



HAL
open science

Characterization of histone modifications inside nucleosome H4K31ac and H4K31me1 in Apicomplexan parasites

Fabien Sindikubwabo

► **To cite this version:**

Fabien Sindikubwabo. Characterization of histone modifications inside nucleosome H4K31ac and H4K31me1 in Apicomplexan parasites. Microbiology and Parasitology. Université Grenoble Alpes, 2017. English. NNT: 2017GREAV047 . tel-01686271

HAL Id: tel-01686271

<https://theses.hal.science/tel-01686271>

Submitted on 17 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTE UNIVERSITE GRENOBLE ALPES

Spécialité : **Physiologie-Physiopathologies-Pharmacologie**

Arrêté ministériel : 25 mai 2016

Présentée par

« Fabien SINDIKUBWABO »

Thèse dirigée par **Mohamed-Ali HAKIMI, DR2, INSERM-
Institute for Advanced Biosciences IAB, Grenoble**

Préparée au sein du **Laboratoire Host-Pathogens interactions
& Immunity to Infections.**

Dans l'École Doctorale Chimie et Sciences du vivant.

Réseau régulateur de HDAC3 pour comprendre les mécanismes de différenciation et de pathogénèse de *Toxoplasma gondii*

Thèse soutenue publiquement le « **12 octobre 2017** », devant le
jury composé de :

M.ARTUR SCHERF

Directeur de recherche Emérite et CNRS Ile de France Ouest et Nord,
Rapporteur

M. THIERRY LAGRANGE

Directeur de recherche et CNRS Délégation Languedoc-Roussillon,
Rapporteur

M.SAADI KHOCHBIN

Directeur de recherche Emérite et CNRS Délégation Alpes, Président et
examineur

M. JEROME GOVIN

Directeur de recherche et CEA Grenoble, Examineur



Acknowledgements

I would like to express my sincere gratitude to my advisor, Dr. Hakimi Mohamed-Ali, who gave me the chance to do my PhD in his laboratory. He was always present and worked hard to train me and encouraged me to strive for success during these past four years. This work would not have been possible without his continuous support and advice in all the time of my research.

My sincere thanks go also to Dr. Thierry Lagrange and Mr. Hassan Belrhali for their pertinent suggestions during my thesis. Many thanks to Dr. Saadi Khochbin, Dr. Artur Scherf, Dr. Thierry Lagrange and Dr. Jérôme Govin who, as thesis committee members, accepted to examine my work.

I am thankful to the people in Hakimi laboratory who participated to my thesis: Dominique, Laurence, Alexandre, Céline, H el ene, Huan, Gabrielle, Dayana, Marie Pierre, Herv e, Tahir, Andr es, Christopher, Rose, Bastien, Aur elie, S everine and Valeria. I also thank Jean Benjamin, Gabrielle and Jason for their encouragements. Many thanks to Isabelle for her advice and contribution to my thesis.

I am extremely thankful to Shuai, Sebastian, Artur, Pieter-Jan, Yohann, Philippe, Mohamed, Lucid and Cyrille who contributed to some of the data presented in this work. I also thank Dr. Stanislas Tomavo and Dr. Marion Sabrina who believed in our collaborations.

A special thanks to my father Emmanuel, my stepmother Christine, my sisters Dyna, Aline, Diane, Joyeuse and my brothers Charles, Claude, Aim e and Hugo who always support me.

I would like to thank my son Erwan and particularly my wife Lucie for her support and encouragement throughout my PhD thesis.

Index

Abbreviations	4
Part I - Introduction	6
Part II - Results	31
Chapter I - <i>T. gondii</i> ApiAP2 transcription factors are embedded in stable multi-subunit complexes gathering the deacetylase TgHDAC3, the MORC protein CRC230 and ELM2-containing scaffolding proteins	32
II- 1.1. Rationale and Preliminary Studies	33
II- 1.2. Results	34
II- 1.2.1. TgCRC230 is a predicted nuclear MORC-related protein in <i>T. gondii</i>	34
II- 1.2.2. TgCRC230 acts as scaffolding platform bridging TgHDAC3 to multiple ApiAP2 transcription factors and ELM2-containing proteins	35
II- 1.2.3. TgCRC230 purification led to the identification of three putative ELM2 domain-containing proteins	37
Chapter II - The chemical inactivation of TgHDAC3 by FR235222 revealed a peculiar acetylome and proteome regulated by the enzyme: a first step toward stage differentiation	40
II- 2.1. Rationale and Preliminary Studies	41
II- 2.2 Results	41
II- 2.2.1. FR235222-mediated inhibition of TgHDAC3 induces bradyzoite- and merozoite-specific proteins expression	41
II- 2.2.2. Monitoring of FR235222-mediated bradyzoite gene expression using new reporter transgenic parasite cell line	46
II- 2.2.3. Proteome-wide mapping of the <i>T. gondii</i> acetylome, a gateway to discovering new TgHDAC3 substrates beyond histones	47
Chapter III - Modifications at K31 on the lateral surface of histone H4 contribute to genome structure and expression in apicomplexan parasites	52
Part III. Discussion	115
III- 1. CRISPR/cas9-mediated gene disruption of TgHDAC3 recapitulates FR235222-mediated phenotypes	116
III- 2. Linking bradyzoite development to the parasite cell cycle ?	118
III- 3. FR235222-mediated inhibition of TgHDAC3 re-programs stage-specific gene expression in tachyzoites	120
III- 4. Targeting of TgHDAC3 to chromatin in a DNA-specific manner : a role for ApiAP2 transcription factors ?	122
III- 5. Deacetylation by TgHDAC3 of ApiAP2 transcription factors: the substrate hypothesis?	127
III- 6. Conclusion	128

Part IV. Materials and methods	129
IV- 1. Parasites and host cells.....	130
IV- 2. HDACi treatments.....	130
IV- 3. Plasmid constructs.....	130
IV- 4. Cas9-mediated gene disruption.....	131
IV- 5. Antibodies.....	132
IV- 6. <i>Toxoplasma gondii</i> transfection.....	133
IV- 7. Immunofluorescence microscopy.....	133
IV- 8. Protein extraction and Trypsin Digestion.....	134
IV- 9. Affinity enrichment of lysine acetylated peptides.....	134
IV- 10. Affinity purification of Flag-tagged proteins.....	134
IV- 11. Mass spectrometry and peptide sequencing.....	135
IV- 12. Histones purification, Immunoblotting and mass spectrometry analysis....	135
IV- 13. Chromatin Immunoprecipitation and Next Generation Sequencing in <i>T.gondii</i>	
IV- 14. Library Preparation, Sequencing and Data analysis (Arraystar).....	136
Part V. References	138
Part VI. Annexes - Publications during the thesis in collaboration	155

Abbreviations

AIDS: acquired immune deficiency syndrome

AP2/ERF: apetala 2 / ethylene responsive factor

ApiAP2: *Apicomplexa* apetala 2

CRC: core repression complex

ChIP: chromatin immunoprecipitation

ChIP-seq: chromatin immunoprecipitation followed by deep sequencing

CRISPR: Clustered Regular Interspaced Short Palindromic Repeats

Cas9: CRISPR-associated protein 9

ELM2: EGL-27 and MTA1 homology 2

ENO: enolase

GCN5: *Toxoplasma gondii* general control non-derepressible 5

GFP: green fluorescent protein

GOI: gene of interest

HATs: histone acetyltransferases

HA tag: hemagglutinin tag

HDACs: Histone deacetylases

HDACi: histone deacetylase inhibitor

HF: halofuginone

HFF: Human Foreskin Fibroblast

HC-toxin: *Helminthosporium carbonum*-toxin

HP1: heteroprotein 1

IDC: intra-erythrocytic development cycle

IMC1: inner membrane complex 1

IFAs: immunofluorescence assays

IFN- γ : interferon gamma

KMT: lysine methyltransferase

LC-MS/MS: Liquid chromatography coupled to tandem mass spectrometry

LDH: lactate dehydrogenase

LIC: ligation independent cloning
MEF: Mouse Embryonic Fibroblast
MORC: microrchidia
MTA1: metastasis-associated protein 1
MYST: MOZ/yeast Ybf2/Sas2/TIP60
Pb: *Plasmodium berghei*
Pf: *Plasmodium falciparum*
Pru: Prugniaud
PTMs: post-translational modifications
SANT: SWI3/ADA2/NCoR/TFIII-B
SRS: SAG (Surface antigen)-Related Sequence
TAREs: telomere-associated repetitive elements
TBL1: transducin beta-like protein1
TFs: transcription factors
Tg: *Toxoplasma gondii*
ToxoDB: *Toxoplasma* database
TSA: trichostatin A
WT: wild type

Part I. Introduction

I-1. *Toxoplasma gondii* belongs to the *Apicomplexa* phylum

Toxoplasma gondii belongs to the phylum *Apicomplexa* which are characterized by a common apical complex involved in host cell attachment and invasion (Blackman MJ *et al.*, 2001). Apicomplexan parasites are important diseases causing organisms that infect both animals and humans, causing extensive health and economic damage to human populations, particularly those in the developing world. Preeminent human pathogens include *Plasmodium spp.* which are responsible for dreadful malaria, which is responsible of an estimated 429 000 deaths worldwide in addition to 212 million new cases of malaria as described by World Health Organization WHO report 2017. The phylum also includes *T. gondii* and *Cryptosporidium spp.* which are leading causes of foodborne and waterborne diseases. As well as *Theileria*, *Eimeria*, *Neospora*, *Babesia* or *Sarcocystis* that are responsible for causing diseases in wild and domesticated animals (Arisue N *et al.*, 2015; Yabsley MJ *et al.*, 2012; Donahoe SL *et al.*, 2015).

Caused by the protozoan *Apicomplexa* parasite *T. gondii*, toxoplasmosis is a widespread foodborne infection in humans that poses significant public health problems. Approximately, more than one third of world human population is infected with *T. gondii*. While the infection remains asymptomatic lifelong in immunocompetent individuals, Toxoplasmosis is a potentially life-threatening chronic disease in people with weakened immune system, such as those suffering from acquired immunodeficiency syndrome or undergoing chemotherapy and graft rejection therapy (Boothroyd, 2009). In addition, outcomes of congenital toxoplasmosis significantly vary with the timing of infection from recurrent eye diseases to adverse motor or neurologic impairments that can cause stillbirth (Weiss LM *et al.*, 2000; Jones, *et al.*, 2003).

Among apicomplexan parasites, *T. gondii* affords many experimental advantages including efficient genetic manipulation in laboratory and availability of many cell markers. While results in *T. gondii* may not always reflect the biology of other apicomplexans, *T. gondii* has been used and remains the best model system to study the cell biology of *Plasmodium spp* as well as other apicomplexan parasites (Kim K *et al.*, 2004). Thus, studies in *T. gondii* may improve our understanding of the apicomplexan pathogenesis and hence, contribute to the drug development against parasites in the phylum.

T. gondii and *P. falciparum* are the most studied parasites of the phylum and are typified by sophisticated mechanisms to invade, multiply within cell they infect, egress, and cause

pathogenesis in intermediate and definitive hosts. All these processes are associated to the multiple developmental stages those parasites go through to persist in their hosts.

I-2. The heteroxenous life cycle of *Toxoplasma gondii*

In response to the environment conditions, *T. gondii* undergoes a complex life cycle alternating between the asexual stage in vast number of intermediate hosts ranging from mammals to birds and the sexual stage in a definitive host the felidae family including domestic cats (Dubey JP, 2009; Black MW *et al.*, 2000). The biology of *T. gondii* follows a heterogeneous life cycle depending on a definitive felid host for sexual transmission while achieving asexual transmission in intermediate hosts (Dubey JP *et al.*, 1998; Robert-Gangneux F *et al.*, 2012).

The sexual parasitic cycle occurs exclusively within the intestine of infected feline. Specifically, the cycle starts when cats are exposed to the parasite by ingesting the infected intermediate hosts such as mice or birds carrying bradyzoites-containing tissue cysts (Dubey JP, 2004). Within the intestine of cats, the bradyzoites differentiate into merozoites which in turn progress into male and female gametes. The two gametes then fuse to form oocyst. When oocysts are shed into the environment with the feces of cats, they sporulate to form sporozoites that are highly infectious orally to intermediate hosts and less infectious to definitive host (Frenkel JK *et al.*, 1970; Dubey JP *et al.*, 1970; Schlüter D *et al.*, 2014). In the environment, oocysts contaminate the soil, grass and water. They resist to different environmental conditions and remain infectious for up to 18 months or longer (Lindsay DS *et al.*, 2009; Innes EA, 2010). To complete the sexual cycle, the cat can be reinfected by consuming oocysts-contaminated water or ingesting infected mice or birds (Figure 1)

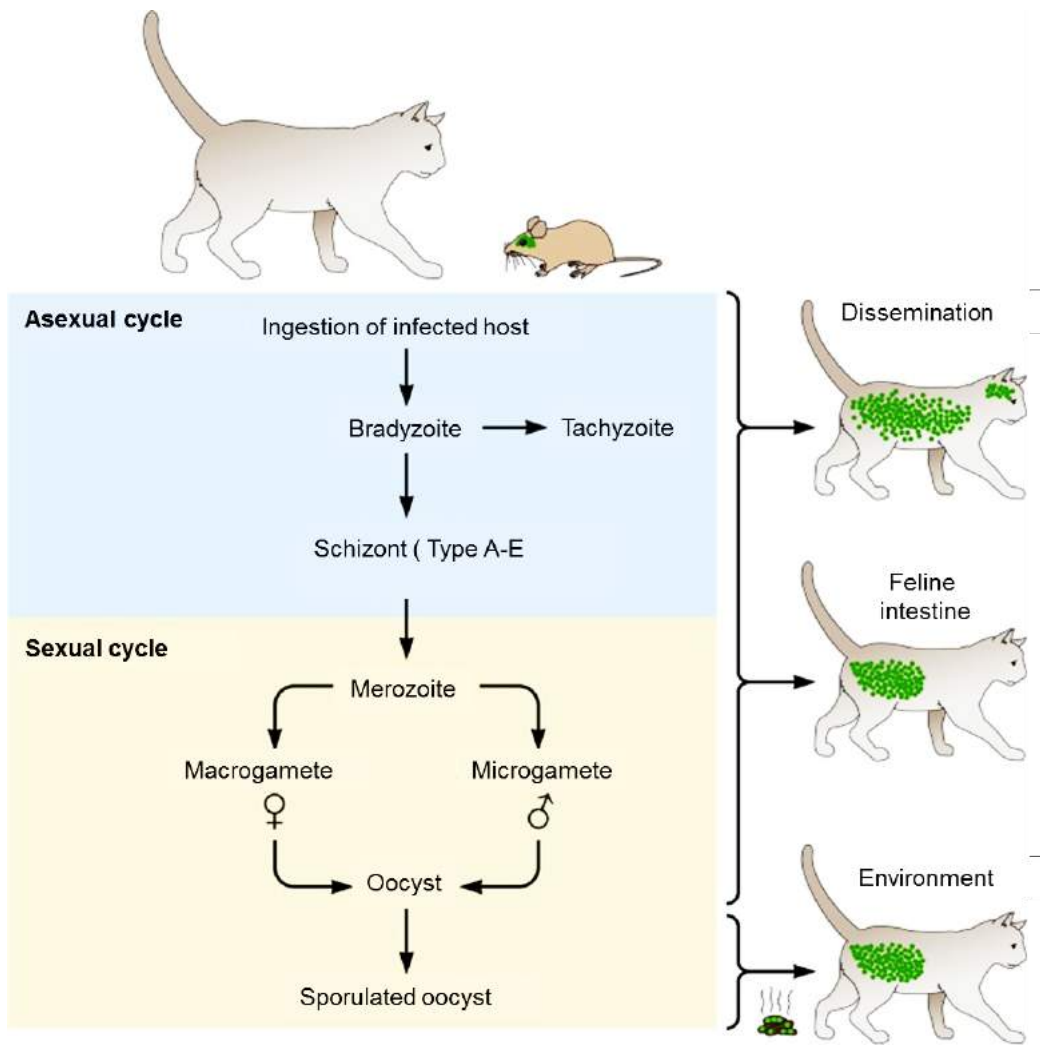


Figure 1. Asexual and sexual cycles of *T.gondii* (from Pittman KJ *et al.*, 2015).

Once ingested, *T. gondii* penetrates the epithelial cells of the feline small intestine and differentiates into tachyzoite and schizont stages. The asexual tachyzoites divide and disseminate throughout the feline. The schizont will remain within the intestinal epithelium and has 5 distinct stages, identified as types A to E. They are classified based on their mode of division, the time post-infection they are observed, and their structural components. Type E schizonts give rise to merozoites, which differentiate into gametes. Gametes can be found throughout the small intestine as soon as 3 days post-infection and can last for a few weeks post-inoculation with tissue cysts. Males (microgametes) fertilize females (macrogametes) to produce oocysts. After fertilization occurs, the oocyst wall forms around the parasite. Sporulation of oocysts occurs 1 to 5 days after being excreted in cat feces. Once sporulation occurs, oocysts are infectious for an extended period of time, depending on environmental conditions.

The asexual cycle takes place in intermediate hosts, i.e. mammals and birds, and is typified by two important stages: tachyzoites and bradyzoites. The cycle starts with sporulated oocysts of the environment orally ingested by new hosts, then sporozoites emerge from oocysts and

invade the intestine epithelial cells before converting into tachyzoites (Schlüter D *et al.*, 2014). To get access to a host cell, the invasive and replicative stage tachyzoites trigger the formation of a peculiar membrane-bound compartment called the Parasitophorous Vacuole (PV). The PV is shaped as a safe niche that supports parasite growth and multiplication being kept hidden from the harmful endocytic pathway. Tachyzoites divide asexually through a typical endodyogeny process (Dubey JP *et al.*, 1998; Mordue DG *et al.*, 1999; Montoya JG *et al.*, 2004). It is during the initial and acute phase of the infection that the tachyzoite-bradyzoite transition occurs in a minor subpopulation while concurring with the massive destruction of the bulk tachyzoite population as a result of a rapid Th1 cell-mediated and short-term proinflammatory response. Thus, few tachyzoites escape destruction and differentiate into bradyzoites (Hunter CA *et al.*, 2012). Bradyzoites cause chronic infection by forming tissue cysts in deep tissue such as brain, heart, retina and skeletal muscles (Di Cristina M *et al.*, 2008; Tenter AM *et al.*, 2000; Weiss LM *et al.*, 2011). Consumption of encysted bradyzoites in infected tissue is another leading cause of infection of livestock animals and humans. During this oral infection, the cyst wall is digested by host digestive enzymes. Bradyzoites are then detected in the mouse intestinal epithelium, and differentiation into tachyzoites occurs rapidly. Once tachyzoites infect migratory cells, such as macrophages and dendritic cells (DCs), they can rapidly disseminate throughout the host. Eventually, pressure from the immune system and other unknown factors induces conversion to the bradyzoite, the asexual stage associated with chronic infection. Cysts containing bradyzoites can persist for decades within striated muscle and tissue of the central nervous system (Black MW *et al.*, 2000; Robert-Gangneux F *et al.*, 2012)

Asexual tachyzoite and bradyzoite stages can be easily discriminate in cell culture or in murine model by their gene expression pattern and to some extent by their proteome. For instance, tachyzoites are typified by the expression of the surface antigens SAG1 or SAG2A and also the metabolic enzymes LDH1 and ENO2, while bradyzoites are characterized by the exclusive expression of SAG2C, SAG2D and SAG4; the bradyzoite-specific recombinant BSR4; the bradyzoite antigen BAG1; the matrix antigen MAG1 and the metabolic enzymes ENO1 and LDH2 (Lyons RE *et al.*, 2002). While tachyzoites are easily spread in cell culture using various cell lines, mature bradyzoites can be exclusively isolated from the brain of mice chronically infected, which limit their study. Moreover not all the isolated strain of *T. gondii* are endowed with the ability to form the chronic stages in mouse model. As such, genotyping studies revealed that the genus has 12 haplotypes and three main clonal lineages, namely type

I, II and III have been documented in North America and Europe. Those lineages differ in their growth rate, their ability to form cysts and their virulence in murine model as well as their pathogenicity in human toxoplasmosis (Sibley LD *et al.*, 1992; Howe DK *et al.*, 1995). Type I strains (e.g. RH) are highly virulent in mice, grow faster and emerged as the experimentally most tractable strain so far. Type II (e.g. Prugniaud, PRU) and type III strains (e.g. VEG) are less virulent, difficult to genetically manipulate however they make a good chronic model for toxoplasmosis as they form mature cysts *in vivo* in mice (Kim K *et al.*, 2004; Saeij JP *et al.*, 2005). Apart from tachyzoites and bradyzoites, culturing methods for merozoites and sexual stages have not been developed, restricting the understanding of parasites biology in the sexual life cycle.

I-3. The complex life cycle of *Plasmodium falciparum*

Similarly to *T. gondii*, *Plasmodium falciparum* has a complex life cycle with two different hosts, the *Anopheles* mosquito and humans. *P. falciparum* completes the life cycle by alternating between asexual and sexual development stages. The infected *Anopheles* mosquito injects the sporozoites into the human host after their release from oocysts to spread in the liver. In the hepatocytes, sporozoites invade and develop asymptotically into merozoites (7-14 days) that are released in the bloodstream to invade red blood cells (RBCs) and thereby entering in asexual intra-erythrocytic development cycle (IDC) for 48 hours. During IDC, merozoites progress successively into ring stage (0-24 hours post-infection), trophozoite stage (24-34 hours post- infection) and schizont stage (34-48 hours post-infection). To complete the cycle, merozoites within schizont are released into the bloodstream. However, some merozoites can form the sexual male and female gametes during a process called gametocytogenesis. The gametocytes can be transmitted to the midgut lumen of mosquito during blood meal and they develop into macro-gametes (male) and micro-gametes (female) which fuse to form first a zygote and second produce a motile ookinete. This ookinete crosses the mosquito midgut epithelium and forms, on the outer face of this gut wall, an oocyst containing sporozoites. These sporozoites can reach the salivary gland and reinfect the new human host. Even though asexual parasites are pathogenic, gametocytogenesis remains the pivotal step in the *Plasmodium* life cycle because it is required for the parasite transmission from infected human to *Anopheles* mosquito and then from mosquito to the new human host (Ménard R *et al.*, 2013; Josling GA *et al.*, 2015; Doerig C *et al.*, 2015) (Figure 2).

Therefore, *P. falciparum* invades different types of cells and self-propagates in very distinct environments in the mosquito (gut, hemolymph and salivary glands) as well as in the human host (skin, liver and erythrocytes). Each of these distinct environments exerts selective pressure related to morphological changes that force *P. falciparum* to exhibit differential gene expression during its life cycle.

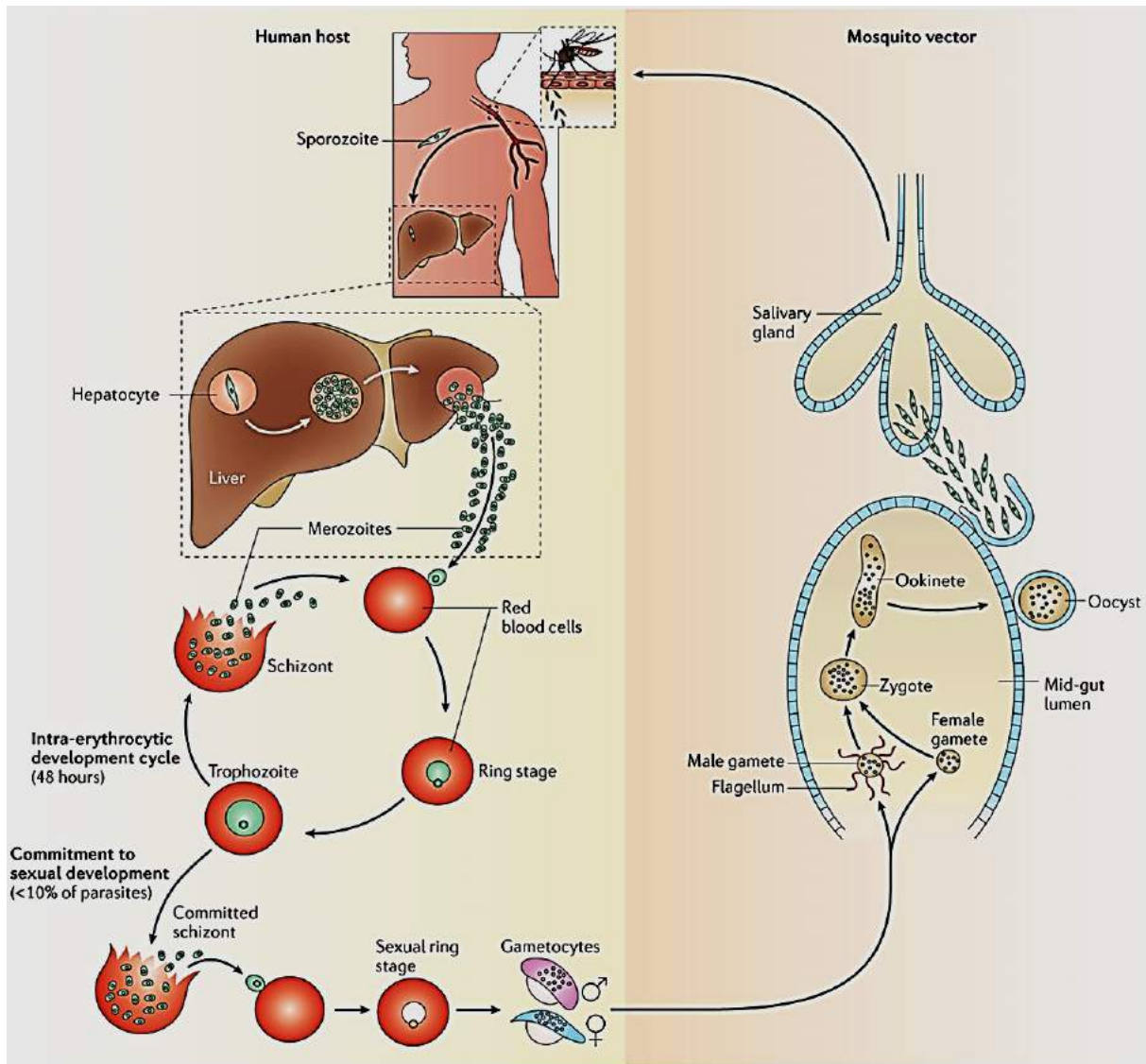


Figure 2. Life cycle of *Plasmodium falciparum*. (From Josling GA, Llinás M. 2015)

I- 4. Gene expression programs in *Toxoplasma* and *Plasmodium*

As above mentioned, both *T. gondii* and *P. falciparum* have complex life cycles that include infections of more than one host organism, differentiation through several morphologically distinct forms, and both sexual and asexual replication. Conceptually, it is possible to divide

the regulation of gene expression into two distinct areas based on parasite development: (1) changes in gene expression as parasites progress through the cell cycle and (2) activation and silencing of specific genes during differentiation of the parasite into different morphological forms, including sexual differentiation.

I- 4.1. Cell cycle progression

Asexual replication in these parasites occurs via either schizogony in *Plasmodium* or endodyogeny in *Toxoplasma*. In either case, progression through the cell cycle involves a strict program of gene expression with distinct subsets of genes expressed in the early, mid and late stages of the cycle. This course of gene expression has been extensively studied for both parasites (Bozdech Z *et al.*, 2003; Cleary MD *et al.*, 2002; Radke JR *et al.*, 2005), and they appear to display a tightly coordinate cascade which utilizes a ‘just-in-time’ manufacturing process whereby induction of any given gene occurs once per cycle and only at a time when it is required (Bozdech Z *et al.*, 2003). Genes located at adjacent positions within the chromosomes are rarely coregulated, suggesting that the expression of each individual gene is independently controlled. Upstream regulatory regions placed on transfected plasmid constructs appear to be properly regulated, indicating that control of transcription is determined by the DNA sequences found close to the coding regions of the genes, however the details of how proper timing is maintained has not yet been deciphered. The global cell cycle transcriptome implies the cyclical expression pattern of *Toxoplasma* genes (~2,833) that accompany parasite replication and present in the two major functional subtranscriptomes associated with S/M phases- subtranscriptome or G1 phase-subtranscriptome. Genes present in G1 phase are involved in the renewal of basal biosynthetic functions and metabolism. S/M periods are characterized by specialized apicomplexan processes of daughter maturation and egress. This indicates that the mechanisms responsible for cyclical gene expression timing are likely crucial to the efficiency of parasite replication (Behnke MS *et al.*, 2010).

I- 4.2. Cellular differentiation

During their transmission from one host to the next, both *Toxoplasma* and *Plasmodium* undergo differentiation into several morphologically distinct forms, a process that includes sexual differentiation and genetic recombination. Differentiation depends on alterations in the

parasites transcriptional profiles. The ‘transcriptomes’ of the parasites at different points in the complete life cycle have been determined (Bozdech Z *et al.*, 2003; Cleary MD *et al.*, 2002; Radke JR *et al.*, 2005) and also studied at the proteomic level (Hall N *et al.*, 2005). This has allowed the identification of many stage specific genes that may eventually help lead to vaccines or intervention strategies resulting in transmission blocking. In addition to transcriptional control, regulation at the level of mRNA translation has also been observed (Mair GR *et al.*, 2006), adding a further complication to understanding how the developmental process is controlled.

One of the crucial step of pathogenesis, transmission and persistence of *T. gondii* is the conversion of tachyzoite into latent bradyzoite. Encystment allows *Toxoplasma* to persist in the host and affords the parasite a unique opportunity to spread to new hosts without proceeding through its sexual stage, which is restricted to felids. Bradyzoite tissue cysts can cause reactivated toxoplasmosis if host immunity becomes impaired. A greater understanding of the molecular mechanisms orchestrating bradyzoite development is needed to better manage the disease. Over the last decade, a wealth of studies has been focusing on the molecular mechanisms driving or regulating tachyzoite-bradyzoite interconversion in order to identify protein factors involved in the switching or the maintenance of the stage-specific genetic program (Sullivan Jr WJ *et al.*, 2009).

I- 4.3. *T. gondii* differentiation: a stress response or a genetic program at work?

It is well documented that conversion to the latent stage is a stress-mediated response, coupled with a slowing of the parasite cell cycle. One of the most commonly used *in vitro* methods to prompt bradyzoite differentiation is alkaline pH 8.0-8.2 (Soète M *et al.*, 1994). Nutrient deprivation was also shown to trigger bradyzoite formation. As such, arginine starvation (Fox BA *et al.*, 2004) or pyrimidine depletion in uracil phosphoribosyltransferase (UPRT)-deficient parasites subjected to ambient (0.03%) CO₂ (Bohne W, Roos DS., 1997; Dziarszinski F *et al.*, 2004) promote significantly tachyzoite-to-bradyzoite conversion *in vitro*.

In vivo, it is during the initial and acute phase of the infection that the tachyzoite-bradyzoite transition occurs in a minor subpopulation while concurring with the massive destruction of the bulk tachyzoite population as a result of a rapid Th1 cell-mediated and short-term proinflammatory response. In this context, IL-12, TNF, and IFN- γ cytokines act at the

frontline of defenses against *T. gondii*, with IFN- γ also guaranteeing long-term persistence. Conversely, decline of IFN- γ level correlates with cerebral toxoplasmosis in AIDS patients (Pereira-Chioccola VL *et al.*, 2009; Meira CS *et al.*, 2014), whereas genetic loss of IFN- γ renders mice extremely susceptible to toxoplasmosis (Yap GS, Sher A., 1999). While *in vivo* IFN- γ was suspected to promote stage conversion, the cytokine failed to induce bradyzoite cultured in human fibroblasts (Soète M *et al.*, 1994), though the IFN- γ signaling pathway was shown to be restricting parasite growth in this cell type (Gay G *et al.*, 2016). On the other hand, treatment with IFN- γ does induce conversion to bradyzoites in murine macrophages and astrocytes, presumably owing to the stimulated releases of NO (Jones TC *et al.*, 1986; Bohne W *et al.*, 1993).

Whether those stimuli act on the parasite directly or indirectly by stressing the host cell is still under consideration. Both the host cell background and the parasite strain type were shown to contribute to spontaneous differentiation *in vitro*, albeit cysts are more frequently present in differentiated long-lived host cells (Ferreira da Silva Mda F *et al.*, 2008; Dubey JP *et al.*, 1998). In line with the environmental hypothesis, Compound 1 was shown to act directly on human host cells slowing tachyzoite replication and thereby inducing bradyzoite-specific gene expression in both type II and III strain parasites (Radke JR *et al.*, 2006). Conversely, early studies argue against it by showing that extracellular tachyzoites briefly exposed to alkaline stress were able to convert at low frequency into bradyzoites following their reinfection of host cells (Weiss LM *et al.*, 1998; Yahiaoui B *et al.*, 1999). Those experiments brought the first evidence that extracellular parasites are able to launch upon stimulation a bradyzoite development program that persists after invasion. Our team had originally suggested that this 'short term memory' driving bradygenesis may be also due to epigenetic-mediated gene regulation (see below).

It was also suggested that stage differentiation may be under the control of translational mechanisms. Thus, bradyzoite development was shown to be accompanied by phosphorylation of the parasite eukaryotic initiation factor 2 alpha subunit (eIF2 α), which dampens global protein synthesis and reprograms gene expression (Narasimhan J *et al.*, 2008). This result was substantiated by the observation that inhibitors of eIF2 α dephosphorylation slow replication and stabilize latency in *T. gondii* (Konrad C *et al.*, 2013; Sullivan Jr WJ *et al.*, 2009). In *Plasmodium*, IF2 α phosphorylation is high in infectious sporozoite form which is quiescent in mosquito salivary glands (Zhang M *et al.*, 2010)

I- 5. Regulation of gene expression in *Apicomplexa*

Transcriptional profiling across asexual development of *P. falciparum* (Bozdech Z *et al.*, 2003; Le Roch KG *et al.*, 2003) and *T. gondii* (Radke JR *et al.*, 2005) suggests a model where a cascade of gene expression results in a ‘just-in-time’ production of products only when needed. How these changes are regulated at the molecular level remains, to a large extent, unknown.

I- 5.1. Apicomplexan ApiAP2: the vegetal face of gene regulation in *Apicomplexa*

An unexpected feature is the apparent lack in *Apicomplexa* of large families of recognizable specific transcription factors (TFs) operating in other eukaryotes (Iyer LM *et al.*, 2008). This observation is paradoxical given the complex life cycle of these parasites, which certainly require a tight regulation of gene expression. What we and others have initially proposed was that although many phenotypic differences in *T. gondii* are genetically encoded, epigenetic control could be part of the parasite developmental programs and adjustments to fluctuant environment (Sullivan WJ Jr, Hakimi MA., 2006; Hakimi MA, Deitsch KW., 2007). Epigenetic regulation, which includes potentially heritable changes in gene expression that do not involve changes in DNA sequence, provides a mechanism by which an organism can maintain a type of short-term memory of its most recent environment, allowing it to respond quickly to changing conditions. The initial hypothesis had to be re-assessed in light of the discovery in *Apicomplexa* of an expanded family of plant-like TFs harboring APETALA2 (AP2) DNA-binding domain (Balaji S *et al.*, 2005; Iyer LM *et al.*, 2008; Oberstaller *et al.*, 2014). The AP2 domain was initially identified in plant kingdom, where it is a defining feature of a major family of AP2/ERF proteins. These proteins control the growth and development of plants by acting as activators or repressors of gene expression (Riechmann JL *et al.*, 1998; Licausi F *et al.*, 2013).

Based on sequence annotation, there are a total of 68 and 27 AP2-like transcription factors in *T. gondii* and *P. falciparum*, respectively (Altschul SF *et al.*, 2010; De Silva EK *et al.*, 2008). Many studies focused on how apicomplexan AP2 domain-containing (ApiAP2) proteins contribute to the regulation of gene expression at each stage transition of parasites. In *Plasmodium spp*, some ApiAP2 proteins were shown to play different functions in parasite differentiation throughout the life cycle. For examples, in *Plasmodium berghei*, a malaria

parasite of rodents, transcription factor PbAP2-Sp (ApiAP2 in sporozoites) which is expressed during the transition from the late oocyst to the salivary gland sporozoite, regulates the expression of all known genes in the sporozoite stage (Yuda M *et al.*, 2010). Similar to PbAP2-Sp, PbAP2-O (ApiAP2 in ookinete stage) regulates also several genes known to be specifically expressed in the ookinete stage including genes necessary for midgut invasion during development in midgut lumen of mosquito (Yuda M *et al.*, 2009). These two PbAP2 share the common feature in that they activate genes by directly binding to the proximal promoter regions. The binding sites at the promoter may be a factor characterizing promoter activities in the ookinete and sporozoite stages (Yuda M *et al.*, 2010). During infection and development of *P.berghei* in hepatocytes, transcription factor AP2-L was shown to play a crucial role for liver- stage development of parasite. The disruption of *PbAP2-L* in parasites does not affect the proliferation of other stages in the life cycle, while it causes the arrest of parasites development in the liver (Iwanaga S *et al.*; 2012). During parasite sexual differentiation, *Plasmodium falciparum* PfAP2-G (ApiAP2 for gametogenesis) may act as a master regulator of sexual-stage development *P.falciparum* because disrupting PfAP2-G results in loss of gametocyte formation in malaria parasites. It was suggested that PfAP2-G may control the decision of cells to commit to gametocyte formation or by default to continue the asexual replication (Kafsack BF *et al.*, 2014). However, in *Plasmodium berghei*, PbAP2-G is essential for the development switch from asexual replicating forms to sexual development in parasite (Sinha A *et al.*, 2014).

Despite a large number of ApiAP2 transcription factors in *T.gondii*, few were shown to control parasites development in the asexual life cycle. The nuclear TgAP2XI-4 is cell-cycle regulated, the detection of protein product is high during the end of cytokinesis and the beginning of G1 phase in tachyzoites. When compared to replicating tachyzoite, the transcript of TgAP2XI-4 is strongly abundant in brain tissue cysts of chronically infected mice. The disruption of *TgAP2XI-4* gene impairs the bradyzoite gene expression known to be expressed under stress condition (pH8.2), and altogether validating the idea that AP2XI-4 controls the induction of bradyzoite gene expression program and cyst formation (Walker R *et al.*, 2013). Another nuclear TgAP2XI-9 which is repressed in tachyzoites and upregulated in the early bradyzoite stage under pH8.2 treatment, was shown to bind functionally sequence elements of bradyzoite promoter. The disruption of *TgAP2IX-9* gene increases the tissue cyst formation while the overexpression of this protein inactivates suggesting its role in restriction of tissue cyst development (Radke JB *et al.*, 2013). While many ApiAP2 transcription factors are cell-

cycle regulated (~ 24 ApiAP2), the nuclear TgAP2XI-5 is constitutively expressed throughout the tachyzoite cell cycle, binds to also to promoters of more than 300 *Toxoplasma* genes including genes coding for virulence factors. This protein may regulate gene expression of virulence factors. However, the direct and inducible knockout of *TgAP2XI-5* gene was unsuccessful, suggesting its essential role in tachyzoite proliferation (Walker R *et al.*, 2013).

The global regulation of stage-specific gene expression mediated by ApiAP2 proteins in *Apicomplexa* remains a path to explore and better understand how parasites manage to control the development and differentiation within their life cycle.

I- 5.2. Contribution of chromatin structure to gene regulation in Eukaryotic cells

As mentioned above, transcriptional regulation clearly plays a major role in bradyzoite development as evidenced by numerous studies showing stage-specific gene expression. How these changes are regulated at the molecular level remains to a large extent unknown. What we and others have initially proposed was that although many phenotypic differences in *T. gondii* are genetically encoded, epigenetic control could be part of the parasite developmental programs and adjustments to changing environmental conditions. The so called ‘epigenetic events’ mainly regulate chromatin structures and gene function.

The packaging of DNA into chromatin in eukaryotic cells is recognized to be a major mechanism by which the access of genomic DNA is restricted and regulated. Regulation of gene expression requires alterations of chromatin architecture modulated by various number of enzymes. This results in temporal and spatial regulation of different cellular processes such as transcription, replication, DNA repair, recombination and chromosome segregation. As the basic functional unit of chromatin structure, nucleosome contains 147 base pairs of genomic DNA wrapping around a core histone octamer. Each octamer is composed of a stable tetramer of histones H3-H4 flanked by two separate H2A-H2B dimers. The interactions between nucleosomal histones involve a structured globular domain of histone. Protruding out from this globular domain of each histone, are the unstructured amino-terminal domain as well as carboxy-terminal domain. Moreover, nucleosomes are connected each other by the linker DNA which in turn binds to the linker histone H1 near the nucleosome entry/exit site, leading to the higher-order chromatin conformation. The chromatin is constantly remodeled to provide either a more accessible and transcriptionally active structure termed ‘euchromatin’ or

a more compact, less accessible and transcriptionally silent structure termed 'heterochromatin' (Li B *et al.*, 2007). A number of molecular processes play a role in the dynamic transition between these two transcriptionally distinct states, including nucleosome assembly and remodelling, histone replacement, reversible modification of histones, methylation of DNA and changes in sub-nuclear localization of chromatin (Goldberg AD *et al.*, 2007; Kouzarides T., 2007; Li B *et al.*, 2007). Chromatin regulators do not have DNA-binding properties of their own and are recruited to genomic loci by specific transcription factors and through the mediation of corresponding co-activators or co-repressors proteins (Young RA, 2011).

I- 5.2. Contribution of histone PTMs to gene regulation in Eukaryotic cells

In Eukaryotes, the timely opening and closing of chromatin required for gene expression, chromosomal organization, DNA repair or replication is governed by histone turnover and their post-translational modifications (PTMs), such as lysine methylation (me) and acetylation (ac) among many others. While DNA methylation is a highly stable silencing mark that extends over long chromosomal regions leading to 'memorized' states of gene expression, PTMs of histone subunits may be more labile and mediate regulation of gene expression over shorter-term periods. The functional consequences of histone PTMs can be direct, causing structural changes to chromatin and thereby affecting the DNA accessibility, or indirect, acting through the recruitment of effector proteins (Goldberg AD *et al.*, 2007; Kouzarides T., 2007). Various models such as the charge-neutralization model (Roth SY, Allis CD.,1992; Shogren-Knaak M *et al.*, 2006) and the 'histone code' (Strahl BD, Allis CD., 2000; Turner BM., 2000), as well as a signalling pathway model (Schreiber SL, Bernstein BE.,2002), have been proposed to explain the role of histone PTMs in transcription. The charge neutralization model, in which histone PTMs directly affect chromatin compaction, include phosphorylation or acetylation on core histones that serves to attenuate the favourable coulombic interactions between basic histone proteins and the negative charge of the DNA (Shogren-Knaak M *et al.*, 2006). The 'histone code' hypothesis predicts that diverse covalent modifications within the highly accessible histone tails are read by effector molecules, which in turn mediate distinct outcomes (Strahl BD, Allis CD., 2000; Turner BM., 2000). In this model, PTMs work in concert, and the cross-talk between different modifications determines the final biological readout. In this context, some modifications can influence others, and it appears that specific combinations of these modifications can form a complex and dynamic code. In addition,

PTMs also act as signals to recruit ATP-dependent remodeling enzymes to either move, eject or reposition nucleosomes.

