicing*. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 2012. **1819**(11): p. 1186-1199.
93. Shankar, J., A. Pradhan, and R. Tuteja, *Isolation and characterization of Plasmodium falciparum UAP56 homolog: Evidence for the coupling of RNA binding and splicing activity by site-directed mutations*. Archives of Biochemistry and Biophysics, 2008. **478**(2): p. 143-153.
94. Eshar, S., et al., *PfSR1 controls alternative splicing and steady-state RNA levels in Plasmodium falciparum through preferential recognition of specific RNA motifs*. Molecular Microbiology, 2015. **96**(6): p. 1283-1297.

95. Ladd, A.N., N. Charlet, and T.A. Cooper, *The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing*. Molecular and cellular biology, 2001. **21**(4): p. 1285-1296.
96. Wongsombat, C., et al., *Molecular characterization of Plasmodium falciparum Bruno/CELF RNA binding proteins* Author. Vol. 198. 2014. 1-10.
97. Reddy, B.P.N., et al., *A bioinformatic survey of RNA-binding proteins in Plasmodium*. BMC genomics, 2015. **16**: p. 890-890.
98. Bunnik, E.M., et al., *The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum*. Genome Biology, 2016. **17**(1): p. 147.
99. Shock, J.L., K.F. Fischer, and J.L. DeRisi, *Whole-genome analysis of mRNA decay in Plasmodium falciparum reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle*. Genome biology, 2007. **8**(7): p. R134-R134.
100. Chêne, A., et al., *PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites*. Nucleic acids research, 2012. **40**(7): p. 3066-3077.
101. Vembar, S.S., et al., *The PfAlba1 RNA-binding protein is an important regulator of translational timing in Plasmodium falciparum blood stages*. Genome Biology, 2015. **16**: p. 212.
102. Muñoz, E.E., et al., *ALBA4 modulates its stage-specific interactions and specific mRNA fates during Plasmodium yoelii growth and transmission*. Molecular microbiology, 2017. **106**(2): p. 266-284.
103. Balu, B., et al., *CCR4-associated factor 1 coordinates the expression of Plasmodium falciparum egress and invasion proteins*. Eukaryotic cell, 2011. **10**(9): p. 1257-1263.
104. Le Roch, K.G., et al., *Global analysis of transcript and protein levels across the Plasmodium falciparum life cycle*. Genome Research, 2004. **14**(11): p. 2308-2318.
105. Foth, B.J., et al., *Quantitative protein expression profiling reveals extensive post-transcriptional regulation and post-translational modifications in schizont-stage malaria parasites*. Genome Biology, 2008. **9**(12): p. R177-R177.
106. Jackson, K.E., et al., *Protein translation in Plasmodium parasites*. Trends in Parasitology, 2011. **27**(10): p. 467-476.
107. Shaw, P.J., et al., *Characterization of human malaria parasite Plasmodium falciparum eIF4E homologue and mRNA 5' cap status*. Molecular and Biochemical Parasitology, 2007. **155**(2): p. 146-155.
108. Tuteja, R. and A. Pradhan, *Isolation and functional characterization of eIF4F components and poly(A)-binding protein from Plasmodium falciparum*. Parasitology International, 2009. **58**(4): p. 481-485.
109. Tuteja, R., *Identification and bioinformatics characterization of translation initiation complex eIF4F components and poly(A)-binding protein from Plasmodium falciparum*. Communicative & integrative biology, 2009. **2**(3): p. 245-260.
110. Tarique, M., et al., *Plasmodium falciparum DOZI, an RNA helicase interacts with eIF4E*. Gene, 2013. **522**(1): p. 46-59.
111. Calvo, S.E., D.J. Pagliarini, and V.K. Mootha, *Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(18): p. 7507-7512.
112. Mok, B.W., et al., *Default Pathway of var2csa Switching and Translational Repression in Plasmodium falciparum*. PLoS ONE, 2008. **3**(4): p. e1982.
113. Mair, G.R., et al., *Universal Features of Post-Transcriptional Gene Regulation Are Critical for Plasmodium Zygote Development*. PLOS Pathogens, 2010. **6**(2): p. e1000767.
114. Guerreiro, A., et al., *Genome-wide RIP-Chip analysis of translational repressor-bound mRNAs in the Plasmodium gametocyte*. Genome Biology, 2014. **15**(11): p. 493.

115. Mair, G.R., et al., *Regulation of sexual development of Plasmodium by translational repression*. Science (New York, N.Y.), 2006. **313**(5787): p. 667-669.
116. Zhang, M., et al., *The Plasmodium eukaryotic initiation factor-2alpha kinase IK2 controls the latency of sporozoites in the mosquito salivary glands*. The Journal of experimental medicine, 2010. **207**(7): p. 1465-1474.
117. Miao, J., et al., *Puf Mediates Translation Repression of Transmission-Blocking Vaccine Candidates in Malaria Parasites*. PLOS Pathogens, 2013. **9**(4): p. e1003268.
118. Silva, P.A.G.C., et al., *Translational Control of UIS4 Protein of the Host-Parasite Interface Is Mediated by the RNA Binding Protein Puf2 in Plasmodium berghei Sporozoites*. PLoS one, 2016. **11**(1): p. e0147940-e0147940.
119. Zhang, M., et al., *UIS2: A Unique Phosphatase Required for the Development of Plasmodium Liver Stages*. PLOS Pathogens, 2016. **12**(1): p. e1005370.
120. Müller, K., K. Matuschewski, and O. Silvie, *The Puf-family RNA-binding protein Puf2 controls sporozoite conversion to liver stages in the malaria parasite*. PLoS one, 2011. **6**(5): p. e19860-e19860.
121. *PfSET10, a Plasmodium falciparum Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division*. Cell Host & Microbe, 2012(1): p. 7.
122. Deshmukh, A.S., et al., *Regulation of Plasmodium falciparum Origin Recognition Complex subunit 1 (PfORC1) function through phosphorylation mediated by CDK-like kinase PK5*. Molecular Microbiology, 2015(1): p. 17.