Most of the tails PTMs do not affect the chromatin structure directly, but rather function as signaling platforms by recruiting or excluding effector proteins to chromatin and subsequently alter nucleosomal structure to regulate DNA accessibility (Jaskelioff M *et al.*, 2003; Tropberger P *et al.*, 2010). One exception is H4K16 acetylation that apparently alters the inter-nucleosomal interactions and therefore contribute to chromatin decondensation and increase gene transcription both *in vivo* and *in vitro* (Shogren-Knaak M *et al.*, 2006).

While PTMs have been primarily detected in the histone tails sticking out from the nucleosome, an ever-growing list of PTMs is now identified in the lateral surface of the histone octamer, that directly contacts DNA, and characterized as critical regulators of the chromatin structure and function (Lawrence M *et al.*, 2016; Tropberger P, Schneider R., 2013) Those ‘core’ histone PTMs promote different outcomes on nucleosome dynamics depending on their precise location. Modifications near the DNA entry-exit region of the nucleosome were shown to favor the local unwrapping of DNA from the histone octamer thereby providing a better exposure of nucleosomal DNA to chromatin-remodeling and DNA-binding proteins. On the other hand, lateral-surface PTMs mapping close to the dyad axis were shown to decrease the affinity of the octamer to DNA and significantly affect nucleosome stability/mobility.

Thus, by sterically altering the intranucleosomal histone-DNA interactions, lateral surface PTMs, such as acetylation and phosphorylation, were shown to influence all DNA-based processes, including transcription, replication and repair. For instance, H3K36ac by increasing DNA unwrapping at the entry-exit point of the nucleosome (Neumann H *et al.*, 2009) promotes chromatin disassembly during transcriptional activation (Williams S. K *et al.*, 2008) and, in response to DNA damage, favors chromatin reassembly after DNA repair has been ended (Chen CC *et al.*, 2008; Das C *et al.*, 2009)

In the nucleosome dyad where histone–DNA interactions reach their maximum strength, H3K122ac is sufficient to reduce nucleosomal density around the transcriptional start site of genes and to stimulate transcription (Tropberger P *et al.*, 2013). Therefore, a slight decrease in the overall charge of the lateral surface by only one unit (e.g. corresponding to the

neutralization of one lysine by acetylation) can functionally contribute to transcription suggesting that H3K122ac is not merely a readout of gene activity (Lawrence M *et al.*, 2016).

Similarly to what has been described for histone tails, different lateral-surface modifications on the same residue could be associated with opposite transcriptional programs. This is the case for the H3K64 residue near the dyad axis when acetylated facilitates nucleosome eviction and thereby gene expression (Di Cerbo V *et al.*, 2014), whereas trimethylation of the same residue acts as a repressive heterochromatic mark (Daujat S *et al.*, 2009). It is well appreciated that lateral-surface modifications were shown to instruct transcriptional changes at gene promoters, however, as described for H3K64ac and H3K122ac, they are able to act at long-distance by altering nucleosomal stability/mobility in the vicinity of enhancer (Pradeepa MM *et al.*, 2016).

Histone modifications serve as an effective way to regulate gene transcription but they do not operate alone; rather, they act in concert with other putative epigenetic information carriers (histone variants, small RNAs) and DNA sequence-specific transcription factors to modulate the higher-order structure of the chromatin fiber and govern the on-time recruitment of the transcriptional machinery to specific genes.

I- 5.3. An unexpected and sophisticated ‘histone code’ at work in *Apicomplexa*

Unlike the majority of higher Eukaryotes, *Apicomplexa* genomes have a unique chromatin architecture typified by an unusually high proportion of euchromatin and only a few heterochromatic islands scattered through the chromosome bodies or embedded at telomeres and centromeres (Vanagas L *et al.*, 2012; Nardelli SC *et al.*, 2013; Coetzee N *et al.*, 2016; Ponts N *et al.*, 2010; Bunnik EM *et al.*, 2014). Although alterations in chromatin structure are acknowledged as important for the transcriptional control of commitment to stage differentiation in several *Apicomplexa* as well as for antigenic variation-mediated immune evasion in *P. falciparum*, yet the molecular mechanisms of chromatin remodeling have not been fully determined (Bougdour A *et al.*, 2010; Scherf A *et al.*, 2008). As *P. falciparum* or *T. gondii* do not appear to have detectable DNA cytosine methylation (Gissot M *et al.*, 2008), remodelling of the chromatin structure particularly through PTMs of histones is potentially a major process that co-ordinates regulation of gene expression. For this reason, histone modifications have rapidly moved to the forefront of gene regulation research in many protozoan parasites. Histones are well conserved across species and apicomplexan parasites

are no exception. In every apicomplexan parasite analyzed so far, the four core histones (H2A, H2B, H3 and H4) have been identified, but the linker histone H1 appears to be absent. Histone H2B is represented by two isoforms H2Ba and H2Bb. These parasites have also five histone variants: centromeric H3 CenH3, H3.3, H2A.X, and H2A.Z ((Dalmasso MC *et al.*, 2011; Vanagas L *et al.*, 2012). Interestingly, apicomplexan parasites contain a novel variant of H2B family named H2Bv that forms in *T. gondii* dimers mainly with H2AZ, but not with H2AX (Dalmasso MC *et al.*, 2006; Dalmasso MC *et al.*, 2009; Bogado SS *et al.*, 2014). These findings reveal that the nucleosomal arrangement is not random in protozoa and may exhibit intriguing differences relative to their mammalian counterparts.

With regard to histone tails PTMS, the epigenome is much more complex than previously thought in *Apicomplexa*. We and others have begun our investigations by looking at the evolutionary conservation of histone PTMs. The proteome-wide study of lysine acetylation in *T. gondii* revealed that acetylation is abundant occurring on histones H2AZ, H2B variant, H2Bb/H2Ba, H3/H3.3, and H4 but also on nonhistone proteins of diverse function and localization in parasite (Jeffers V *et al.*, 2012; Xue B *et al.*, 2013). The lysines acetylation of histone tails that were documented in *T. gondii* comprise K9, K14, and K18 on histone H3 as well as K5, K8, K12, and K16 on histone H4. Other novel lysines acetylation within globular domain of histones have been reported, H3K79 and H4K31, although the acetylation role of these residues has yet to be investigated (Bougdour A *et al.*, 2010; Jeffers V *et al.*, 2012). However, few divergences were observed so far between apicomplexan and other species although there were substantial differences in their abundances (Bougdour A *et al.*, 2010). The most remarkable finding was that H4K16 is not the major acetylation site in *P. falciparum* histone H4, which contrasts with what is observed in most other organisms (Smith CM *et al.*, 2003; Zhang K *et al.*, 2002). Instead, acetylation of H4K8 and H4K12 was frequently observed in both *P. falciparum* (Trelle MB *et al.*, 2009) and *T. gondii* (Hakimi MA, unpublished data). In general, parasites appear to harbour more activation (H3K4me or H3K9ac) than silencing marks (H3K9me) when compared with human cells (Trelle MB *et al.*, 2009). The histone variants are no exception to this rule, since they are extensively acetylated in their tails probably to maintain open chromatin structure (Trelle MB *et al.*, 2009). This observation mirrors the studies on *Saccharomyces cerevisiae* and *Tetrahymena thermophila* that exhibit more modifications associated with transcriptional activation than repressive marks, whereas the reverse situation is observed in mammals (Garcia BA *et al.*, 2007). This correlates quite well with the fact that the majority of the *P. falciparum* genome is

transcriptionally competent (Bozdech Z *et al.*, 2003; Le Roch KG *et al.*, 2003), whereas more than 60% of the mammalian genome is permanently silenced (Jenuwein T *et al.*, 1998; Whitfield ML *et al.*, 2002). Few divergences were observed so far between *T. gondii* and *P. falciparum*. The one known exception is the serine 10 phosphorylation of histone H3, which is involved in both transcription and chromosome condensation during mitosis in eukaryotic cells (Prigent C, Dimitrov S., 2003). While in *T. gondii* H3S10 is phosphorylated (H3S10ph) prior to mitosis in tachyzoites (Sautel C *et al.*, 2007), the mark is apparently absent or too low to be detected in *P. falciparum* (Trelle MB *et al.*, 2009). This discrepancy is interesting and calls for further investigation.

Studies mapping specific PTMs over the *T. gondii* or *P. falciparum* genome using chromatin immunoprecipitation (ChIP) suggest that the interplay between chromatin and transcription is dynamic and more complex than previously appreciated (Gissot M *et al.*, 2007; Lopez-Rubio JJ *et al.*, 2009; Salcedo-Amaya AM *et al.*, 2009). Genome-wide data obtained from *T. gondii* revealed that H3K9me2 and H3K9me3 repressive marks tend to spread over large regions of heterochromatin, particularly near the centromeres (Brooks CF *et al.*, 2011). In contrast, in *P. falciparum* H3K9me3 exhibits a restricted pattern at subtelomeric regions with no enrichment at the pericentromeric chromatin like in *T. gondii* (Lopez-Rubio JJ *et al.*, 2009) TgSet8 that methylates H4K20me1,3 appears also to associate with the *T. gondii* centromeres themselves (Sautel CF *et al.*, 2007). Those genome-wide approaches clearly showed that histone PTMs are sequestered to distinct regions of the parasite genome and they eventually correlate with sites of transcription *in vivo*, though whether they contribute functionally to gene induction remains elusive.

I- 6. Overview of the histone modifying enzymes

Qualitative examination of *Apicomplexa* genomes reveals a rich but still largely unexplored repertoire of ATP- dependent remodelers and histone-modifying enzymes that move, replace or decorate histones (Sullivan WJ Jr, Hakimi MA., 2006; Hakimi MA, Deitsch KW., 2007; Bougdour A *et al.*, 2010). What we learn so far is that these parasites possess a sophisticated capacity to modify histones, rivaling the system observed in higher eukaryotic cells. For instance, there was significant duplication and divergence in the *Apicomplexa* lineage of the lysine methyltransferases (KMT) harboring a conserved Suv(39)-E(z)-TRX (SET) catalytic domain (Sautel CF *et al.*, 2007, Sautel CF *et al.*, 2009). In *T. gondii*, there is 20 SET domain-

containing KMTs and 2 Dot1 homologs. Some histone-modifying enzymes have acquired gain-of-function mutations that confer broader or enhanced activity on substrates. It is the case of parasite SET8-related proteins endowed with H4K20 mono-, di-, and trimethylase activities that contrast with the mono-methylase-restricted metazoan Set8, and that derive from a single-amino-acid change in the substrate-specific channel (Sautel CF *et al.*, 2007). Beyond lysine methylation, a more extensive repertoire of arginine methylation machinery is present in *T. gondii* compared with yeast and *C. elegans*; each possessing a unique PRMT related enzyme (Saksouk N *et al.*, 2005).

In general, duplication and divergence of histone modifiers is more pronounced in *T. gondii* than in *P. falciparum*. Notably, *T. gondii* appears to be unique inside the phylum in harbouring more than one GCN5 family histone acetyl-transferase (HAT) (Bhatti MM *et al.*, 2006) (Duplication does not mean redundancy since TgGCN5-B acetylates H3K9, K14 and K18, the expected substrate profile for archetypical GCN5 HATs (Bhatti MM *et al.*, 2006), while TgGCN5-A exhibits a strong tendency to acetylate in vitro H3K18 (Saksouk N *et al.*, 2005). In terms of diversity, the chromatin-modifying apparatus of *T. gondii* goes beyond what was expected for a unicellular eukaryotic cell and this was consistent with the initial hypothesis that the apparent lack of traditional transcription factors may be somehow compensated at an epigenetic level.

I- 7. The HAT/HDAC interplay: multilevel control of chromatin structure and gene regulation

Dynamic acetylation of the ϵ -amino group of specific lysine residues is a reversible PTM, evolutionary conserved from bacteria to humans and one of the well characterized modifications. The role of histone acetylation in the regulation of chromatin structure in higher eukaryotes involves neutralization of the positive charge of the histone N-terminal tails. This attenuation between the histone proteins and the DNA leads to chromatin decondensation, thereby enhancing transcriptional activity. On the other hand, histone hypoacetylation restores the positive charge of histone N-tails, which tightens the binding of DNA and histones, leading to condensed chromatin and gene silencing. Histone acetylation is dynamically regulated by antagonistic actions of two families of enzymes, histone acetyltransferases HATs and histone deacetylases HDACs. Histone acetylation is almost associated with activation of transcription, while conversely histone deacetylation modulates

transcriptional repression (Haberland M *et al.*, 2009). The interplay between HDACs and HATs results in dynamic of chromatin structure and regulation.

Histone acetylation, in *T. gondii* is catalyzed by five distinct HATs which are important for parasite survival and proliferation. Based on subcellular localization and structural homology in the primary sequence, *T. gondii* HATs include two GCN5 families (GCN5-A and -B), two MYST families (MYST-A and MYST-B), and Hat1 (Dixon SE *et al.*, 2010).

HDAC enzymes have been organized into four phylogenetic classes based on sequence homology, inhibitor sensitivity, and cofactor necessity (de Ruijter AJ *et al.*, 2003). Class I of HDACs, which is homologous to Rpd3 in yeast, consists of HDAC1, 2, 3, and 8 and localizes in the nucleus to regulate histone and nonhistone acetylation. Class II of HDACs is homologous to yeast Hda1 and has two subtypes, class IIa which includes HDAC4, 5, 7, and 9 and class IIb comprises HDAC6 and 10. In response to different cellular stresses, class IIa of HDACs can shuttle between the cytoplasm and the nucleus, whereas class IIb locates predominantly in cytoplasm. Class III of HDACs are also called sirtuins which include seven members Sir1-7. They localize in the nucleus, cytoplasm, mitochondria and share homologous sequence with the yeast Sir2 family of proteins. Class IV of HDACs has only HDAC11 mainly located in the nucleus and highly conserved from *C.elegans*, *D.melanogaster* to humans (de Ruijter AJ *et al.*, 2003; Xu K *et al.*, 2011; Yang XJ *et al.*, 2008). Several HDACs from class I were shown to form multi-subunit complexes that predetermine their enzymatic activity as well as their function in cells, e.g. apoptosis, cell cycle progression or development (Guenther MG *et al.*, 2000; Guenther MG *et al.*, 2001; Zhang J *et al.*, 2002; Li J *et al.*, 2000; Fischle W *et al.*, 2002; Lazar MA, 2003; Jayne S *et al.*, 2006; Telles E *et al.*, 2012; Ahringer J, 2000).

Toxoplasma genome encodes genes for six putative HDACs based on similarity with the human counterparts that branch into the four main classes; namely HDAC1, HDAC2, HDAC3, HDAC4, HDAC5 and Sir2 (Saksouk N *et al.*, 2005; Vanagas L *et al.*, 2012). In *Plasmodium falciparum*, many chromatin modifier enzymes such as histone deacetylases HDACs are essential for *Plasmodium falciparum* development. For instance, the Sir2 orthologue has been noted to have a critical role in mediating var gene expression in *P. falciparum* (Tonkin CJ *et al.*, 2009). Another example is the depletion of HDAC2, known to deacetylate the euchromatin mark H3K9ac, that results in the transcriptional of genes (example: var genes) located in previously defined heterochromatin regions and thus leading

to the failure of blood-stage development and an increased gametocyte (Coleman BI *et al.*, 2014). These activated genes (*var* genes) after HDAC2 depletion, were shown to be associated with HP1, a protein which binds to heterochromatic repressive mark H3K9me3. HP1 protein acts as a recruitment platform for histone lysine methyltransferase to maintain and spread H3K9me3 marks (Lomberk G *et al.*, 2006). Likewise, HP1 knockdown causes the loss of *var* gene expression, altogether suggesting that HDAC2 may be an upstream regulator of HP1 required for heterochromatin repression (Coleman BI *et al.*, 2014).

I- 8. The TgHDAC3 regulatory network defines a new ‘epigenetic’ path to explore *Toxoplasma gondii* gene expression regulation and differentiation

Our team took an early interest in histone acetylation (Saksouk N *et al.*, 2005) and found that there is a strong link between differentiation and the rate of histone acetylation (unpublished work by *Alexandre Bougdour*, Figure 3). Thus, switching the pH from 7.1 to 8.1 causes both tachyzoite-to-bradyzoite conversion (Figure 3A) and also an unexpected hyperacetylation of histone H4 in the parasite nuclei, without any change of the acetylation status in the host cell nuclei (Figures 3B and 3C). This hyperacetylation is not mediated by any host cell signaling cascade as free parasites maintained in culture medium buffered at pH 8.1 also exhibit a drastic H4 acetylation (Figure 3D). The absence of this phenomenon in the host cell (Figure 3C) has encouraged us to further investigate the mechanisms involved.

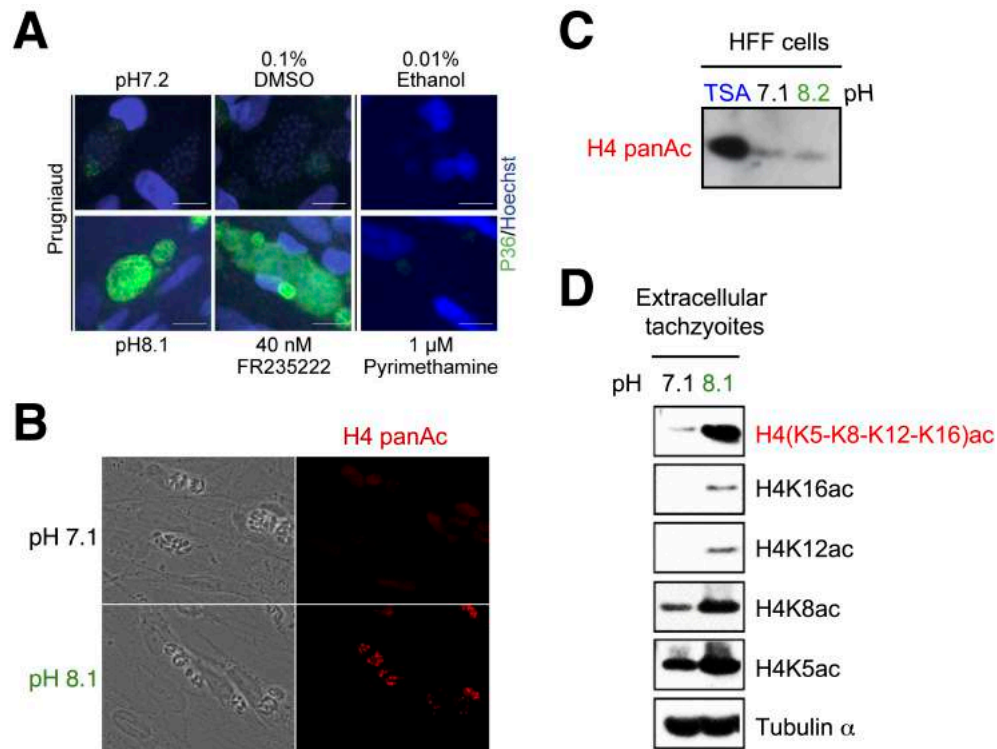


Figure 3 - Switching the pH triggers histone hyperacetylation exclusively in parasite nuclei. (A) FR235222 treatment induces *T. gondii* Prugniaud strain differentiation. Intracellular parasites (12h after infection) were incubated in the presence of either alkaline growth media to induce in vitro differentiation or 40nM FR235222 for 2–3 d. As a control, parasites were treated with 1 μ M pyrimethamine. Bradyzoite-differentiated parasites were identified by IFA using an anti-P36 (in green) antibody. Bars, 10 μ m. (B) HFFs cells were infected with tachyzoites and hyperacetylation of histone H4 tail in parasites nuclei was detected by immunofluorescence assay using pan-acetyl lysine histone H4 antibody (in red). (C) Primary human fibroblast (HFF) cells were treated with pH 7.1, pH 8.1 and trichostatin A (TSA). Histones were extracted and analyzed by immunoblotting using a pan-acetyl H4 antibody (in red). (D) Extracellular parasites were treated with pH7.1 and pH8.1. Histones were then extracted and analyzed in western blot by using antibodies against the acetylated lysines within histone H4 tail. Images from Bougdour A *et al.* 2009 and unpublished work from the team.

Acetylation of core histones is mediated by histone acetyltransferases (HATs) and, in many instances, results in relaxation of chromatin structure and transcriptional activation of associated genes. Histone deacetylases (HDACs) counteract HAT activity by catalyzing the removal of acetyl moieties from lysine residues in histone tails, thereby inducing chromatin condensation and transcriptional repression (Kurdistani SK *et al.*, 2003). Part of our research has been devoted to identifying new drug targets for anti-Apicomplexa. HDAC inhibitors (HDACi) have been widely evaluated for their cytotoxic, anticancer and other potential therapeutic properties (Bertrand P, 2010). Our team investigated the mode of action of FR235222, a novel cyclic tetrapeptide HDACi and showed that the drug is active against a

wide range of *Apicomplexa* and is more potent than other HDACis such as trichostatin A and the clinically relevant compound pyrimethamine (Figure 4A) (Bougdour A *et al.* 2009).

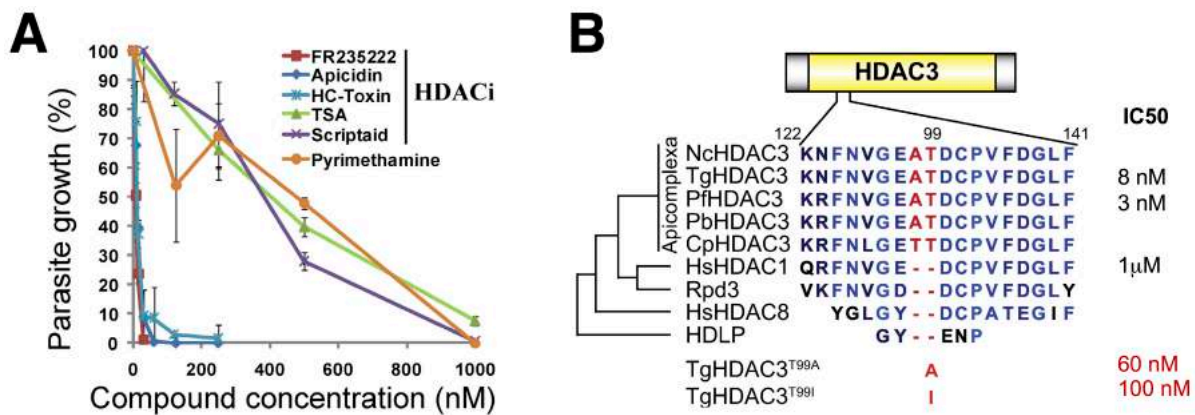


Figure 4. FR235222 targets TgHDAC3 in *T. gondii* (A) Effect of FR235222 and other HDACi's on *T. gondii* RH strain growth in HFF monolayer. Pyrimethamine (non-HDACi compound) was used as a clinically relevant control. Means \pm SD of parasite growth (percentages) are shown (n = 3 experiments). (B) Sequence alignment of TgHDAC3 homologues in Apicomplexan parasites and other organisms. The region (from amino acids 122–141 of TgHDAC3) surrounding the point mutation identified in TgHDAC3-resistant mutants is shown. Rpd3 and HDLP are the HDAC homologues in *Saccharomyces cerevisiae* and the hyperthermophilic bacterium *Aquifex aeolicus*, respectively. Point mutations identified in the *T. gondii* FR235222-resistant mutants are shown at the bottom and as well the corresponding IC50. Images from Bougdour A *et al.* 2009.

In the same study, we isolated *Toxoplasma* parasite lines resistant to the molecule. Single-point mutations found in mutagenized parasites target the amino acid T99 in TgHDAC3 and are sufficient to confer resistance to FR235222 (Bougdour *et al.*, 2009). Interestingly, the residue T99 along with the amino acid A98 creates an insertion within the catalytic site of the enzyme that is exclusively conserved in apicomplexan HDAC3 family of proteins and absent in any other eukaryotic HDAC (Figure 4B). Phenotypic analysis of parasite mutants and their responsiveness to the drug suggests that TgHDAC3 is a regulator in *T. gondii* of both cell cycle and differentiation (Bougdour A *et al.*, 2009).

We think that TgHDAC3 is a key regulator acting within the regulatory pathway leading to parasite differentiation. Several arguments support this assertion. First, inhibition of TgHDAC3 by FR235222 mimics the pH stress by inducing the expression of bradyzoite markers, which we believe marks the beginning of stage conversion (e.g. P36, a bradyzoite marker Figure 3A). Additionally, even in the absence of drug, P36 expression became constitutively expressed in the TgHDAC3^{T99A} mutant compared to wild-type parasites (Figure 5A). This suggests that a simple point mutation is able to affect the basal activity of the

enzyme that was confirmed in the vicinity of our gene model *20.m00351* (Figure 5B, compare the yield of H4 enrichment in the absence of FR235222 in WT versus mutant background).

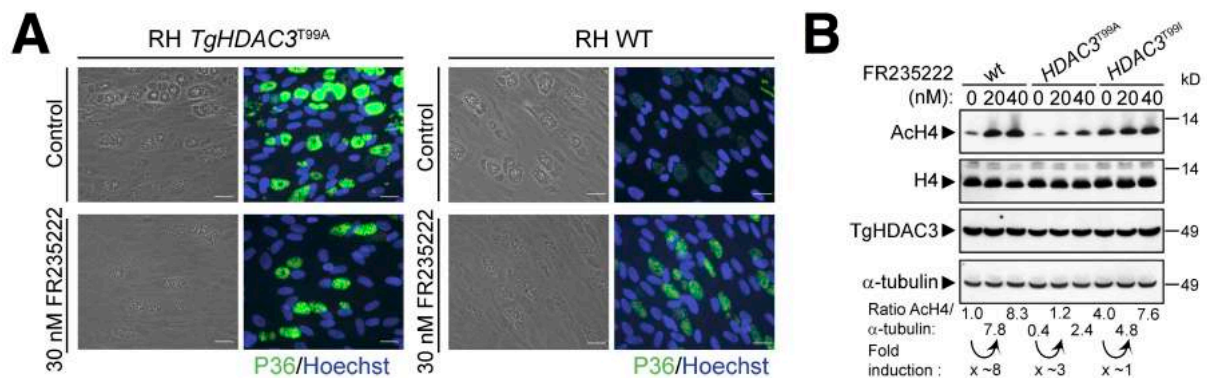


Figure 5. **The T99A mutation affects the regulatory role of TgHDAC3 on P36/SRS9 expression.** (A) FR235222 treatment induces *T. gondii* RH strain differentiation, as revealed by the expression of the bradyzoite-specific antigen P36 (SRS9). Intracellular parasites (6 h after infection) were treated with low concentrations of FR235222 (30 nM) for 24 h. Effect of the TgHDAC3^{T99A} allele on the expression of the bradyzoite-specific antigen P36. *T. gondii* RH strain WT and R20D9 mutant (TgHDAC3^{T99A}) were fixed 24 h after infection and analyzed by IFA for P36 expression levels (n = 3 experiments). Representative data are shown. Bars, 50 μ m. (B) Scanning ChIP experiments showing the effects of FR235222 on AcH4 levels in the presence of the TgHDAC3 WT, TgHDAC3^{T99A}, and TgHDAC3^{T99I} alleles at the promoter region of the bradyzoite-specific gene *20.m00351*. The ratios of *20.m00351* and control DHFR signals present in input samples were used to calculate the relative precipitated fold enrichment shown below each lane. Images from Bougdour *et al.* 2009.

Nonetheless, all these elements are scattered and do not provide conclusive evidence for the impact of histone acetylation on the expression of stage-specific genes. During my four-year thesis, I intend to investigate the contribution of TgHDAC3 to transcriptional regulation and parasite differentiation in *Toxoplasma gondii*.

This work is organized into three chapters:

Chapter 1. *Toxoplasma* ApiAP2 transcription factors are embedded in stable multi-subunit complexes gathering the deacetylase TgHDAC3, the MORC protein CRC230 and ELM2-containing scaffolding proteins.

Chapter 2. The chemical inactivation of TgHDAC3 by FR235222 revealed a peculiar acetylome and proteome regulated by the enzyme: a first step toward stage differentiation.

Chapter 3. A versatile acetylation-methylation switch at K31 on the lateral surface of histone H4 dictates chromosomal organization and expression in *Apicomplexan* parasites.

Part II. Results

Chapter 1.

***T. gondii* ApiAP2 transcription factors are embedded in stable multi-subunit complexes gathering the deacetylase TgHDAC3, the MORC protein CRC230 and ELM2-containing scaffolding proteins**

II- 1.1. Rationale and Preliminary Studies:

Most of the Eukaryotic HDAC enzymes act in megadalton complexes containing corepressors and scaffolding proteins. Most often these enzymes are embedded into a network of proteins, which modify or directly support their enzymatic activity. Usually, histone deacetylases are unable to access their histone substrates, unless DNA-bound activators or repressors target them there. *Apetala2* (AP2)-related proteins, which contain plant-like DNA-binding domain, represent a major lineage of transcription factors in *Apicomplexa* (Balaji S *et al.*, 2005; Iyer LM *et al.*, 2008; Oberstaller *et al.*, 2014). They are serious candidates for targeting the TgHDAC3 to DNA in a sequence specific manner.

In Eukaryotes, the timely opening and closing of chromatin is required for gene expression. Chromatin modifying and remodeling enzymatic activities serve as an effective way to regulate gene transcription but they do not operate alone. They do not themselves bind DNA and their recruitment by specific transcription factor to specific genomic *loci* is ordinarily mediated by co-activator or co-repressor. Interestingly, our team purified a core repression complex (CRC) in *Toxoplasma*, using RH tachyzoites expressing ectopically a second copy of TgHDAC3 fused to HA-Flag (Saksouk *et al.*, 2005). TgHDAC3 was shown to co-purify with the transcription factor AP2VIII-4 (CRC350), the transducin beta-like protein1 TBL1 and the CRC230 protein. Another study reports the interaction of the HAT GCN5b with two distinct transcription factors AP2IX-7 and AP2X-8 (Wang J *et al.*, 2014). Collectively, these observations indicate that the *Toxoplasma* chromatin enzymes TgHDAC3 and GCN5b may involve, for their recruitment to DNA, AP2 domain-containing proteins.

A real breakthrough in the understanding of TgHDAC3 came during my PhD with the discovery demonstrating that TgHDAC3 was able to form multiple complexes in addition to CRC.

II- 1.2. Results

II- 1.2.1. TgCRC230 is a predicted nuclear MORC-related protein in *Toxoplasma gondii*

Toxoplasma gondii CRC230 belongs to the microorchidia (MORC) family. The protein sequence analysis with Simple Modular Architecture Research Tool SMART shows that CRC230 consists of six kelch repeats, histidine kinase-like ATPase domain (HATPase_C) and putative nuclear localization signals (NLS) (Saksouk N *et al.*, 2005) (Figure 6A). Moreover, previous studies reported that the *Apicomplexa* phylum evolves a unique version of the MORC ATPase fused to kelch-type β - propellers (Iyer LM *et al.*, 2008). The MORC proteins are nuclear and highly conserved in both Prokaryotes and Eukaryotes. They are characterized by a gyrase, histidine kinase, and MutL (GHKL) domain combined with a C-terminal S5 domain that together form an ATPase module (Inoue *et al.*, 1999; Iyer LM *et al.*, 2008). They were first identified in mice and involved in male meiosis and spermatogenesis (Lorković ZJ, 2012).

In human, there are five predicted MORC members including MORC1, MORC2, MORC3, MORC4, and the divergent SMCHD1 some of which are upregulated in breast cancer while others regulate P53 activity and induce cellular senescence (Li DQ *et al.*, 2013). In addition, members of the MORC family in fusions with other protein domains emerged as new regulators of transcription and DNA damage response. For example, the phosphorylation of MORC2 by P21-activated kinase1 (PAK1) orchestrates the interplay between chromatin dynamics and the maintenance of genomic integrity during DNA damage response (Li DQ *et al.*, 2012). In plant, MORC ATPase proteins found in *Arabidopsis thaliana* are involved in heterochromatin condensation and gene silencing (Moissiard G *et al.*, 2012). Another interesting study has shown the involvement of MORC2 and histone deacetylase 4 in transcriptional repression of CAIX gene in gastric cancer cells (Shao Y *et al.*, 2010).

Consistent with the above observations, we hypothesized that TgCRC230 protein that was previously found in TgHDAC3-containing complex may play a key role in gene repression in tachyzoites.

II- 1.2.2. TgCRC230 acts as scaffolding platform bridging TgHDAC3 to multiple ApiAP2 transcription factors and ELM2-containing proteins

When we monitored CRC230 dynamics in *T. gondii* lines expressing the endogenous protein in fusion with the HA-Flag tags, we found that the protein selectively accumulates in the parasite nuclei of tachyzoites (Figure 6B). Flag affinity of CRC230 followed by western blotting confirmed the partnership between CRC230 and TgHDAC3 in freshly egressed parasites or in fast-growing intracellular parasites in both RH and Pru strains (Figure 6D). This finding is in agreement with a previous study showing that HDAC3 is embedded in a remarkably robust molecular mass complex with CRC230 that resists to high stringent washing conditions (500mM KCl and 0.1% NP-40) (Saksouk N *et al.*, 2005).

Silver stain analysis of the CRC230 Flag eluates, however, revealed a more complex protein pattern than previously seen with the HDAC3-purified CRC (Figure 6E), suggesting that CRC230 may bind to more partners and those interactions are quite stable considering the stringency of washing. These partnerships were subsequently resolved by mass spectrometry-based proteomics that identified CRC230 and HDAC3 as part of a multisubunit complex encompassing ten ApiAP2 transcription factors, three ELM2/PHD domains-containing proteins and six hypothetical proteins (Table 1). Our findings indicate that CRC230 may act as a scaffold protein connecting ApiAP2 transcription factors with HDAC3 that, in turn, may be involved in recruiting the repressor complex to the target genes for its activity.

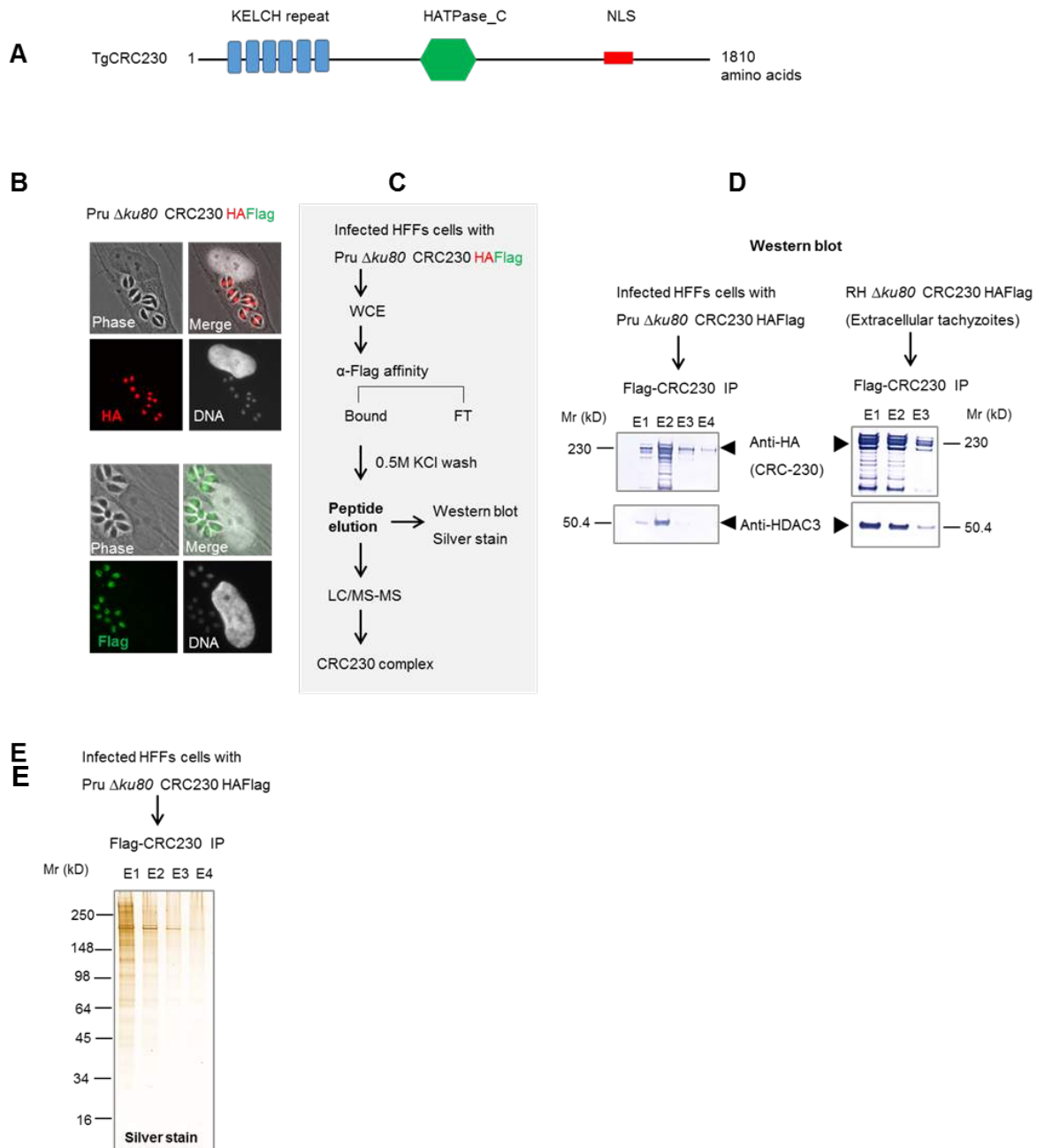


Figure 6. TgCRC230 is a nuclear MORC protein that binds directly to TgHDAC3 and forms a high stable molecular weight complex with transcription factors and ELM2 domain-containing proteins. (A) Schematic representation of structural domain of TgCRC230 highlighting kelch (in blue), HATPase_C (in green) and NLS (in red) domains. (B) TgCRC230 protein is located to the nucleus of tachyzoites parasites. Human foreskin fibroblasts HFFs cells were infected with Pru $\Delta ku80$ CRC230 HAFlag for 24h and fixed with 3% formaldehyde for 20 min and subjected to immunofluorescence assay using antibodies anti-HA and anti-FLAG. (C) Schematic strategy to obtain TgCRC230-associated polypeptides from intracellular Pru $\Delta ku80$ CRC230 HAFlag parasites. (D) Copurification of TgCRC230 and TgHDAC3 in both intracellular and extracellular tachyzoites. Isolated components, as shown in (C), were verified by western blot using antibodies anti-HA corresponding to CRC230 and anti-HDAC3. (E) Silver staining analysis of TgCRC230-associated polypeptides from intracellular tachyzoites Pru $\Delta ku80$ CRC230 HAFlag.

Table 1. Proteins associating with TgCRC230 in tachyzoites and identified by LC/MS-MS

Infected HFF cells with Pru $\Delta ku80$ CRC230 HAFlag

Mr (kDa)	ToxoDB Accession number	Gene description	Found in intracellular RH $\Delta ku80$ CRC230 HAFlag
230	TGME49_305340	CRC230	Yes
50.4	TGME49_227290	HDAC3	Yes
103	TGME49_288950	AP2IX-4	Yes
365	TGME49_285895	AP2V-2	Yes
350	TGME49_272710	AP2VIII-4	Yes
105	TGME49_205650	AP2VIIa-3	Yes
137	TGME49_282220	AP2VIIa-9	Yes
133	TGME49_227900	APX-1	Yes
346	TGME49_315760	AP2XI-4	Yes
247	TGME49_218960	AP2XII-1	Yes
179	TGME49_217700	AP2XII-2	Yes
161	TGME49_247730	AP2XII-5	Yes
48	TGME49_232650 ^a	ELM2 domain-containing protein	Yes
69	TGME49_275680 ^b	ELM2 domain – containing protein	Yes
84	TGME49_230890 ^c	PHD/ELM2 domains – containing protein	Yes
490	TGME49_234230	Hypothetical protein	Yes
328	TGME49_209500	Hypothetical protein	Yes
205	TGME49_214140	Hypothetical protein	Yes
326	TGME49_248720	Hypothetical protein	Yes
339	TGME49_242640	Hypothetical protein	Yes
118	TGME49_314220	Hypothetical protein	Yes

II- 1.2.3. TgCRC230 purification led to the identification of three putative ELM2 domain-containing proteins

Three ELM2 (EGL-27 and MTA1 homology 2)-containing proteins were also co-purified with CRC230 and TgHDAC3 (ELM2a, b, c; Table1). ELM2 domains were initially shown to interact with histone deacetylase in several studies (Ding Z *et al.*, 2003; Wang L *et al.*, 2006; Millard CJ *et al.*, 2013). The ELM2-SANT domain is one of the best-characterized regions of MTA1, which recruits histone deacetylase 1 (HDAC1) and activates the enzyme in the presence of inositol phosphate. More recently, the structure of the ELM2 and SANT domains from MTA1 bound to HDAC1 revealed that the ELM2-SANT domains wrap completely around HDAC1 occupying both sides of the active site such that the adjacent BAH domain is

ideally positioned to recruit nucleosomes to the active site of the enzyme (Millard CJ *et al.*, 2013). In the early stages of the project, we have attempted to solve the structure of the dimeric complex TgHDAC3/ELM2a (*TGME49_232650*) using various approaches and following the protocol of others (Watson PJ *et al.* 2012), but have encountered severe solubility problems. This led us to initiate a computational approach, as a close homologue model of HDAC-ELM2 was available, that enables the building of homology model. The 3D structure of TgHDAC3-TgELM2 was built through a classical homology/comparative modeling by using the Task “Homology Modeling” of the Maestro-Schrödinger Suite in collaboration with Dr. Amit Sharma (ICGEB, New Delhi). The TgHDAC3/ELM2a model was built using, as a template, the crystallographic structure of ELM2 of MTA1 complexed with human HDAC1 (pdb Id: 4BKX) (Millard CJ *et al.*, 2013) (Figure 7).

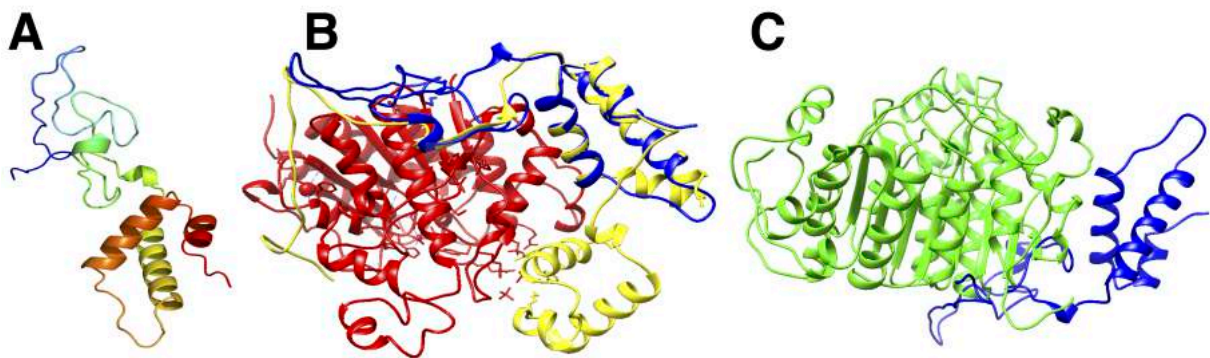


Figure 7. Structure prediction of *T. gondii* ELM2a in complex compared to human ELM2 in MTA1-HDAC1 complex. (A) Homology modeling of *T. gondii* ELM2a. (B) Structural alignment of *T. gondii* ELM2a (in blue) and *Homo sapiens* ELM2 of MTA1 (in yellow) complex to *Homo sapiens* HDAC1 (in red) (C) Docking model of *T. gondii* HDAC3 (in green) and ELM2a (in blue).

Structure prediction of the ELM2a domain suggests that the domain is largely intrinsically disordered (i.e. lacking an intrinsically fixed structure), although there are predicted helical regions at the carboxy-terminus (Figure 7A). As for human ELM2 (in yellow, Figure 7B), the amino-terminal part of *T. gondii* ELM2a domain adopts an extended conformation that wraps around the human HDAC1 (Figure 7B) and TgHDAC3 model (Figure 7C) making multiple interactions. The carboxy-terminal region of human ELM2 domain (in yellow, Figure 7B) forms a four-helix bundle with a small hydrophobic core and enlarges the interacting surface with HDAC1 as it completes a path around the ‘back’ of the enzyme. The four-helix bundle were shown to form a homodimer by bridging two HDAC1:MTA1 dimeric complex. *T.*

gondii ELM2 is typified by an exclusive two-helix structure and we hypothesize that this non-polar side chains-containing interface could act as a homodimerisation interface to connect two histone deacetylases molecules as described in NuRD complex.

Here, we thoroughly determine the TgHDAC3 interactome by purifying multiple TgHDAC3-containing holoenzymes using affinity and conventional chromatography in both virulent (RH) and cystogenic (Prugniaud) *Toxoplasma* strains. While we uncover a core set of ApiAP2 (10 over 68) as serious candidates for targeting the deacetylase to DNA in a sequence specific manner, we still need to address this hypothesis by co-mapping (ChIP-seq) the genome-wide location of the aforementioned ApiAP2 and TgHDAC3. Nonetheless, our data highlight for the first time a role in *Apicomplexa* for a MORC-like protein as a putative scaffolding protein bridging ApiAP2 to TgHDAC3. Moreover, by revealing the presence of ELM2/HDAC core complex in *Toxoplasma gondii*, we brought the first evidence that this partnership between the two families arose in eukaryotic cell long before the split between apicomplexan, plant and metazoan.

Chapter 2.

The chemical inactivation of TgHDAC3 by FR235222 revealed a peculiar acetylome and proteome regulated by the enzyme: a first step toward stage differentiation

II- 2.1. Rationale and Preliminary Studies:

Apart its central role in epigenetic regulation of histones, lysine acetylation has recently emerged as one of the major reversible post-translational modifications involved in diverse cellular processes such as cell migration, cellular metabolism, cytoskeleton dynamics or protein folding (Close P *et al.*, 2010). Identification of acetylation sites in targeted proteins is the first essential step for functional characterization of acetylation in physiological regulation. The recent development of anti-acetyl lysine antibodies dedicated to enrichment of acetylated peptides has rendered feasible the study of protein acetylation at the whole proteome level (Kim SC *et al.* 2006). High-throughput mass spectrometry analyses of acetylomes using such a strategy have revealed that protein acetylation is nearly as widespread as and probably more conserved than phosphorylation (Choudhary C *et al.*, 2009; Weinert BT *et al.*, 2011).

Previous results on TgHDAC3 role indicate that acetylation might be a major way to tightly regulate parasite differentiation (Saksouk N *et al.*, 2005). We initially speculated that FR235222 acts on TgHDAC3 primarily promoting the deregulation of histone acetylation rates, which in turn altered the expression of genes, especially those known to be stage-dependent (Bougdour A *et al.*, 2009). The overall effect on gene expression, however, did not identify genes coding for master regulators/transcription factors whose alteration of their expression could result in parasitic conversion. We then hypothesized that the action of TgHDAC3 could go beyond nucleosomes by deacetylating non-histone proteins. While the proteome-wide acetylome of *Toxoplasma* tachyzoite was recently published (Jeffers V *et al.*, 2012; Nardelli SC *et al.*, 2013; Bouchut A *et al.*, 2015), its regulation and alteration during stage conversion is nearly unknown. In collaboration with Yohann Coute and Pieter-Jan Debock (CEA, Grenoble) we produced a snapshot of *Toxoplasma* acetylome and analyzed its regulation in the context of TgHDAC3 activity modulation.

II- 2.2 Results

II- 2.2.1. FR235222-mediated inhibition of TgHDAC3 induces bradyzoite- and merozoite-specific proteins expression

Before starting the study of the acetylome we have mapped the proteome changes in response to FR235222 in order to define the appropriate conditions for the expression, if any, of stage-specific proteins. These results can be compared to those published by our team to see whether the expression deduced from changes in chromatin acetylation (Bougdour A *et al.*, 2009) correlates with protein levels. To investigate proteome alterations induced by TgHDAC3 inhibition, we used the type II strain, Pru $\Delta ku80$. This strain is typified by its ability to form cyst in murine model of toxoplasmosis and to convert well *in vitro* into bradyzoite stage under different stress conditions such as alkaline induction or chemical compound FR235222 (Bougdour A *et al.*, 2009). In this work, we first carried out mass spectrometry-based proteome profiling to identify changes in protein levels. FR235222-mediated changes were compared to untreated control parasites. The key steps in this strategy include protein extraction from comparing intracellular Pru $\Delta ku80$ with or without FR235222 treatment; trypsin digestion to yield peptides; LC/MS-MS analysis to determine protein ratios; database search and bioinformatic analysis (Figure 8A). As result, we identified a total of 5039 proteins. Among them, 2001 proteins were identified in *Toxoplasma* and 3039 proteins in human HFFs cells. When we analyzed statistically the *Toxoplasma* proteins from the complete dataset, we identified a total of 284 proteins which could be significantly upregulated or downregulated. Therefore, the volcano plot of the normalized *Toxoplasma* proteins indicates that more proteins are upregulated in the FR235222-treated samples. Surprisingly, the majority of these upregulated proteins correspond to the bradyzoite proteins for instance glycolytic enzymes enolase 1 (ENO1) and lactate dehydrogenase 2 (LDH2) both involved in parasite metabolism; surface antigens including SAG-related sequence (SRS) proteins and family A proteins. This finding shows clearly that the compound FR235222 induces the expression of bradyzoite-specific proteins in tachyzoites parasites (Figure 8B, right side). Surprisingly, we also identified 56 proteins whose gene expression was restricted to enteroepithelial stages in cats (see examples in Figure 9A). We suspected that FR235222 altered the genetic program dedicated to repress bradyzoite genes in tachyzoites stimulating bradyzoite conversion but also stimulate neo-converted bradyzoites to reach the next step, i.e. merozoites and the sexual stages. Since our data are representative of a population and not based on single-cell experiment, we cannot evaluate properly the conversion rate. That is said we estimate that the number of converted merozoites is quite low. One way to improve the conversion will be to trigger the differentiation process starting with brain cyst bradyzoites

instead of tachyzoites. Indeed bradyzoites are the main source of Type E schizonts that give rise to merozoites, which in turn differentiate into gametes (Figure 1).

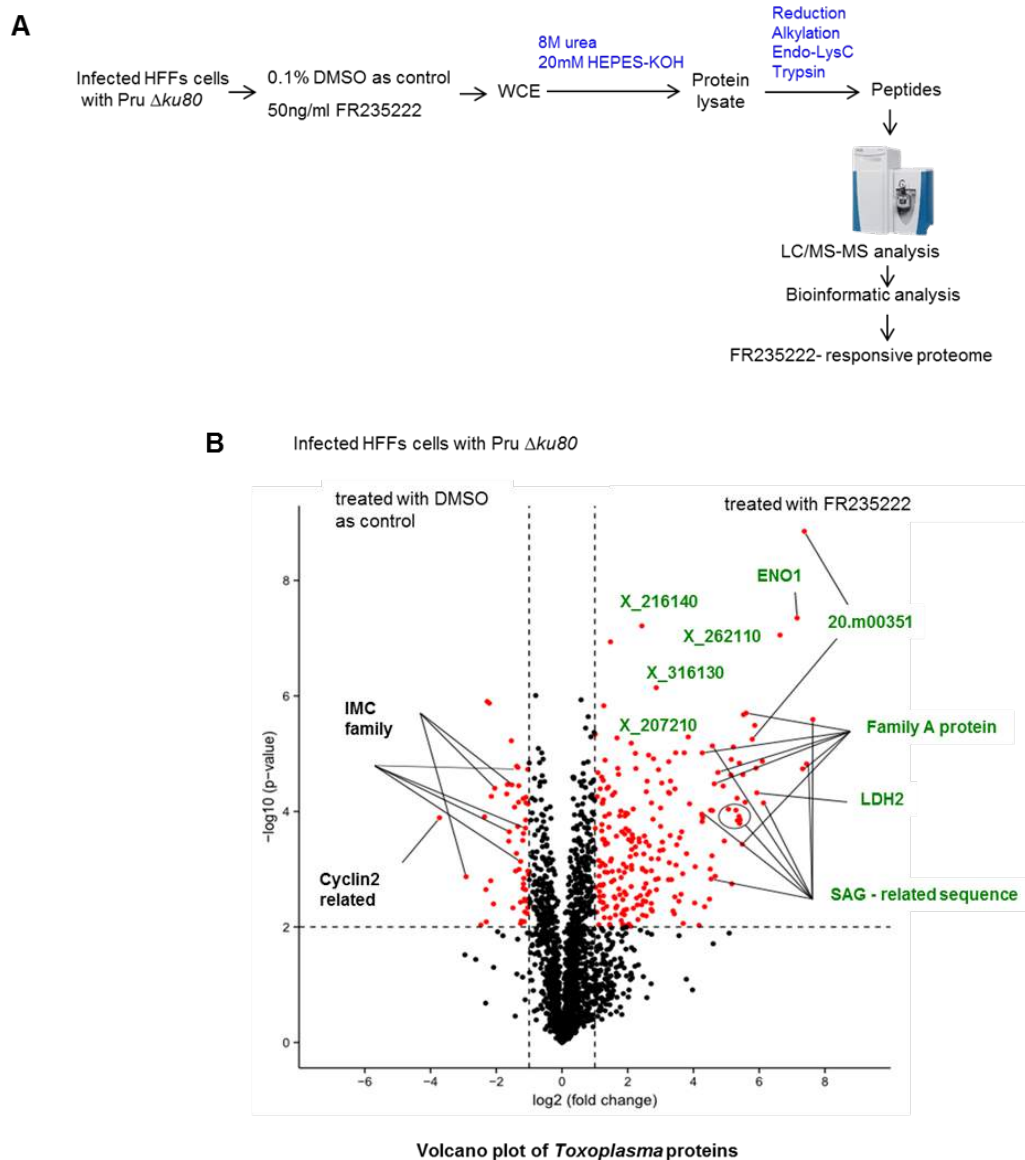


Figure 8. Quantitative proteome analysis after TgHDAC3 inhibition with FR235222 reveals the expression of bradyzoite-specific proteins in parasites. (A) Experimental design for proteome-wide characterization in Pru $\Delta ku80$ parasites by using LC-MS/MS. (B) Volcano plot of the over-expressed proteins in *Toxoplasma gondii*. The upregulated proteins after 0.1% DMSO treatment are shown in red spots on left side. On the right side, the red spots indicate the upregulated proteins after 90nM FR235222 treatment corresponding to bradyzoite-specific proteins. The log₂ ratios for proteins counts were obtained by dividing intensities of FR235222-treated samples by intensities of DMSO-treated samples (control). The negative log₂ ratio means the abundant proteins in control samples which correspond to the cell-cycle proteins of *Toxoplasma*. The two vertical black lines indicate log₂ fold changes values. The horizontal dashed black line shows at least a 2-fold abundance change with an adjusted p-value of maximum 0.01.

A third group contains proteins with levels decreasing in presence of FR235222 treatment, when compared to mock DMSO. While most of them have unknown function, we observed that they displayed a typical cell cycle-regulated gene expression pattern (source ToxoDB, Figure 8B, left side) when they are not themselves involved in cell cycle progression, e.g. the inner membrane complex proteins, the cyclin2 related protein and several mitotic proteins (Figure 8B, left side). Those results confirmed our initial observations (i) that parasites treated with FR235222 were vacuolated and lacked IMC1-delineated daughter cells, or contained aberrant progeny, and (ii) that the majority of the drug-treated parasites displayed massive DNA over- replication (>1N DNA content per cell), indicating that FR235222 interferes directly or indirectly with *T. gondii* cell-cycle progression (Figure 9B) (Bougdour *et al.*, 2009).

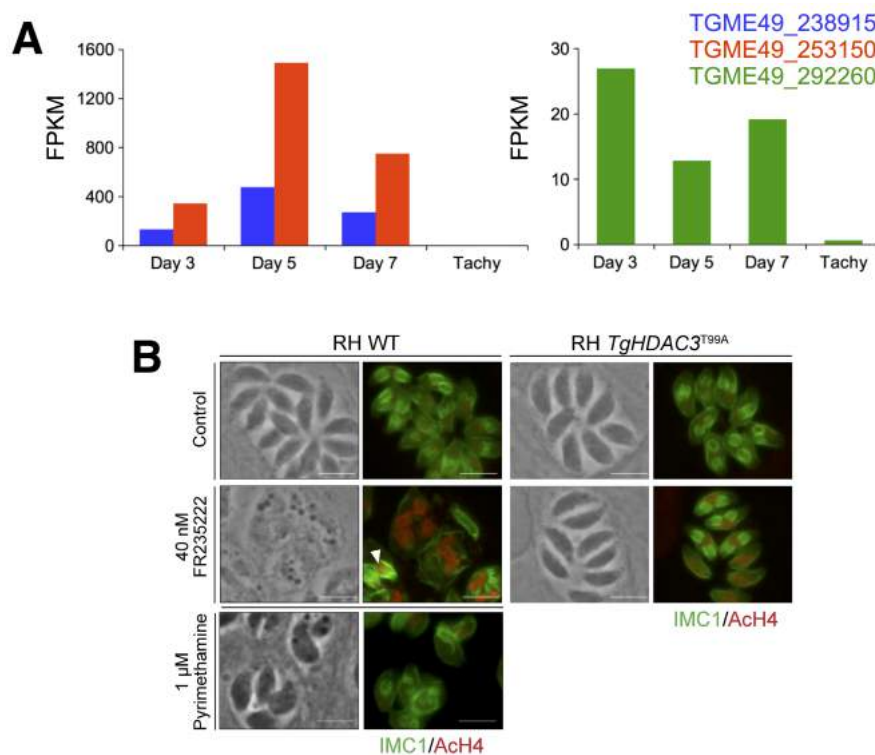


Figure 9. (A) RNAseq analysis (source : ToxoDB) from cultured tachyzoites, 3 days post-infection in cat (merozoites), 5 days post-infection in cat (merozoite and sexual stages) and 7 days post-infection in cat (sexual stages and oocysts). Three selected genes whose protein expression was increased upon FR235222 treatment. (B) Images from Bougdour *et al.* (2009) : Effects of FR235222 on histone H4 acetylation in intracellular *T. gondii* parasites. Confluent monolayers of HFF cells were infected with *T. gondii* RH WT and R20D9 (*TgHDAC3*^{T99A}) strains in the presence of 40 nM FR235222 and 0.1% DMSO as a control. As a control, parasites were treated with 1 μ M pyrimethamine. After 24 h of growth, cells were fixed and stained for ACh4 (red) and IMC1 (green). The arrowhead indicates aberrant progeny. Bars, 5 μ m.

Taken together, our data show that FR232222 mediates down-regulation of the level of regulators of cell cycle progression. Clearly and as evidenced by hundreds of proteins, FR2325222 was also able to up-regulated bradyzoite- and merozoite proteins whose function range from metabolism to surface antigen. While our proteomic data revealed that FR235222 was able to shift the tachyzoite proteome into a mixed stages (tachyzoite, bradyzoite and merozoite) proteome, we failed to discover FR235222-mediated alteration of master regulators protein levels, whose function could have been to regulate the process of differentiation. If ApiAP2 transcription factors have been extensively associated with stage differentiation or sexual commitment either in *Toxoplasma gondii* (Walker R *et al.*, 2013; Radke JB *et al.*, 2013) or in *Plasmodium* spp. (Kafsack *et al.*, 2014; Sinha *et al.*, 2014; Yuda *et al.*, 2010; Yuda *et al.*, 2015), none of them were found affected by the drug treatment, at least at the protein level. Since FR235222 was clearly shown to target TgHDAC3 (Bougdour A *et al.*, 2009), we confirmed the role of the deacetylase in signaling pathway leading to the induction of stage-specific proteins but we still have no clue of how the enzyme triggers its effect.

As mentioned in the introduction, we initially observed that switching the pH from 7 to 8 causes an unexpected hyperacetylation of histone H4 in the parasite nuclei, without any change of the acetylation status in the host cell nuclei (Figure 3B and 3D). This hyperacetylation is not mediated by any host cell signaling cascade as free parasites maintained in culture medium buffered at pH 8 also exhibit a drastic H4 acetylation (Figure 3C). On the other hand, alkaline pH was recognized for a long time as a *bona fide* inducer of bradyzoite-specific protein expression. Interferon-gamma (IFN γ) was also recognized as triggering stage conversion, at least *in vivo* the cytokine represses parasite proliferation and spreading which, in turn, forces the parasites to convert and encyst in deep tissue (e.g. brain) in order to be unseen par the immune system.

Therefore, we decided to repeat the proteome analysis using the same parasite strain but this time studying the effects of alkaline pH and IFN γ . The extracted proteins from Pru $\Delta ku80$ parasites, after treatment with IFN γ for 9 hours and with alkaline pH 8.1 for 3 days, were analyzed in mass spectrometry for proteome-wide by using the same strategy shown in figure 8A. Surprisingly, IFN γ or alkaline treatments were shown to induce quite a few bradyzoite-specific proteins when compared to FR235222 stimulation. Altogether, those data provide strong evidence that inhibition of TgHDAC3 activity has dramatic consequences on stage-specific proteins, much more so than what has been reported with high pH or IFN γ .

II- 2.2.2. Monitoring of FR235222-mediated bradyzoite gene expression using new reporter transgenic parasite cell line

In order to validate the proteome data, we generated in tachyzoites parasites Pru $\Delta ku80$ stable transgenic cell lines in which we endogenously C-terminal epitope tagged (HAFlag) candidate gene encoding protein which level increases upon FR235222 treatment. Using transcriptomic data from ToxoDB, we selected 7 genes coding for known or predicted bradyzoite proteins and one gene that is exclusively expressed in cat enteroepithelial merozoite and sexual stages. Following FR235222 induction, we were able to monitor by immunofluorescence both induction of the proteins and their subcellular localization (Figure 10A). We were able also to validate by immunoblotting analysis the drastic increase of the proteins upon FR235222 treatment (data not shown).

To determine whether those genes were selectivity regulated by TgHDAC3, we selected 7 HDACi that cover the entire selectivity range for class I and II deacetylases, albeit with varying specificity profiles against apicomplexan parasites, as determined previously (Bougdour A *et al.*, 2009) and included as a control the compound halofuginone a specific inhibitor of protein translation in Apicomplexa (Jain V *et al.*, 2015). We then showed that cyclopeptides HDACi strongly enhance ENO1 levels in parasite nuclei, whereas other inhibitors have no effect (Figure 10B). These results are coherent with those showing that distinct point mutations at a single locus in apicomplexan conserved region of TgHDAC3 abolishes the enzyme sensitivity to the cyclic tetrapeptide compounds (Bougdour A *et al.*, 2009).

Collectively, our data showed here that the bradyzoite and partly the merozoite genetic programs are under the control of TgHDAC3-dependent signaling pathway. While we have a strong correlation between the mRNA (source ToxoDB) and protein (our present data) levels for a majority of predicted bradyzoite-coding genes, we also identified a cluster of genes that are regulated at the level of translation in a TgHDAC3-dependent fashion. This results prompt us to consider that TgHDAC3 may also have alternative non-histone substrate some of which being involved in translational control.

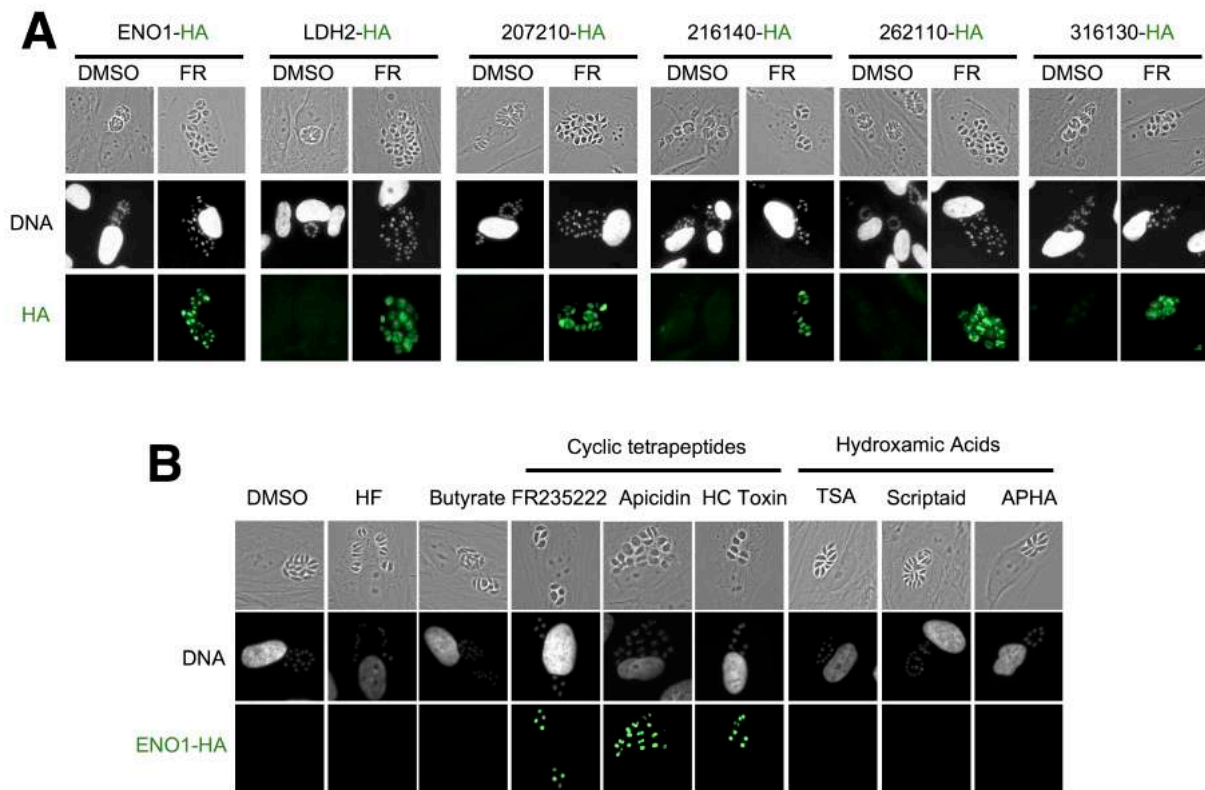


Figure 10 FR235222-mediated bradyzoite gene expression. (A) Bradyzoite genes whose protein product was detected in tachyzoites only upon FR235222 treatment were epitope (HAFLAG)-tagged in Pru Δ ku80 strain. Stable transgenic parasites were treated with a low concentration of FR235222 and expression of each gene of interest was monitored by following the HA fluorescence (in green). (B) The histone deacetylase inhibitors cyclotetrapeptides induce bradyzoite-specific genes. HFFs cells were infected with Pru Δ ku80 parasites expressing ENO1 HAflag for 18 hours. Parasites were then treated with different histone deacetylase inhibitors for 18h. The protein expression of ENO1 was analyzed in immunofluorescence assay with anti-HA antibody (in green).

II- 2.2.3. Proteome-wide mapping of the *Toxoplasma* acetylome, a gateway to discovering new TgHDAC3 substrates beyond histones

Our proteome-wide analysis uncovered TgHDAC3 as a key enzyme involved in repressing hundred of stage-specific proteins, though it did not allowed us to identify any master regulator (e.g. chromatin-modifying enzyme) or any transcription factor (e.g. ApiAP2) which could explain the dramatic reshuffling of a tachyzoite genome subjected to FR235222 stimulation. Since we also observed that not all the effects mediated by the drug were at the transcriptional level (data not shown), it has become increasingly evident that TgHDAC3, as for its human counterparts, may have a more diverse repertoire of substrates that extended beyond the histone tails modifications. At the time of our studies, several competing manuscripts have reported the existence of a complex acetylome in *T. gondii* (Jeffers V *et al.*,

2012; Xue B *et al.*, 2013). Their results have showed that lysine acetylation was abundant in the actively proliferating tachyzoite form of the parasite. They successfully identified known acetylation marks on *T. gondii* histones and α -tubulin while detecting over 411 novel acetylation sites on a wide variety of additional proteins, including those with roles in transcription, translation, metabolism, and stress responses (Jeffers V and Sullivan WJ Jr, 2012; Xue B *et al.*, 2013).

The aforementioned reports of *T. gondii* acetylome were determined in RH strain that lost the ability to develop into mature cysts (Jeffers V and Sullivan WJ Jr, 2012; Xue B *et al.*, 2013). We opted during my PhD for infections with a type II strain (Pru $\Delta ku80$) which is quite more relevant as it does readily develop tissue cysts and latent infections in animals. We then applied large-scale mass spectrometry-based acetylome after the specific inactivation of TgHDAC3 with the goal to uncover new enzyme substrates. By the combination of high affinity enrichment and high resolution LC-MS/MS analysis, we identify 810 acetylated peptides after TgHDAC3 inhibition. Contrary to the reported broad inhibitory effects of HDACi on deacetylases, FR235222 increased lysine acetylation of small but specific subproteome of *T. gondii* which is presumably regulated by TgHDAC3.

We then analyzed local sequence context around the acetylation sites in order to find an acetylation consensus motif. The amino acids flanking the targeted acetyl lysine in *T. gondii* proteome exhibit stretches of lysines and commonly conserved position across species, the aromatic residues phenylalanine (F) at -2 and tyrosine (Y) at +1 position (Figure 11A). Both F and Y at their respective positions are similar to those of acetylated proteins in prokaryotes (e.g. *Escherichia coli*) and in mitochondria of *T. gondii* (Zhang J *et al.*, 2009; Jeffers V *et al.*, 2012). Lysine-acetylated peptides from FR235222-treated samples show different preferences in their flanking sequences, namely the proclivity for glutamic acid (E) at the -1 position, but also a general lack of lysine and a bias towards serine (S) (Figure 11B).

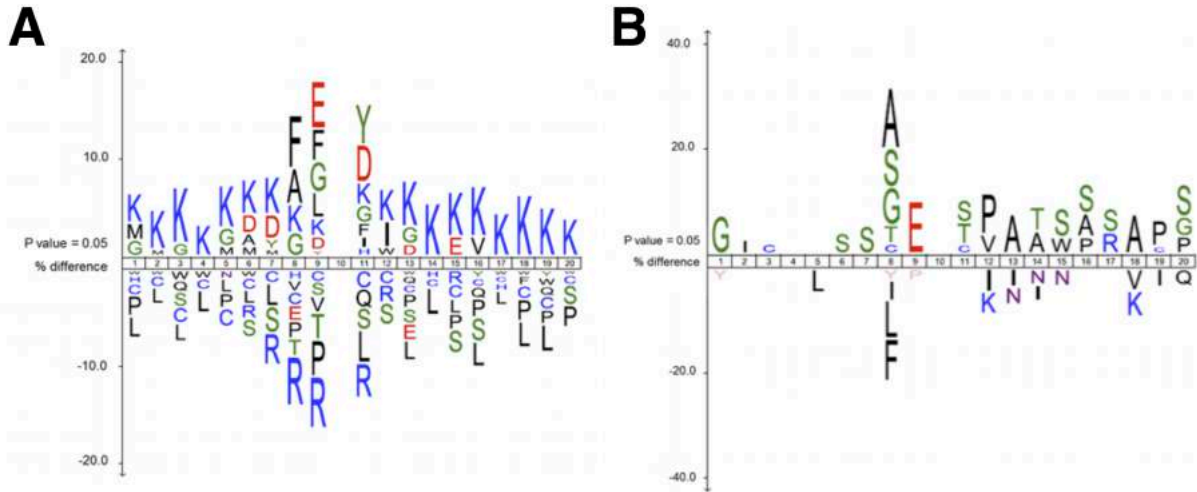


Figure 11 (A) IceLogo sequence motif of all acetylated peptides (n = 1376). The acetylated lysine is located at position 10 and surrounded by ± 9 amino acids. The amino acids that were more frequently identified in the proximity of the acetylation site are indicated over the middle line, while the amino acids with lower frequency are indicated below the line. The lysine motif is overrepresented, whereas the leucine, serine and arginine are underrepresented on this IceLogo. (B) IceLogo sequence motif of acetylated peptides after FR235222 treatment. The acetylated lysine is located at position 10 and surrounded by ± 9 amino acids. The amino acids that were more frequently identified in the proximity of the acetylation site are indicated over the middle line, while the amino acids with lower frequency are indicated below the line. The serines (S) at position 6, 7, 8, 11, 15, 16 are overrepresented in this IceLogo, whereas they are significantly underrepresented in the previous IceLogo (figure 11A). Proline (P) at position 12 and alanine (A) at position 13 could be overrepresented.

From the IceLogo analysis (Figure 11B), we inferred the following consensus sequence surrounding the lysine being acetylated upon FR235222 inhibition of TgHDAC3: SSxE**K**(ac)(S/T)PAXSS.

Our *T. gondii* acetylome study also revealed the acetylation of non-histone proteins including particularly AP2 transcription factors, translation factors, Gcn5b and ATP-dependent chromatin factors (e.g. SNF2h). When comparing our data to previously reported *T. gondii* acetylome (Jeffers V and Sullivan WJ Jr., 2012), we found 112 novel lysine being acetylated upon FR235222 treatment which are presumably putative TgHDAC3 substrates in tachyzoite stage. When looking at AP2 family (68 members), we identified 11 AP2 harboring lysine residue being acetylated following FR235222 stimulation (Table 2). Five of them were even exclusively acetylated upon drug treatment and were not identified in previous acetylomes (Jeffers V and Sullivan WJ Jr., 2012) (Table 2). Some AP2 domain proteins were cell-cycle

regulated (e.g. AP2X-5, AP2X-7, AP2VII-4 and AP2VIIa-4) and other constitutively expressed in tachyzoites (e.g. AP2VIIb-3, AP2VIIa-3, AP2VIIa-7, AP2VIIb-3, AP2X-8, AP2XI-2, AP2IX-7). We still need to validate those modifications and their relevance in the context of gene expression during stage conversion. To do so, we are now generating specific antibodies against several acetylated lysine embedded in AP2 factor. With those tools in hand, we will be able to confirm by IFA/western that TgHDAC3 play a role (or not) in the steady state of those PTMs. This can be done by monitoring their levels in parasites where TgHDAC3 activity is altered chemically (FR235222-mediated) or genetically (through cas9-gene editing).

We also identified the lysine K811 of the acetylase Gcn5b as a residue exclusively acetylated upon drug treatment as well as the previously reported acetylated lysine residues (K857, K989, K997 and K1027) (Jeffers V and Sullivan WJ Jr., 2012). Of note, Gcn5b was shown to bind strongly to AP2X-8, AP2XI-2, AP2IX-7 and AP2VII-4 (Wang J *et al.*, 2014) that are all acetylated under FR235222 stimulation. This suggests that TgHDAC3 might regulate the steady state of acetylation of the Gcn5b-associated AP2s and thereby may impede or favor the AP2-mediated recruitment of the acetylase to chromatin. However, it is not known if the acetylated lysines within these AP2 domain factors could influence the DNA recognition or their interactions to their cognate partners. In other organisms, lysine acetylation within the DNA-binding domain (e.g HMGI(Y) transcription factor) can affect the DNA binding as well the lysine acetylation adjacent to a DNA-binding domain was shown to stimulate DNA interactions (Kouzarides T., 2000). Overall, we uncover here a complex network in which TgHDAC3 may cross-talk with its opposite enzyme Gcn5b likely modulating its ability to bind DNA by altering the acetylation states of Gcn5b-associated AP2 transcription factors.

AP2 Transcription factors		RH tachyzoites	Pru $\Delta ku80$ tachyzoites under FR235222 treatment
		Already known lysine acetylation sites (ToxoDB)	New lysine acetylation sites
HDAC3 partners	AP2XI-2	-	1024
	AP2VIIa-3	-	291
HDAC3 substrates	AP2VIIa-4	-	1959
	AP2VIIb-3	-	619
	AP2VIII-6	-	1054
	AP2IX-7	236,1379	453, 855
	AP2X-8	1927	445
	AP2XII-4	1277, 1547, 1874, 2655, 2812	1473, 2833
	AP2X-7	152, 570	30, 1785, 152*
	AP2VIIa-7	2267, 2278, 2524	2278*
	AP2VIIb-1	272, 1168	272*

Table 2. Detection of new lysine acetylation sites in AP2 transcription factors in FR235222-treated Prugniaud tachyzoites compared with other known lysine acetylation sites in untreated RH tachyzoites from Jeffers V and Sullivan WJ Jr., 2012. Eukaryot Cell. An asterisk indicates that the reported acetylated lysine sites in RH parasites are also found in treated prugniaud parasites.

Translation factors and ribosomal proteins represent the second major group of proteins harboring lysine residues acetylated upon FR235222 treatment in line with previous report showing that these families represent 15% of the all parasite acetylome (Jeffers V and Sullivan WJ Jr., 2012). Finally and more importantly our data uncover novel histone PTMs and during my stay in the lab I had the challenge to characterize a versatile acetylation-methylation switch at K31 on the lateral surface of histone H4 in *Apicomplexa* (Sindikubwabo *et al.*, eLife, in revision).

Chapter 3.

Modifications at K31 on the lateral surface of histone H4 contribute to genome structure and expression in apicomplexan parasites.

Modifications at K31 on the lateral surface of histone H4 contribute to genome structure and expression in apicomplexan parasites

Tracking no: 07-06-2017-RA-eLife-29391

Fabien Sindikubwabo (INSERM), Shuai DING (Pasteur Institute), Tahir Hussain (INSERM), Philippe ORTET (CEA), Mohamed Barakat (CEA), Sebastian Baumgarten (Pasteur Institute), Dominique Cannella (CNRS), Andres Palencia (INSERM), Alexandre Bougdour (INSERM), Lucid Belmudes (CEA), Yohann Couté (Univ. Grenoble Alpes), Isabelle Tardieux (CNRS), Cyrille Botte (CNRS), Artur Scherf (Pasteur Institute), and Mohamed-ali Hakimi (INSERM)

Abstract:

A striking unusual genome architecture characterizes the two related human parasitic pathogens *Plasmodium falciparum* and *Toxoplasma gondii*. A major fraction of the bulk parasite genome is packaged as transcriptionally permissive euchromatin with few loci embedded in silenced heterochromatin. Primary chromatin shapers include histone modifications at the nucleosome lateral surface close to the DNA but their mode of action remains unclear. We identify versatile modifications at Lys31 within the globular domain of histone H4 as key determinants of genome organization and expression in *Apicomplexa* parasites. H4K31 acetylation promotes a relaxed chromatin state at the promoter of active genes through nucleosome disassembly in both parasites. In contrast, monomethylated H4K31 is enriched in the core body of *Toxoplasma* active genes but inversely correlates with transcription while being surprisingly enriched at transcriptionally inactive pericentromeric heterochromatin in *Plasmodium*, a region that is lacking H3K9me3 and heterochromatin protein 1 in this parasite.

Impact statement: This study shows that the chromosomal organization and genic expression in *Toxoplasma gondii* and *Plasmodium falciparum* is critically dictated by a versatile acetylation-methylation switch at lysine 31 on the lateral surface of histone H4

Competing interests: No competing interests declared

Author contributions:

Fabien Sindikubwabo: Conceptualization; Resources; Formal analysis; Validation; Investigation; Visualization; Methodology Shuai DING: Resources; Investigation; Visualization; Methodology Tahir Hussain: Resources; Investigation; Visualization Philippe ORTET: Data curation; Software; Formal analysis; Visualization Mohamed Barakat: Data curation; Software; Formal analysis; Visualization Sebastian Baumgarten: Resources; Data curation; Software; Formal analysis; Visualization Dominique Cannella: Resources; Investigation Andres Palencia: Investigation; Visualization Alexandre Bougdour: Resources; Investigation; Visualization Lucid Belmudes: Investigation Yohann Couté: Conceptualization; Investigation; Visualization Isabelle Tardieux: Conceptualization; Writing—original draft; Writing—review and editing Cyrille Botte: Conceptualization; Investigation Artur Scherf: Conceptualization; Supervision; Funding acquisition; Writing—review and editing Mohamed-ali Hakimi: Conceptualization; Formal analysis; Supervision; Funding acquisition; Investigation; Methodology; Writing—original draft; Project administration; Writing—review and editing

Funding:

European Commission (EC): Fabien Sindikubwabo, Dominique Cannella, Mohamed-ali Hakimi, ERC Consolidator Grant N{degree sign}614880; European Commission (EC): Artur Scherf, ERC AdG N{degree sign} 670301; Agence Nationale de la Recherche (ANR): Lucid Belmudes, Yohann Couté, ANR-10-INBS-08-01; Agence Nationale de la Recherche (ANR): Fabien Sindikubwabo, Shuai DING, Tahir Hussain, Dominique Cannella, Andres Palencia, Alexandre Bougdour, Cyrille Y Botte, Artur Scherf, Mohamed-ali Hakimi, LABEX PARAFRAP ANR-11-LABX-0024; Agence Nationale de la Recherche (ANR): Dominique Cannella, Andres Palencia, Alexandre Bougdour, Mohamed-ali Hakimi, ANR-12-BSV3-0009-01 The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Datasets:

Datasets Generated: Chromosomal organization and genic expression in Apicomplexan parasites is critically dictated by a versatile acetylation-methylation switch at K31 on the lateral surface of histone H4: Sindikubwabo F, Ding S, Hussain T, Ortet P, Barakat M, Baumgarten S, Cannella D, Palencia A, Bougdour A, Belmudes L, Couté Y, Tardieux I, Botte CY, Scherf A, Hakimi MA, 2017, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98806>, GSE98806; *Plasmodium falciparum* histone post-translational modifications: Sebastian Baumgarten, 2017, <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA386433>, PRJNA386433 Reporting Standards: N/A Previously Published Datasets: Chromosomal organization and genic expression in Apicomplexan parasites is critically dictated by a versatile acetylation-methylation switch at K31 on the lateral surface of histone H4: Sindikubwabo F, Ding S, Hussain T, Ortet P, Barakat M, Baumgarten S, Cannella D, Palencia A, Bougdour A, Belmudes L, Couté Y, Tardieux I, Botte CY, Scherf A, Hakimi MA, 2017, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98806>, GSE98806; *Plasmodium falciparum* histone post-translational

modifications: Sebastian Baumgarten, 2017, <https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA386433>, PRJNA386433

Ethics:

Human Subjects: No Animal Subjects: No

Author Affiliation:

Fabien Sindikubwabo(Institute for Advanced Biosciences - U1209,INSERM,France) Shuai DING(Unité de Biologie des Interactions Hôte-Parasite,Pasteur Institute,France) Tahir Hussain(Institute for Advanced Biosciences - U1209,INSERM,France) Philippe ORTET(UMR 7265,CEA,France) Mohamed Barakat(UMR 7265,CEA,France) Sebastian Baumgarten(Unité de Biologie des Interactions Hôte-Parasite,Pasteur Institute,France) Dominique Cannella(Institute for Advanced Biosciences - U1209,CNRS,France) Andres Palencia(Institute for Advanced Biosciences - U1209,INSERM,France) Alexandre Bougdour(Institute for Advanced Biosciences - U1209,INSERM,France) Lucid Belmudes(BIG-BGE,CEA,France) Yohann Cauté(IRT SV-BGE,Univ. Grenoble Alpes,France) Isabelle Tardieux(Institute for Advanced Biosciences - U1209,CNRS,France) Cyrille Botte(Institute for Advanced Biosciences - U1209,CNRS,France) Artur Scherf(,Pasteur Institute,) Mohamed-ali Hakim(Institute for Advanced Biosciences - U1209,INSERM,France)

Dual-use research: No

Permissions: Have you reproduced or modified any part of an article that has been previously published or submitted to another journal?
No

1 Modifications at K31 on the lateral surface of histone H4
2 contribute to genome structure and expression
3 in apicomplexan parasites

4
5 Fabien Sindikubwabo¹, Shuai Ding², Tahir Hussain¹, Philippe Ortet³, Mohamed
6 Barakat³, Sebastian Baumgarten², Dominique Cannella¹, Andres Palencia¹, Alexandre
7 Boudgour¹, Lucid Belmudes⁴, Yohann Coute⁴, Isabelle Tardieux⁵, Cyrille Y. Botté⁶,
8 Artur Scherf² and Mohamed-Ali Hakimi¹ *

9 ¹ Institute for Advanced Biosciences (IAB), Team Host-pathogen interactions &
10 immunity to infection, INSERM U1209, CNRS UMR5309, Université Grenoble
11 Alpes, Grenoble, F-38700.

12 ² Unité de Biologie des Interactions Hôte-Parasite, Institut Pasteur, CNRS, ERL 9195,
13 INSERM, Unit U1201, Paris, F-75724.

14 ³ Aix-Marseille Univ, CEA, CNRS, UMR 7265, BIAM-LEMIRE, F-13108 St-Paul-
15 lez-Durance, France

16 ⁴ Univ. Grenoble Alpes, CEA, INSERM, BIG-BGE, F-38000 Grenoble, France

17 ⁵ Institute for Advanced Biosciences (IAB), Team Membrane and Cell Dynamics of
18 Host Parasite Interactions, INSERM U1209, CNRS UMR5309, Université Grenoble
19 Alpes, Grenoble, F-38700.

20 ⁶ Institute for Advanced Biosciences (IAB), Team ApicoLipid, INSERM U1209,
21 CNRS UMR5309, Université Grenoble Alpes, Grenoble, F-38700.