123. Ponts, N., et al., *Deciphering the Ubiquitin-Mediated Pathway in Apicomplexan Parasites: A Potential Strategy to Interfere with Parasite Virulence*. PLoS ONE, 2008. **3**(6): p. e2386.
124. Reiter, K., et al., *Identification of Biochemically Distinct Properties of the Small Ubiquitin-related Modifier (SUMO) Conjugation Pathway in Plasmodium falciparum*. The Journal of Biological Chemistry, 2013. **288**(39): p. 27724-27736.
125. Collart, M.A. and O.O. Panasenko, *The Ccr4-Not complex*. Gene, 2012. **492**(1): p. 42-53.
126. Collart, M.A. and K. Struhl, *CDC39, an essential nuclear protein that negatively regulates transcription and differentially affects the constitutive and inducible HIS3 promoters*. The EMBO journal, 1993. **12**(1): p. 177-186.
127. Tucker, M., et al., *The Transcription Factor Associated Ccr4 and Caf1 Proteins Are Components of the Major Cytoplasmic mRNA Deadenylation in Saccharomyces cerevisiae*. Cell, 2001. **104**(3): p. 377-386.
128. Teixeira, D. and R. Parker, *Analysis of P-body assembly in Saccharomyces cerevisiae*. Molecular biology of the cell, 2007. **18**(6): p. 2274-2287.
129. Panasenko, O.O. and M.A. Collart, *Presence of Not5 and ubiquitinated Rps7A in polysome fractions depends upon the Not4 E3 ligase*. Molecular Microbiology, 2012. **83**(3): p. 640-653.
130. Bai, Y., et al., *The CCR4 and CAF1 Proteins of the CCR4-NOT Complex Are Physically and Functionally Separated from NOT2, NOT4, and NOT5*. Molecular and Cellular Biology, 1999. **19**(10): p. 6642-6651.
131. Liu, H.Y., et al., *The NOT proteins are part of the CCR4 transcriptional complex and affect gene expression both positively and negatively*. The EMBO journal, 1998. **17**(4): p. 1096-1106.
132. Chen, J., et al., *Purification and characterization of the 1.0 MDa CCR4-NOT complex identifies two novel components of the complex*¹¹Edited by D. Draper. Journal of Molecular Biology, 2001. **314**(4): p. 683-694.
133. Morel, A.P., et al., *BTG2 antiproliferative protein interacts with the human CCR4 complex existing in vivo in three cell-cycle-regulated forms*. Journal of Cell Science, 2003. **116**(14): p. 2929-2936.

134. Collart, M.A., *Global control of gene expression in yeast by the Ccr4-Not complex*. Gene, 2003. **313**: p. 1-16.
135. Mauxion, F., B. Prève, and B. Séraphin, *C2ORF29/CNOT11 and CNOT10 form a new module of the CCR4-NOT complex*. RNA biology, 2013. **10**(2): p. 267-276.
136. Temme, C., M. Simonelig, and E. Wahle, *Deadenylation of mRNA by the CCR4-NOT complex in Drosophila: molecular and developmental aspects*. Frontiers in genetics, 2014. **5**: p. 143-143.
137. Mittal, S., et al., *The Ccr4a (CNOT6) and Ccr4b (CNOT6L) deadenylase subunits of the human Ccr4-Not complex contribute to the prevention of cell death and senescence*. Molecular biology of the cell, 2011. **22**(6): p. 748-758.
138. Morita, M., et al., *Depletion of mammalian CCR4b deadenylase triggers elevation of the p27Kip1 mRNA level and impairs cell growth*. Molecular and cellular biology, 2007. **27**(13): p. 4980-4990.
139. Aslam, A., et al., *The Ccr4-NOT deadenylase subunits CNOT7 and CNOT8 have overlapping roles and modulate cell proliferation*. Molecular biology of the cell, 2009. **20**(17): p. 3840-3850.
140. Lau, N.-C., et al., *Human Ccr4-Not complexes contain variable deadenylase subunits*. Biochemical Journal, 2009. **422**(3): p. 443-453.
141. Nasertorabi, F., et al., *Insights into the structure of the CCR4-NOT complex by electron microscopy*. FEBS letters, 2011. **585**(14): p. 2182-2186.
142. Basquin, J., et al., *Architecture of the Nuclease Module of the Yeast Ccr4-Not Complex: the Not1-Caf1-Ccr4 Interaction*. Molecular Cell, 2012. **48**(2): p. 207-218.
143. Bhaskar, V., et al., *Structure and RNA-binding properties of the Not1-Not2-Not5 module of the yeast Ccr4-Not complex*. Nature Structural & Molecular Biology, 2013. **20**: p. 1281.
144. Bawankar, P., et al., *NOT10 and C2orf29/NOT11 form a conserved module of the CCR4-NOT complex that docks onto the NOT1 N-terminal domain*. RNA biology, 2013. **10**(2): p. 228-244.
145. Boland, A., et al., *Structure and assembly of the NOT module of the human CCR4-NOT complex*. Nature Structural & Molecular Biology, 2013. **20**: p. 1289.
146. Petit, A.-P., et al., *The structural basis for the interaction between the CAF1 nuclease and the NOT1 scaffold of the human CCR4-NOT deadenylase complex*. Nucleic acids research, 2012. **40**(21): p. 11058-11072.
147. Collart, M.A., *The Ccr4-Not complex is a key regulator of eukaryotic gene expression*. Wiley interdisciplinary reviews. RNA, 2016. **7**(4): p. 438-454.
148. Xu, K., et al., *Insights into the structure and architecture of the CCR4-NOT complex*. Frontiers in genetics, 2014. **5**: p. 137-137.
149. Ponting, C.P., *Novel eIF4G domain homologues linking mRNA translation with nonsense-mediated mRNA decay*. Trends in Biochemical Sciences, 2000. **25**(9): p. 423-426.
150. Sgromo, A., et al., *A CAF40-binding motif facilitates recruitment of the CCR4-NOT complex to mRNAs targeted by Drosophila Roquin*. Nature Communications, 2017. **8**: p. 14307.
151. Temme, C., et al., *Subunits of the Drosophila CCR4-NOT complex and their roles in mRNA deadenylation*. RNA (New York, N.Y.), 2010. **16**(7): p. 1356-1370.
152. Albert, T.K., et al., *Isolation and characterization of human orthologs of yeast CCR4-NOT complex subunits*. Nucleic acids research, 2000. **28**(3): p. 809-817.
153. Albert, T.K., et al., *Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex*. The EMBO journal, 2002. **21**(3): p. 355-364.
154. Peng, W., et al., *Regulators of cellular levels of histone acetylation in Saccharomyces cerevisiae*. Genetics, 2008. **179**(1): p. 277-289.

155. Mulder, K.W., et al., *Regulation of histone H3K4 tri-methylation and PAF complex recruitment by the Ccr4-Not complex*. Nucleic acids research, 2007. **35**(7): p. 2428-2439.
156. Mersman, D.P., et al., *Polyubiquitination of the demethylase Jhd2 controls histone methylation and gene expression*. Genes & development, 2009. **23**(8): p. 951-962.
157. Gulshan, K., B. Thommandru, and W.S. Moye-Rowley, *Proteolytic degradation of the Yap1 transcription factor is regulated by subcellular localization and the E3 ubiquitin ligase Not4*. The Journal of biological chemistry, 2012. **287**(32): p. 26796-26805.
158. Sun, H.-Y., et al., *Protein Degradation of RNA Polymerase II-Association Factor 1(PAF1) Is Controlled by CNOT4 and 26S Proteasome*. PLOS ONE, 2015. **10**(5): p. e0125599.
159. Winkler, G.S., et al., *Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription*. The EMBO journal, 2006. **25**(13): p. 3089-3099.
160. Liu, H.-Y., et al., *Characterization of CAF4 and CAF16 Reveals a Functional Connection between the CCR4-NOT Complex and a Subset of SRB Proteins of the RNA Polymerase II Holoenzyme*. Journal of Biological Chemistry, 2001. **276**(10): p. 7541-7548.
161. Lee, T.I., et al., *Interplay of Positive and Negative Regulators in Transcription Initiation by RNA Polymerase II Holoenzyme*. Molecular and Cellular Biology, 1998. **18**(8): p. 4455-4462.
162. Deluen, C., et al., *The Ccr4-not complex and yTAF1 (yTaf(II)130p/yTaf(II)145p) show physical and functional interactions*. Molecular and cellular biology, 2002. **22**(19): p. 6735-6749.
163. Badarinarayana, V., Y.C. Chiang, and C.L. Denis, *Functional interaction of CCR4-NOT proteins with TATAA-binding protein (TBP) and its associated factors in yeast*. Genetics, 2000. **155**,3: p. 1045-54.
164. Lemaire, M. and M.A. Collart, *The TBP associated factor "yTaf II 19p" functionally interacts with components of the global transcriptional regulator "the Ccr4-Not complex" and physically interacts with the Not5 subunit*. Journal of Biological Chemistry, 2000.
165. Benson, J.D., et al., *Association of distinct yeast Not2 functional domains with components of Gcn5 histone acetylase and Ccr4 transcriptional regulatory complexes*. The EMBO journal, 1998. **17**(22): p. 6714-6722.
166. Cui, Y., et al., *Genome wide expression analysis of the CCR4-NOT complex indicates that it consists of three modules with the NOT module controlling SAGA-responsive genes*. Molecular Genetics and Genomics, 2008. **279**(4): p. 323-337.
167. Kruk, J.A., et al., *The multifunctional Ccr4-Not complex directly promotes transcription elongation*. Genes & development, 2011. **25**(6): p. 581-593.
168. Babbarwal, V., J. Fu, and J.C. Reese, *The Rpb4/7 module of RNA polymerase II is required for carbon catabolite repressor protein 4-negative on TATA (Ccr4-not) complex to promote elongation*. The Journal of biological chemistry, 2014. **289**(48): p. 33125-33130.
169. Villanyi, Z., et al., *The Not5 subunit of the ccr4-not complex connects transcription and translation*. PLoS genetics, 2014. **10**(10): p. e1004569-e1004569.
170. Wind, M. and D. Reines, *Transcription elongation factor SII*. BioEssays : news and reviews in molecular, cellular and developmental biology, 2000. **22**(4): p. 327-336.
171. Dutta, A., et al., *Ccr4-Not and TFIIS Function Cooperatively To Rescue Arrested RNA Polymerase II*. Molecular and cellular biology, 2015. **35**(11): p. 1915-1925.
172. Erkmann, J.A. and U. Kutay, *Nuclear export of mRNA: from the site of transcription to the cytoplasm*. Experimental Cell Research, 2004. **296**(1): p. 12-20.

173. Kerr, S.C., et al., *The Ccr4-Not complex interacts with the mRNA export machinery*. PloS one, 2011. **6**(3): p. e18302-e18302.
174. Strambio-de-Castillia, C., G. Blobel, and M.P. Rout, *Proteins connecting the nuclear pore complex with the nuclear interior*. The Journal of cell biology, 1999. **144**(5): p. 839-855.
175. Tseng, S.S., et al., *Dbp5p, a cytosolic RNA helicase, is required for poly(A)⁺ RNA export*. The EMBO journal, 1998. **17**(9): p. 2651-2662.
176. Chávez, S., et al., *A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae*. The EMBO journal, 2000. **19**(21): p. 5824-5834.
177. Cooke, A., A. Prigge, and M. Wickens, *Translational repression by deadenylases*. The Journal of biological chemistry, 2010. **285**(37): p. 28506-28513.
178. Jeske, M., et al., *Rapid ATP-dependent Deadenylation of nanos mRNA in a Cell-free System from Drosophila Embryos*. Journal of Biological Chemistry, 2006. **281**(35): p. 25124-25133.