22
23 *Correspondence to: Mohamed-ali Hakimi, Phone +33 4 76 63 74 69, e.mail:

24 mohamed-ali.hakimi@inserm.fr

25 ORCID : [0000-0002-2547-8233](https://orcid.org/0000-0002-2547-8233)

26
27 **Running title:** Chromatin regulation by H4K31 modifications

28 **Key words:** *Toxoplasma gondii*, *Plasmodium falciparum*, histone core modifications,
29 acetylation, methylation, chromatin, gene expression

30

31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57

Abstract

A striking unusual genome architecture characterizes the two related human parasitic pathogens *Plasmodium falciparum* and *Toxoplasma gondii*. A major fraction of the bulk parasite genome is packaged as transcriptionally permissive euchromatin with few loci embedded in silenced heterochromatin. Primary chromatin shapers include histone modifications at the nucleosome lateral surface close to the DNA but their mode of action remains unclear. We identify versatile modifications at Lys31 within the globular domain of histone H4 as key determinants of genome organization and expression in *Apicomplexa* parasites. H4K31 acetylation promotes a relaxed chromatin state at the promoter of active genes through nucleosome disassembly in both parasites. In contrast, monomethylated H4K31 is enriched in the core body of *Toxoplasma* active genes but inversely correlates with transcription while being surprisingly enriched at transcriptionally inactive pericentromeric heterochromatin in *Plasmodium*, a region that is lacking H3K9me3 and heterochromatin protein 1 in this parasite.

58 **Introduction**

59 The phylum of *Apicomplexa* clusters thousands of single-celled eukaryotes identified
60 as parasites of metazoans including humans in who they cause or put at risk for major
61 public health problems. Preeminent human pathogens include *Plasmodium spp.* which
62 are responsible for dreadful malaria as well as *Toxoplasma gondii* and
63 *Cryptosporidium spp.* which are leading causes of food-borne and water-borne
64 diseases. A shared characteristic of apicomplexan life cycles is the multiplicity of
65 developmental stages that progress from one to the other along with precise genetic
66 reprogramming to ensure survival and transmission of parasite populations. The
67 emerging concept of a remarkably dynamic nature of gene expression in *Apicomplexa*
68 has risen from the observation that large numbers of mRNAs are exclusively
69 expressed in a given developmental stage (Bozdech et al., 2003; Radke et al., 2005).

70

71 Unlike the majority of higher Eukaryotes, *Apicomplexa* genomes have a unique
72 chromatin architecture typified by an unusually high proportion of euchromatin and
73 only a few heterochromatic islands scattered through the chromosome bodies or
74 embedded at telomeres and centromeres. Although alterations in chromatin structure
75 are acknowledged as important for the transcriptional control of commitment to stage
76 differentiation in several *Apicomplexa* as well as for antigenic variation-mediated
77 immune evasion in *P. falciparum*, yet the molecular mechanisms of chromatin
78 remodeling have not been fully determined (Bougdour et al., 2010; Scherf et al.,
79 2008).

80

81 In Eukaryotes, the timely opening and closing of chromatin required for gene
82 expression, chromosomal organization, DNA repair or replication is governed by
83 histone turnover and their post-translational modifications (PTMs), such as lysine
84 methylation (me) and acetylation (ac) among many others. PTMs on histone tails were
85 indeed shown, alone or in combination, to alter the accessibility of effector proteins to
86 nucleosomal DNA and thereby impact chromatin structure, according to the “histone
87 code” hypothesis (Strahl and Allis, 2000; Turner, 2000). In addition PTMs also act as
88 signals to recruit ATP-dependent remodeling enzymes to either move, eject or
89 reposition nucleosomes. Accounting for the PTMs versatile activity onto chromatin
90 are enzymes carrying antagonist activities: it is the opposite, yet well concerted

91 activities of histone acetyltransferases (HATs) and histone deacetylases (HDACs) that
92 determine acetylation levels in cells.

93

94 *Apicomplexa* have evolved unique ways to modify histones that rival the strategies
95 adopted by higher Eukaryotes and provide them with remarkable capacities to
96 differentiate and multiply (Bougdour et al., 2010). Some histone-modifying enzymes
97 have acquired gain-of-function mutations that confer broader or enhanced activity
98 on substrates. It is the case of parasite Set8-related proteins endowed with H4K20
99 mono-, di-, and trimethylase activities that contrast with the mono-methylase-
100 restricted metazoan Set8, and that derive from a single-amino-acid change in the
101 substrate-specific channel (Sautel et al., 2007). Another example is provided by the
102 *Apicomplexa* HDAC3 family which is typified by an AT motif insertion at the
103 entrance of the active-site tunnel of the conserved catalytic domain causing additional
104 substrate/inhibitor recognition and binding properties (Bougdour et al., 2009).

105

106 In Eukaryotes, while PTMs have been primarily detected in the histone tails
107 sticking out from the nucleosome, an ever-growing list of PTMs is now identified in
108 the lateral surface of the histone octamer, that directly contacts DNA, and
109 characterized as critical regulators of the chromatin structure and function (Lawrence
110 et al., 2016; Tropberger and Schneider, 2013). Those “core” histone PTMs promote
111 different outcomes on nucleosome dynamics depending on their precise location.
112 Modifications near the DNA entry-exit region (e.g., H3K36ac) of the nucleosome
113 were shown to favor local unwrapping of DNA from histone octamer thereby
114 providing a better exposure of nucleosomal DNA to chromatin-remodeling and DNA-
115 binding proteins (Neumann et al., 2009). On the other hand, lateral-surface PTMs
116 mapping close to the dyad axis (e.g., H3K122ac) were shown to decrease the affinity
117 of the octamer to DNA and significantly affect nucleosome stability/mobility
118 (Tropberger et al., 2013). Similarly to what has been described for histone tails,
119 different lateral-surface modifications on the same residue can be associated with
120 opposite transcriptional programs. This is the case for the H3K64 residue near the
121 dyad axis that facilitates nucleosome eviction and thereby gene expression when
122 acetylated (Di Cerbo et al., 2014), whereas trimethylation of the same residue acts as a
123 repressive heterochromatic mark (Daujat et al., 2009).

124

125 In both *T. gondii* and *P. falciparum* unbiased mass spectrometry has led to uncover the
126 repertoire of the most prevalent histone PTMs including singular marks. However,
127 only few of them were mapped at the outer surface of the octamer (Saraf et al., 2016;
128 Trelle et al., 2009). In this study, we investigated in depth how histone H4 PTMs
129 could influence chromosome organization and gene regulation in apicomplexan
130 parasites. We reported versatile modifications at lysine 31 of histone H4, which lies at
131 the protein-DNA interface close to the dyad axis of the nucleosome. Genome-wide
132 mapping revealed that H4K31 could either be acetylated or methylated and the mark
133 enrichment occurred in a mutually exclusive manner. In *T. gondii* H4K31 residue
134 tended to be acetylated at the promoter of a nearby active gene and to be mono-
135 methylated in the core body of the gene. H4K31me1 occupancy was inversely
136 correlated with gene expression suggesting that the mark acts as a repressive mark
137 impeding RNA polymerase progression. In *P. falciparum*, H4K31ac was also seen at
138 the promoter whereas H4K31me1 occupancy was highly enriched at peri-centromeric
139 heterochromatin possibly compensating the absence of H3K9me3 and HP1 in this
140 atypical chromosome structure in order to maintain a constitutive heterochromatin
141 environment.

142 **Results**

143 **H4K31 maps at the Dyad Axis of the Nucleosome**

144 While studying the protein content of nucleosomes from *T. gondii* infected cells, we
145 and others mapped an acetylation site on histone H4 lysine 31 (H4K31ac) that was
146 largely underestimated thus far (Fig. 1a) (Jeffers and Sullivan, 2012; Xue et al., 2013).
147 H4K31 residue is located at the N-terminus of the H4 α 1 helix and its side chain is
148 extended in the major groove of the DNA (Fig. 1b and c). The closed state of
149 chromatin is contributed by interaction of K31 and R35 residues to DNA by a water
150 mediated hydrogen bond. Addition of acetyl group to -NH₂ group of lysine side chain
151 (K31) leads to replacement of this water molecule, thus abolishing its interaction with
152 DNA (Fig. 1c, panel 2). Acetylation may therefore destabilize the protein-DNA
153 interface close to the dyad axis of the nucleosome where the residue lies and thus
154 presumably open the chromatin. Although the residue H4K31 is well conserved across
155 species (Fig. 1b), mass spectrometry initially indicates its acetylation to be restricted

156 to metazoans, as an unexpected mark in inflammatory and auto-immune contexts
157 (Garcia et al., 2005; Soldi et al., 2014) since it was found neither in yeast nor in the
158 ciliated protozoan *Tetrahymena* (Garcia et al., 2006). Our present data contradict this
159 view as they show this PTM to also arise in the phylum of *Apicomplexa*.

160 **H4K31ac marks euchromatin in mammalian cells and apicomplexan parasites**

161 Although H4K31ac was unequivocally identified by mass spectrometry in histone
162 extracts from *T. gondii*-infected cells (Fig. 1a), its abundance/frequency was quite
163 low, a situation potentially reflecting high dynamics of the modification. To further
164 probe *in situ* the kinetics of this histone mark in apicomplexans, we raised an antibody
165 against a synthetic peptide acetylated at the H4K31 position. Antibody specificity was
166 assessed by western blot and dot blot analyses (data not shown). Using human
167 primary fibroblasts infected by *T. gondii*, we found H4K31ac to be uniformly
168 distributed within the nuclei of both parasite and human cell and excluded from their
169 cytoplasm (Fig. 1d). We also found that exposing cells to histone deacetylase inhibitor
170 (HDACi), e.g., FR235222 significantly increased H4K31ac signal intensity otherwise
171 moderate in parasite nucleosomes (Fig. 1e), in strong support for specific detection of
172 acetylation. Similarly, response to HDACi treatment was observed by immunoblot
173 analysis on cell population samples (Fig. 1f). Interestingly, H4K31ac localization
174 pattern sharply contrasts with those of tails PTMs (e.g., H4K8ac, H3K14ac, H3K18ac
175 and H3K27ac), which are typified by a high and already maximal signal in parasite
176 nuclei, that do not increase upon FR235222 treatment under our conditions (Fig. 1e).

177

178 To gain insights on the behavior of H4K31ac during its intraerythrocytic
179 developmental cycle (IDC) of *P. falciparum*, immunofluorescence assays were
180 conducted over 48 hours of culture to probe the ring, trophozoite and schizont stages
181 (Fig. 2a). H4K31ac was typified by low nuclear signal throughout the IDC that
182 increased upon HDACi treatment (Fig. 2b). Overall, H4K31ac showed a nuclear
183 punctate pattern, reminiscent of active loci cluster in specialized ‘transcription
184 factories’ (Fig. 2b) (Mancio-Silva and Scherf, 2013).

185

186 Strikingly, H4K31ac has remained understudied in higher eukaryotes thus far. To gain
187 better resolution of any nuclear or chromatin structures with which H4K31ac might be

188 associating, we co-stained murine embryonic fibroblasts (MEFs) for DNA and various
189 chromatin marks. H4K31ac was observed scattered through the nucleoplasm of MEFs
190 but excluded from nucleoli and segregated away from heterochromatic foci similarly
191 to the transcription-associated PTMs H3K4ac, H3K9ac and H3K27ac (Fig. 2c). This
192 pattern is typically euchromatic and opposed to the one revealed by the repressive
193 marks H3K9me3 and H4K20me3, rather associated with regions of highly condensed
194 pericentromeric heterochromatin (Fig. 2d). Taken together, these experiments show
195 that H4K31ac displays an euchromatic pattern in both metazoans and apicomplexans.
196

197 **GCN5b and HDAC3 enzymes fine-tune the H4K31ac levels in *Toxoplasma gondii***

198 To our knowledge, the enzymes that acetylate and deacetylate H4K31 remain
199 unknown. The relatively low level of H4K31ac found in both human and parasite
200 nuclei compared to other acetylated residues in histone tails (Fig. 1e) prompted us to
201 analyze potential deacetylases targeting H4K31ac. To this end, we selected 7 HDACi
202 that cover the entire selectivity range for class I and II deacetylases, albeit with
203 varying specificity profiles against apicomplexan parasites, as determined previously
204 (Bougdour et al., 2009) and included as a control the compound halofuginone a
205 specific inhibitor of protein translation in *Apicomplexa* (Jain et al., 2015). Here, we
206 show that cyclopeptides HDACi strongly enhance H4K31ac levels in parasite nuclei,
207 whereas other inhibitors have no effect (Fig. 3a). These results are coherent with those
208 showing that distinct point mutations at a single locus in apicomplexan conserved
209 region of TgHDAC3 abolishes the enzyme sensitivity to the cyclic tetrapeptide
210 compounds (Bougdour et al., 2009). *T. gondii* is particularly suited for a single gene
211 perturbation strategy, since its genome does not contain extensive HAT and HDAC
212 paralogs unlike mammalian genomes. To investigate further which of the five class-
213 I/II HDAC homologues in *T. gondii* may be responsible for the deacetylation of
214 this residue, we systematically performed CRISPR-mediated gene disruption.
215 Notably, the inactivation of the *TgHDAC3* gene, unlike other *TgHDACs*, causes
216 hyperacetylation of H4K31 in parasite nuclei (Fig. 3b and c), thereby mimicking the
217 effect of the cyclic tetrapeptide HDACi on the enzyme (Fig. 3a) (Bougdour et al.,
218 2009).

219

220 Reciprocally, we next sought to identify the one or more responsible HATs targeting
221 this residue. We used a candidate-based approach by systematically depleting the
222 parasite for key members of the three main HAT classes. Apicomplexans possess
223 homologues of the Type A GCN5 and MYST family nuclear HATs as well as the Type
224 B cytoplasmic HAT1 (Vanagas et al., 2012) while missing the mammalian restricted
225 PCAF (p300/CBP-associating factor) family. Intriguingly, *T. gondii* is unique among
226 fellow apicomplexan parasites and other invertebrates in possessing two GCN5 HATs,
227 designated TgGCN5a and b, that exhibit different histone acetylation activities
228 (Vanagas et al., 2012). We first thought that *T. gondii* HAT1 made promising
229 candidate enzyme as its human counterpart HAT4 was shown to acetylate core PTMs
230 *in vivo* and even H4K31 although under *in vitro* conditions (Yang et al., 2011).
231 However, cas9-mediated gene disruption of HAT1 had no effect whatsoever on
232 H4K31 acetylation (Fig. 4a) whereas *TgGCN5b* but not other HAT gene disruption
233 resulted in a drastic drop of H4K31ac signals in the parasite nuclei (Fig. 4a).
234 *TgGCN5b* is the prototypical GCN5 HAT in *T. gondii* because it is capable of
235 targeting H3K9, K14 and K18 (Bhatti et al., 2006). Furthermore, we noticed that the
236 amino acid sequence surrounding H4K31 was not homologous to preferred Gcn5
237 consensus sites of acetylation found at H3K14 (Kuo et al., 1996) or H3K36 (Morris et
238 al., 2006), yet the depletion of *TgGCN5b* led to a reduction of both H3K14ac (Fig. 4b)
239 and H4K31ac (Fig. 4a) signals *in vivo*, suggesting that the repertoire of the lysine
240 residues being acetylated by the GCN5 family is more diverse in *T. gondii*. Altogether
241 these data clearly uncover *TgGCN5b* as H4K31 acetyltransferase whose activity is
242 counteracted by *TgHDAC3*.

243

244 **H4K31me1 associates *in vivo* with distinct chromatin patterns**

245 It is well appreciated that the targeting of lysine residues by acetylation and
246 methylation cannot occur simultaneously. As such, previous studies have
247 characterized H4K31 as a site of monomethylation in plant, budding yeast and
248 metazoan cells but not in *Tetrahymena* (Garcia et al., 2006; Moraes et al., 2015). The
249 possibility of “dual” modifications occurring on H4K31 has not been yet explored in
250 apicomplexans nor in any other species besides proteomic identifications. Therefore,
251 we raised an antibody against a synthetic peptide containing monomethylated H4K31
252 and assessed its selectivity and specificity (data not shown). H4K31me1 was detected

253 by Western analysis in *T. gondii*, but not in the recombinant unmodified H4 form (Fig.
254 1f and data not shown).

255

256 *In situ*, the H4K31me1 modification appeared uniformly distributed within the nuclei
257 of dividing parasites but surprisingly no signal was detected in the nucleus of the
258 infected human cell (Fig. 5a), despite that this PTM had been previously detected by
259 mass spectrometry in human samples (Garcia et al., 2006). However, while
260 H4K31me1 was not (or barely) detected in interphase nuclei of either quiescent
261 infected HFFs (Fig. 5a) or uninfected MEFs (Fig. 5b), it decorated mitotic
262 chromosomes providing more pronounced signals in the chromosome arms than the
263 usual mitotic marker H3S10 phosphorylation (Fig. 5b). These observations argue for a
264 possible role of H4K31 methylation during cell division in mammalian cells.

265

266 In *P. falciparum*, H4K31me1 displayed a peculiar condensed punctate pattern (Fig.
267 1c), similar to the H3K9me3 mark (Lopez-Rubio et al., 2009) at the nuclear periphery,
268 which is reminiscent of heterochromatin/sub-telomeric regions clustering (Freitas-
269 Junior et al., 2000). *P. falciparum* centromeres also clustered prior to and throughout
270 mitosis and cytokinesis leading to single nuclear location from early trophozoites to
271 mature schizonts (Hoeijmakers et al., 2012). Therefore, H4K31me1-containing foci
272 could be associated with sub-telomeric or/and centromeric regions. The number of
273 foci observed per parasite however varies signifying its dynamic nature through parasite
274 developmental stages. The mark is observed in all asexual forms and remains
275 unaffected by treatment of FR235222 (data not shown).

276

277 Because H4K31 is also a site of methylation, the transition between H4K31ac and
278 H4K31me1 may represent a novel “chromatin switch” contributing to chromatin
279 structure and function in eukaryotic cells. Yet, a different readout is expected from
280 one species to another. In metazoan, H4K31me1 was temporally regulated during the
281 cell cycle and interplay, if any, with H4K31ac should be restricted to mitotic
282 chromosomes. In *P. falciparum*, H4K31me1 forms discrete immuno-fluorescent foci
283 around the nucleus, a pattern quite distinct from that of H4K31ac typified by a
284 diffused signal distributed throughout the parasite nuclei. Since they have distinct
285 nuclear locations and different stoichiometry, H4K31ac being a low abundant species,
286 the transition between H4K31me1 and H4K31ac may be not an issue in *P. falciparum*

287 as it can be in *T. gondii* where both modifications are concomitantly distributed
288 throughout chromatin.

289

290 **Nucleosomes with H4K31ac and H4K31me1 are enriched at the promoter and** 291 **the core body of active genes, respectively**

292 To further explore whether H4K31ac and H4K31me1 are indeed alternative
293 antagonistic PTMs on the same H4 molecules in *T. gondii*, we examined their
294 genome-wide distributions using chromatin immunoprecipitation coupled with next-
295 generation sequencing (ChIP-seq). We first investigated the relative performance of
296 our home-made antibodies in terms of specificity, sensitivity and the number and
297 distribution of peaks. We observed low variability and a high degree of similarity in
298 read coverage between technical replicates, regardless of antibodies used for
299 immunoprecipitation (Figure 6 – figure supplement 1). We next sought to compare the
300 locations of the peaks from each antibody type. Visual display of the chromosomal
301 distribution indicated that H4K31ac and H4K31me1 exhibit distinct patterns of
302 enrichment across the chromosomes and are mutually exclusive genome-wide (e.g.,
303 Chr. Ib, Fig. 6a). Zooming into detailed gene level revealed that H4K31ac was
304 enriched in distinct peaks at intergenic regions (IGRs) (Fig. 6b), of which 75%
305 percent mapped outside the gene body (Fig. 6c), in line with the euchromatic *in situ*
306 localization (Fig. 1 and Fig. 2). The calculated average profile of H4K31ac showed a
307 pattern strikingly similar to H3K14ac and H3K4me3, characterized by high signals at
308 5'UTR/promoter that drop sharply after the translational initiation site (Fig. 6d and
309 Figure 7 – figure supplement 1). Conversely, H4K31me1 showed a distinct pattern of
310 enrichment, best discernable at large genes, spanning from the ATG to the entire gene
311 body (Fig. 6b) while being absent at IGRs (Fig. 6c and d). Its computed average
312 density profile fully matched with gene prediction making this mark useful to
313 explicitly detect unannotated genes (Figure 8 – figure supplement 1). Remarkably,
314 these data allowed identifying H4K31me1 as the first PTM to our knowledge whose
315 spreading is restricted by the translational initiation and stop sites at active genes (Fig.
316 6b and d).

317

318 **Interplay between H4K31ac and H4K31me1 predicts distinctive patterns of gene** 319 **expression in *Toxoplasma gondii***

320 A closer view of H4K31ac and H4K31me1 chromosomal binding revealed that, at
321 some loci (e.g., *SRS* gene family, Fig. 7a and *GRA1*, Fig. 7b), the latter was absent
322 while the former was enriched at 5'UTR/promoter and unexpectedly spread over a
323 much larger area, overlapping the gene body (see *GRA1* and *MAG1* examples, Fig.
324 7c). The restricted H4K31ac enrichment at the vicinity of *GRA1* or *SRS* genes
325 contrasts with the H4K31ac and H4K31me1 location at their neighboring genes
326 (*TGME49_233490* or *TGME49_270230*) and this discrepancy may be explained by
327 the higher level of *SRS* or *GRA1* gene expression (Fig. 7a and b). We therefore
328 interrogated whether enrichment patterns of modified H4K31 could specify levels of
329 gene expression in *T. gondii*. We first conducted a global transcriptome analysis by
330 RNA-Seq of tachyzoites during growth phase in murine bone marrow-derived
331 macrophages (BMDMs). Cluster analysis revealed varying levels of gene expression
332 with cluster Q1 displaying the highest level, clusters Q2 and Q3 defining intermediate
333 levels and cluster Q4 displaying the lowest (Fig. 7d). High mRNA level (Q1 that
334 includes *GRA1*, *MAG1* and *SRS* genes) was associated with high level of H4K31ac
335 upstream of the ATG together with an enrichment along the gene body which
336 coincided with the expected lack of H4K31me1 (Fig. 7e and f). In highly expressed
337 relatively long or intron-containing genes (e.g., *MAG1*), H4K31ac spread but did not
338 extend over the entire gene body as observed for *GRA1* (Fig. 7c), indicating a limited
339 spreading of H4K31 acetylation around the translational initiation site. Strikingly,
340 moderate mRNA levels (Q2 and Q3 that include *TGME49_233490* or
341 *TGME49_270230*) related to a relatively high level of H4K31me1 in the gene body
342 and a restricted mapping of H4K31ac at the promoter, thereby arguing for an inverse
343 correlation between the yield of expression and the level of H4 methylation (Fig. 7e
344 and f).

345

346 Finally, transcriptionally repressed genes clustered in Q4 such as *TGME49_207730*
347 and *TGME49_270240* showed no significant enrichment of either H4K31ac or
348 H4K31me1 but were highly enriched in the repressive mark H3K9me3 and
349 unexpectedly in H3K14ac, a hallmark of gene activation (Fig. 8a). The co-enrichment
350 of H3K14ac and H3K9me3 at repressed genes might be due to population averaging
351 therefore reflecting heterogeneity within the parasite population. An alternative
352 explanation is that H3K14ac and H3K9me3 form the so-called bivalent chromatin

353 domain capable of silencing developmental genes while keeping them poised for rapid
354 activation upon cell differentiation (Voigt et al., 2013). As such, H3K14ac along with
355 the repressive mark H3K27me3 were shown to be enriched at a subset of inactive
356 promoters in mouse embryonic stem cells (Karmodiya et al., 2012). Bivalent domains
357 have gathered wide attention, because they might contribute to the precise unfolding
358 of gene expression programs during cell differentiation. Likely both
359 *TGME49_207730* and *TGME49_270240* remain in a state poised for transcription in
360 tachyzoite stage and H3K14ac/H3K9me3 bookmark these genes for reactivation in
361 the feline enterocytic stages where their transcripts were detected. Apparently, those
362 bivalent marks are restricted to inactive stage-specific promoters since both
363 pericentromeric (Fig. 8b and Figure 8 – figure supplement 2) and telomeric (Fig. 8c)
364 heterochromatin, while being decorated by H3K9me3 displayed no H3K14ac
365 enrichment.

366

367 Remarkably, these findings highlight unique chromatin signatures associated with the
368 transcription rate in *T. gondii*. Genes clustered in Q1 are primarily defined by
369 H4K31me1 low enrichment and enhanced acetylation at both promoter and 5'
370 proximal gene body, while those from Q2 and Q3 are markedly typified by the
371 presence of H4K31 methylation in the gene body and an acetylation mark restricted to
372 the promoter. In this context, H4K31ac would be predicted to disrupt histone-DNA
373 interaction thereby affecting nucleosome stability while promoting RNA polymerase
374 progression across transcribed units. Conversely, H4K31me1 is likely acting as a
375 transcription-linked repressive mark that substantially slows the progress of the RNA
376 polymerase on active genes, likely by modulating the transcription-dependent histone
377 turnover. However, the mark does not elicit its repressive effect on constitutively
378 repressed genes which are displaying typical bivalent chromatin domains
379 characterized by H3K9me3 and H3K14ac enrichments.

380

381 **H4K31me1 enrichment, a blueprint for unannotated genes and uncharacterized** 382 **long non-coding RNAs**

383 As mentioned previously, H4K31me was mainly detected throughout the body of
384 active genes with translation start and stop codons as boundaries and its enrichment
385 was inversely correlated to the yield of mRNA. These features should allow this mark

386 to explicitly predict unannotated genes even when the low level of expression
387 impedes detection by RNA profiling (Figure 8 – figure supplement 1).

388

389 Although H4K31me1 rarely covered IGRs (Fig. 6c), the mark was found enriched
390 occasionally in chromosomal regions devoid of any predicted protein-coding genes
391 (Fig. 8c and Figure 8 – figure supplement 3). This enrichment correlated with
392 extensive transcription of large RNA transcripts ranging from 20 to 70 kB that we
393 termed “long noncoding RNAs” (lncRNAs). Those *T. gondii* lncRNAs are stand-alone
394 transcription units with a proper chromatin signature, i.e., H4K31ac and H3K4me3 at
395 the promoter and H4K31me1 along the transcribed length (Fig. 8c and Figure 8 –
396 figure supplement 3). Considering their distribution at both telomere-adjacent regions
397 (Fig. 8c) and chromosome arms (Figure 8 – figure supplement 3), those lncRNAs may
398 work in *cis* near the site of their production (e.g., functions in telomere homeostasis)
399 or act in *trans* to alter chromatin shape and gene expression at distant loci, as reported
400 in other model organisms (Azzalin and Lingner, 2015).

401

402 **Distribution of H4K31 modifications across the *P. falciparum* genome reveals** 403 **H4K31me1 as a novel pericentromeric PTM**

404 The *P. falciparum* genome was shown to be primarily maintained in a decondensed
405 euchromatic state with perinuclear heterochromatin islands. Those heterochromatin-
406 based gene silencing regions are used for the regulation of monoallelic expression of
407 clonally variant genes (e.g. *var* and *rifin*) and are enriched in H3K9me3 which binds
408 HP1 (Voss et al., 2014). We observe an apparent non-overlapping staining for
409 acetylated and methylated H4K31 and more specifically a discrete focal distribution
410 of H4K31me1 at the nuclear periphery (Fig. 2b and Fig. 5c). To get a comprehensive
411 view of the genomic distribution of those PTMs across *P. falciparum* genome, we also
412 performed ChIP-seq analyses during the IDC. As for *T. gondii*, we observed low
413 variability and high similarity in read coverage between technical replicates for all the
414 antibodies used (Figure 9 – figure supplement 1).

415

416 We next compared peak location for each antibody type. H4K31ac displayed a rather
417 even distribution throughout the genome similarly to the euchromatic mark H3K4me3
418 (Fig. 9a and enhanced view at Figure 9 – figure supplement 2). As for *T. gondii*,
419 H4K31ac matched with the gene annotation, i.e., high at promoter and low at gene

420 body of active genes (e.g., *GAPDH*, Fig. 9b). Consistent with this, the H4K31ac and
421 the repressive mark H3K9me3 were found inversely correlated (Fig. 9a). However,
422 while H4K31ac displayed a relatively narrow enrichment restricted to transcribed
423 promoters, H3K4me3 was instead enriched in a large fraction of the genome (Fig. 9c)
424 as already described (Salcedo-Amaya et al., 2009).

425

426 Interestingly, the methylation of H3K9 and the properties to bind HP1 which have
427 emerged as hallmarks of pericentromeric heterochromatin in model systems,
428 including *T. gondii* (Fig. 8b and Figure 8 – figure supplement 2) have not been
429 detected in *P. falciparum*, leading to the view that the parasite may lack pericentric
430 heterochromatin (Flueck et al., 2009; Lopez-Rubio et al., 2009; Salcedo-Amaya et al.,
431 2009). While our ChIP-seq analysis confirmed the absence of pericentric enrichment
432 of both H3K9me3 and HP1, it clearly highlighted a remarkable enrichment of
433 H4K31me1 at pericentromeric regions that flank the cenH3-enriched centromeres
434 (Fig. 9d, Figure 10 – figure supplement 1 and Fig.10a). It is therefore possible that
435 H4K31me1 constrains PfCENH3 to the centromeres in *P. falciparum* thus replacing
436 the H3K9me3/HP1 functions in most of the species. In addition to the pericentromeric
437 localization H4K31me1 was also enriched to few sub-telomeric regions and more
438 specifically at telomere-associated repetitive element (TARE, Fig. 10b) repeat blocks
439 shown to encode lncRNAs (Fig. 10c) (Sierra-Miranda et al., 2012). The presence of
440 the mark at 6 out of 60 silenced *var* genes (Fig. 9f and 10a) combined with its absence
441 at transcriptionally permissive loci (e.g. *GAPDH*, Fig. 9b) suggest H4K31me1 as a
442 novel hallmark of heterochromatin in *P. falciparum*, but not similar to H3K9me3/HP1
443 in subtelomeric regions, which have significantly higher coverage over *var* genes.
444 Moreover, while HP1 mediates H3K9 methylation spreading to next-door nucleosome
445 until the *var* genes-containing region is fully recovered by the methyl mark,
446 H4K31me1 apparently nucleates at specific DNA elements but was not able to spread
447 over a series of nucleosomes (Fig. 10a, lower panel), a situation that could result from
448 the lack of HP1-like reader.

449

450 Discussion

451 In this study, we provide in depth understanding of the interaction between the core
452 histone H4 and the template DNA by functionally characterizing novel modifications

453 of H4K31, a residue exposed on the outer surface of the nucleosome in close
454 proximity to the DNA entry-exit point. Proteome-wide mapping of
455 acetylome/methylome as well as nucleosome protein content analyses allowed
456 identifying H4K31 as a site for both acetylation and methylation across a wide range
457 of species including those from the apicomplexan parasitic phylum. The K31 residue
458 lies at the N-terminus of the histone H4 α 1 helix and its side chain forms a water-
459 mediated salt bridge with the DNA phosphate backbone (Fig. 1c). Its acetylation was
460 predicted to trigger substantial conformational changes in the nucleosome by shifting
461 the side chain of lysine from unacetylated to acetylated state and causing a loss of the
462 water-mediated interactions K31 establishes with DNA and the residue R35 (Fig. 1c).
463 However, this prediction was not validated since X-ray crystallography did not
464 indicate large structural changes into nucleosomes when glutamine was substituted to
465 lysine to mimic the acetylated state (H4Q31, Fig. 1c) (Iwasaki et al., 2011).
466 Alternatively, H4K31ac may increase DNA unwrapping at the entry-exit point of the
467 nucleosome thus giving access to the ATP-dependent chromatin remodelers that act on
468 nucleosome disassembly and turnover as proposed by Chatterjee et al., 2015. The
469 latter assumption would fit with the “regulated nucleosome mobility” model
470 (Cosgrove et al., 2004), which predicts that outer surface PTMs (e.g., H3K36ac,
471 Williams et al., 2008) regulate the equilibrium between mobile and relatively
472 stationary nucleosomes by altering histone-DNA molecular interplay.

473
474 In both *T. gondii* and *P. falciparum*, genome wide studies pinpointed a local
475 enrichment of H4K31ac at active gene promoters, in line with the cooperative
476 contribution of acetylation and other PTMs to shape a transcriptionally permissive
477 chromatin state. While H4K31ac relieves nucleosomal repression thus facilitating
478 the access of the transcriptional machinery to the DNA template, H4K31
479 monomethylation not only locks the nucleosome in a repressed conformation
480 which maintains chromatin in a closed or semi-closed state also called poised-state,
481 but it also prevents GCN5-related HAT to catalyze acetylation of the residue.
482 Interestingly, in *T. gondii*, apart from its direct effects on the nucleosome mobility and
483 chromatin state, we found that H4K31ac also prevents methylation at the body of
484 highly expressed genes, thereby ensuring maximal efficacy of the RNA polymerase
485 progression and activity. Indeed it is only in the transcribed coding sequence of a
486 subset of genes typically associated with limited activity of the RNA polymerase II

487 that we found enrichment in H4K31me1. In a model where the nucleosome
488 disassembles in front of transcribing RNA polymerase II to allow its physical
489 progression across transcribed units, it is plausible that H4K31me1 by stabilizing the
490 wrapping of DNA around the histone octamer slows down the RNA Pol II
491 processing along the fiber hence reducing the level of transcription.

492
493 Aside from specific patterns of PTMs, histone chaperones significantly contribute to
494 control how the RNA polymerase II engages the nucleosome in and around a
495 promoter and during the elongation step. For instance, the FACT (Facilitates
496 Chromatin Transactions) histone chaperone was shown to assist first in the removal
497 of nucleosomes ahead of the transcribing RNA Pol II and next in their reassembly
498 after polymerase passage. While we showed that H4K31 modifications, at least in *T.*
499 *gondii*, play a key role in gene regulation, studies in other Eukaryotes have underlined
500 H4K31 as instrumental in the recruitment/mobilization of histone chaperone at
501 transcribed genes. In budding yeast, H4K31 along with two proximal residues on the
502 side of the nucleosome (i.e., H4R36 and H3L61) promotes the recruitment of the
503 yFACT subunit Spt16 across transcribed genes as assessed by the typical change in
504 Spt16 distribution which occupancy shifts toward the 3' ends of transcribed genes in
505 the H4K31E yeast mutant ([Nguyen et al., 2013](#)).

506
507 The versatility of H4K31 goes even beyond these modifications since
508 ubiquitylation of H4K31 has been reported in human cells as an additional
509 regulatory PTM for transcription elongation ([Kim et al., 2013](#)). Indeed, it was
510 shown that the histone H1.2 subtype while localized at target genes interacts
511 with the elongating RNA Polymerase II, typified by phosphorylation of Ser2 on its
512 carboxy terminal domain (CTD). Indeed, it was shown that upon interaction with
513 the Ser2-phosphorylated carboxy terminal domain CTD of the active RNA Pol II,
514 the histone H1.2 subtype becomes able to recruit the Cul4A E3 ubiquitin ligase
515 and PAF1 elongation complexes. Those stimulates, in turn, H4K31 ubiquitylation
516 that influences positively the accumulation of the H3K4me3/H3K79me2
517 signature, thereby leading to more productive elongation phase of transcription.
518 Importantly, blocking H4K31 ubiquitylation by K31R mutation markedly reduces
519 H3K4 and H3K79 methylation and consequently impairs gene transcription ([Kim
520 et al., 2013](#)).

521

522 In order to test the functional significance of H4K31 modifications *in vivo*, we tried
523 but remained unsuccessful at substituting H4K31 in *T. gondii* genome with alanine or
524 glutamine to mimic acetyl lysine or with arginine to mimic nonacetylated lysine ([data](#)
525 [not shown](#)). Engineered budding yeast with those substitutions did not significantly
526 affect cell viability but led to an unexpected increase of telomeric and ribosomal DNA
527 silencing ([Hyland et al., 2005](#)) that both argue for the mutations driving a non-
528 permissive chromatin state. This does not fit with our working model in which,
529 H4K31Q should promote open chromatin. It is however plausible that the
530 substitutions did not faithfully mimic the effects of the modifications in these
531 instances. In sharp contrast with the aforementioned substitutions, glutamic acid
532 (E) that mimics succinylated lysine was shown to severely compromise the growth in
533 budding yeast ([Xie et al., 2012](#)) maybe as a consequence of the alteration in the
534 distribution of Spt16 across yeast genes ([Nguyen et al., 2013](#)). The succinylation on
535 H4K31 has also been detected by mass spectrometry in *T. gondii* ([Li et al., 2014](#);
536 [Nardelli et al., 2013](#)). The modification could drastically impact intranucleosomal
537 structure and induce “abnormal” histone-DNA interactions ([Fig. 1c](#)).

538

539 In this context, H4K31 methylation would counteract the activating effect of H4K31
540 acetylation and succinylation, by preventing the nucleosome from adopting an
541 open conformation permissive to gene expression. The analysis in *P. falciparum*
542 revealed remarkable features of H4K31me1 by stressing a much more
543 pronounced repressive character as the modification was exclusively restricted to
544 non-permissive silenced chromosomal zones. Originally, *P. falciparum*
545 heterochromatin in which clusters of genes are maintained in a silent state was
546 singularly defined by increased nucleosomal occupancy, histone deacetylation,
547 H3K9me3 and the binding of PfHP1 ([Scherf et al., 2008](#)). While most of the
548 genome can be characterized as euchromatin, those silenced regions were
549 organized towards the periphery of the nucleus and contain among others the
550 *var*, *rif* and *stevor* families that cluster *together*, proximal to each telomere. The
551 repression of the *var* genes for instance involves the trimethylation of H3K9 and
552 its spreading to the next-door nucleosome by the action of HP1 ([Scherf et al.,](#)
553 [2008](#)). H4K31me1 enrichment was detected, yet unevenly and at low rates in the

554 vicinity of *var* and *rifin* genes (Fig. 9f and 10c). However, the modification does
555 not spread while its enrichment fades quickly and remains likely limited to the
556 site of heterochromatin initiation where both H3K9me3 and HP1 levels
557 culminate (Fig. 10C). The lack of spread of H4K31me1 along a series of
558 nucleosomes may be explained by the absence of a competent protein reader that
559 specifically recognizes the PTM and recruits the H4K31me1-catalyzing
560 methyltransferase. So far, no H4K31me1-reading protein was identified although
561 the PTM is not buried and hence accessible for regulatory factor binding. In fact,
562 the bromodomain of BRD4 is able to recognize the acetylated isoform of H4K31
563 (Filippakopoulos et al., 2012).

564

565 While H4K31me1 occupancy is overall limited across *P. falciparum* genome, the
566 modification is by far the most promiscuous PTMs found at pericentromeric
567 zones of all chromosomes (Fig. 10a and Figure 10 – figure supplement 1). As such,
568 both H4K31me1 (Fig. 5c) and centromeres (Hoeijmakers et al., 2012) were found
569 to be clustered towards nuclear periphery. *P. falciparum* centromeres were
570 originally described as displaying a unique epigenetic status typified by the
571 noteworthy absence of the canonical pericentromeric PTM H3K9me3
572 (Hoeijmakers et al., 2012) present in all species including *T. gondii* (Fig. 8C and
573 Figure 8 – figure supplement 2). Clearly this study has emphasized a pivotal role of
574 H4K31me1 in pericentromeric heterochromatin in *P. falciparum* and have
575 provided new insights on the mechanism of transcriptional regulation in *T. gondii*.

576

577 In metazoan, H4K31me1 was shown to decorate the mitotic chromosome arms (Fig.
578 5b). The PTM is in this regard a novel mitotic marker that targets newly
579 synthesized H4 hence regulating chromosomal condensation and segregation during
580 mitosis. H4K31 is structurally very close to H3K56 (Fig. 1c), the acetylation of
581 which reported to increase the binding affinity of H3 toward histone chaperones,
582 thereby promoting nucleosome assembly during S phase of the cell cycle (Li et al.,
583 2008). Collectively, our results argue for a similar role for H4K31me1 in chromatin
584 assembly during DNA replication in metazoan. However the picture appears more
585 complex since H4K31 methylation, unlike H3K56ac, is predicted to prevent

586 histone exchange, thereby slowing histone turnover rate behind the replication forks
587 which overall contributes to stabilize newly incorporated nucleosomes into chromatin.

588

589 In conclusion, we demonstrate that H4K31 acetylation and methylation are
590 associated to very distinct nuclear functions in *T. gondii* and *P. falciparum*.
591 Moreover, we demonstrate the evolution of distinct epigenetic strategies in
592 these closely linked parasites to organize chromosome regions that are essential
593 for cell division and gene expression.

594

595 **Materials and Methods**

596 **Parasites and host cells culture**

597 HFF primary cells were cultured in Dulbecco's Modified Eagle Medium (DMEM)
598 (Invitrogen) supplemented with 10% heat inactivated Fetal Bovine Serum (FBS)
599 (Invitrogen), 10mM (4-(2-hydroxyethyl)-1-piperazine ethanesulphonic acid) (HEPES)
600 buffer pH 7.2, 2 mM L-glutamine and 50 µg ml of penicillin and streptomycin
601 (Invitrogen). Cells were incubated at 37°C in 5% CO₂. Type I (RH wild type and RH
602 $\Delta ku80$) and type II strains (Pru $\Delta ku80$) of *T. gondii* were maintained *in vitro* by serial
603 passage on monolayers of HFFs. *P. falciparum* 3D7 strain was grown in RPMI 1640
604 media supplemented with 0.5% Albumax II, 0.1mM Hypoxanthine and Gentamicin
605 10 mcg/ml. The culture was maintained at 2% hematocrit and 5% parasitemia. The
606 parasites were grown at 37°C and at 1% O₂, 5% CO₂ and 94% N₂ gas mixture
607 concentration.

608

609 **Immunofluorescence microscopy**

610 *T. gondii* infecting HFF cells grown on coverslips were fixed in 3% formaldehyde for
611 20 min at room temperature, permeabilized with 0.1% (v/v) Triton X-100 for 15 min
612 and blocked in Phosphate buffered saline (PBS) containing 3% (w/v) BSA. The cells
613 were then incubated for 1 hour with primary antibodies followed by the addition of
614 secondary antibodies conjugated to Alexa Fluor 488 or 594 (Molecular Probes).
615 Nuclei were stained for 10 min at room temperature with Hoechst 33258. Coverslips
616 were mounted on a glass slide with Mowiol mounting medium, and images were
617 acquired with a fluorescence ZEISS ApoTome.2 microscope and images were

618 processed by ZEN software (Zeiss). *P. falciparum* asexual blood life stages were
619 washed with phosphate-buffered saline (PBS) and fixed in solution with 4%
620 paraformaldehyde and 0.0075% glutaraldehyde in PBS for 30 min. After one wash
621 with PBS, cells were permeabilized with 0.1% Triton X-100 in PBS for 10 min. Cells
622 were washed twice with PBS, blocked with 3% bovine serum albumin (BSA) in PBS
623 for 1 hour. The cells were then incubated for 1 hour with primary antibodies followed
624 by the addition of secondary antibodies conjugated to Alexa Fluor 488 or 594
625 (Molecular Probes). Nuclei were stained for 30 min at room temperature with Hoechst
626 33258. The parasites were finally washed 2-3 times before loading on to glass slides
627 mixed with fluoro-gel (Electron Microscopy Sciences). Images were acquired with a
628 fluorescence ZEISS ApoTome.2 microscope and images were processed by ZEN
629 software (Zeiss).