179. Waghray, S., et al., *Xenopus CAF1 requires NOT1-mediated interaction with 4E-T to repress translation in vivo*. RNA (New York, N.Y.), 2015. **21**(7): p. 1335-1345.
180. Collier, J. and R. Parker, *General translational repression by activators of mRNA decapping*. Cell, 2005. **122**(6): p. 875-886.
181. Sweet, T., C. Kovalak, and J. Collier, *The DEAD-box protein Dhh1 promotes decapping by slowing ribosome movement*. PLoS biology, 2012. **10**(6): p. e1001342-e1001342.
182. Boulon, S., E. Bertrand, and B. Pradet-Balade, *HSP90 and the R2TP co-chaperone complex: Building multi-protein machineries essential for cell growth and gene expression*. RNA Biology, 2012. **9**(2): p. 148-154.
183. Panasenko, O., et al., *The Yeast Ccr4-Not Complex Controls Ubiquitination of the Nascent-associated Polypeptide (NAC-EGD) Complex*. Journal of Biological Chemistry, 2006. **281**(42): p. 31389-31398.
184. Dimitrova, L.N., et al., *Nascent peptide-dependent translation arrest leads to Not4p-mediated protein degradation by the proteasome*. The Journal of biological chemistry, 2009. **284**(16): p. 10343-10352.
185. Halter, D., M.A. Collart, and O.O. Panasenko, *The Not4 E3 ligase and CCR4 deadenylase play distinct roles in protein quality control*. PloS one, 2014. **9**(1): p. e86218-e86218.
186. Preissler, S., et al., *Not4-dependent translational repression is important for cellular protein homeostasis in yeast*. The EMBO journal, 2015. **34**(14): p. 1905-1924.
187. Fabian, M.R., et al., *Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin*. Nature structural & molecular biology, 2013. **20**(6): p. 735-739.
188. Bhandari, D., et al., *Structural basis for the Nanos-mediated recruitment of the CCR4-NOT complex and translational repression*. Genes & development, 2014. **28**(8): p. 888-901.
189. Doidge, R., et al., *The anti-proliferative activity of BTG/TOB proteins is mediated via the Caf1a (CNOT7) and Caf1b (CNOT8) deadenylase subunits of the Ccr4-not complex*. PloS one, 2012. **7**(12): p. e51331-e51331.
190. Funakoshi, Y., et al., *Mechanism of mRNA deadenylation: evidence for a molecular interplay between translation termination factor eRF3 and mRNA deadenylases*. Genes & development, 2007. **21**(23): p. 3135-3148.
191. Fabian, M.R. and N. Sonenberg, *The mechanics of miRNA-mediated gene silencing: a look under the hood of miRISC*. Nature Structural & Molecular Biology, 2012. **19**: p. 586.

192. Mathys, H., et al., *Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in MicroRNA Repression*. Molecular Cell, 2014. **54**(5): p. 751-765.
193. Chen, Y., et al., *A DDX6-CNOT1 Complex and W-Binding Pockets in CNOT9 Reveal Direct Links between miRNA Target Recognition and Silencing*. Molecular Cell, 2014. **54**(5): p. 737-750.
194. Loh, B., S. Jonas, and E. Izaurralde, *The SMG5-SMG7 heterodimer directly recruits the CCR4-NOT deadenylase complex to mRNAs containing nonsense codons via interaction with POP2*. Genes & development, 2013. **27**(19): p. 2125-2138.
195. Shuster, E.O. and B. Byers, *Pachytene Arrest and Other Meiotic Effects of the Start Mutations in Saccharomyces Cerevisiae*. Genetics, 1989. **123**(1): p. 29-43.
196. de Barros Lopes, M., J.Y. Ho, and S.I. Reed, *Mutations in cell division cycle genes CDC36 and CDC39 activate the Saccharomyces cerevisiae mating pheromone response pathway*. Molecular and Cellular Biology, 1990. **10**(6): p. 2966-2972.
197. Neiman, A.M., et al., *CDC36 and CDC39 are negative elements in the signal transduction pathway of yeast*. Cell Regulation, 1990. **1**(5): p. 391-401.
198. Collart, M.A. and K. Struhl, *NOT1(CDC39), NOT2(CDC36), NOT3, and NOT4 encode a global-negative regulator of transcription that differentially affects TATA-element utilization*. Genes & Development, 1994. **8**(5): p. 525-537.
199. James, N., E. Landrieux, and M.A. Collart, *A SAGA-Independent Function of SPT3 Mediates Transcriptional Deregulation in a Mutant of the Ccr4-Not Complex in Saccharomyces cerevisiae*. Genetics, 2007. **177**(1): p. 123-135.
200. Sandler, H., et al., *Not1 mediates recruitment of the deadenylase Caf1 to mRNAs targeted for degradation by tristetraproline*. Nucleic acids research, 2011. **39**(10): p. 4373-4386.
201. Nishimura, T., et al., *The eIF4E-Binding Protein 4E-T Is a Component of the mRNA Decay Machinery that Bridges the 5' and 3' Termini of Target mRNAs*. Cell Reports, 2015. **11**(9): p. 1425-1436.
202. Du, H., et al., *YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex*. Nature Communications, 2016. **7**: p. 12626.
203. Suzuki, A., et al., *NANOS2 interacts with the CCR4-NOT deadenylation complex and leads to suppression of specific RNAs*. Proceedings of the National Academy of Sciences, 2010. **107**(8): p. 3594-3599.
204. Ozgur, S., et al., *Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex*. Cell Reports, 2015. **13**(4): p. 703-711.
205. Bartlam, M. and T. Yamamoto, *The structural basis for deadenylation by the CCR4-NOT complex*. Protein & cell, 2010. **1**(5): p. 443-452.
206. Daugeron, M.C., F. Mauxion, and B. Séraphin, *The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation*. Nucleic acids research, 2001. **29**(12): p. 2448-2455.