630

631 **HDACi treatments**

632 The final concentration of histone deacetylase inhibitors dissolved in DMSO was, as
633 described (Bougdour et al., 2013), FR-235222 (90nM), apicidin (100nM), HC-toxin
634 (100nM), trichostatin A (100nM), scriptaid (100nM), APHA (100mM) and sodium
635 butyrate (5mM). They were added to infected HFF cells for 18 hours. Halofuginone
636 (10 nM) was shown to inhibit prolyl-tRNA synthetase (Jain et al., 2015) and was used
637 as a control.

638

639 **Plasmid constructs**

640 To construct the vector pLIC-ENO1-HAFlag, the coding sequence of
641 ENO1(TGME49_268860) was amplified using primers LIC-268860_Fwd
642 (TACTTCCAATCCAATTTAGCgaacatgcaggcaatggcttgctcttc) and LIC-268860_Rev
643 (TCCTCCACTTCCAATTTTAGCttttgggtgtcgaagctctctcccgcg) using PruKu80
644 genomic DNA as template. The resulting PCR product was cloned into the pLIC-HF-
645 dhfr vector using the LIC cloning method as reported previously (Bougdour et al.,
646 2013).

647

648 **Cas9-mediated gene disruption in *Toxoplasma gondii***

649 The plasmid pTOXO_Cas9-CRISPR was described in (Sangaré et al., 2016). For gene
650 disruption using CRISPR/Cas9 system, the genes of interests (GOI) were: GCN5A
651 (*TGGT1_254555*), GCN5B (*TGGT1_243440*), MYST-A (*TGGT1_318330*), MYST-B

652 (*TGGT1_207080*), HAT1 (*TGGT1_293380*), HDAC1 (*TGGT1_281420*), HDAC2
653 (*TGGT1_249620*), HDAC3 (*TGGT1_227290*), HDAC4 (*TGGT1_257790*) and
654 HDAC5 (*TGGT1_202230*). Twenty mers-oligonucleotides corresponding to specific
655 GOI were cloned using Golden Gate strategy. Briefly, primers TgGOI-CRISP_FWD
656 and TgGOI-CRISP_REV containing the sgRNA targeting TgGOI genomic sequence
657 were phosphorylated, annealed and ligated into the linearized pTOXO_Cas9-CRISPR
658 plasmid with BsaI, leading to pTOXO_Cas9-CRISPR::sgTgGOI. *T. gondii*
659 tachyzoites were then transfected with the plasmid and grown on HFF cells for 18-36
660 hours. Cloning oligonucleotides used in this study:

661 TgHDAC1-CRISP-FWD : 5'- AAGTTGCGTCGCCGTTCTCTCACGCG -3'
662 TgHDAC1-CRISP-REV : 5'- AAAACGCGTGAGAGAACGGCGACGCA -3'
663 TgHDAC2-CRISP-FWD : 5'- AAGTTGCGCCCGTCGCCTCCCCCGCG -3'
664 TgHDAC2-CRISP-REV : 5'- AAAACGCGGGGGAGGCGACGGGCGCA -3'
665 TgHDAC3-CRISP-FWD : 5'- AAGTTGATATCGGAAGTTACTACTAG -3'
666 TgHDAC3-CRISP-REV : 5'- AAAACTAGTAGTAACTTCCGATATCA -3'
667 TgHDAC4-CRISP-FWD : 5'- AAGTTGCTGTTGCTGAAGCCCAGGCG -3'
668 TgHDAC4-CRISP-REV : 5'- AAAACGCCTGGGCTTCAGCAACAGCA -3'
669 TgHDAC5-CRISP-FWD : 5'- AAGTTGGCGAGACCGGGGCAGCCGCG -3'
670 TgHDAC5-CRISP-REV : 5'- AAAACGCGGCTGCCCCGGTCTCGCCA -3'
671 TgGCN5A-CRISP-FWD : 5'- AAGTTGCGTGACGAACGACAGGCAAG -3'
672 TgGCN5A-CRISP-REV : 5'- AAAACTTGCCTGTCGTTTCGTCACGCA -3'
673 TgGCN5B-CRISP-FWD : 5'- AAGTTGGGTTTCCTGTGTCGAGACCG -3'
674 TgGCN5B-CRISP-REV : 5'- AAAACGGTCTCGACACAGGAAACCCA -3'
675 TgMYSTA-CRISP-FWD : 5'- AAGTTGGCTGCTCCGCGACTCAGCGG -3'
676 TgMYSTA-CRISP-REV : 5'- AAAACCGCTGAGTCGCGGAGCAGCCA -3'
677 TgMYSTB-CRISP-FWD : 5'- AAGTTGCGCGAAGAAGGGAGAGAGCG -3'
678 TgMYSTB-CRISP-REV : 5'- AAAACGCTCTCTCCCTTCTTCGCGCA -3'
679 TgHAT1-CRISP-FWD : 5'- AAGTTGCCGACGGGTCACGGAGACTG -3'
680 TgHAT1-CRISP-REV : 5'- AAAACAGTCTCCGTGACCCGTCGGCA -3'

681

682 ***Toxoplasma gondii* transfection**

683 *T. gondii* RH, RH $\Delta ku80$ and Pru $\Delta ku80$ were electroporated with vectors in cytomix
684 buffer (120mM KCl, 0.15mM CaCl₂, 10mM K₂HPO₄/ KH₂PO₄ pH7.6, 25mM HEPES
685 pH7.6, 2mM EGTA, 5mM MgCl₂) using a BTX ECM 630 machine (Harvard

686 Apparatus). Electroporation was performed in a 2mm cuvette at 1.100V, 25Ω and
687 25μF. Stable transgenic parasites were selected with 1μM pyrimethamine, single-
688 cloned in 96 well plates by limiting dilution and verified by immunofluorescence
689 assay.

690

691 **Antibodies**

692 Primary antibodies : rabbit home-made anti-TgHDAC3 described in ([Bougdour et al.,](#)
693 [2009](#)), mouse anti-HA (3F10, Roche), rabbit anti-H4K8ac (Upstate 06-760), rabbit
694 anti-H4K12ac (Upstate 06-761), rabbit anti-H3K4ac (Diagenode C15410165), rabbit
695 anti-H3K9ac (Diagenode C15410004), rabbit anti-H3K14ac (C15210005), rabbit anti-
696 H3K18ac (Diagenode C15410193), rabbit anti-H3K27ac (Millipore, 04-1044-S) and
697 mouse anti-H3K27ac (Diagenode C15200184), H4K20me3 (C15410207), H3K9me3
698 (Millipore, 17-625), H3K4me1 (C15410194) and H3K4me3 (C15410003-50).
699 Western blot secondary antibodies were conjugated to alkaline phosphatase
700 (Promega), while immunofluorescence secondary antibodies were coupled with Alexa
701 Fluor 488 or Alexa Fluor 594 (Thermo Fisher Scientific). We also raised homemade
702 H4K31acetylation and H4K31monomethylation-specific antibodies in rabbit against
703 linear peptides corresponding to amino acid residues 23-35 of histone H4 and carrying
704 modified residue K31: C-DNIQGITKme1PAIR; C-DNIQGITKacPAIR and C-
705 RDNIQGITKacPAIR. They were produced by Eurogentec and used for
706 immunofluorescence, immunoblotting and chromatin immunoprecipitation.

707

708 **Histones purification, Immunoblotting and mass spectrometry-based proteomic** 709 **analysis**

710 For histone purification, HFF cells were grown to confluence and infected with
711 PruΔ*ku80* parasites. Intracellular tachyzoites were treated with histone deacetylase
712 HDAC3 inhibitor, 90nM FR235222 for 18 hours. As appropriate control, we treated
713 tachyzoites with 0.1% DMSO. Histones were extracted and purified using histone
714 purification kit (Active motif) according to manufacturer's protocol. For western
715 blotting, histone proteins were run on a NuPAGE 4-12% Bis-Tris polyacrylamide gels
716 in MES-SDS running buffer (Invitrogen) and transferred to a polyvinylidene fluoride
717 PVDF membrane (Immobilon-P; Millipore) using NuPAGE transfer buffer
718 (Invitrogen). The blots were probed using primary antibodies: pan acetyl H4,
719 H4K31ac and H4K31me1, followed by phosphatase-conjugated goat secondary

720 antibodies (*Promega*). The expected band of histones were detected using NBT-BCIP
721 (*Amresco*). Nucleosomes from *T. gondii*-infected cells were purified and proteins
722 separated by SDS-PAGE. The band corresponding to H4 was excised and its protein
723 content digested using trypsin. Resulting peptides were submitted to mass
724 spectrometry-based proteomic analysis (U3000 RSLCnano coupled to Q-Exactive HF,
725 Thermo Scientific). Peptides and proteins were identified using Mascot software
726 (Matrix Science).

727

728 **Chromatin Immunoprecipitation and Next Generation Sequencing in** 729 ***Toxoplasma gondii***

730 HFF cells were grown to confluence and infected with type II (Pru Δ ku80) strain.
731 Harvested intracellular parasites were crosslinked with formaldehyde (final
732 concentration 1%) for 8 min at room temperature and the crosslinking was stopped by
733 addition of glycine (final concentration 0.125M) for 5 min at room temperature.
734 Crosslinked chromatin was lysed in ice-cold lysis buffer (50mM HEPES KOH pH7.5,
735 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5%NP-40, 0.125% triton X-100,
736 protease inhibitor cocktail) and sheared in shearing buffer (1mM EDTA pH8.0,
737 0.5mM EGTA pH8.0, 10mM Tris pH8.0, protease inhibitor cocktail) by sonication
738 using a Diagenode Biorupter. Samples were sonicated, for 16 cycles (30 seconds ON
739 and 30 seconds OFF), to 200-500 base-pair average size. Immunoprecipitation was
740 carried out using sheared chromatin, 5% BSA, protease inhibitor cocktail, 10% triton
741 X-100, 10% deoxycholate, DiaMag Protein A-coated magnetic beads (Diagenode) and
742 antibodies (H4K31ac, H4K31me1, pan acetyl H4, H4K20me3, H3K9me3, H3K4me3,
743 H3K4me1, H3K14ac). A rabbit IgG antiserum was used as a control mock. After
744 overnight incubation at 4°C on rotating wheel, chromatin-antibody complexes were
745 washed and eluted from beads by using iDeal ChIP-seq kit for Histones (Diagenode)
746 according to the manufacturer's protocol. Samples were decrosslinked by heating for
747 4 hours at 65°C. DNA was purified by using IPure kit (Diagenode) and quantified by
748 using Qubit Assays (*Thermo Fisher Scientific*) according to the manufacturer's
749 protocol. For ChIP-seq, purified DNA was used to prepare libraries and then
750 sequenced by Arraystar (USA, <http://www.arraystar.com/>).

751

752 **Library Preparation, Sequencing and Data analysis (Arraystar)**

753 ChIP-Sequencing library preparation was performed according to Illumina's protocol
754 Preparing Samples for ChIP Sequencing of DNA. **Library Preparation:** 10 ng DNA
755 of each sample was converted to phosphorylated blunt-ended with T4 DNA
756 polymerase, Klenow polymerase and T4 polymerase (NEB); An 'A' base was added
757 to the 3' end of the blunt phosphorylated DNA fragments using the polymerase
758 activity of Klenow (exo minus) polymerase (NEB); Illumina's genomic adapters were
759 ligated to the A tailed DNA fragments; PCR amplification was performed to enrich
760 ligated fragments using Phusion High Fidelity PCR Master Mix with HF Buffer
761 (Finnzymes Oy). The enriched product of ~200-700 bp was cut out from gel and
762 purified. **Sequencing:** The library was denatured with 0.1M NaOH to generate
763 single-stranded DNA molecules, and loaded onto channels of the flow cell at 8pM
764 concentration, amplified in situ using TruSeq Rapid SR cluster kit (#GD-402-4001,
765 Illumina). Sequencing was carried out by running 100 cycles on Illumina HiSeq 4000
766 according to the manufacturer's instructions. **Data Analysis:** After the sequencing
767 platform generated the sequencing images, the stages of image analysis and base
768 calling were performed using Off-Line Basecaller software (OLB V1.8). After
769 passing Solexa CHASTITY quality filter, the clean reads were aligned to *T. gondii*
770 reference genome (Tgo) using BOWTIE (V2.1.0). Aligned reads were used for peak
771 calling of the ChIP regions using MACS V1.4.0. Statistically significant ChIP-
772 enriched regions (peaks) were identified by comparison of two samples, using a p-
773 value threshold of 10^{-5} . Then the peaks in each sample were annotated by the
774 overlapped gene using the newest *T. gondii* database. The EXCEL/BED format file
775 containing the ChIP-enriched regions was generated for each sample. **Data**
776 **visualization:** The mapped 100 bp reads represent enriched DNA fragments by ChIP
777 experiment. Any region of interest in the raw ChIP-seq signal profile can be directly
778 visualized in genome browser. We use 10-bp resolution intervals (10-bp bins) to
779 partition the stacked reads region, and count the number of reads in each bin. All the
780 10 bp resolution ChIP-seq profiles of each sample are saved as UCSC wig format
781 files, which can be visualized in *T. gondii* Genome Browser
782 (http://protists.ensembl.org/Toxoplasma_gondii/Info/Index). All these raw and
783 processed files can be found at [Series GSE98806](#).

784

785 **Chromatin Immunoprecipitation and Next Generation Sequencing in *P.***
786 ***falciparum***

787 Chromatin from synchronous rings stage parasites of 3D7 clone G7 was prepared and
788 3×10^8 cells per ChIP used for the previously described protocol (Lopez-Rubio et al.,
789 2013). Briefly, chromatin was crosslinked in 1% formaldehyde for 10 min (Sigma-
790 Aldrich, #SZBD1830V), sheared to an average length of 300 bp using the BioRuptor
791 Pico and individual histone modifications were pulled down using 0.5 μ g of antibody
792 for H3K4me3 (Diagenode, cat # K2921004), H3K9me3 (Millipore, cat # 257833),
793 and home-made rabbit polyclonal anti-PfHP1. 5 μ l rabbit polyclonal anti-H4K31me1
794 and 15 μ l anti-H4K31ac were used for each experiment. To generate Illumina-
795 compatible sequencing libraries, the immunoprecipitated DNA and input was
796 processed using the MicroPlex Library Preparation Kit (Diagenode C05010014)
797 according to manufacturer's instructions. The optimized library amplification step
798 was used KAPA Biosystems HIFI polymerase (KAPA Biosystems KK2101). Pooled,
799 multiplexed libraries were sequenced on an Illumina NextSeq® 500/550 system as a
800 150 nucleotide single-end run. The raw data were demultiplexed using bcl2fastq2
801 (Illumina) and converted to fastq format files for downstream analysis. Two
802 biological replicates were analyzed for each antibody.

803

804 ***Plasmodium falciparum* ChIP-seq Data Analysis**

805 Sequencing reads were mapped to the *P. falciparum* 3D7 genome assembly
806 (PlasmoDB v3.0) with Burrows-Wheeler Alignment tool (BWA) using default
807 settings, and then sequences were quality filtered at Q20 Phred quality score. ChIP-
808 seq peak calling was performed using the MACS2 algorithm. For genome-wide
809 representation of each histone mark's distribution, the coverage was calculated as
810 average reads per million over bins of 1000 nucleotides using bamCoverage from the
811 package deepTools. Correlation of the different biological replicates were calculated
812 by performing Pearson's and Spearman's correlation analysis of pairwise comparison
813 of BAM alignment files, and ChIP-seq peak enrichment scores (\log_2) using MACS2
814 and deepTools. Circular and linear coverage plots were generated using Circos and
815 Integrated Genomics Viewer, respectively. All these raw and processed files can be
816 found at [NCBI Bioproject ID PRJNA386433](#).

817

818 **Acknowledgments**

819 This work was supported by the Agence Nationale Recherche, France [grant ANR
820 Blanc 2012 TOXOHDAC – grant number ANR-12-BSV3-0009-01]; the Laboratoire
821 d’Excellence ParaFrap, France [grant number ANR-11-LABX-0024]; Atip-Avenir and
822 Finovi programs [Apicolipid projects to CYB]; the European Research Council [ERC
823 Consolidator Grant N°614880 Hosting TOXO to M.A.H] and [ERC AdG N° 670301
824 PlasmoSilencing to A.S.). Proteomic experiments were partly supported by the ProFi
825 grant (ANR-10-INBS-08-01).

826

827 **Author contributions**

828 F. S., D. C., A. B., I. T. and M. A. H conceived and designed the experiments,
829 analysed data and wrote paper; S. D., T. H., C.Y. B. and A. S. contributed to *P.*
830 *falciparum* experiments; P. O., M. B. and S. B. contributed to ChIP-seq analysis; A. P.
831 contributed to structural modeling and L. B. and Y. C. to mass spectrometry analysis.

832

833 **References**

- 834 Azzalin, C.M., and Lingner, J. (2015). Telomere functions grounding on TERRA
835 firma. *Trends Cell Biol.* 25, 29–36.
- 836 Bhatti, M.M., Livingston, M., Mullapudi, N., and Sullivan, W.J. (2006). Pair of
837 Unusual GCN5 Histone Acetyltransferases and ADA2 Homologues in the Protozoan
838 Parasite *Toxoplasma gondii*. *Eukaryot. Cell* 5, 62–76.
- 839 Bougdour, A., Maubon, D., Baldacci, P., Ortet, P., Bastien, O., Bouillon, A., Barale,
840 J.-C., Pelloux, H., Ménard, R., and Hakimi, M.-A. (2009). Drug inhibition of HDAC3
841 and epigenetic control of differentiation in Apicomplexa parasites. *J. Exp. Med.* 206,
842 953–966.
- 843 Bougdour, A., Braun, L., Cannella, D., and Hakimi, M.-A. (2010). Chromatin
844 modifications: implications in the regulation of gene expression in *Toxoplasma*
845 *gondii*. *Cell. Microbiol.* 12, 413–423.
- 846 Bougdour, A., Durandau, E., Brenier-Pinchart, M.-P., Ortet, P., Barakat, M., Kieffer,
847 S., Curt-Varesano, A., Curt-Bertini, R.-L., Bastien, O., Coute, Y., et al. (2013). Host
848 cell subversion by *Toxoplasma* GRA16, an exported dense granule protein that targets
849 the host cell nucleus and alters gene expression. *Cell Host Microbe* 13, 489–500.
- 850 Bozdech, Z., Llinás, M., Pulliam, B.L., Wong, E.D., Zhu, J., and DeRisi, J.L. (2003).
851 The Transcriptome of the Intraerythrocytic Developmental Cycle of *Plasmodium*
852 *falciparum*. *PLoS Biol.* 1, e5.
- 853 Chatterjee, N., North, J.A., Dechassa, M.L., Manohar, M., Prasad, R., Luger, K.,
854 Ottesen, J.J., Poirier, M.G., and Bartholomew, B. (2015). Histone Acetylation near the

- 855 Nucleosome Dyad Axis Enhances Nucleosome Disassembly by RSC and SWI/SNF.
856 Mol. Cell. Biol. 35, 4083–4092.
- 857 Cobbold, S.A., Santos, J.M., Ochoa, A., Perlman, D.H., and Llinás, M. (2016).
858 Proteome-wide analysis reveals widespread lysine acetylation of major protein
859 complexes in the malaria parasite. Sci. Rep. 6, 19722.
- 860 Cosgrove, M.S., Boeke, J.D., and Wolberger, C. (2004). Regulated nucleosome
861 mobility and the histone code. Nat. Struct. Mol. Biol. 11, 1037–1043.
- 862 Daujat, S., Weiss, T., Mohn, F., Lange, U.C., Ziegler-Birling, C., Zeissler, U., Lappe,
863 M., Schübeler, D., Torres-Padilla, M.-E., and Schneider, R. (2009). H3K64
864 trimethylation marks heterochromatin and is dynamically remodeled during
865 developmental reprogramming. Nat. Struct. Mol. Biol. 16, 777–781.
- 866 Di Cerbo, V., Mohn, F., Ryan, D.P., Montellier, E., Kacem, S., Tropberger, P., Kallis,
867 E., Holzner, M., Hoerner, L., Feldmann, A., et al. (2014). Acetylation of histone H3 at
868 lysine 64 regulates nucleosome dynamics and facilitates transcription. Elife 3, e01632.
- 869 Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J.-P., Barsyte-
870 Lovejoy, D., Felletar, I., Volkmer, R., Müller, S., Pawson, T., et al. (2012). Histone
871 Recognition and Large-Scale Structural Analysis of the Human Bromodomain
872 Family. Cell 149, 214–231.
- 873 Flueck, C., Bartfai, R., Volz, J., Niederwieser, I., Salcedo-Amaya, A.M., Alako,
874 B.T.F., Ehlgén, F., Ralph, S.A., Cowman, A.F., Bozdech, Z., et al. (2009).
875 Plasmodium falciparum Heterochromatin Protein 1 Marks Genomic Loci Linked to
876 Phenotypic Variation of Exported Virulence Factors. PLoS Pathog. 5, e1000569.
- 877 Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F.,
878 Nehrbass, U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination
879 of virulence factor genes in telomeric chromosome clusters of P. falciparum. Nature
880 407, 1018–1022.
- 881 Garcia, B.A., Busby, S.A., Shabanowitz, J., Hunt, D.F., and Mishra, N. (2005).
882 Resetting the Epigenetic Histone Code in the MRL- *lpr* / *lpr* Mouse Model of Lupus
883 by Histone Deacetylase Inhibition. J. Proteome Res. 4, 2032–2042.
- 884 Garcia, B.A., Hake, S.B., Diaz, R.L., Kauer, M., Morris, S.A., Recht, J., Shabanowitz,
885 J., Mishra, N., Strahl, B.D., Allis, C.D., et al. (2006). Organismal Differences in Post-
886 translational Modifications in Histones H3 and H4. J. Biol. Chem. 282, 7641–7655.
- 887 Hoeijmakers, W.A.M., Flueck, C., François, K.-J., Smits, A.H., Wetzels, J., Volz,
888 J.C., Cowman, A.F., Voss, T., Stunnenberg, H.G., and Bartfai, R. (2012). *Plasmodium*
889 *falciparum* centromeres display a unique epigenetic makeup and cluster prior to and
890 during schizogony: Centromeres of *P. falciparum*. Cell. Microbiol. 14, 1391–1401.
- 891 Hyland, E.M., Cosgrove, M.S., Molina, H., Wang, D., Pandey, A., Cottee, R.J., and
892 Boeke, J.D. (2005). Insights into the Role of Histone H3 and Histone H4 Core
893 Modifiable Residues in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 25, 10060–10070.

- 894 Iwasaki, W., Tachiwana, H., Kawaguchi, K., Shibata, T., Kagawa, W., and
895 Kurumizaka, H. (2011). Comprehensive Structural Analysis of Mutant Nucleosomes
896 Containing Lysine to Glutamine (KQ) Substitutions in the H3 and H4 Histone-Fold
897 Domains. *Biochemistry (Mosc.)* 50, 7822–7832.
- 898 Jain, V., Yogavel, M., Oshima, Y., Kikuchi, H., Touquet, B., Hakimi, M.-A., and
899 Sharma, A. (2015). Structure of Prolyl-tRNA Synthetase-Halofuginone Complex
900 Provides Basis for Development of Drugs against Malaria and Toxoplasmosis. *Struct.*
901 *Lond. Engl.* 1993 23, 819–829.
- 902 Jeffers, V., and Sullivan, W.J. (2012). Lysine Acetylation Is Widespread on Proteins
903 of Diverse Function and Localization in the Protozoan Parasite *Toxoplasma gondii*.
904 *Eukaryot. Cell* 11, 735–742.
- 905 Karmodiya, K., Krebs, A.R., Oulad-Abdelghani, M., Kimura, H., and Tora, L. (2012).
906 H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while
907 H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem
908 cells. *BMC Genomics* 13, 424.
- 909 Kim, K., Lee, B., Kim, J., Choi, J., Kim, J.-M., Xiong, Y., Roeder, R.G., and An, W.
910 (2013). Linker Histone H1.2 Cooperates with Cul4A and PAF1 to Drive H4K31
911 Ubiquitylation-Mediated Transactivation. *Cell Rep.* 5, 1690–1703.
- 912 Kuo, M.H., Brownell, J.E., Sobel, R.E., Ranalli, T.A., Cook, R.G., Edmondson, D.G.,
913 Roth, S.Y., and Allis, C.D. (1996). Transcription-linked acetylation by Gcn5p of
914 histones H3 and H4 at specific lysines. *Nature* 383, 269–272.
- 915 Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone
916 Modifications Regulate Gene Expression. *Trends Genet.* 32, 42–56.
- 917 Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A., and Zhang,
918 Z. (2008). Acetylation of Histone H3 Lysine 56 Regulates Replication-Coupled
919 Nucleosome Assembly. *Cell* 134, 244–255.
- 920 Li, X., Hu, X., Wan, Y., Xie, G., Li, X., Chen, D., Cheng, Z., Yi, X., Liang, S., and
921 Tan, F. (2014). Systematic Identification of the Lysine Succinylation in the Protozoan
922 Parasite *Toxoplasma gondii*. *J. Proteome Res.* 13, 6087–6095.
- 923 Lopez-Rubio, J.-J., Mancio-Silva, L., and Scherf, A. (2009). Genome-wide Analysis
924 of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear
925 Repressive Centers in Malaria Parasites. *Cell Host Microbe* 5, 179–190.
- 926 Lopez-Rubio, J.-J., Siegel, T.N., and Scherf, A. (2013). Genome-wide chromatin
927 immunoprecipitation-sequencing in Plasmodium. *Methods Mol. Biol. Clifton NJ* 923,
928 321–333.
- 929 Mancio-Silva, L., and Scherf, A. (2013). In situ fluorescence visualization of
930 transcription sites and genomic Loci in blood stages of *Plasmodium falciparum*.
931 *Methods Mol. Biol. Clifton NJ* 923, 335–351.
- 932 Miao, J., Lawrence, M., Jeffers, V., Zhao, F., Parker, D., Ge, Y., Sullivan, W.J., and
933 Cui, L. (2013). Extensive lysine acetylation occurs in evolutionarily conserved

- 934 metabolic pathways and parasite-specific functions during *Plasmodium falciparum*
935 intraerythrocytic development: The malaria parasite acetylome. *Mol. Microbiol.* *89*,
936 660–675.
- 937 Moraes, I., Yuan, Z.-F., Liu, S., Souza, G.M., Garcia, B.A., and Casas-Mollano, J.A.
938 (2015). Analysis of histones H3 and H4 reveals novel and conserved post-
939 translational modifications in sugarcane. *PloS One* *10*, e0134586.
- 940 Morris, S.A., Rao, B., Garcia, B.A., Hake, S.B., Diaz, R.L., Shabanowitz, J., Hunt,
941 D.F., Allis, C.D., Lieb, J.D., and Strahl, B.D. (2006). Identification of Histone H3
942 Lysine 36 Acetylation as a Highly Conserved Histone Modification. *J. Biol. Chem.*
943 *282*, 7632–7640.
- 944 Nardelli, S.C., Che, F.-Y., Silmon de Monerri, N.C., Xiao, H., Nieves, E., Madrid-
945 Aliste, C., Angel, S.O., Sullivan, W.J., Angeletti, R.H., Kim, K., et al. (2013). The
946 Histone Code of *Toxoplasma gondii* Comprises Conserved and Unique
947 Posttranslational Modifications. *mBio* *4*, e00922-13-e00922-13.
- 948 Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J.,
949 Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J.W. (2009). A Method for
950 Genetically Installing Site-Specific Acetylation in Recombinant Histones Defines the
951 Effects of H3 K56 Acetylation. *Mol. Cell* *36*, 153–163.
- 952 Nguyen, H.-T.T., Wharton, W., Harper, J.A., Dornhoffer, J.R., and Duina, A.A.
953 (2013). A Nucleosomal Region Important for Ensuring Proper Interactions Between
954 the Transcription Elongation Factor Spt16 and Transcribed Genes in *Saccharomyces*
955 *cerevisiae*. *G3amp58 GenomesGenetics* *3*, 929–940.
- 956 Radke, J.R., Behnke, M.S., Mackey, A.J., Radke, J.B., Roos, D.S., and White, M.W.
957 (2005). The transcriptome of *Toxoplasma gondii*. *BMC Biol.* *3*, 26.
- 958 Salcedo-Amaya, A.M., van Driel, M.A., Alako, B.T., Trelle, M.B., van den Elzen,
959 A.M., Cohen, A.M., Janssen-Megens, E.M., van de Vegte-Bolmer, M., Selzer, R.R.,
960 Iniguez, A.L., et al. (2009). Dynamic histone H3 epigenome marking during the
961 intraerythrocytic cycle of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci.* *106*, 9655–
962 9660.
- 963 Sangaré, L.O., Alayi, T.D., Westermann, B., Hovasse, A., Sindikubwabo, F.,
964 Callebaut, I., Werkmeister, E., Lafont, F., Slomianny, C., Hakimi, M.-A., et al.
965 (2016). Unconventional endosome-like compartment and retromer complex in
966 *Toxoplasma gondii* govern parasite integrity and host infection. *Nat. Commun.* *7*,
967 11191.
- 968 Saraf, A., Cervantes, S., Bunnik, E.M., Ponts, N., Sardu, M.E., Chung, D.-W.D.,
969 Prudhomme, J., Varberg, J.M., Wen, Z., Washburn, M.P., et al. (2016). Dynamic and
970 Combinatorial Landscape of Histone Modifications during the Intraerythrocytic
971 Developmental Cycle of the Malaria Parasite. *J. Proteome Res.* *15*, 2787–2801.
- 972 Sautel, C.F., Cannella, D., Bastien, O., Kieffer, S., Aldebert, D., Garin, J., Tardieux,
973 I., Belrhali, H., and Hakimi, M.-A. (2007). SET8-mediated methylations of histone

- 974 H4 lysine 20 mark silent heterochromatic domains in apicomplexan genomes. *Mol.*
975 *Cell. Biol.* *27*, 5711–5724.
- 976 Scherf, A., Lopez-Rubio, J.J., and Riviere, L. (2008). Antigenic Variation in
977 *Plasmodium falciparum*. *Annu. Rev. Microbiol.* *62*, 445–470.
- 978 Sierra-Miranda, M., Delgadillo, D.M., Mancio-Silva, L., Vargas, M., Villegas-
979 Sepulveda, N., Martínez-Calvillo, S., Scherf, A., and Hernandez-Rivas, R. (2012).
980 Two long non-coding RNAs generated from subtelomeric regions accumulate in a
981 novel perinuclear compartment in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.*
982 *185*, 36–47.
- 983 Soldi, M., Cuomo, A., and Bonaldi, T. (2014). Improved bottom-up strategy to
984 efficiently separate hypermodified histone peptides through ultra-HPLC separation on
985 a bench top Orbitrap instrument. *Proteomics* *14*, 2212–2225.
- 986 Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications.
987 *Nature* *403*, 41–45.
- 988 Trelle, M.B., Salcedo-Amaya, A.M., Cohen, A.M., Stunnenberg, H.G., and Jensen,
989 O.N. (2009). Global histone analysis by mass spectrometry reveals a high content of
990 acetylated lysine residues in the malaria parasite *Plasmodium falciparum*. *J. Proteome*
991 *Res.* *8*, 3439–3450.
- 992 Tropberger, P., and Schneider, R. (2013). Scratching the (lateral) surface of chromatin
993 regulation by histone modifications. *Nat. Struct. Mol. Biol.* *20*, 657–661.
- 994 Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li,
995 G., Mittler, G., Liu, E.T., Bühler, M., et al. (2013). Regulation of Transcription
996 through Acetylation of H3K122 on the Lateral Surface of the Histone Octamer. *Cell*
997 *152*, 859–872.
- 998 Turner, B.M. (2000). Histone acetylation and an epigenetic code. *BioEssays News*
999 *Rev. Mol. Cell. Dev. Biol.* *22*, 836–845.
- 1000 Vanagas, L., Jeffers, V., Bogado, S.S., Dalmaso, M.C., Sullivan, W.J., and Angel,
1001 S.O. (2012). *Toxoplasma* histone acetylation remodelers as novel drug targets. *Expert*
1002 *Rev. Anti Infect. Ther.* *10*, 1189–1201.
- 1003 Voigt, P., Tee, W.-W., and Reinberg, D. (2013). A double take on bivalent promoters.
1004 *Genes Dev.* *27*, 1318–1338.
- 1005 Voss, T.S., Bozdech, Z., and Bártfai, R. (2014). Epigenetic memory takes center stage
1006 in the survival strategy of malaria parasites. *Curr. Opin. Microbiol.* *20*, 88–95.
- 1007 Williams, S.K., Truong, D., and Tyler, J.K. (2008). Acetylation in the globular core of
1008 histone H3 on lysine-56 promotes chromatin disassembly during transcriptional
1009 activation. *Proc. Natl. Acad. Sci.* *105*, 9000–9005.
- 1010 Xie, Z., Dai, J., Dai, L., Tan, M., Cheng, Z., Wu, Y., Boeke, J.D., and Zhao, Y.
1011 (2012). Lysine Succinylation and Lysine Malonylation in Histones. *Mol. Cell.*
1012 *Proteomics* *11*, 100–107.

1013 Xue, B., Jeffers, V., Sullivan, W.J., and Uversky, V.N. (2013). Protein intrinsic
1014 disorder in the acetylome of intracellular and extracellular *Toxoplasma gondii*. *Mol.*
1015 *Biosyst.* 9, 645.

1016 Yang, X., Yu, W., Shi, L., Sun, L., Liang, J., Yi, X., Li, Q., Zhang, Y., Yang, F., Han,
1017 X., et al. (2011). HAT4, a Golgi Apparatus-Anchored B-Type Histone
1018 Acetyltransferase, Acetylates Free Histone H4 and Facilitates Chromatin Assembly.
1019 *Mol. Cell* 44, 39–50.

1020

1021 **Figure legends**

1022

1023 **Figure 1. The residue K31 on the lateral surface of histone H4 is a novel PTM.**

1024 (a) The high resolution of MS/MS spectrum of H4K31ac peptide generated from
1025 histone H4. H4K31ac was identified using Mascot search engine in the
1026 DNIQGITK(ac)PAIR peptide. (b) Sequence alignment of the first 42 residues of
1027 histone H4 from the indicated organisms. Yellow boxes highlight the conserved
1028 residue H4K31. (c) Structural analysis of H4K31 modifications. Nucleosome core
1029 particle with key H3 and H4 lysine residues that are known to be modified shown
1030 in ball-and-sphere representation. The histone proteins of the nucleosome (PDB
1031 code: 3AFA) are colour coded as follows: H2A cyan, H2B grey, H3 orange and H4
1032 blue. The H4K31 residue, highlighted in red, is placed at the dyad axis and
1033 mediates key interactions to the DNA (in green). The bottom panel is rotated 90
1034 degrees around the molecular dyad axis. On the right, close-up of the interactions
1035 established by H4K31 with a water molecule (red sphere) and residue R35; and
1036 impact of the modifications: methylation, acetylation and succinylation
1037 (mimicked by mutant K31E). The mutant H4K31Q (PDB code: 3AZI) partially
1038 mimics lysine acetylation. (d) Immunofluorescence analysis of H4K31ac (in red) in
1039 both human foreskin fibroblast cells and parasites nuclei. DNA (top) was stained
1040 with Hoechst. Scale bar, 10 μm . (e) Confluent monolayers of HFF cells were infected
1041 with *T. gondii* parasites in the presence of FR235222 and DMSO as a control. IFAs
1042 were carried out with antibodies against specific histone 3 and 4 lysine residue
1043 acetylations as indicated. All modifications, showed specific and distinct localization
1044 in both parasite and host cell nuclei (in red). Scale bar, 20 μm . Data are representative
1045 of two independent experiments. (f) Immunoblots of native purified nucleosomes
1046 from *T. gondii* parasites treated with FR235222 or DMSO for 18 hours. Data are
1047 representative of two independent experiments.

1048

1049 **Figure 2 Immunofluorescence analysis of histone modifications in *P. falciparum***
1050 **and MEF cells. (a)** The blood stages of *P. falciparum* characterized by initial Ring,
1051 followed by mature trophozoite and segmented schizont stage. The three
1052 developmental stages represent the predominant asexual phase of the malaria parasite.
1053 **(b)** Immunofluorescence analysis of H4K31ac (in red) asexual stages following 12
1054 hours of treatment with DMSO (vehicle) or FR235222 HDACi. Parasite nuclear
1055 DNA was stained with Hoechst (blue). Scale bar, 10 μ m. Data are representative of
1056 four independent experiments. **(c)** and **(d)** Immunofluorescence analysis of H4K31ac
1057 in mouse embryonic fibroblasts. DNA was stained with DAPI (blue); the bright foci
1058 mark pericentromeric heterochromatin. The signal for H4K31ac along with H3K27ac,
1059 H3K4ac or H3K9ac are enriched in euchromatic regions as shown in the merge. The
1060 mark is excluded from the DAPI dense foci that are associated with H3K9me3 and
1061 H4K20me3. Scale bar, 10 μ m. Data are representative of three independent
1062 experiments.

1063

1064 **Figure 3 Chemical and genetic inactivation of TgHDAC3 promotes H4K31ac**
1065 **accumulation in *Toxoplasma gondii* nuclei. (a)** Immunofluorescence analysis of
1066 H4K31ac (in red) in HFF cells infected with parasites expressing a HAFlag (HF)-
1067 tagged copy of the bradyzoite gene *ENO1* treated for 18 hours with vehicle (DMSO)
1068 or individual HDAC inhibitors, including short-chain fatty acids (sodium butyrate),
1069 cyclic tetrapeptides and hydroxamic acids classes. Halofuginone (HF), a non-HDACi
1070 anticoccidial compound was used as a relevant control. ENO1 expression was
1071 detected by IFA in parasite nuclei (anti-HA, in green). Host-cell and parasite nuclei
1072 were stained with Hoechst. Scale bar, 20 μ m. **(b)** Representative micrographs
1073 showing intracellular tachyzoites in which the *TgHDAC3* gene was disrupted by
1074 transient transfection of CRISPR/Cas9. The efficiency of TgHDAC3 disruption in
1075 Cas9-expressing parasites was monitored by the anti-TgHDAC3 staining (in pink) and
1076 cas9-GFP expression (in green). The levels of H4K31ac (in red) were monitored in
1077 *TgHDAC3*-disrupted parasites (GFP positive) and compared to untransfected parasites
1078 (GFP negative). Scale bar, 10 μ m. **(c)** Levels of H4K31ac (in red) were monitored in
1079 *TgHDAC* knockout parasites. Scale bar, 10 μ m. Data are representative of four
1080 independent experiments.

1081

1082 **Figure 4 TgGCN5b acetylates H4K31 in *Toxoplasma gondii*** (a) Levels of
1083 H4K31ac (in red) were monitored in both host cell and parasite nuclei following
1084 CRISPR/Cas9-mediated disruption of individual *T. gondii* HAT enzymes. Transfected
1085 vacuole in which H4K31 acetylation was impaired are indicated by a white arrow.
1086 Scale bar, 10 μ m. (b) Levels of H3K14 acetylation (in red) were monitored in
1087 *TgGCN5b* knockout parasites. Scale bar, 10 μ m. Data are representative of three
1088 independent experiments.

1089

1090 **Figure 5 Distribution of H4K31me1 in host cell and parasites nuclei** (a)
1091 Immunofluorescence analysis of H4K31me1 (in green) in intracellular parasite nuclei.
1092 DNA was stained with Hoechst (blue). Scale bar, 10 μ m. (b) The localization of
1093 H4K31me1 (in red), H3K27ac (in green) and H3S10P (in green) were tested by
1094 immunofluorescence in prophase, metaphase, anaphase and telophase MEFs. DNA
1095 was stained with Hoechst (blue). Images are representative of at least three
1096 independent experiments. Scale bar, 10 μ m. (c) Immunofluorescence analysis of
1097 H4K31me1 (in red) or H3K9me3 in asexual stages of Pf-3D7. Scale bar, 5 μ m. Data
1098 are representative of three independent experiments.

1099

1100 **Figure 6 Genome-wide analysis of H4K31ac and H4K31me1 chromatin**
1101 **occupancy in *Toxoplasma gondii*.** (a) A genome browser (IGB) snapshot showing
1102 normalized reads for different histone marks across *T. gondii* chromosome 1b reveals
1103 peak-like distribution of H4K31ac and H4K31me1 ChIP-seq enrichments. The y-axis
1104 depicts read density. Genes are depicted above the profiles in black. (b) A zoomed-in
1105 view of Chr. 1b region (yellow box in a) showing the distribution of the
1106 aforementioned PTMs. (c) Distribution of PTMs occupied regions relative to the *T.*
1107 *gondii* reference genome annotation. (d) Correlation of H4K31 modifications
1108 enrichment with other marks. The average signal profiles of each histone modification
1109 was plotted over a -2-kb to +10-kb region with respect to *T. gondii* genes ATG. The y-
1110 axis shows the average tag count of the enrichment. The vertical dashed line
1111 indicates the position of the ATG.

1112

1113 **Figure 7 The enrichment of H4K31ac and H4K31me1 at transcribed genes**
1114 **correlates with gene expression levels in *Toxoplasma gondii*.** (a) and (b) IGB
1115 screenshots of *T. gondii* Chr. VIII genomic regions showing reads for various histone
1116 marks as well as RNA-seq data (in black). (c) A zoomed-in view of *T. gondii* *GRA1-*
1117 *MAG1* locus. The y-axis depicts read density. (d) Boxplot showing the normalized
1118 expression distribution of *T. gondii* genes in tachyzoite stage subdivided into four
1119 ranges of expression (cluster Q1 to Q4). Genomewide H4K31ac (e) and H4K31me1
1120 (f) occupancy profiles at peri-ATG regions are plotted for the gene groups ranked
1121 by their mRNA levels. The y-axis shows the average tag count of the enrichment. The
1122 vertical dashed line indicates the position of the ATG.

1123

1124 **Figure 8 H4K31me1 marks long non-coding RNA but not pericentromeric**
1125 **heterochromatin in *Toxoplasma gondii*.** (a) IGB screenshot of *T. gondii* Chr. Ib
1126 genomic region showing reads for various histone marks as well as RNA-seq data (in
1127 black). The y-axis depicts read density. (b) A zoomed-in view of *T. gondii* Chr. Ib
1128 peri-centromeric region. (c) IGB screenshot of *T. gondii* Chr. III genomic region
1129 showing reads for various histone marks as well as RNA-seq data. A predicted
1130 lncRNA of 72-kb is indicated in pink.

1131

1132 **Figure 9 Genome-wide analysis of H4K31ac and H4K31me1 chromatin**
1133 **occupancy in *Plasmodium falciparum*.** (a) Chromosomal projection of H4K31ac,
1134 H3K9me3 and H3K4me3 occupancies in *P. falciparum*. The full set of chromosomes
1135 is represented as the circular plot. (b) Zoomed-in view of PTMs and HP1 enrichment
1136 along the *gapdh* locus. (c) Genomewide H4K31ac and H3K4me3 occupancy profiles
1137 at peri-TSS (Transcription Start Sites) and –TTS (Transcription Termination Sites)
1138 regions were plotted. (d) Chromosomal projection of H4K31me1, H3K9me3 and HP1
1139 occupancies in *P. falciparum*. The full set of chromosomes is represented as the
1140 circular plot where CenH3 locations (black arrow) and var genes (in green) are
1141 indicated. (e) Genomewide H4K31me1 occupancy profiles at peri-TSS (Transcription
1142 Start Sites) and –TTS (Transcription Termination Sites) regions were plotted. (f) IGB
1143 view of a section of chromosome 10 showing enrichment of H4K31me1, H3K9me3
1144 and HP1 at *var* gene.

1145

1146 **Figure 10 H4K31me1 singularly marks peri-centromeric heterochromatin in**
1147 *Plasmodium falciparum*. (a) (top) Chromosome-wide coverage plot of histone
1148 modifications and PfHP1 on *P. falciparum* Chr. 7. CenH3 was mapped according to
1149 [Hoeijmakers et al., 2012](#) and *var* genes were indicated. (bottom) Zoomed-in views of
1150 *var* genes-containing internal locus (left panel) and centromeric (right panel) from *P.*
1151 *falciparum* Chr. 7. (b) Genomic organization and nuclear position of *var* genes and
1152 telomere-associated repeat elements (TAREs) in *P. falciparum*. (c) IGB screenshots
1153 of *P. falciparum* sub-telomeric regions of chromosome 10. Rifin and *var* genes as well
1154 as TAREs are highlighted.

1155

1156 **Figure 6 – figure supplement 1 ChIP-seq enrichments between biological**
1157 **replicates are highly correlated.** Scatterplot comparing the enrichment difference of
1158 H4K31ac (a) or H4K31me1 (b) measured in the two independent replicate
1159 experiments. The x- and y-axis show the average tag count of the enrichment. IGB
1160 screenshot of *T. gondii* Chr. X (c) and Ia (d) genomic regions showing reads for
1161 H4K31ac (replicates R1 and R2) and H4K31me1 (replicates R1 and R2).

1162

1163 **Figure 7 – figure supplement 1 PTMs distribution and gene expression in**
1164 *Toxoplasma gondii*. Genomewide PTM occupancy profiles at peri-ATG regions are
1165 plotted for the gene groups ranked by their mRNA levels (a). H3K14ac (b),
1166 H3K4me3 (c), H3K4me1 (d), H4K20me3 (e) and H3K9me3 (f) are shown. The y-axis
1167 shows the average tag count of the enrichment. The vertical dashed line indicates
1168 the position of the ATG.

1169

1170 **Figure 8 – figure supplement 1 H4K31me1 explicitly predicts unannotated**
1171 **genes.** IGB screenshot of *T. gondii* Chr. VIIb (a) and VIII (b) genomic regions
1172 showing reads for various histone marks as well as RNA-seq data. Predicted genes are
1173 indicated in pink along with their putative translated sequence.

1174

1175 **Figure 8 – figure supplement 2 *Toxoplasma gondii* peri-centromeric regions (a)**
1176 IGB screenshots of *T. gondii* peri-centromeric region of chromosomes Ia, II, III, VI,
1177 VIIa, VIII, IX and X.

1178

1179 **Figure 8 – figure supplement 3 H4K31me1 marks long non-coding RNAs.** IGB
1180 screenshot of *T. gondii* Chr. VIIa (a) and XI (b) genomic regions showing reads for
1181 various histone marks as well as RNA-seq data. Predicted lncRNAs of (a) 70-kb and
1182 (b) 22-kb are indicated in pink.

1183

1184 **Figure 9 – figure supplement 1 Correlation matrix between *Plasmodium***
1185 ***falciparum* ChIP-seq experiments.** Heatmap displaying (a) Pearson and (b)
1186 Spearman rank correlations between all pairwise comparisons for all *P.*
1187 *falciparum* ChIPs. Spearman correlations were calculated using the normalized
1188 read depth across the entire set of binding sites identified for all ChIP-seq
1189 experiments.

1190

1191 **Figure 9 – figure supplement 2 Chromosomal projection of H4K31ac, H3K9me3**
1192 **and H3K4me3 occupancies in *P. falciparum*.** The full set of chromosomes is
1193 represented as the circular plot.

1194

1195 **Figure 9 – figure supplement 3 Chromosomal projection of H4K31me1,**
1196 **H3K9me3 and HP1 occupancies in *P. falciparum*.** The full set of chromosomes is
1197 represented as the circular plot, where centromeric regions are marked by a black
1198 arrow.

1199

1200 **Figure 10 – figure supplement 1 H4K31me1 marks pericentromeric chromatin in**
1201 ***Plasmodium falciparum*.** Zoomed-in views of centromeric and peri-centromeric
1202 chromatin from all *P. falciparum* chromosomes.

1203

1204

1205

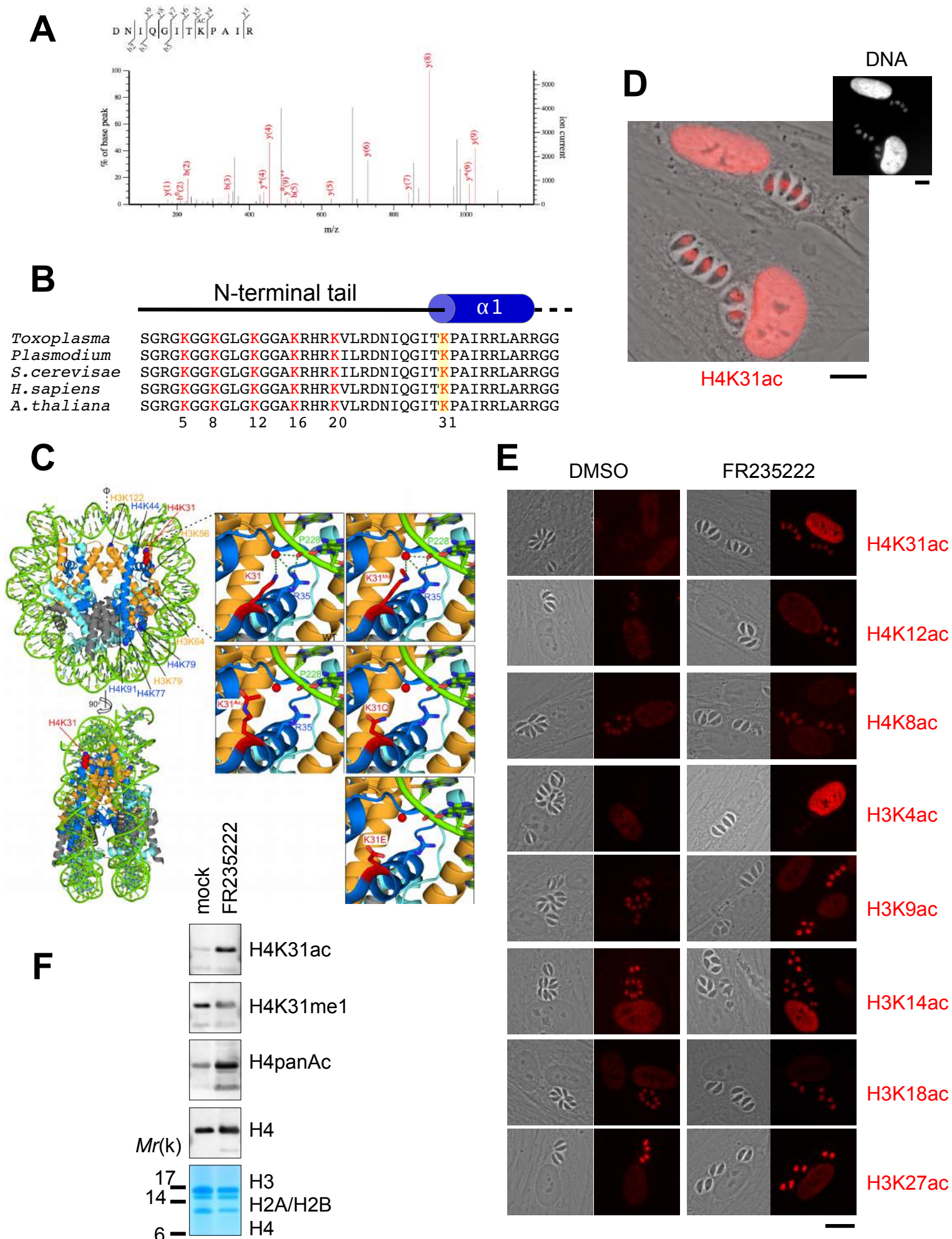


Figure 1

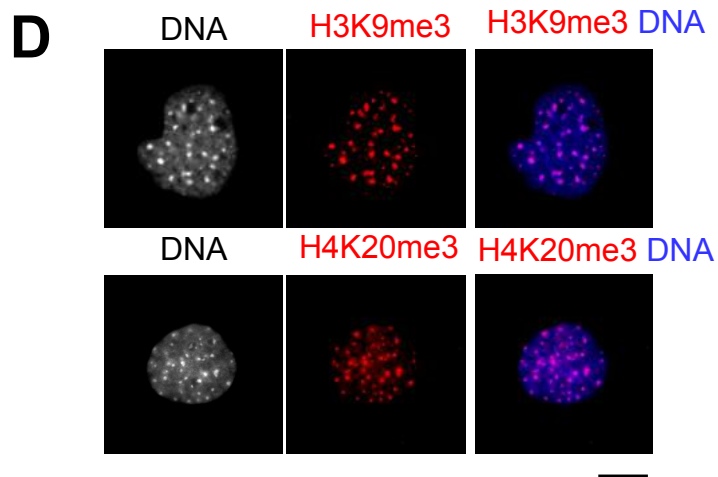
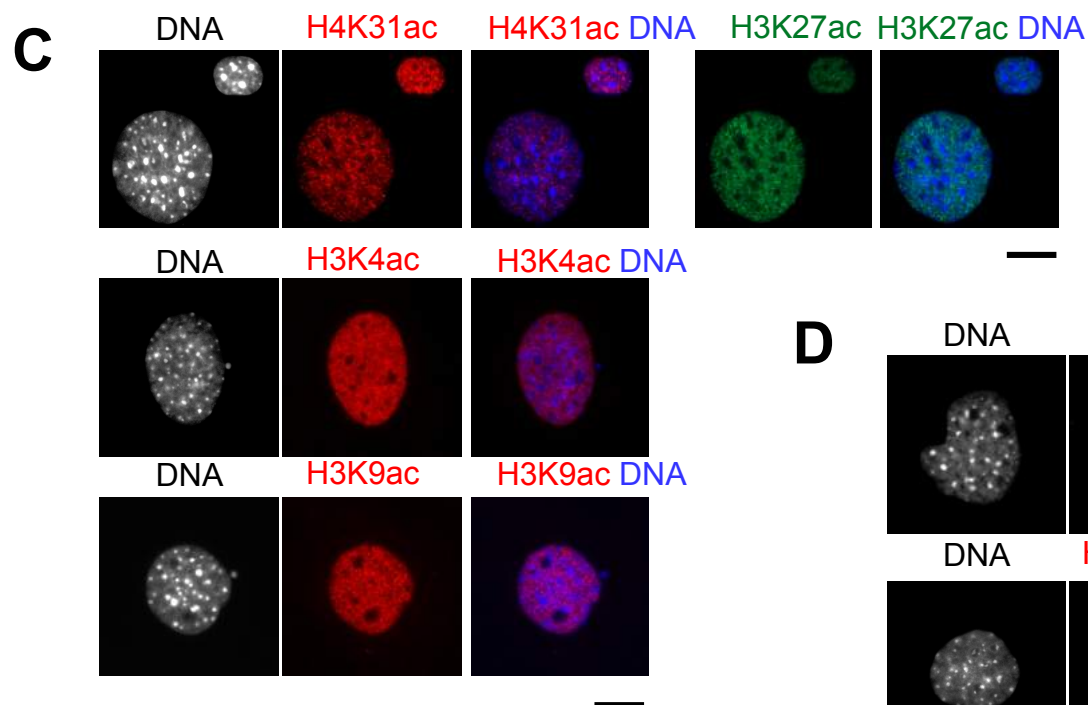
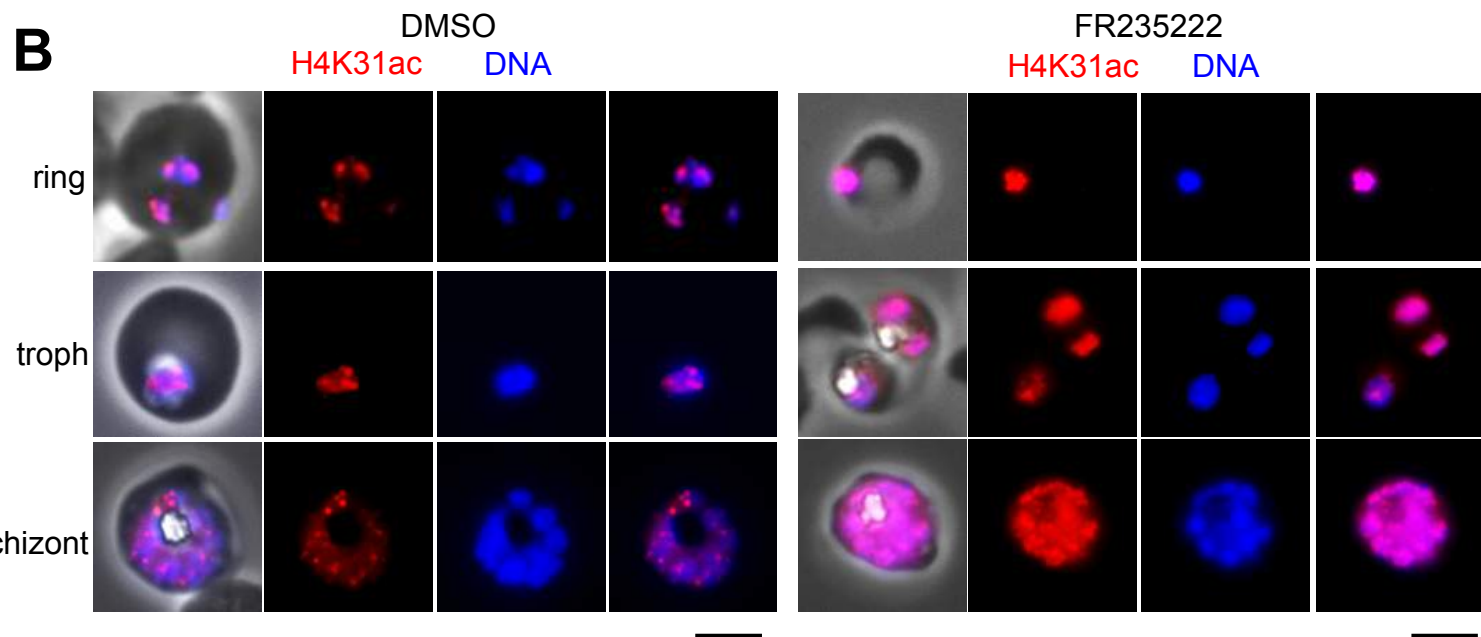
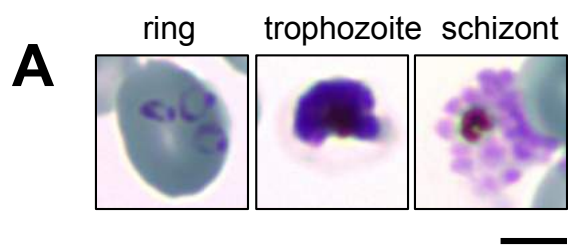


Figure 2

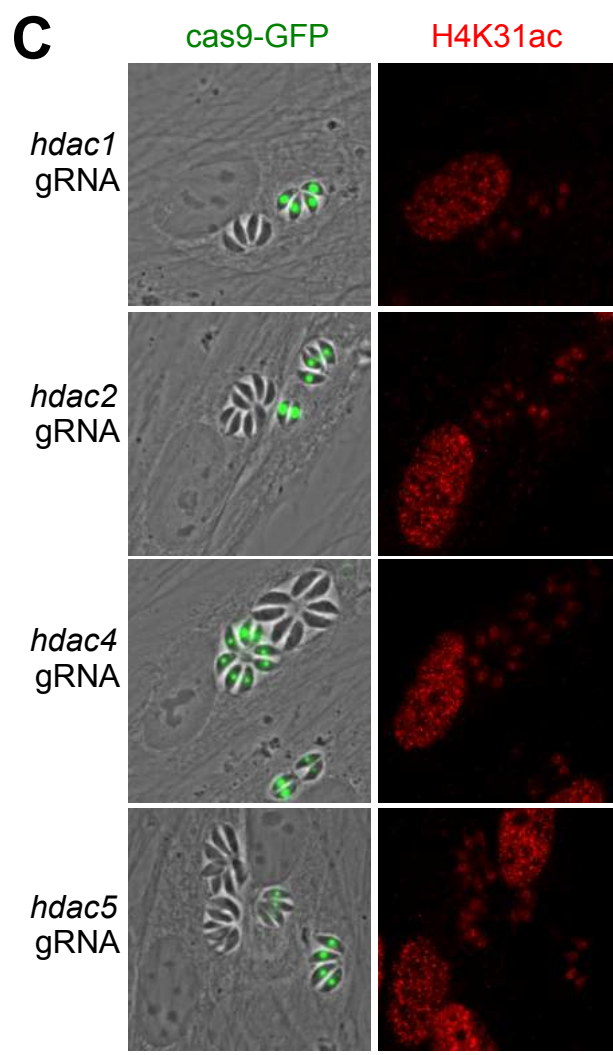
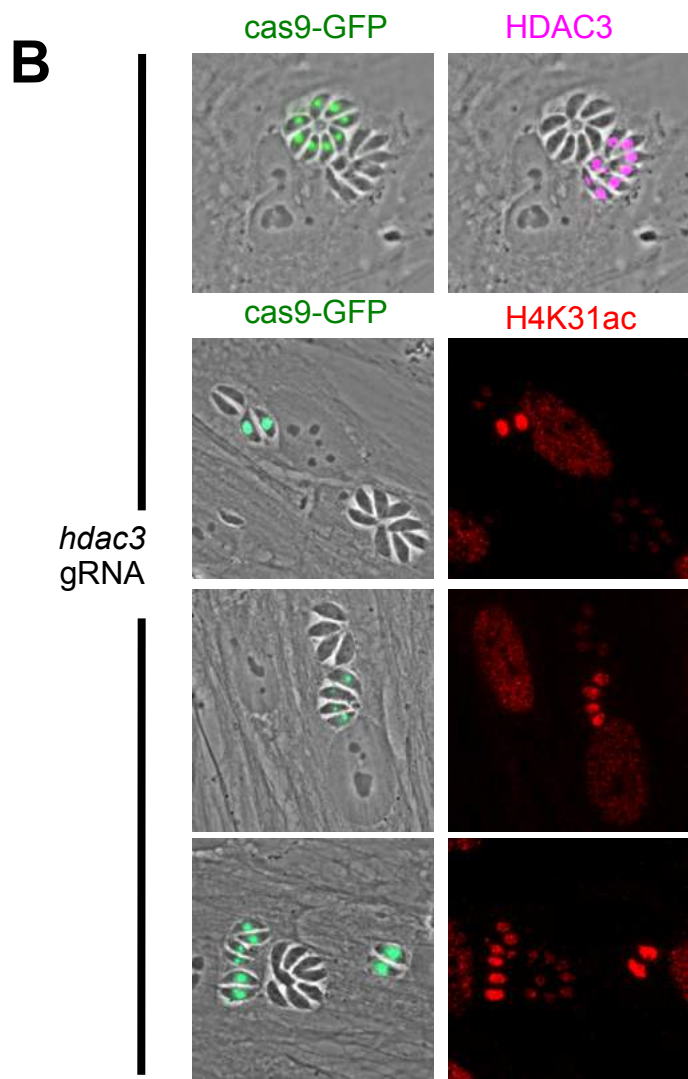
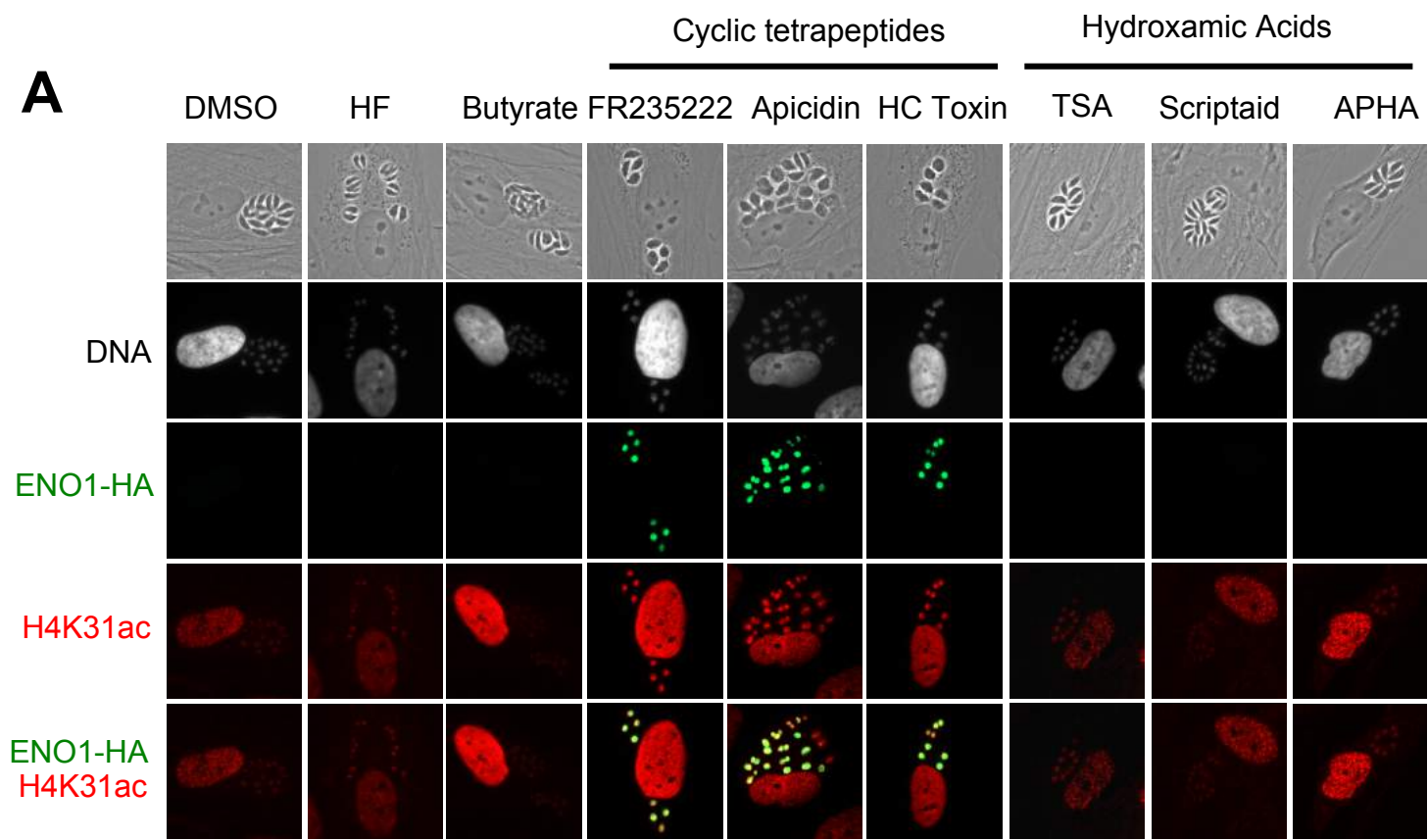
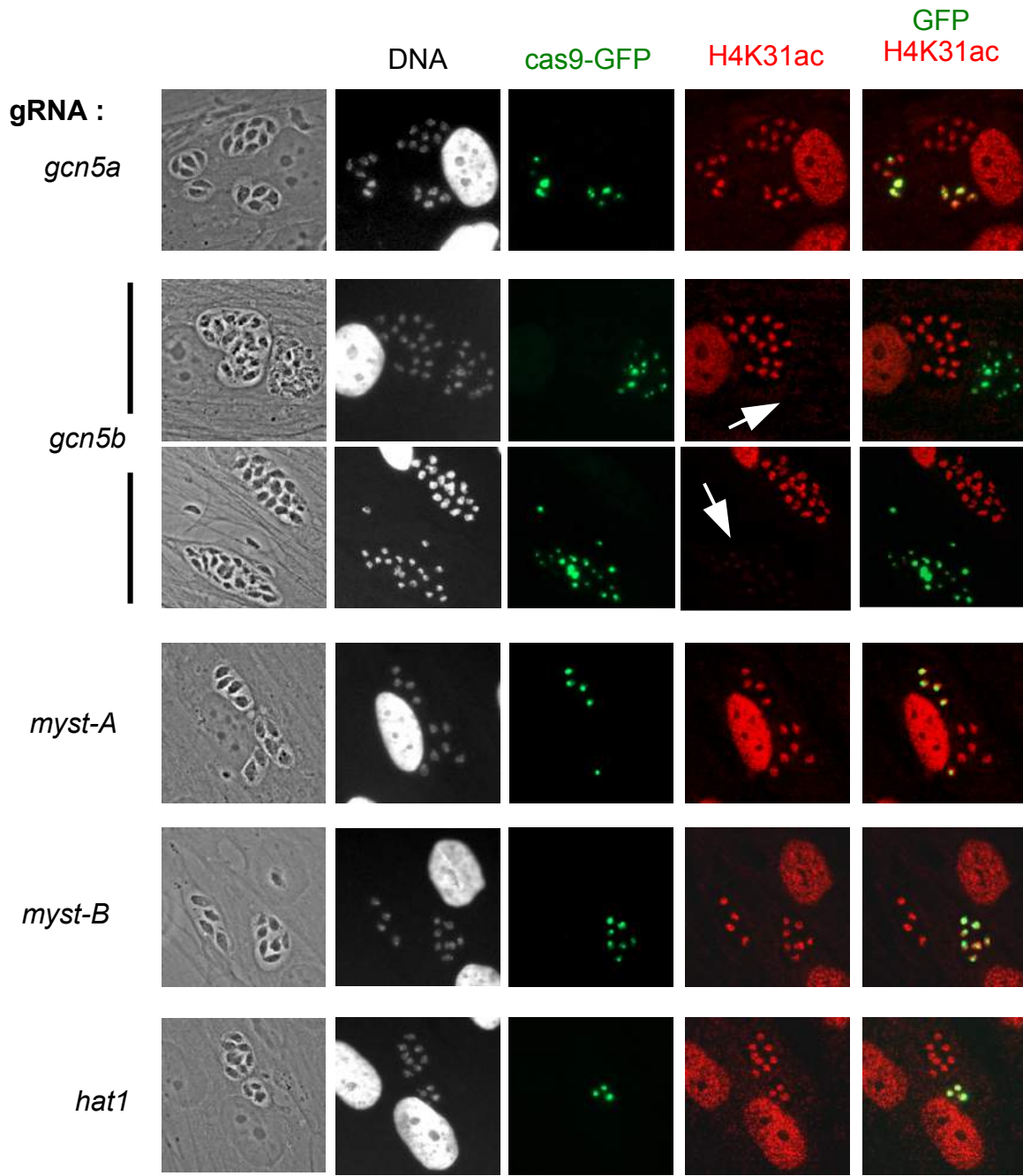
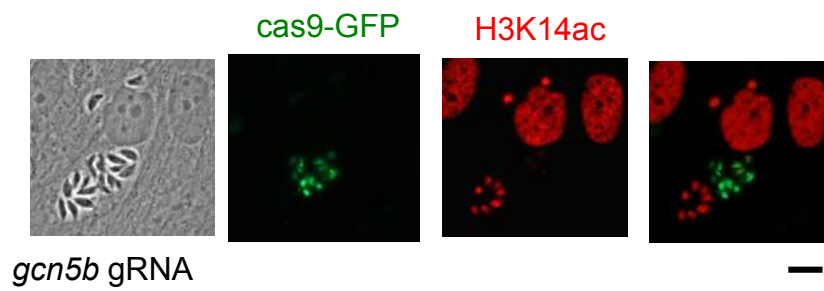


Figure 3

A**B****Figure 4**

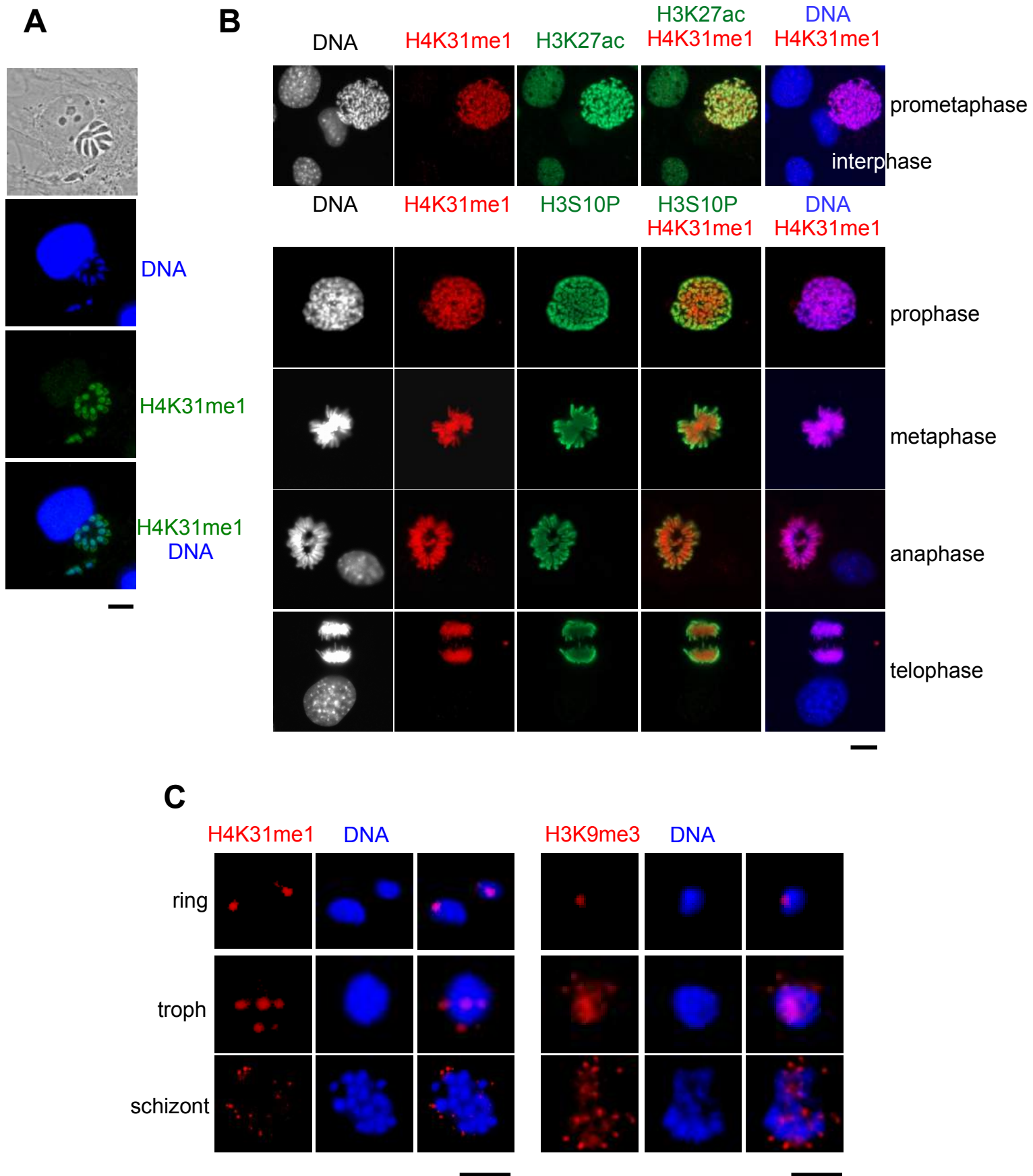


Figure 5

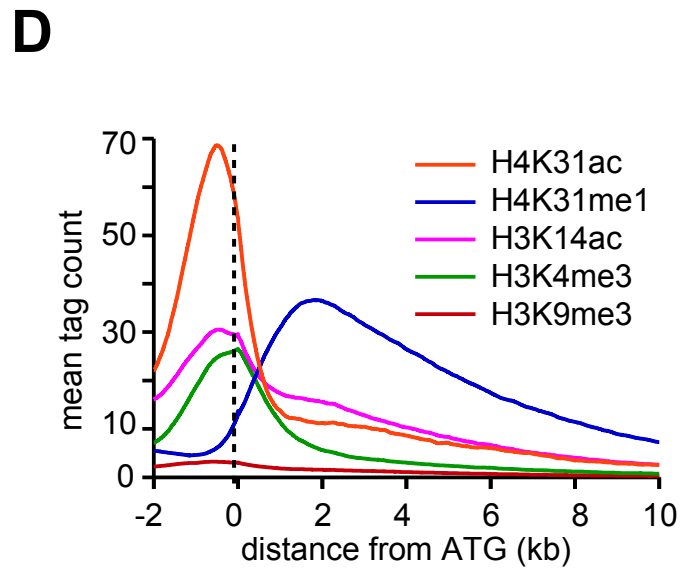
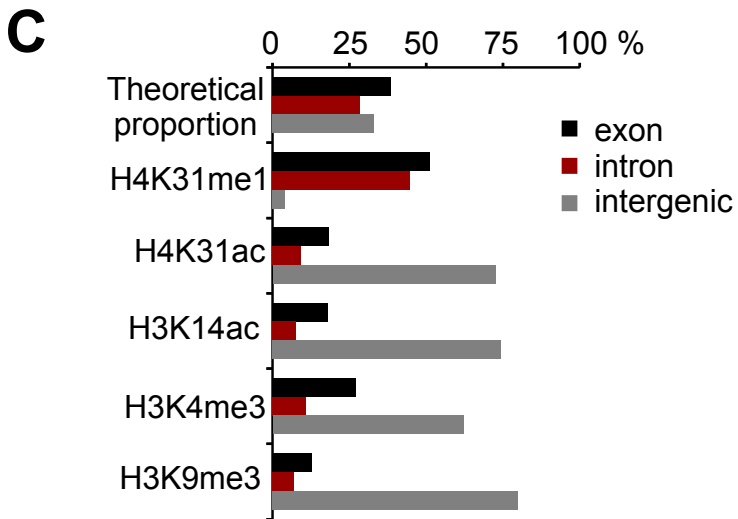
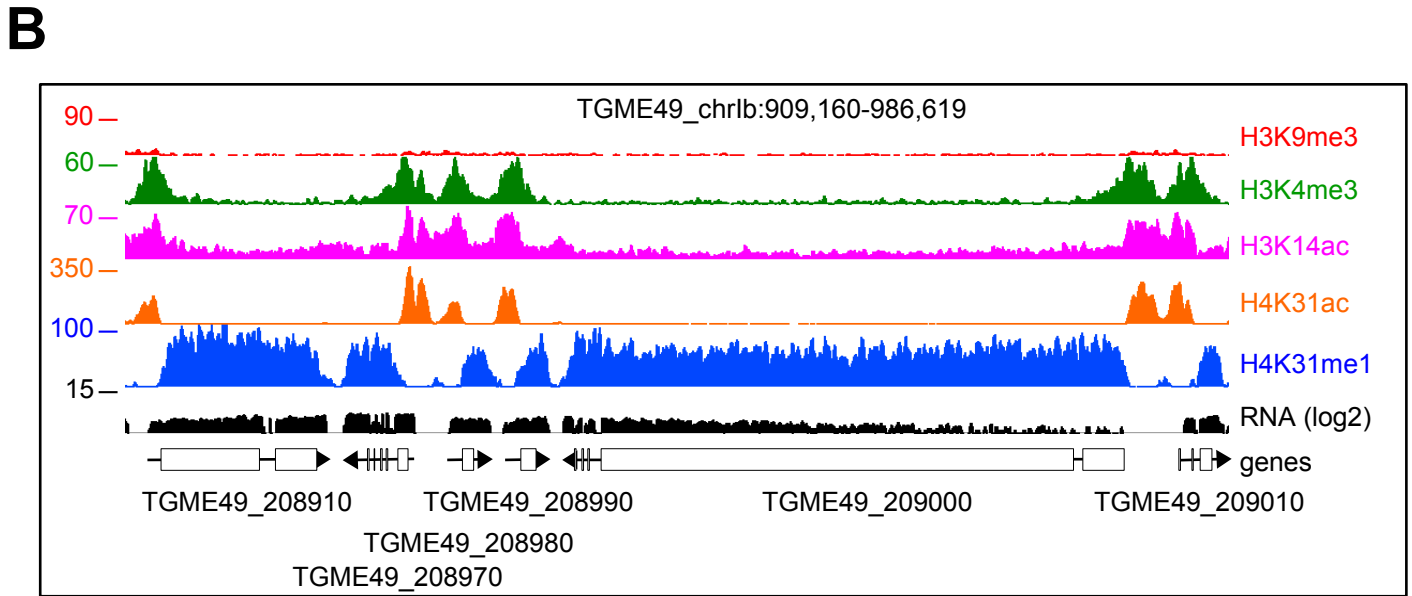
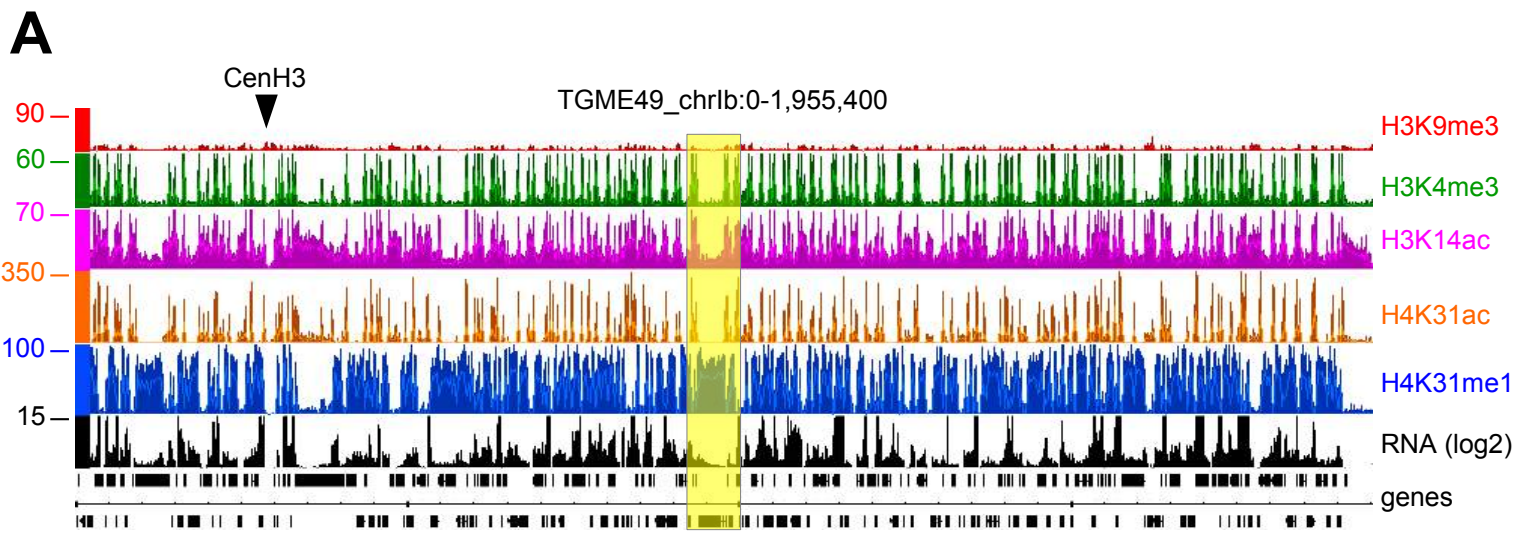