207. Tucker, M., et al., *The Transcription Factor Associated Ccr4 and Caf1 Proteins Are Components of the Major Cytoplasmic mRNA Deadenylase in Saccharomyces cerevisiae*. Cell. **104**(3): p. 377-386.
208. Rouault, J.-P., et al., *Interaction of BTG1 and p53-regulated BTG2 Gene Products with mCaf1, the Murine Homolog of a Component of the Yeast CCR4 Transcriptional Regulatory Complex*. Journal of Biological Chemistry, 1998. **273**(35): p. 22563-22569.
209. Hanzawa, H., et al., *The Structure of the C4C4RING Finger of Human NOT4 Reveals Features Distinct from Those of C3HC4 RING Fingers*. Journal of Biological Chemistry, 2001. **276**(13): p. 10185-10190.

210. Chen, H., et al., *The conserved RNA recognition motif and C3H1 domain of the Not4 ubiquitin ligase regulate in vivo ligase function*. Scientific Reports, 2018. **8**(1): p. 8163.
211. Kirstein-Miles, J., et al., *The nascent polypeptide-associated complex is a key regulator of proteostasis*. The EMBO journal, 2013. **32**(10): p. 1451-1468.
212. Laribee, R.N., et al., *CCR4/NOT complex associates with the proteasome and regulates histone methylation*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(14): p. 5836-5841.
213. Haworth, J., et al., *Ubc4 and Not4 regulate steady-state levels of DNA polymerase- α to promote efficient and accurate DNA replication*. Molecular biology of the cell, 2010. **21**(18): p. 3205-3219.
214. Cooper, K.F., et al., *Oxidative-stress-induced nuclear to cytoplasmic relocalization is required for Not4-dependent cyclin C destruction*. Journal of cell science, 2012. **125**(Pt 4): p. 1015-1026.
215. Panasencko, O.O. and M.A. Collart, *Not4 E3 ligase contributes to proteasome assembly and functional integrity in part through Ecm29*. Molecular and cellular biology, 2011. **31**(8): p. 1610-1623.
216. Zwartjes, C.G.M., et al., *Repression of Promoter Activity by CNOT2, a Subunit of the Transcription Regulatory Ccr4-Not Complex*. Journal of Biological Chemistry, 2004. **279**(12): p. 10848-10854.
217. Venturini, G., et al., *CNOT3 is a modifier of PRPF31 mutations in retinitis pigmentosa with incomplete penetrance*. PLoS genetics, 2012. **8**(11): p. e1003040-e1003040.
218. Garces, R.G., W. Gillon, and E.F. Pai, *Atomic model of human Rcd-1 reveals an armadillo-like-repeat protein with in vitro nucleic acid binding properties*. Protein science : a publication of the Protein Society, 2007. **16**(2): p. 176-188.
219. Leppek, K., et al., *Roquin Promotes Constitutive mRNA Decay via a Conserved Class of Stem-Loop Recognition Motifs*. Cell, 2013. **153**(4): p. 869-881.
220. Sgromo, A., et al., *Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4-NOT complex to elicit repression of mRNA targets*. RNA (New York, N.Y.), 2018. **24**(3): p. 381-395.
221. Haas, M., et al., *c-Myb Protein Interacts with Rcd-1, a Component of the CCR4 Transcription Mediator Complex*. Biochemistry, 2004. **43**(25): p. 8152-8159.
222. Hata, H., et al., *Dhh1p, a putative RNA helicase, associates with the general transcription factors Pop2p and Ccr4p from Saccharomyces cerevisiae*. Genetics, 1998. **148**(2): p. 571-579.
223. Collier, J.M., et al., *The DEAD box helicase, Dhh1p, functions in mRNA decapping and interacts with both the decapping and deadenylase complexes*. RNA (New York, N.Y.), 2001. **7**(12): p. 1717-1727.
224. Ayache, J., et al., *P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes*. Molecular biology of the cell, 2015. **26**. **10.1091/mbc.E15-03-0136**.
225. Liu, H.Y., et al., *DBF2, a cell cycle - regulated protein kinase, is physically and functionally associated with the CCR4 transcriptional regulatory complex*. The EMBO Journal, 1997. **16**(17): p. 5289-5298.
226. Komarnitsky, S.I., et al., *DBF2 protein kinase binds to and acts through the cell cycle-regulated MOB1 protein*. Molecular and cellular biology, 1998. **18**(4): p. 2100-2107.
227. Miller, J.E. and J.C. Reese, *Ccr4-Not complex: the control freak of eukaryotic cells*. Critical reviews in biochemistry and molecular biology, 2012. **47**(4): p. 315-333.
228. Pease, B.N., et al., *Global analysis of protein expression and phosphorylation of three stages of Plasmodium falciparum intraerythrocytic development*. Journal of proteome research, 2013. **12**(9): p. 4028-4045.

229. Katoh, K., J. Rozewicki, and K.D. Yamada, *MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization*. Briefings in Bioinformatics, 2017: p. bbx108-bbx108.
230. Balu, B., et al., *High-efficiency transformation of Plasmodium falciparum by the lepidopteran transposable element piggyBac*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(45): p. 16391-16396.
231. Hart, K.J., et al., *Plasmodium male gametocyte development and transmission are critically regulated by the two putative deadenylases of the CAF1/CCR4/NOT complex*. PLoS pathogens, 2019. **15**(1): p. e1007164-e1007164.
232. Banaszynski, L.A., et al., *A rapid, reversible, and tunable method to regulate protein function in living cells using synthetic small molecules*. Cell, 2006. **126**(5): p. 995-1004.
233. Zhang, M., et al., *Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis*. Science (New York, N.Y.), 2018. **360**(6388): p. eaap7847.
234. Bushell, E., et al., *Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes*. Cell, 2017. **170**(2): p. 260-272.e8.
235. Kruk, J.A., et al., *The multifunctional Ccr4–Not complex directly promotes transcription elongation*. Genes & Development, 2011. **25**(6): p. 581-593.
236. Jeong, J.Y., et al., *One-step sequence- and ligation-independent cloning as a rapid and versatile cloning method for functional genomics studies*. Appl Environ Microbiol, 2012. **78**(15): p. 5440-3.