Figure 6

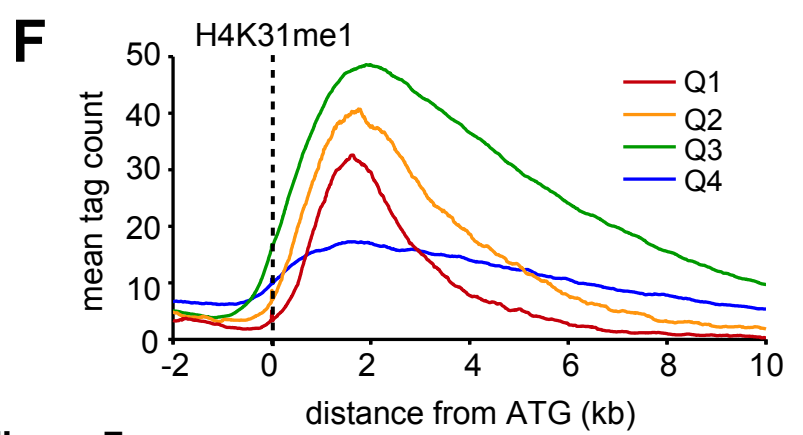
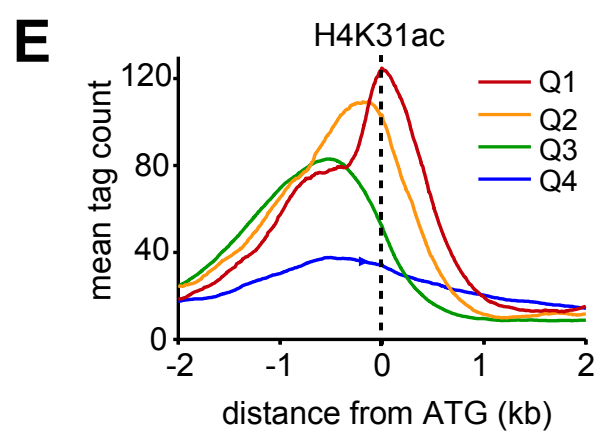
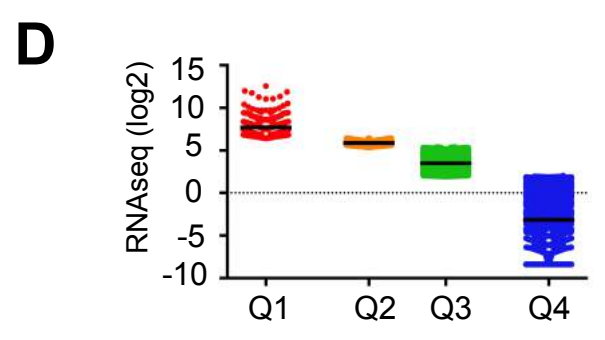
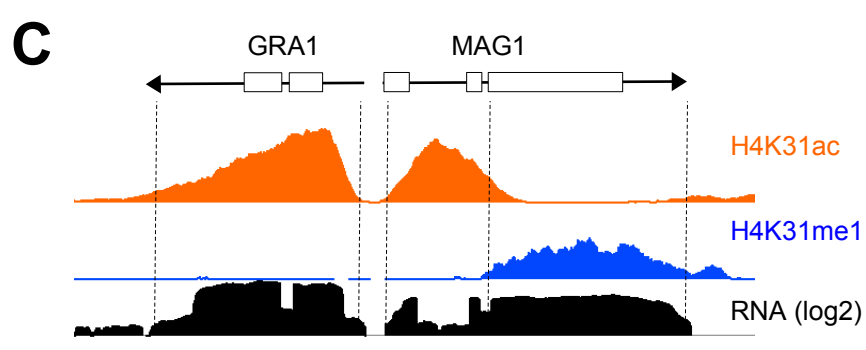
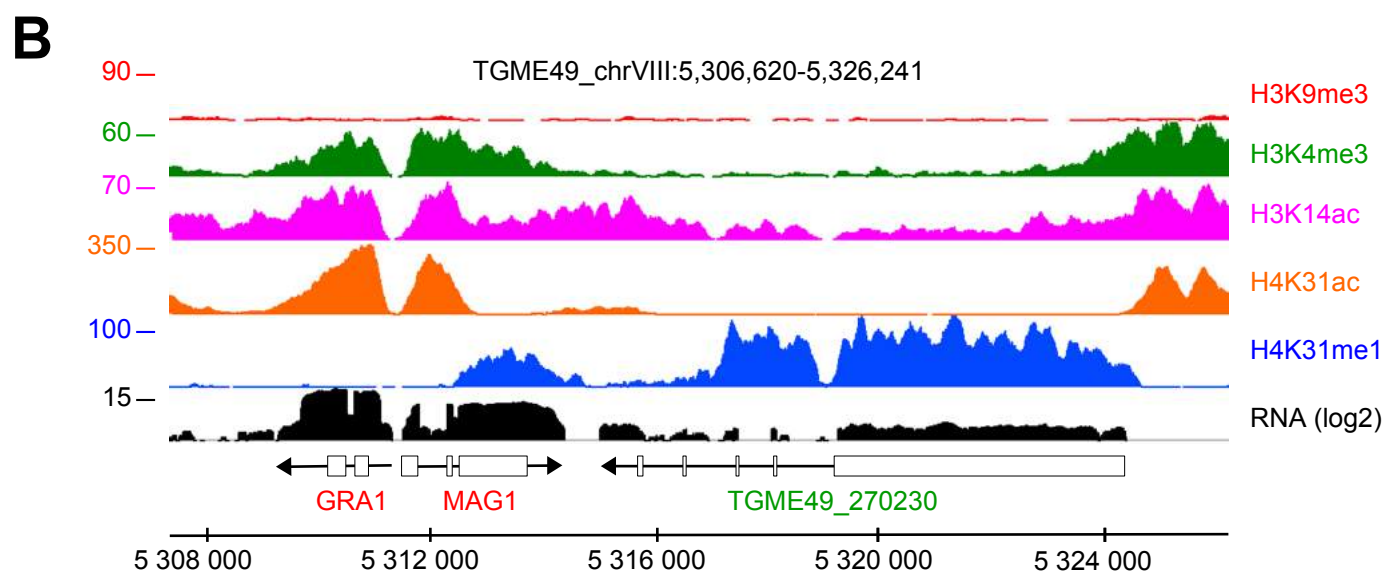
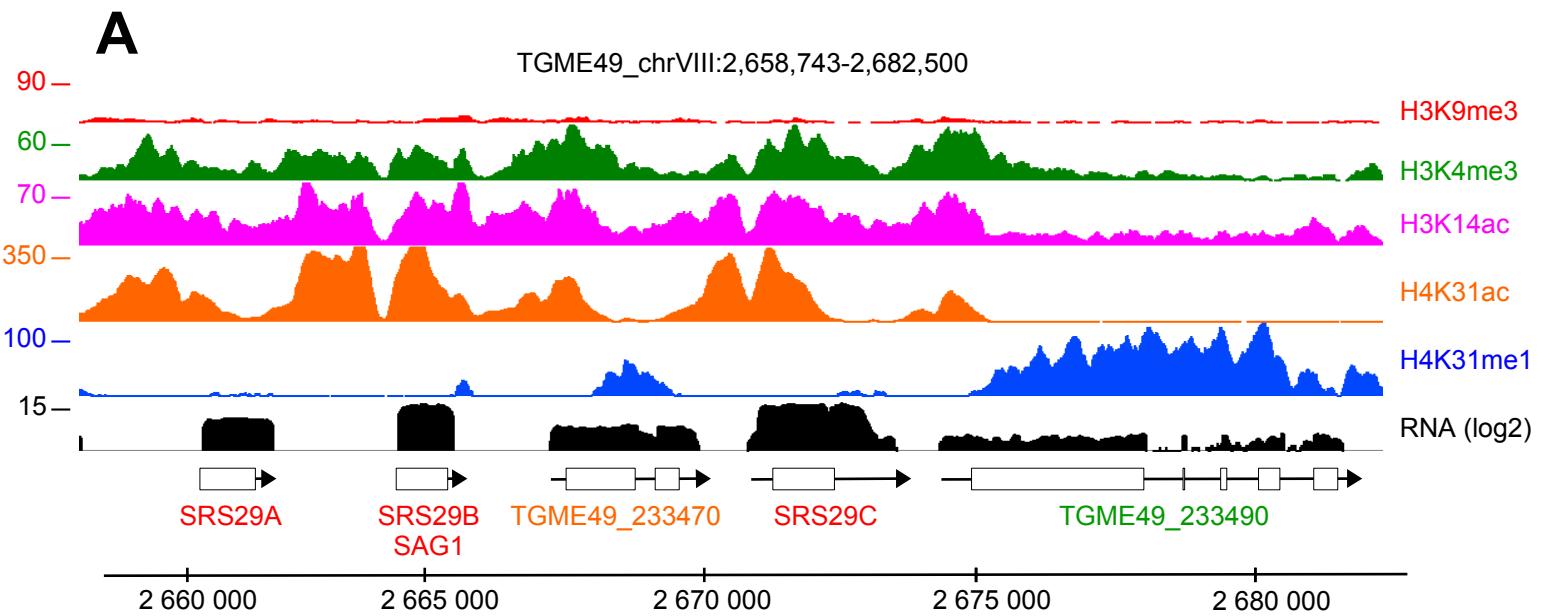


Figure 7

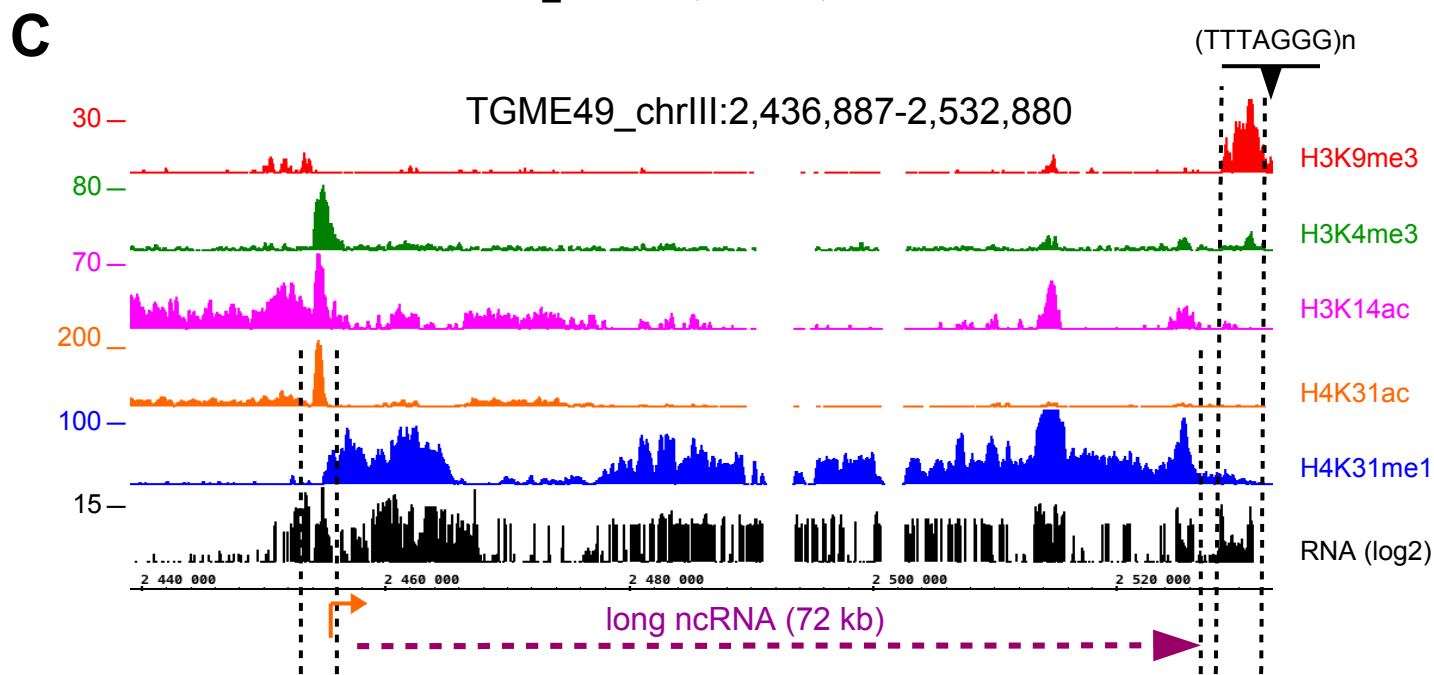
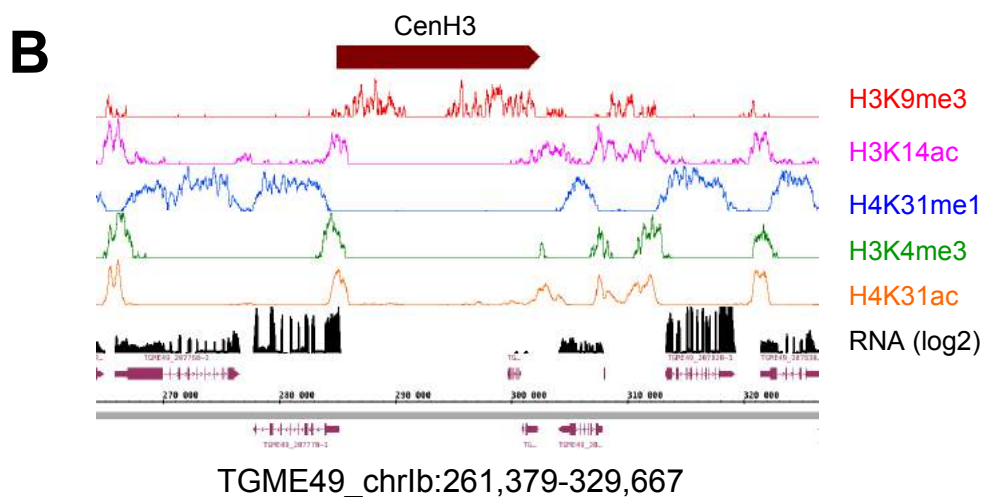
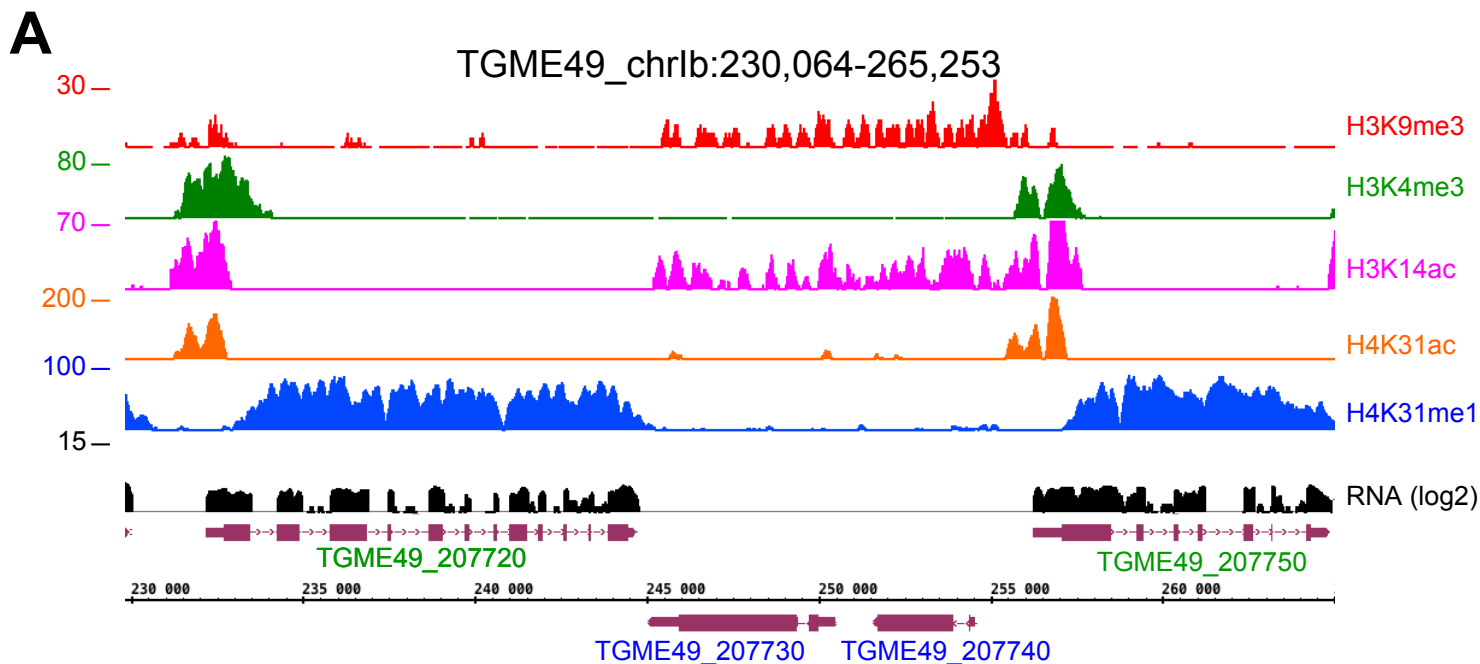


Figure 8

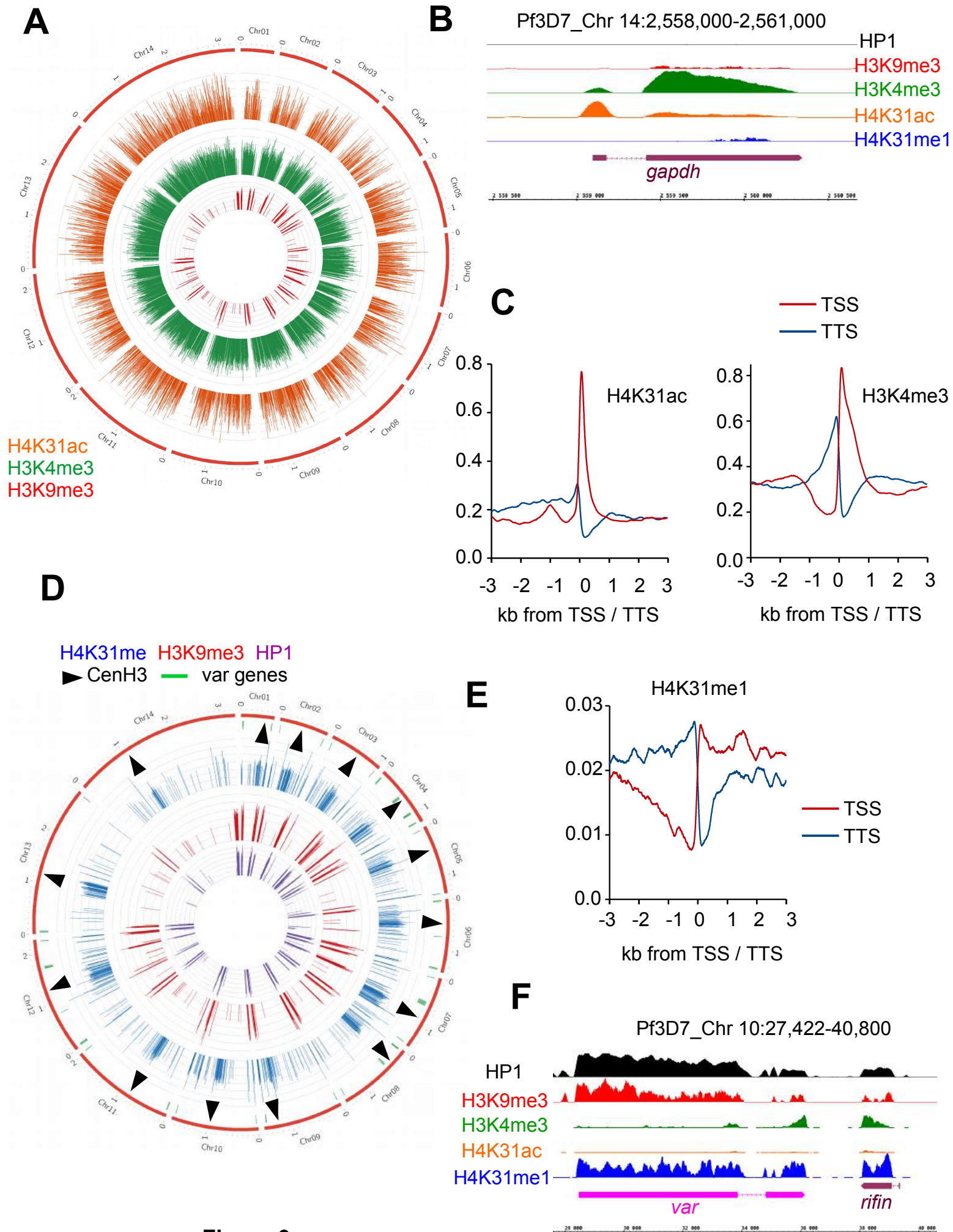


Figure 9

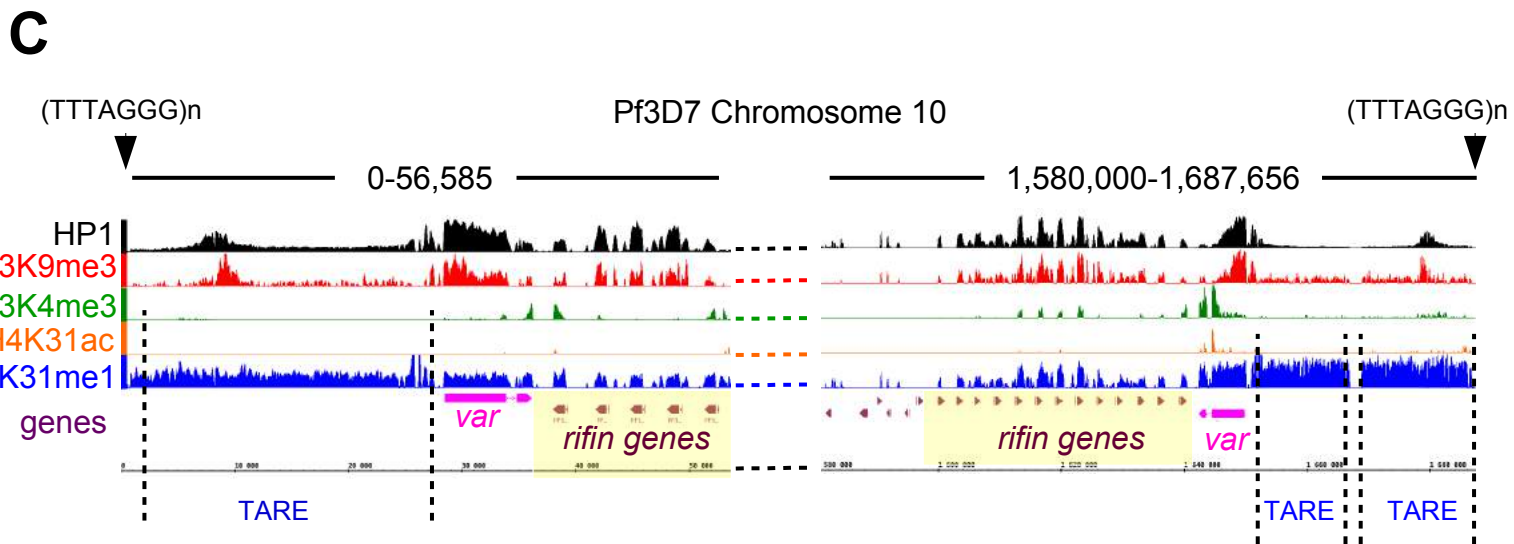
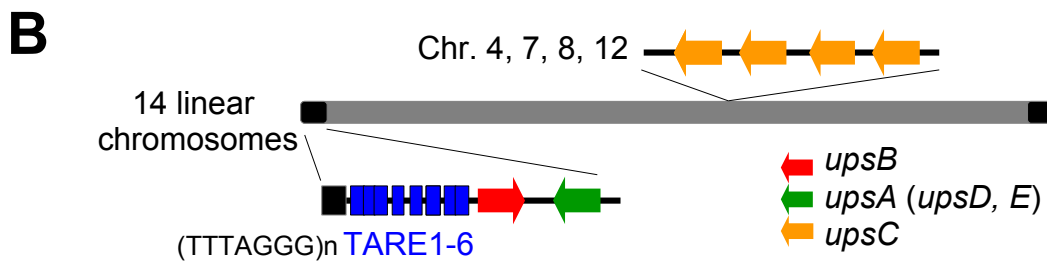
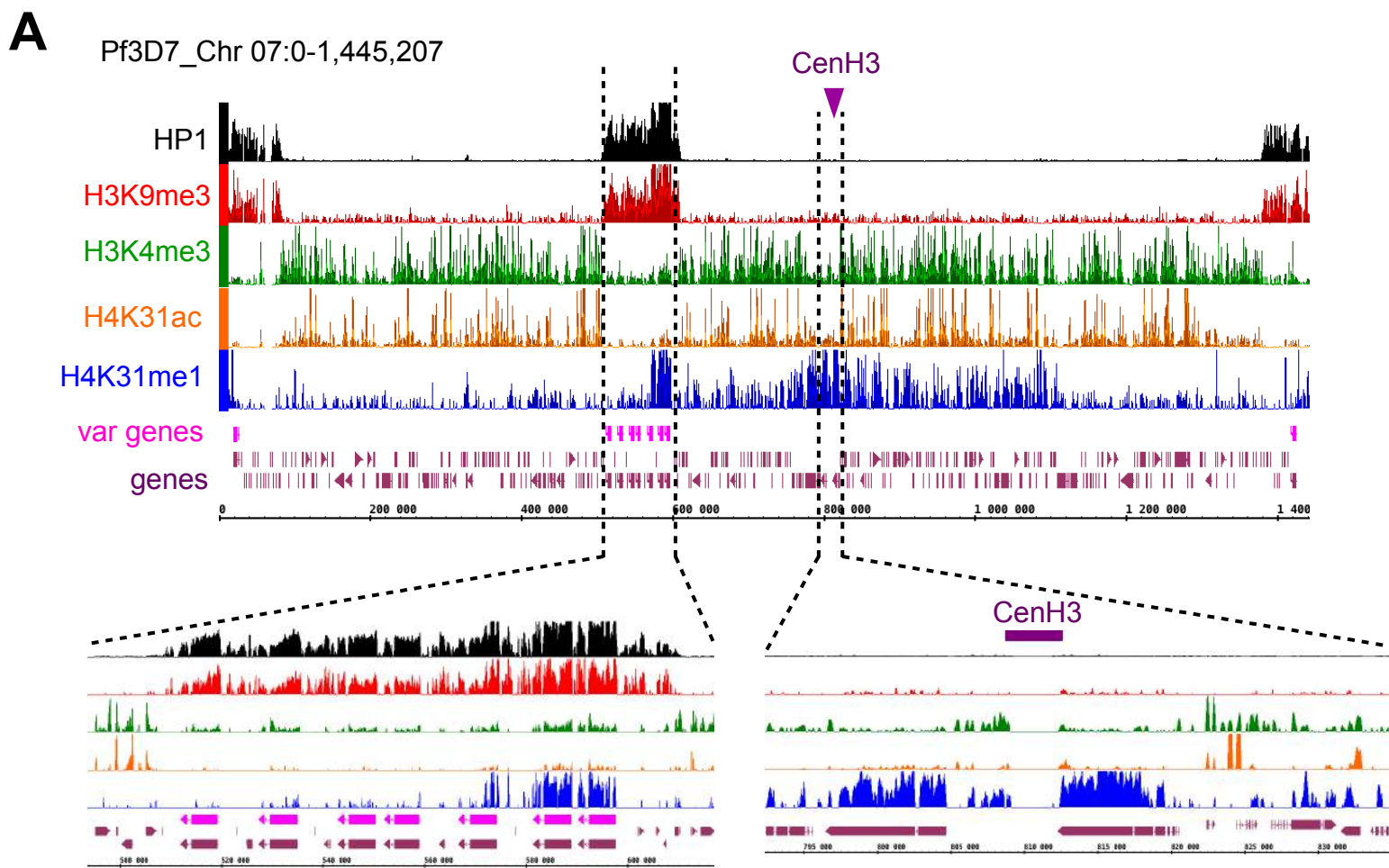
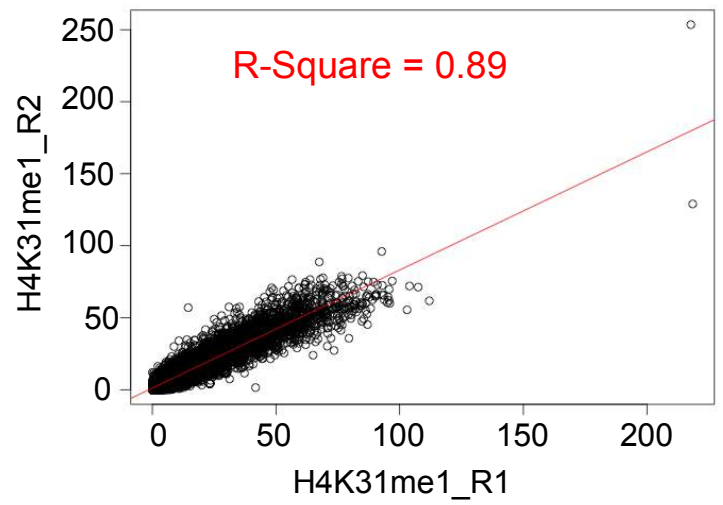
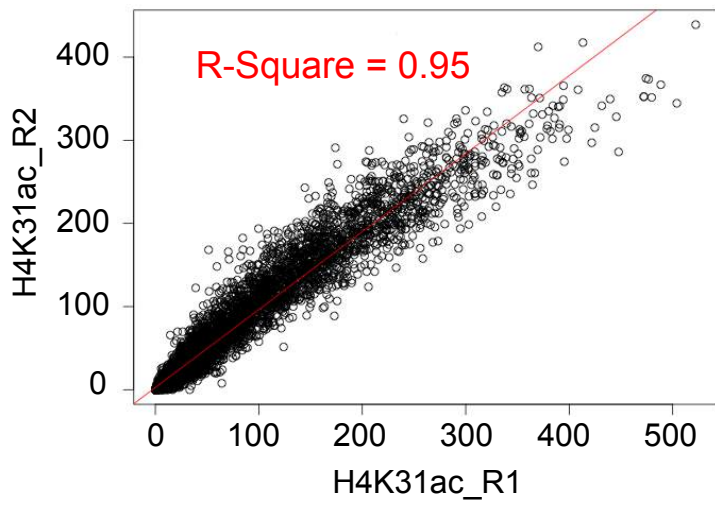
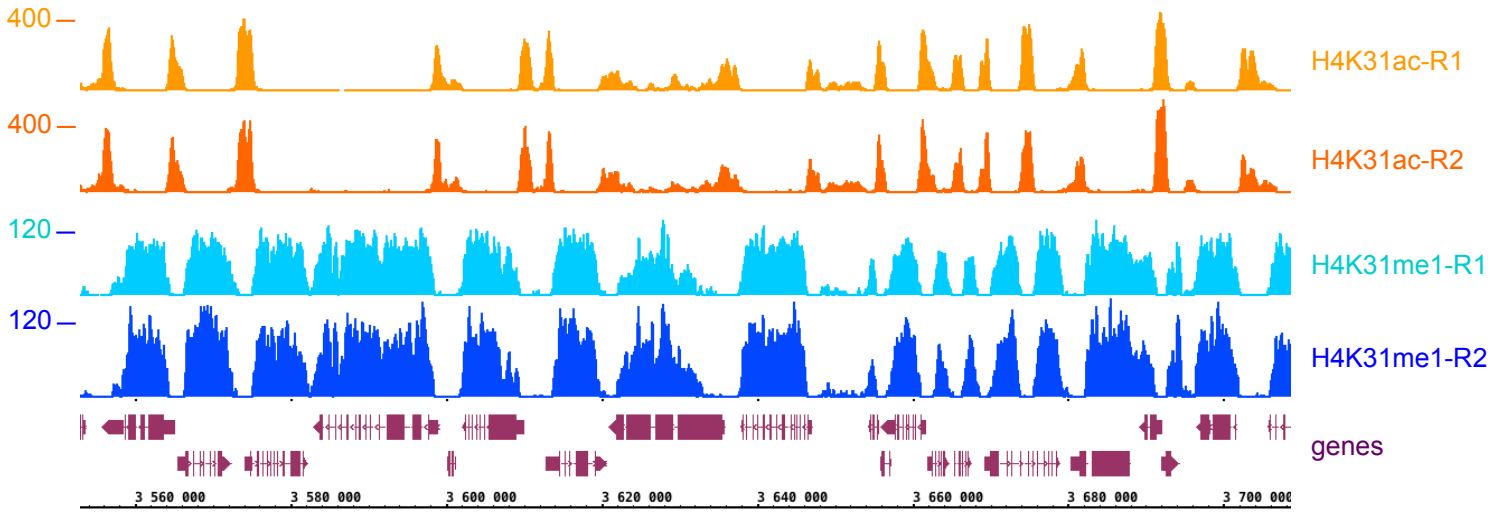


Figure 10



TGME49 chrX:3.545.280-3.708.645



TGME49_chrla:1,613,215-1,841,673

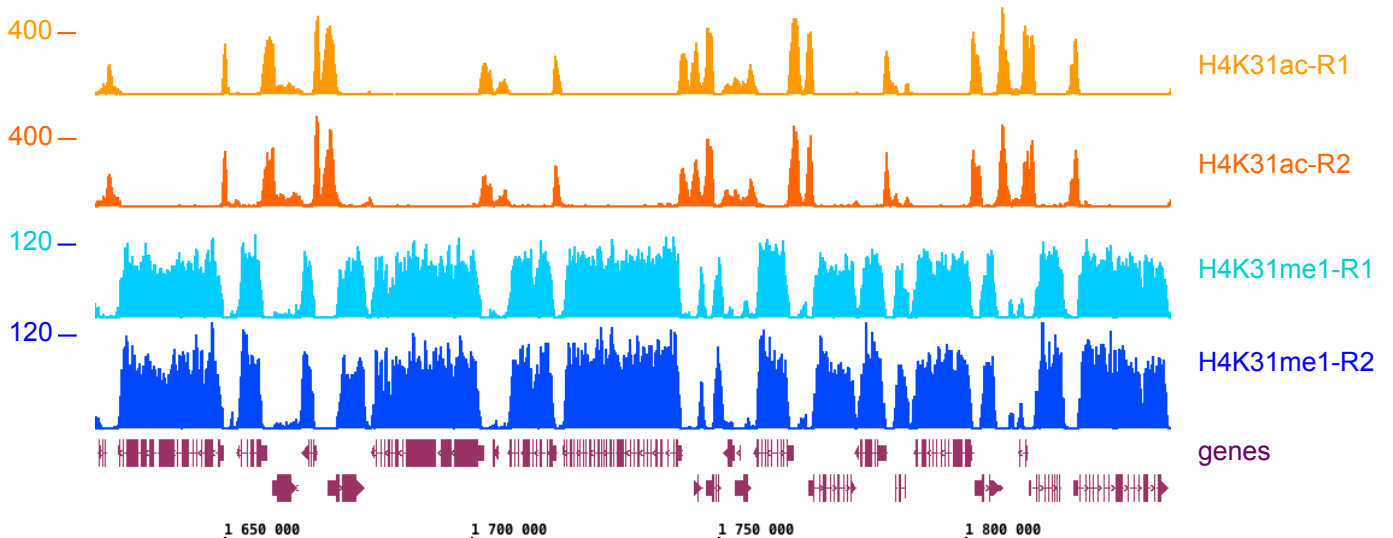


Figure 6 – figure supplement 1

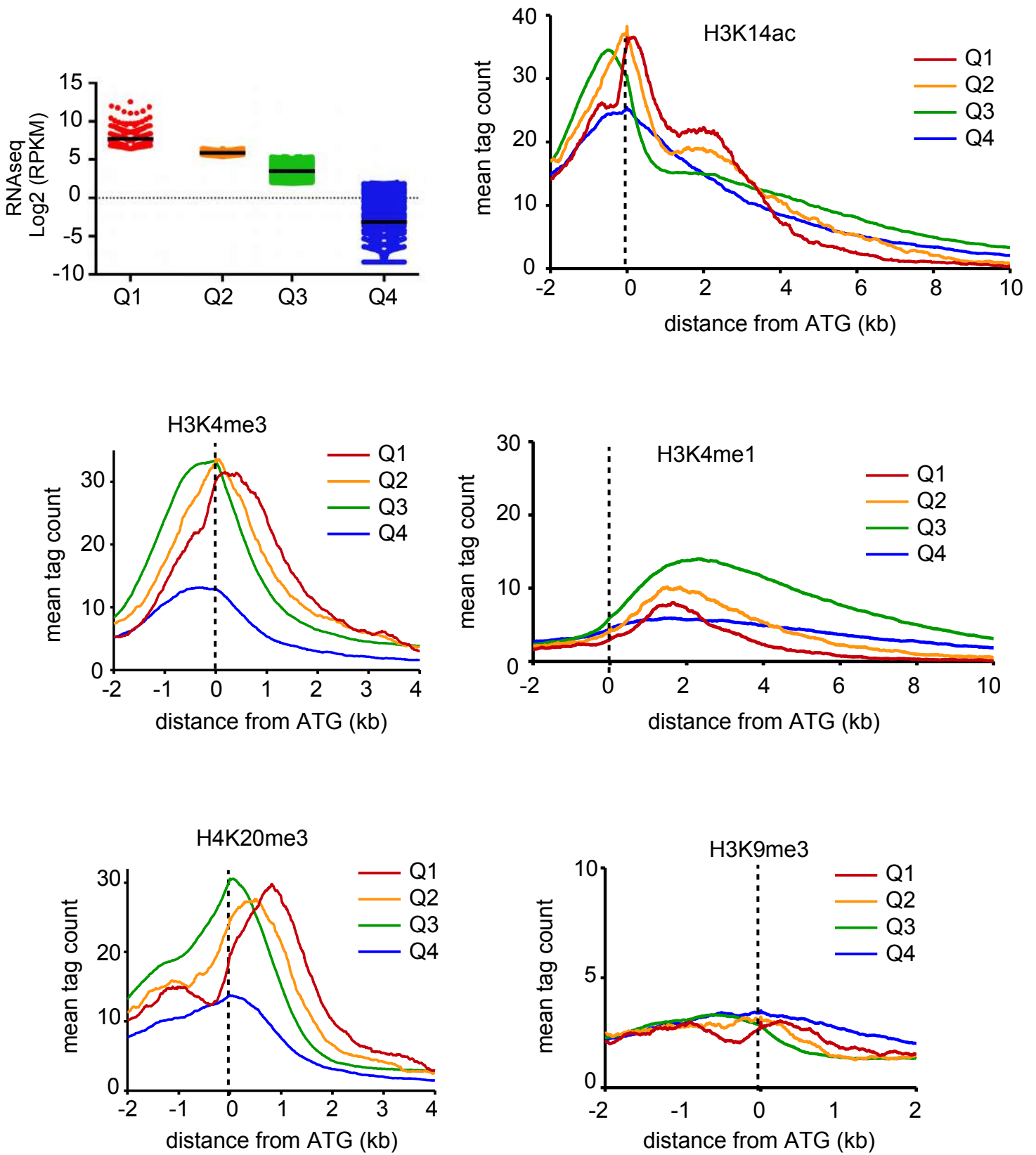
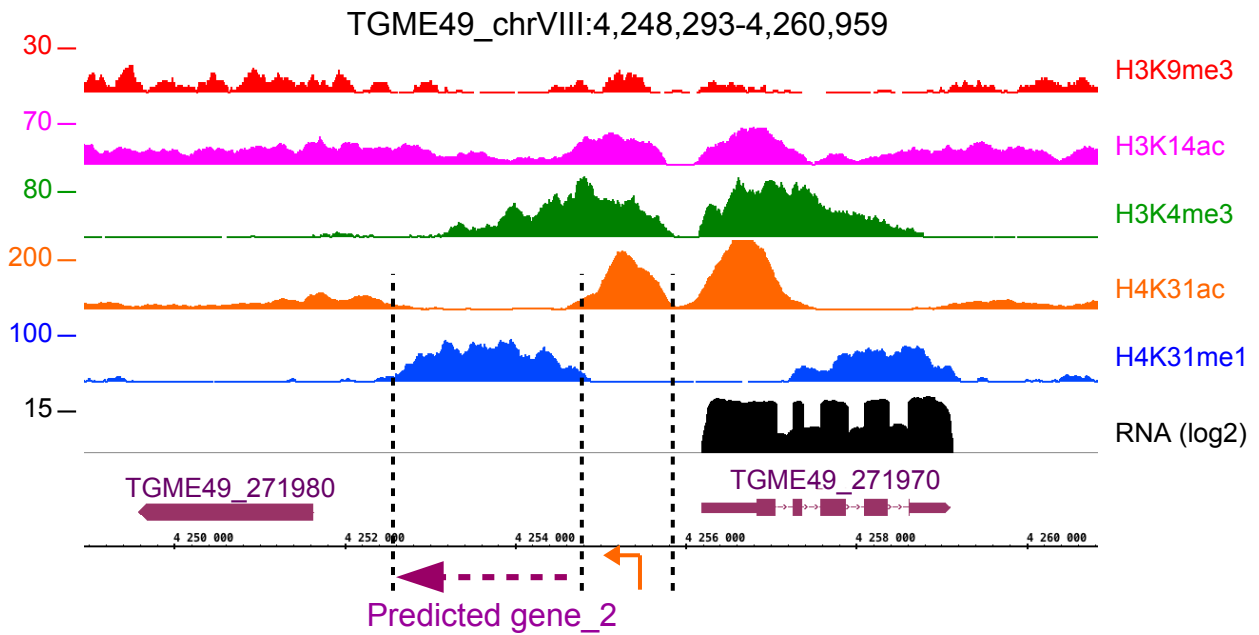
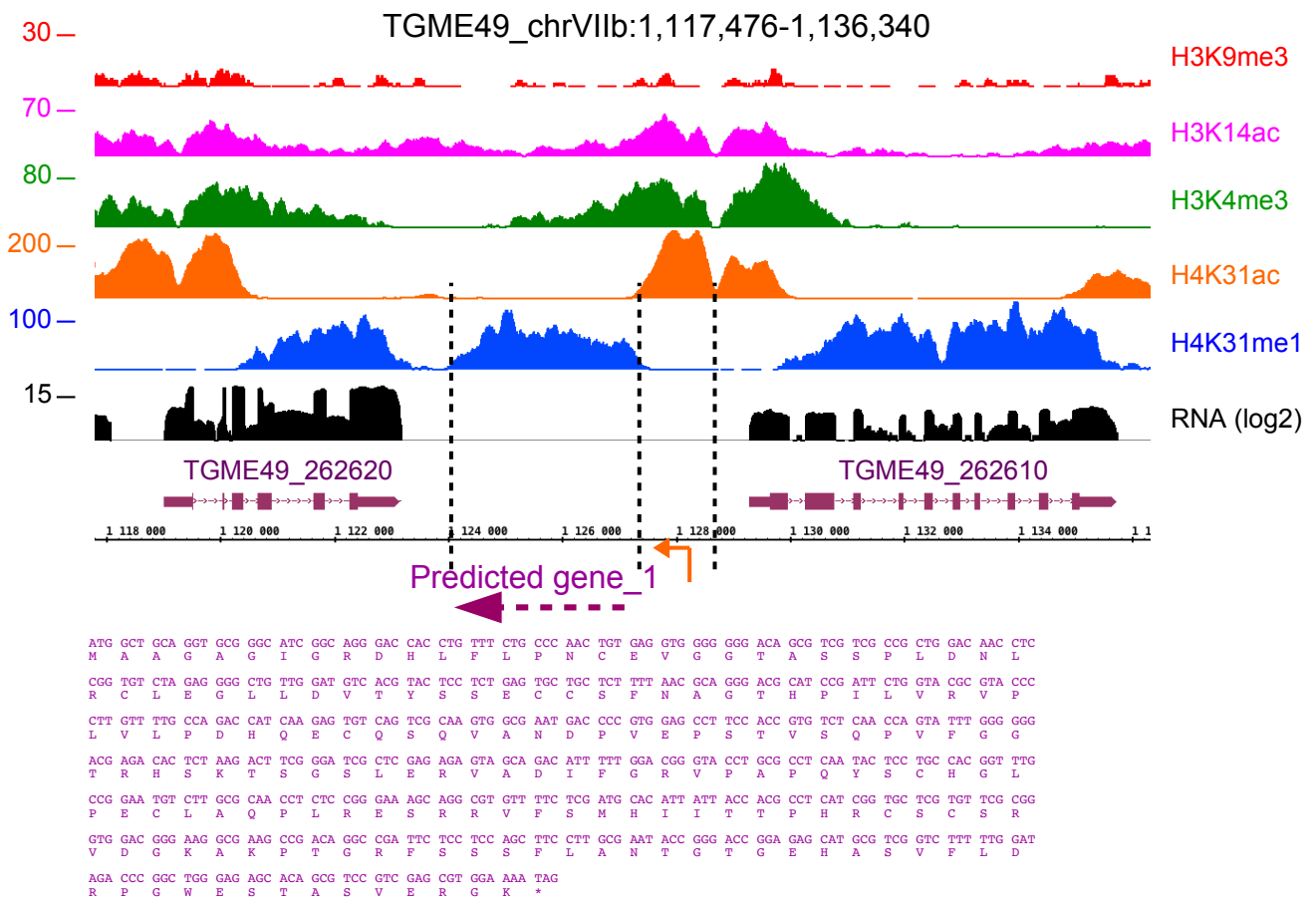