237. Chan, W.-T., et al., *A comparison and optimization of methods and factors affecting the transformation of Escherichia coli*. Bioscience Reports, 2013. **33**(6): p. e00086.
238. Meissner, M., et al., *Tetracycline analogue-regulated transgene expression in Plasmodium falciparum blood stages using Toxoplasma gondii transactivators*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(8): p. 2980-2985.
239. Maier, A.G., et al., *Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in Plasmodium falciparum for targeted gene deletion by double crossover recombination*. Mol Biochem Parasitol, 2006. **150**(1): p. 118-21.
240. Beck, J.R., et al., *PTEX component HSP101 mediates export of diverse malaria effectors into host erythrocytes*. Nature, 2014. **511**(7511): p. 592-5.
241. Cranmer, S.L., et al., *An alternative to serum for cultivation of Plasmodium falciparum in vitro*. Transactions of the Royal Society of Tropical Medicine and Hygiene, 1997. **91**(3): p. 363-365.
242. Lambros, C. and J.P. Vanderberg, *Synchronization of Plasmodium falciparum Erythrocytic Stages in Culture*. The Journal of Parasitology, 1979. **65**(3): p. 418-420.
243. Diggs, C., et al., *Protein Synthesis in Vitro by Cryopreserved Plasmodium Falciparum*. The American Journal of Tropical Medicine and Hygiene, 1975. **24**(5): p. 760-763.
244. Christophers, S.R. and J.D. Fulton, *Experiments with Isolated Malaria Parasites (Plasmodium Knowlesi) Free from Red Cells*. Annals of Tropical Medicine & Parasitology, 1939. **33**(2): p. 161-170.
245. Wu, Y., et al., *Transfection of Plasmodium falciparum within human red blood cells*. Proceedings of the National Academy of Sciences of the United States of America, 1995. **92**(4): p. 973-977.
246. Trager, W., et al., *Clones of the malaria parasite Plasmodium falciparum obtained by microscopic selection: their characterization with regard to knobs, chloroquine sensitivity, and formation of gametocytes*. Proceedings of the National Academy of Sciences of the United States of America, 1981. **78**(10): p. 6527-6530.

247. Bozdech, Z., S. Mok, and A. Patkar, *DNA Microarray-Based Genome-Wide Analyses of Plasmodium Parasites*. Vol. 923. 2013. 189-211.
248. Perkins, D.N., et al., *Probability-based protein identification by searching sequence databases using mass spectrometry data*. ELECTROPHORESIS, 1999. **20**(18): p. 3551-3567.
249. Treeck, M., et al., *The phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries*. Cell host & microbe, 2011. **10**(4): p. 410-419.
250. Sgromo, A., et al., *Drosophila Bag-of-marbles directly interacts with the CAF40 subunit of the CCR4–NOT complex to elicit repression of mRNA targets*. RNA, 2018. **24**(3): p. 381-395.
251. Reddy, B.P.N., et al., *A bioinformatic survey of RNA-binding proteins in Plasmodium*. BMC Genomics, 2015. **16**(1): p. 890.
252. Bernstein, P., S.W. Peltz, and J. Ross, *The poly(A)-poly(A)-binding protein complex is a major determinant of mRNA stability in vitro*. Molecular and cellular biology, 1989. **9**(2): p. 659-670.
253. DERRY, M.C., et al., *Regulation of Poly(A)-binding Protein through PABP-interacting Proteins*. Cold Spring Harbor Symposia on Quantitative Biology, 2006. **71**: p. 537-543.
254. Bi, X. and D.J. Goss, *Wheat Germ Poly(A)-binding Protein Increases the ATPase and the RNA Helicase Activity of Translation Initiation Factors eIF4A, eIF4B, and eIF-iso4F*. Journal of Biological Chemistry, 2000. **275**(23): p. 17740-17746.
255. Ivanov, A., et al., *PABP enhances release factor recruitment and stop codon recognition during translation termination*. Nucleic acids research, 2016. **44**(16): p. 7766-7776.
256. Martineau, Y., et al., *Poly(A)-binding protein-interacting protein 1 binds to eukaryotic translation initiation factor 3 to stimulate translation*. Molecular and cellular biology, 2008. **28**(21): p. 6658-6667.
257. Fischer, K., et al., *Characterization and cloning of the gene encoding the vacuolar membrane protein EXP-2 from Plasmodium falciparum*1Note: Nucleotide sequence data reported in this paper are available in the EMBL, GenBank™ and DDJB databases under the accession number AJ000652.1. Molecular and Biochemical Parasitology, 1998. **92**(1): p. 47-57.
258. de Koning-Ward, T.F., et al., *A newly discovered protein export machine in malaria parasites*. Nature, 2009. **459**(7249): p. 945-949.
259. Garten, M., et al., *EXP2 is a nutrient-permeable channel in the vacuolar membrane of Plasmodium and is essential for protein export via PTEX*. Nature Microbiology, 2018. **3**(10): p. 1090-1098.
260. Hakamada, K., et al., *Expression and characterization of the Plasmodium translocon of the exported proteins component EXP2*. Biochemical and Biophysical Research Communications, 2017. **482**(4): p. 700-705.
261. Garten, M., et al., *Exp2 in the Role of the Small Molecule Pore in the Parasitophorous Vacuole Membrane of Plasmodium falciparum*. Biophysical Journal, 2018. **114**(3): p. 492a.
262. Shakya, B., et al., *The Plasmodium falciparum exported protein PF3D7_0402000 binds to erythrocyte ankyrin and band 4.1*. Molecular and biochemical parasitology, 2017. **216**: p. 5-13.
263. Parish, L.A., et al., *A member of the Plasmodium falciparum PHIST family binds to the erythrocyte cytoskeleton component band 4.1*. Malaria journal, 2013. **12**: p. 160-160.
264. Lamarque, M., et al., *Food vacuole proteome of the malarial parasite Plasmodium falciparum*. PROTEOMICS – Clinical Applications, 2008. **2**(9): p. 1361-1374.

265. Chandran, V. and B.F. Luisi, *Recognition of Enolase in the Escherichia coli RNA Degradosome*. Journal of Molecular Biology, 2006. **358**(1): p. 8-15.
266. Ghosh, A.K. and M. Jacobs-Lorena, *Surface-expressed enolases of Plasmodium and other pathogens*. Memorias do Instituto Oswaldo Cruz, 2011. **106 Suppl 1**(0 1): p. 85-90.
267. Pal Bhowmick, I., et al., *Plasmodium falciparum enolase: stage-specific expression and sub-cellular localization*. Malaria Journal, 2009. **8**(1): p. 179.
268. Das, S., et al., *Plasmodium falciparum enolase complements yeast enolase functions and associates with the parasite food vacuole*. Molecular and biochemical parasitology, 2011. **179**(1): p. 8-17.
269. Holland, Z., et al., *Functional analysis of protein kinase CK2 of the human malaria parasite Plasmodium falciparum*. Eukaryotic cell, 2009. **8**(3): p. 388-397.
270. Li, Y., et al., *Dcp2 decapping protein modulates mRNA stability of the critical interferon regulatory factor (IRF) IRF-7*. Molecular and cellular biology, 2012. **32**(6): p. 1164-1172.
271. Sonoiki, E., et al., *A potent antimalarial benzoxaborole targets a Plasmodium falciparum cleavage and polyadenylation specificity factor homologue*. Nature communications, 2017. **8**: p. 14574-14574.
272. Wang, Z., et al., *The yeast TFB1 and SSL1 genes, which encode subunits of transcription factor IIH, are required for nucleotide excision repair and RNA polymerase II transcription*. Molecular and Cellular Biology, 1995. **15**(4): p. 2288-2293.
273. Matsui, P., J. DePaulo, and S. Buratowski, *An interaction between the Tfb1 and Ssl1 subunits of yeast TFIIH correlates with DNA repair activity*. Nucleic Acids Research, 1995. **23**(5): p. 767-772.
274. Mahajan, B., et al., *Protein disulfide isomerase assisted protein folding in malaria parasites*. International Journal for Parasitology, 2006. **36**(9): p. 1037-1048.
275. Libri, D., et al., *Multiple roles for the yeast SUB2/yUAP56 gene in splicing*. Genes & Development, 2001. **15**(1): p. 36-41.
276. Kistler, A.L. and C. Guthrie, *Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for Sub2, an essential spliceosomal ATPase*. Genes & Development, 2001. **15**(1): p. 42-49.
277. Taniguchi, I. and M. Ohno, *ATP-dependent recruitment of export factor Aly/REF onto intronless mRNAs by RNA helicase UAP56*. Molecular and cellular biology, 2008. **28**(2): p. 601-608.
278. Kammel, C., et al., *Arabidopsis DEAD-box RNA helicase UAP56 interacts with both RNA and DNA as well as with mRNA export factors*. PloS one, 2013. **8**(3): p. e60644-e60644.
279. LaCount, D.J., et al., *A protein interaction network of the malaria parasite Plasmodium falciparum*. Nature, 2005. **438**: p. 103.
280. Fidock, D.A. and T.E. Wellems, *Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil*. Medical Sciences, 1997. **94**: p. 10931-36.
281. Maier, A.G., et al., *Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in Plasmodium falciparum for targeted gene deletion by double crossover recombination*. Molecular and Biochemical Parasitology, 2006. **150**(1): p. 118-121.
282. Maier, A.G., et al., *Exported proteins required for virulence and rigidity of Plasmodium falciparum-infected human erythrocytes*. Cell, 2008. **134**(1): p. 48-61.

283. Mundwiler-Pachlatko, E. and H.-P. Beck, *Maurer's clefts, the enigma of Plasmodium falciparum*. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(50): p. 19987-19994.
284. Pavithra, S.R., R. Kumar, and U. Tatu, *Systems analysis of chaperone networks in the malarial parasite Plasmodium falciparum*. PLoS computational biology, 2007. **3**(9): p. 1701-1715.
285. Raman, J., et al., *Plasmodium falciparum hypoxanthine guanine phosphoribosyltransferase*. The FEBS Journal, 2005. **272**(8): p. 1900-1911.
286. Whittingham, J.L., et al., *dUTPase as a Platform for Antimalarial Drug Design: Structural Basis for the Selectivity of a Class of Nucleoside Inhibitors*. Structure, 2005. **13**(2): p. 329-338.
287. Kikuchi, G., et al., *Glycine cleavage system: reaction mechanism, physiological significance, and hyperglycinemia*. Proceedings of the Japan Academy. Series B, Physical and biological sciences, 2008. **84**(7): p. 246-263.
288. Shastri, S., et al., *Plasmodium CDP-DAG synthase: An atypical gene with an essential N-terminal extension*. International Journal for Parasitology, 2010. **40**(11): p. 1257-1268.
289. Hayashi, M., et al., *Vacuolar H⁺-ATPase Localized in Plasma Membranes of Malaria Parasite Cells, Plasmodium falciparum, Is Involved in Regional Acidification of Parasitized Erythrocytes*. Journal of Biological Chemistry, 2000. **275**(44): p. 34353-34358.
290. Lim, L. and G.I. McFadden, *The evolution, metabolism and functions of the apicoplast*. Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 2010. **365**(1541): p. 749-763.
291. van Dooren, G.G., L.M. Stimmler, and G.I. McFadden, *Metabolic maps and functions of the Plasmodium mitochondrion*. FEMS Microbiology Reviews, 2006. **30**(4): p. 596-630.
292. Vaidya, A.B. and M.W. Mather, *Mitochondrial Evolution and Functions in Malaria Parasites*. Annual Review of Microbiology, 2009. **63**(1): p. 249-267.
293. van Dooren, G.G., et al., *Development of the endoplasmic reticulum, mitochondrion and apicoplast during the asexual life cycle of Plasmodium falciparum*. Molecular Microbiology, 2005. **57**(2): p. 405-419.
294. Kobayashi, T., et al., *Mitochondria and apicoplast of Plasmodium falciparum: Behaviour on subcellular fractionation and the implication*. Mitochondrion, 2007. **7**(1): p. 125-132.
295. Rug, M., et al., *The role of KAHRP domains in knob formation and cytoadherence of P falciparum-infected human erythrocytes*. Blood, 2006. **108**(1): p. 370-378.
296. Rodriguez, C.R., et al., *Kin28, the TFIIH-associated carboxy-terminal domain kinase, facilitates the recruitment of mRNA processing machinery to RNA polymerase II*. Molecular and cellular biology, 2000. **20**(1): p. 104-112.
297. Phatnani, H.P. and A.L. Greenleaf, *Phosphorylation and functions of the RNA polymerase II CTD*. Genes & Development, 2006. **20**(21): p. 2922-2936.
298. Buratowski, S., *Progression through the RNA polymerase II CTD cycle*. Molecular cell, 2009. **36**(4): p. 541-546.
299. Banaszynski, L.A., et al., *A Rapid, Reversible, and Tunable Method to Regulate Protein Function in Living Cells Using Synthetic Small Molecules*. Cell, 2006. **126**(5): p. 995-1004.
300. Iwamoto, M., et al., *A general chemical method to regulate protein stability in the mammalian central nervous system*. Chemistry & biology, 2010. **17**(9): p. 981-988.
301. Armstrong, C.M. and D.E. Goldberg, *An FKBP destabilization domain modulates protein levels in Plasmodium falciparum*. Nature Methods, 2007. **4**(12): p. 1007-1009.

302. Muralidharan, V., et al., *Asparagine repeat function in a Plasmodium falciparum protein assessed via a regulatable fluorescent affinity tag*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(11): p. 4411-4416.
303. de Azevedo, M.F., et al., *Systematic Analysis of FKBP Inducible Degradation Domain Tagging Strategies for the Human Malaria Parasite Plasmodium falciparum*. PLoS ONE, 2012. **7**(7): p. e40981.
304. Beck, J.R., et al., *HSP101/PTEX mediates export of diverse malaria effector proteins into the host erythrocyte*. Nature, 2014. **511**(7511): p. 592-595.
305. Turowski, T.W. and D. Tollervey, *Transcription by RNA polymerase III: insights into mechanism and regulation*. Biochemical Society transactions, 2016. **44**(5): p. 1367-1375.
306. Mair, G.R., et al., *Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development*. PLoS pathogens, 2010. **6**(2): p. e1000767-e1000767.
307. Wong, W., et al., *Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine*. eLife, 2014. **3**: p. e03080.
308. Cuchalová, L., et al., *The RNA Recognition Motif of Eukaryotic Translation Initiation Factor 3g (eIF3g) Is Required for Resumption of Scanning of Posttermination Ribosomes for Reinitiation on GCN4 and Together with eIF3i Stimulates Linear Scanning*. Molecular and Cellular Biology, 2010. **30**(19): p. 4671-4686.
309. Pestova, T.V., et al., *Molecular mechanisms of translation initiation in eukaryotes*. Proceedings of the National Academy of Sciences, 2001. **98**(13): p. 7029-7036.
310. Vembar, S.S., et al., *The PfAlba1 RNA-binding protein is an important regulator of translational timing in Plasmodium falciparum blood stages*. Genome Biology, 2015. **16**(1): p. 212.
311. Bunnik, E.M., et al., *The mRNA-bound proteome of the human malaria parasite Plasmodium falciparum*. Genome biology, 2016. **17**(1): p. 147-147.
312. Kaplan, D.L. and M. O'Donnell, *Rho Factor: Transcription Termination in Four Steps*. Current Biology, 2003. **13**(18): p. R714-R716.
313. Iyer, G.R., et al., *Calcium-dependent phosphorylation of Plasmodium falciparum serine repeat antigen 5 triggers merozoite egress*. Journal of Biological Chemistry, 2018. **293**(25): p. 9736-9746.
314. De Zoysa, M.D. and Y.-T. Yu, *Posttranscriptional RNA Pseudouridylation*. The Enzymes, 2017. **41**: p. 151-167.
315. Jack, K., et al., *rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells*. Molecular cell, 2011. **44**(4): p. 660-666.
316. Chook, Y. and G. Blobel, *Karyopherins and nuclear import*. Current Opinion in Structural Biology, 2001. **11**(6): p. 703-715.
317. Panchal, M., et al., *Plasmodium falciparum signal recognition particle components and anti-parasitic effect of ivermectin in blocking nucleo-cytoplasmic shuttling of SRP*. Cell Death & Disease, 2014. **5**: p. e994.
318. Liang, X., et al., *Puf3 participates in ribosomal biogenesis in malaria parasites*. Journal of Cell Science, 2018. **131**(6).
319. Dahan-Pasternak, N., et al., *PfSec13 is an unusual chromatin-associated nucleoporin of Plasmodium falciparum that is essential for parasite proliferation in human erythrocytes*. Journal of Cell Science, 2013. **126**(14): p. 3055-3069.
320. Fidock, D.A., et al., *Cloning and characterization of a novel Plasmodium falciparum sporozoite surface antigen, STARP*. Molecular and Biochemical Parasitology, 1994. **64**(2): p. 219-232.

321. Leneghan, D. and A. Bell, *Immunophilin–protein interactions in Plasmodium falciparum*. *Parasitology*, 2015. **142**(11): p. 1404-1414.
322. Mok, S., et al., *Drug resistance. Population transcriptomics of human malaria parasites reveals the mechanism of artemisinin resistance*. *Science (New York, N.Y.)*, 2015. **347**(6220): p. 431-435.
323. Shock, J.L., K.F. Fischer, and J.L. DeRisi, *Whole-genome analysis of mRNA decay in Plasmodium falciparum reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle*. *Genome Biology*, 2007. **8**(7): p. R134.