Figure 7 – figure supplement 1

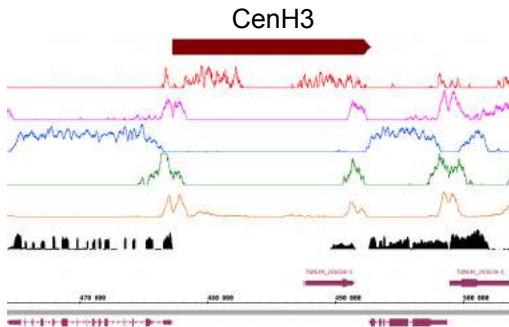


ATG GCA CTT TCT TGT GAG GGC CAG GCC CCG CCG AGA GAC ACA GTG TGG CGA AGC CCC ATG TCC AGT CGC TTT ACC TCG CTG TTC AGC AAG
M A L S C E G Q A P P R D T V W R S P M S S R F T S L F S K
CGA GGA ATT AGC ACG CAT TCA TCC GCG TTG GAC ACT CCG CCA AGA GAA GCT GGC GAC AAT GGA CGG ACC CGA CCG GAA GCC GTG GCG ACA
R G I S T H S S A L D T P P R E A G D N G R T R P E A V A T
GCA GAG AGC CTG GCG AGC AGC GCA GCG GAG GCA GCC CCG CCT CGC GAA AAC TTG AGA GGA ACA CCT GTG CGC GCG AGA CAG GAC TCA CTG
A E S L A S S A A E A A P P R E N L R G T P V R A R Q D S L
CTG CCA CGA CCT TCA CGT CCG CTG TCT TTC AGA GAG ATG TAC AGG CCG CTG CTG AGG AAA CAA AAG GGA AGA AAG CCG AAA GAG GCG ACA
L P R F S R L R S F R L S F R L L R K Q K G R K R K E A T
ACA ACT GTT CAC AGT TCG AAC CCC AGG GCA CCC GTT ACG GCT ACG CAG TCG GAT CCG GCT TTA TCG AGA GTG AAG TCT GTG CAC TTT GGG
T T V H S S N P R A P V T A T Q S D P A L S R V K S V H F G
ACC ATG GTT GTC CAC ACA TAT CCA CGT GAC GCG CAG CAC GGT AGT CCA GAA GCA CCT ATG CTC CTT TGC AAG TCC GAC CCT TCA ATT GTA
T M V D V H T Y P R D A Q H G S P E A P M L V C K S D P S I V
TTC GTT ACG GAG AGT GAT GAG GAC AGT GAG GAC AGT GAA AAC GAC GAC TGC ATC TTT GCC AAC AGG CGT GCA GCG CCA GGA CTG
F V T E S D E D S E E D S D L R E E G G P D D S C P P K S
CAC CCG GTT TGC AGT ACG TCA AGT TTG AGT GAC TCG GCA GGG ATT CGA GGT CCG AGT TCC CCG GTT TAT ACG AAG GAC AAG
H P V C S T S S S T L S D S A G I R G P S S L A V Y T K D K
TCT TCT GCA CGT TAT CTA CAG TCC GGT CCG TTT TTG TCG GAC TTA CGC GAG GAA GGT GGT CCC GAT GAT GAT TCC TGT CCC CCC AAA TCG
S S A R Y L Q S G R F L S D L R E E G G P D D S C P P K S
TTT TTG GAG TAC CAC GGT GAC GAA TTC GGC GTC TGC CAA TAC TTC AAG GAG GGC AGC GCC GGC AGA GGG AGT GCG GTT TCG CAC CAA TCG
F L E Y H G R G V C Q Y F K E G S A G R G S A V S H Q S
ACT TCG AAT TCG CTT CTT AAT TCG CTG TCG TCT GTT GGC CTT GCT AAA CCT CCA GGC CCA AAC GAT AAA GAA CTG GAG AAT GTT CTA ATG
T S N S L L N S L S S V G L A K P P G P N D K E L E N V L M
GGC TGG TGC TCG TAG
G W C S *

Figure 8 – figure supplement 1

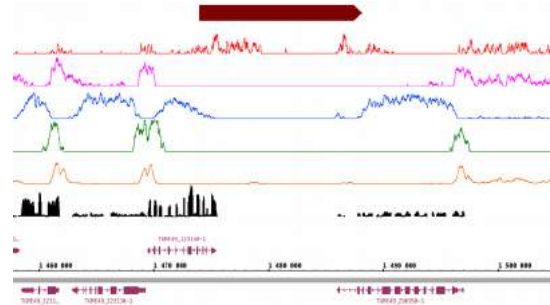
TGME49_chrIa:462,421-508,192

H3K9me3
H3K14ac
H4K31me1
H3K4me3
H4K31ac
RNA (log2)



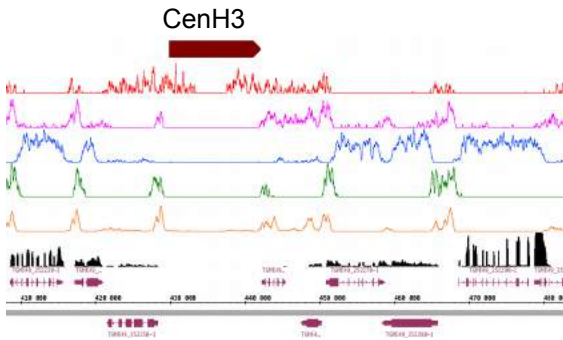
TGME49_chrII:1,454,969-1,506,206

CenH3



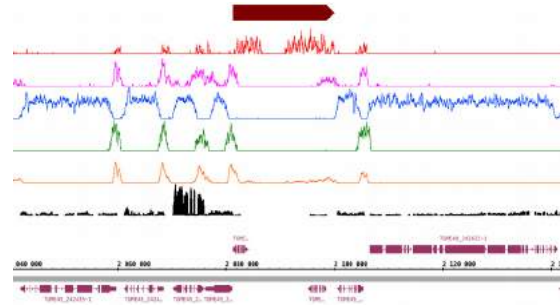
TGME49_chrIII:405,000-484,132

H3K9me3
H3K14ac
H4K31me1
H3K4me3
H4K31ac
RNA (log2)



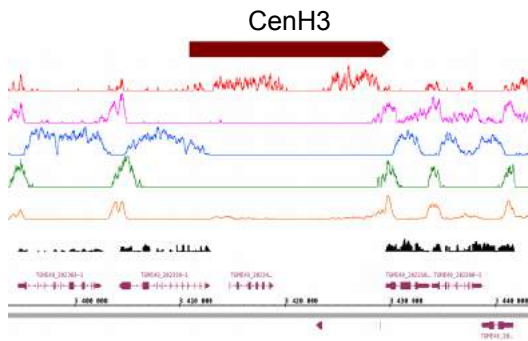
TGME49_chrVI:2,034,000-2,142,325

CenH3



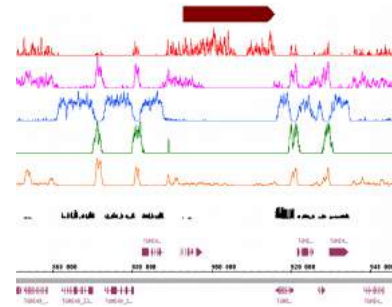
TGME49_chrVIIa:3,389,916-3,445,691

H3K9me3
H3K14ac
H4K31me1
H3K4me3
H4K31ac
RNA (log2)



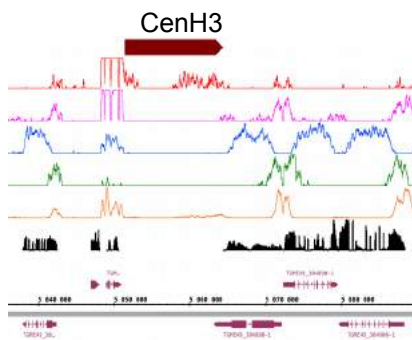
TGME49_chrVIII:845,681-954,681

CenH3



TGME49_chrIX:5,030,000-5,090,499

H3K9me3
H3K14ac
H4K31me1
H3K4me3
H4K31ac
RNA (log2)



TGME49_chrX:3,881,001-3,944,179

CenH3

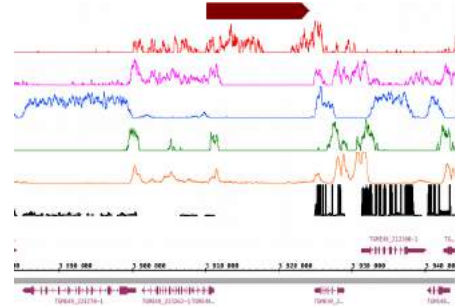


Figure 8 – figure supplement 2

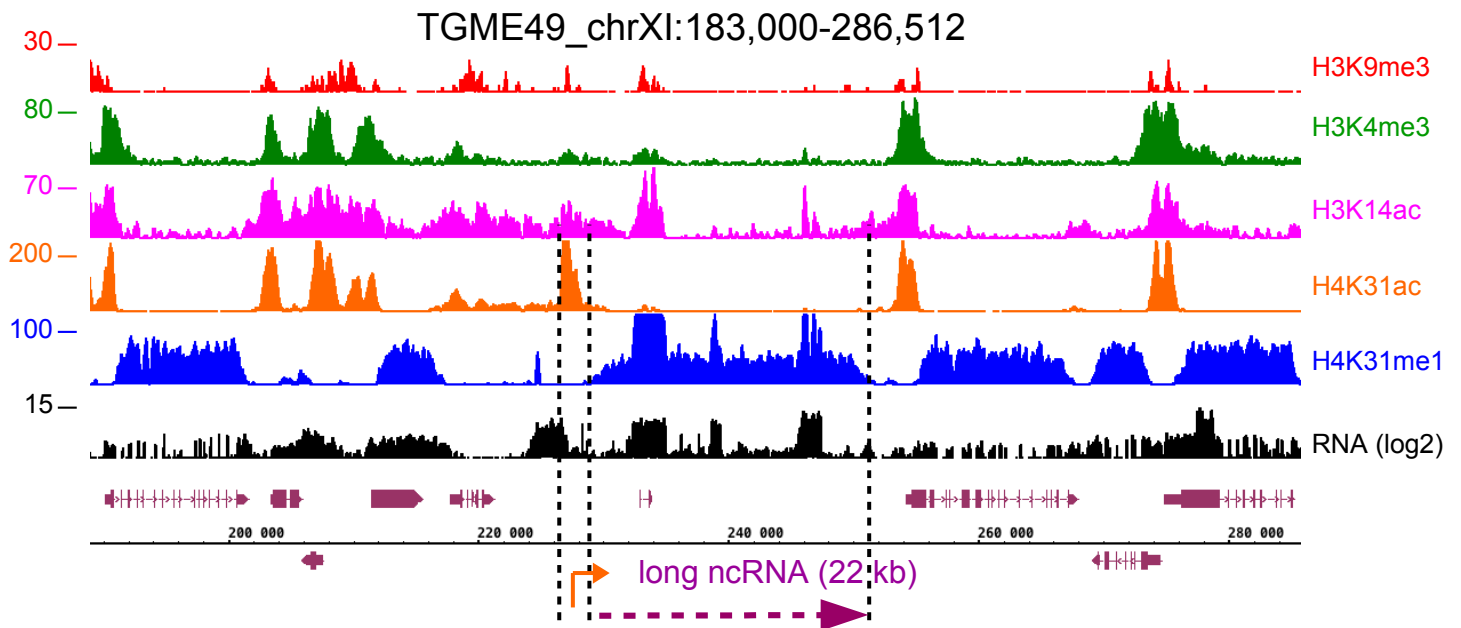
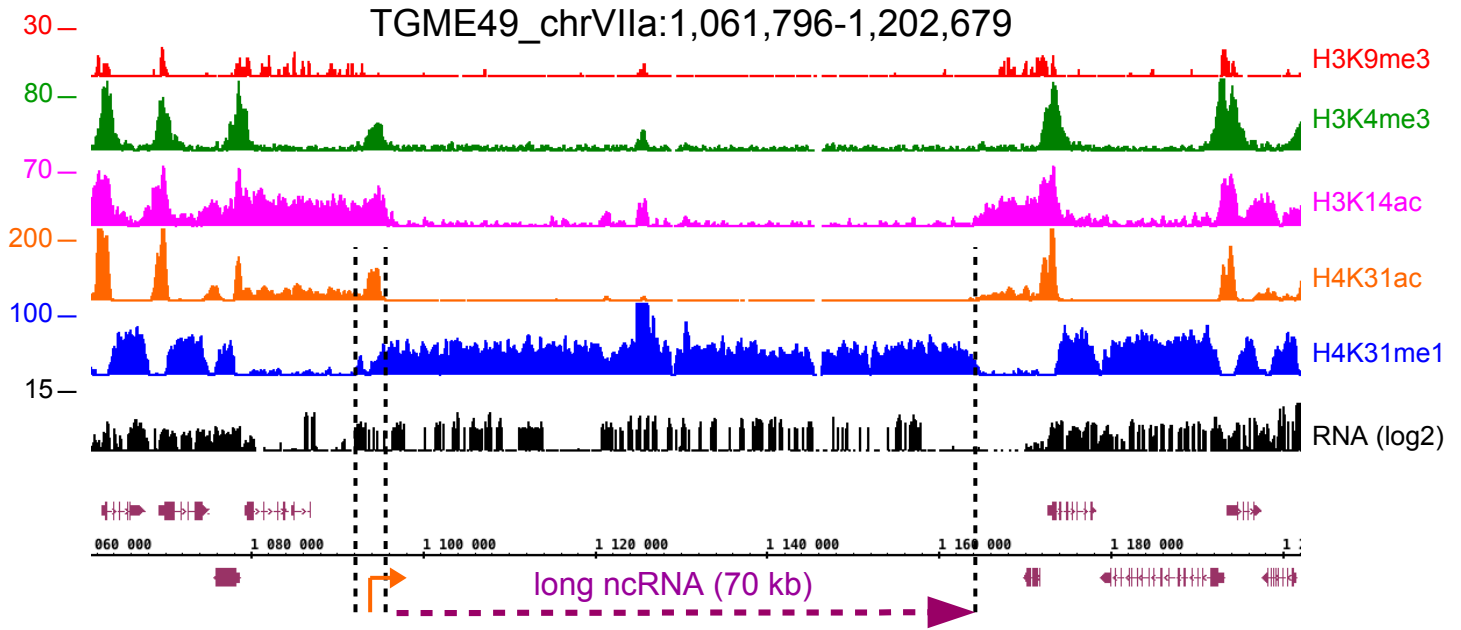
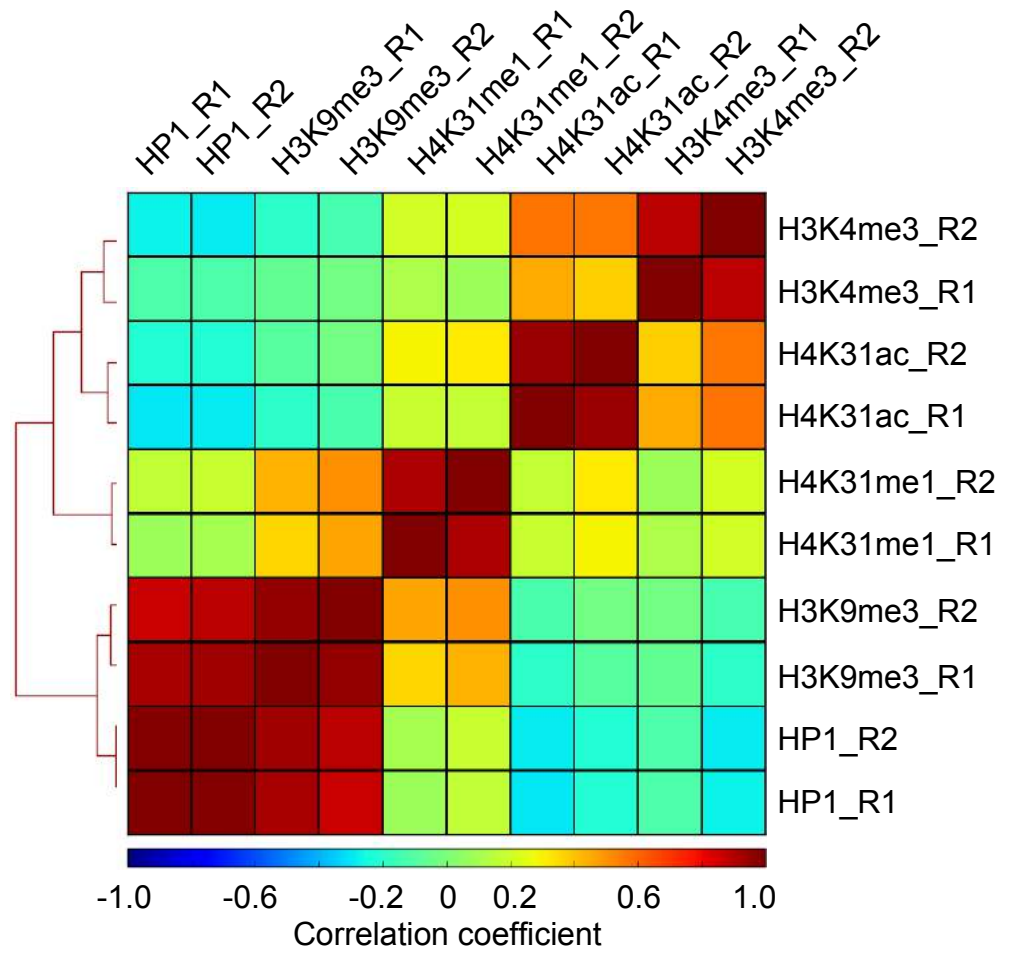


Figure 8 – figure supplement 3

Pearson correlation



Spearman correlation

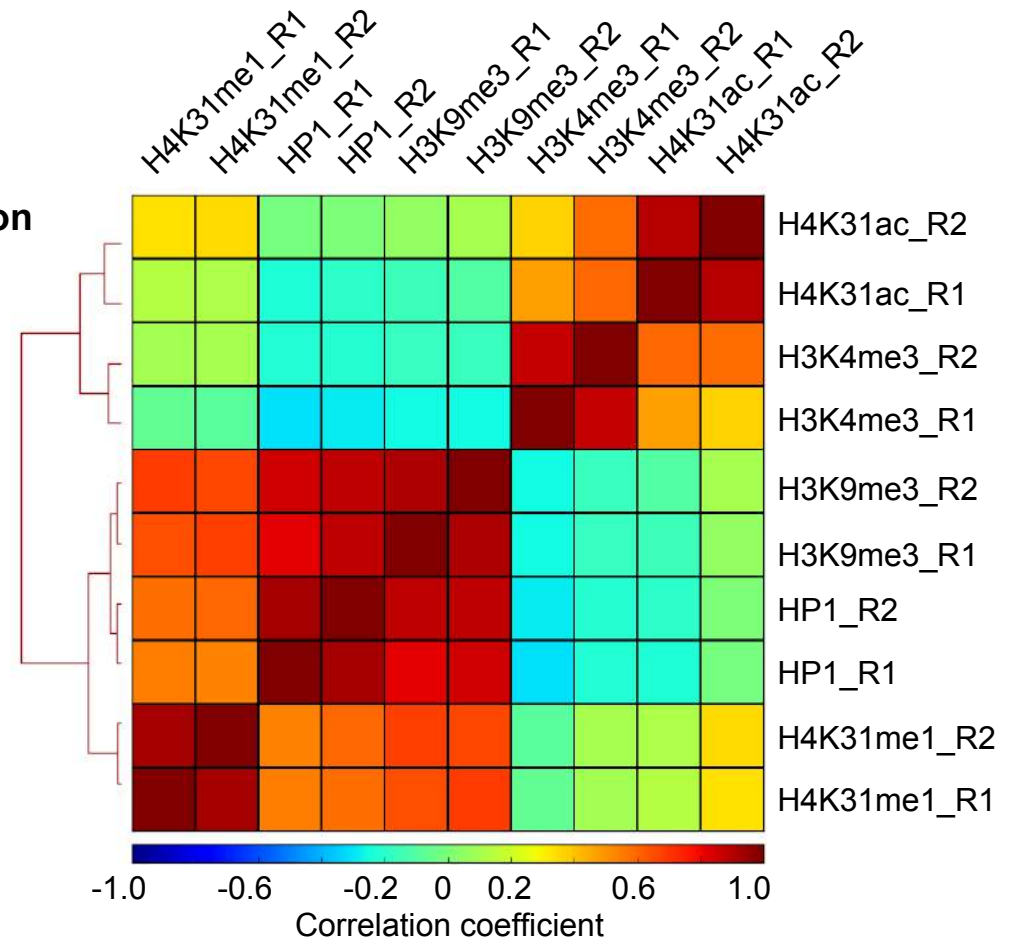


Figure 9 – figure supplement 1

H4K31ac H3K4me3 H3K9me3

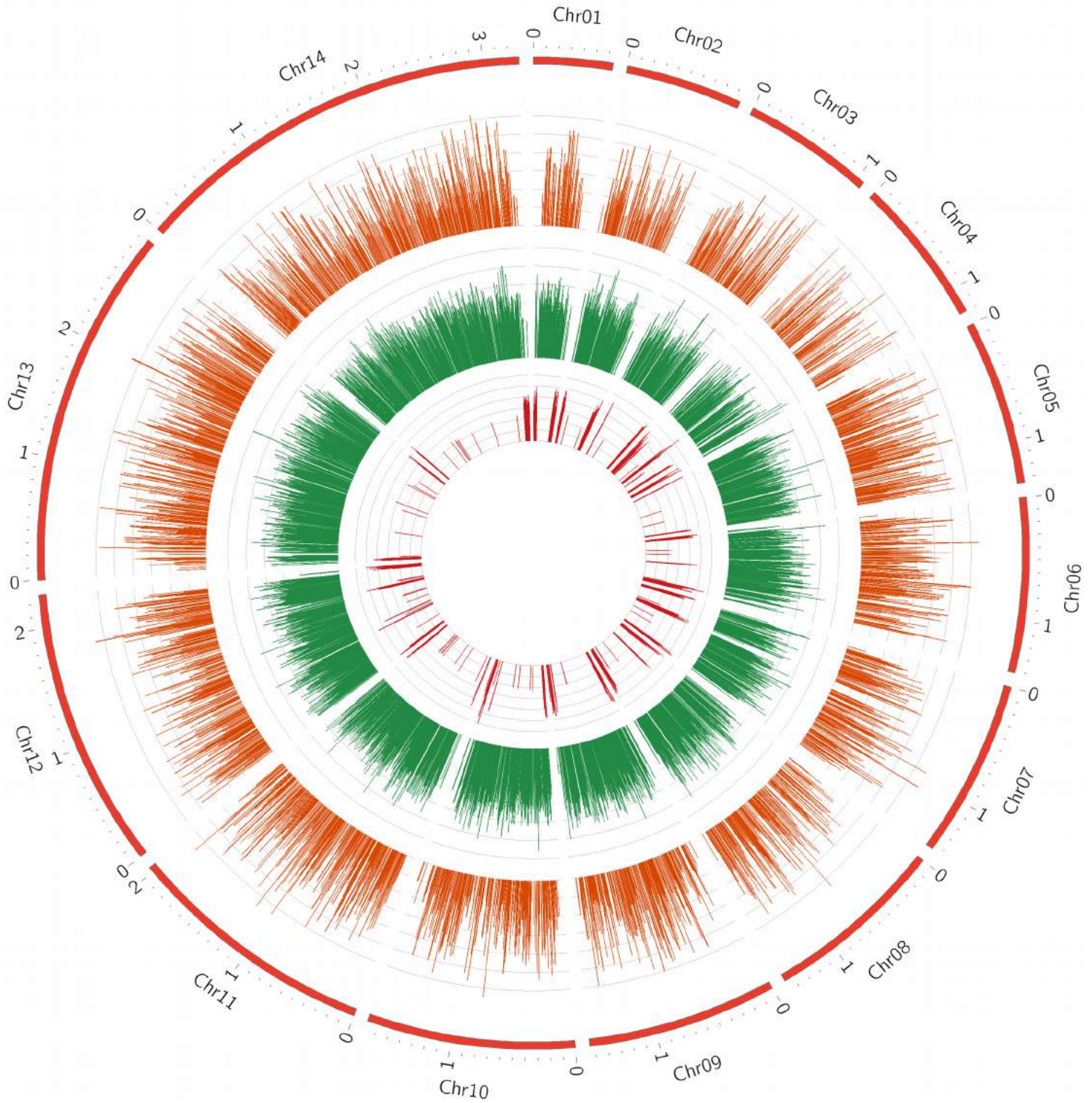


Figure 9 – figure supplement 2

H4K31me H3K9me3 HP1
▶ CenH3 — var genes

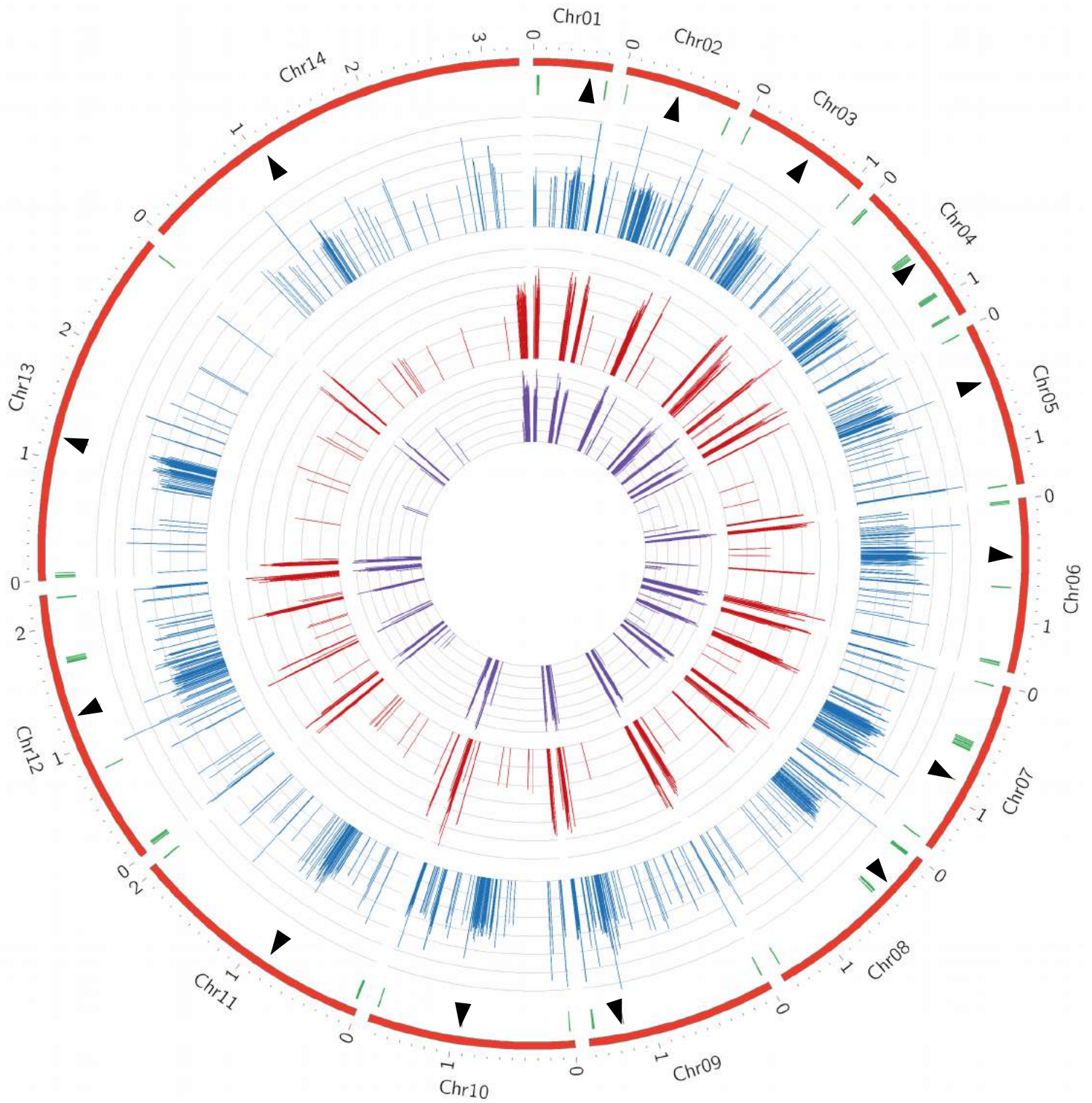
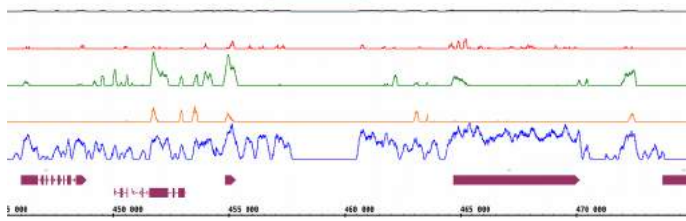


Figure 9 – figure supplement 3

Pf3D7_Chr 01:445,000-475,000

CenH3

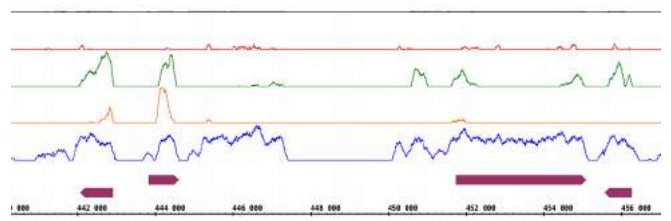


Pf3D7_Chr 02:439,793-457,562

CenH3

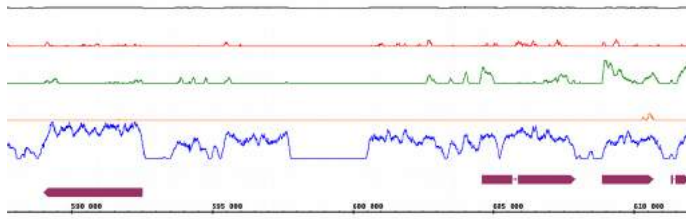


HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
genes



Pf3D7_Chr 03:587,000-612,000

CenH3

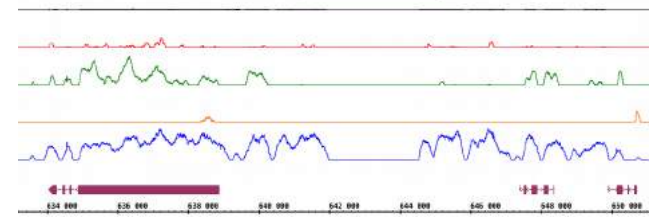


Pf3D7_Chr 04:632,247-652,133

CenH3

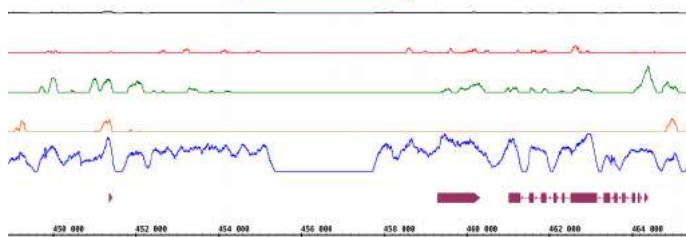


HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
genes



Pf3D7_Chr 05:448,292-465,722

CenH3

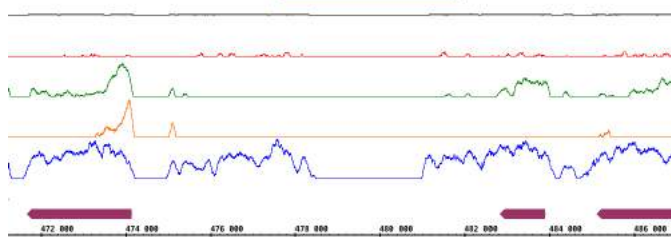


Pf3D7_Chr 06:470,000-488,000

CenH3

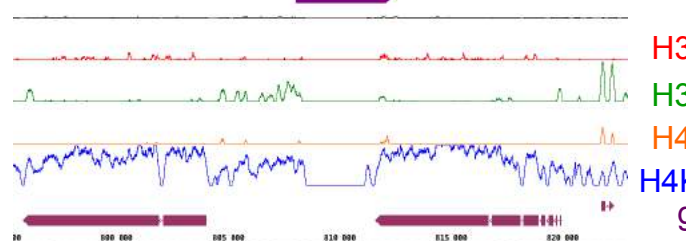


HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
genes



Pf3D7_Chr 07:792,105-825,008

CenH3



Pf3D7_Chr 08:292,000-309,000

CenH3



HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
genes

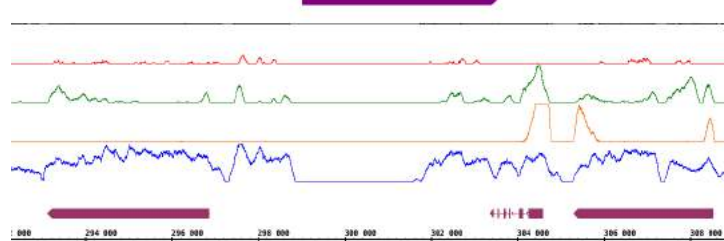
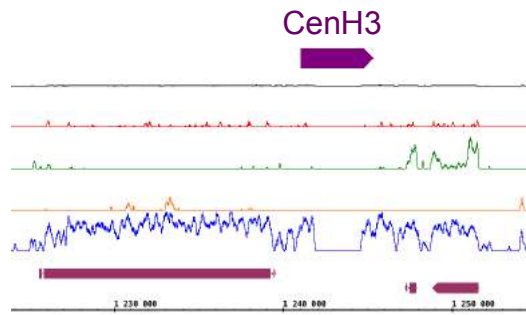


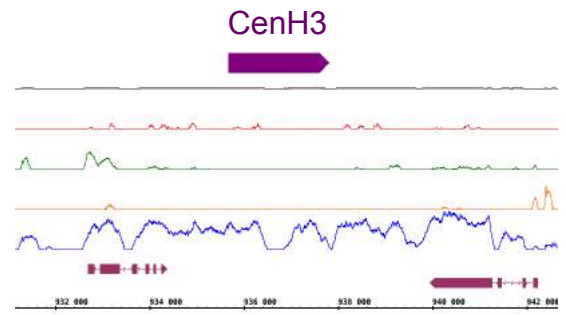
Figure 10 – figure supplement 1

Pf3D7_Chr 09:1,221,335-1,255,984

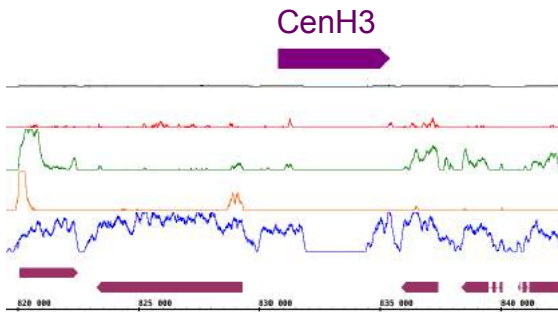


HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
genes

Pf3D7_Chr 10:930,550-942,669

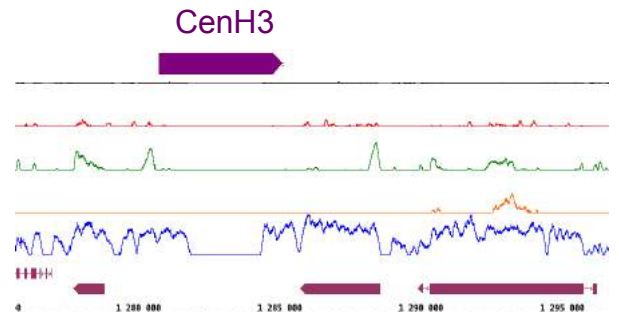


Pf3D7_Chr 11:818,788-843,000

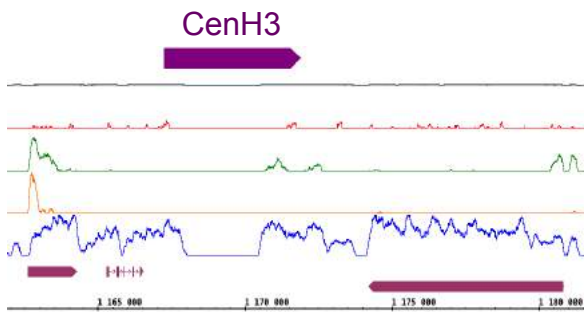


HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
genes

Pf3D7_Chr 12:1,275,979-1,297,490



Pf3D7_Chr 13:1,161,340-1,181,799



HP1
H3K9me3
H3K4me3
H4K31ac
H4K31me1
genes

Pf3D7_Chr 14:1,062,788-1,081,621

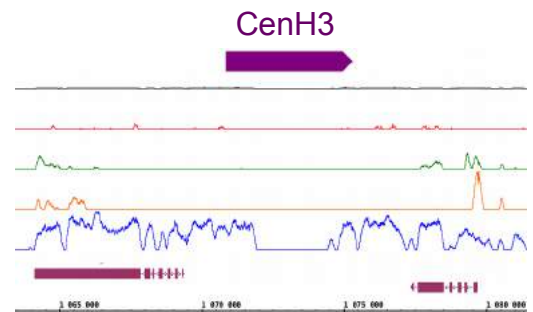
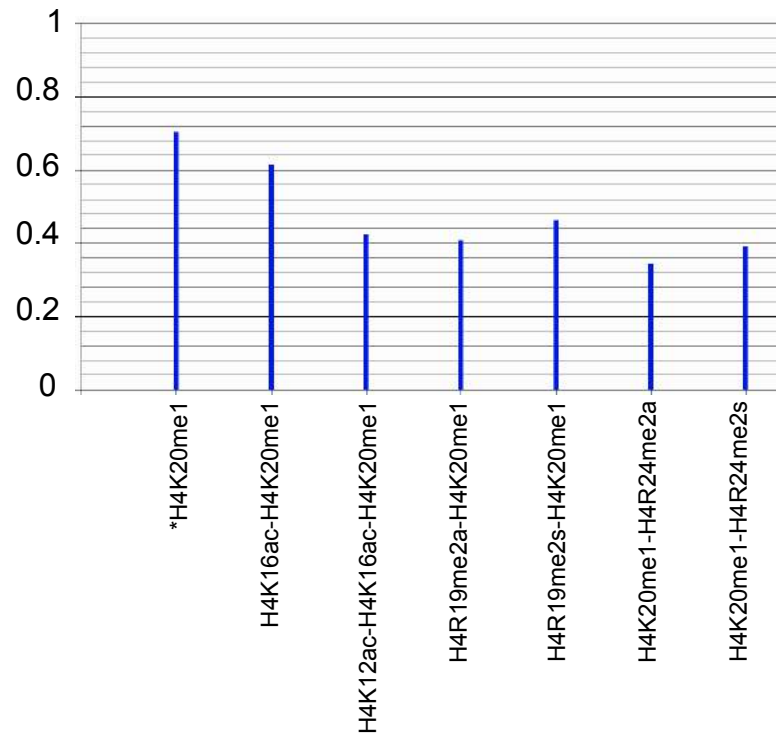
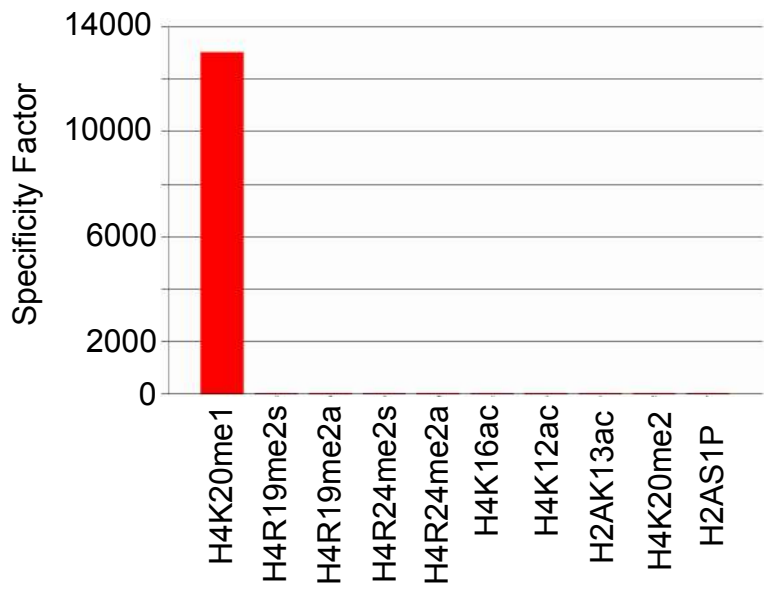
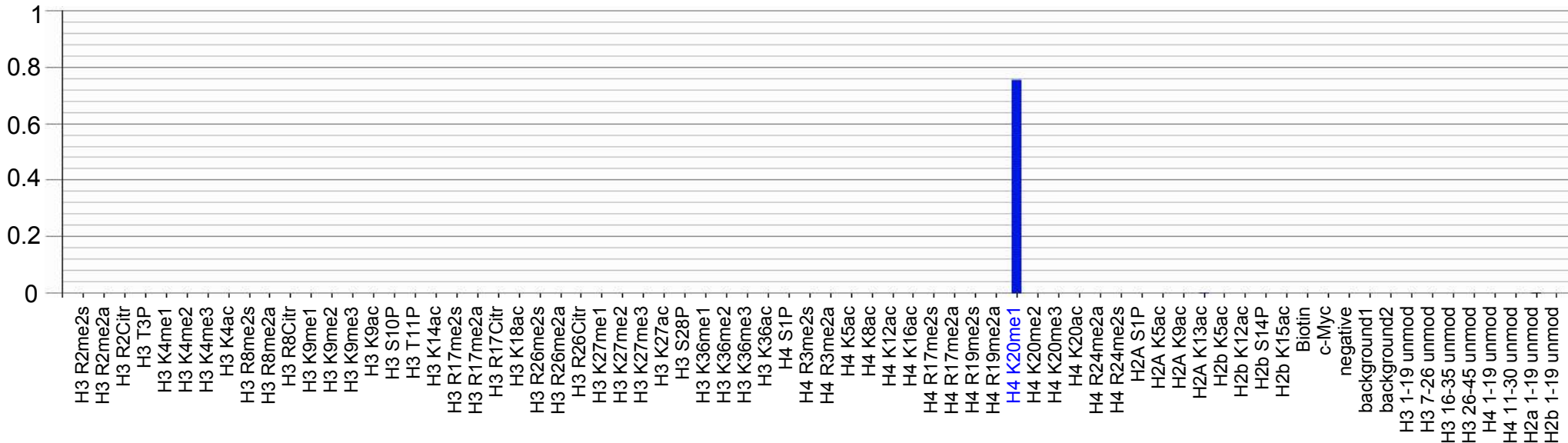
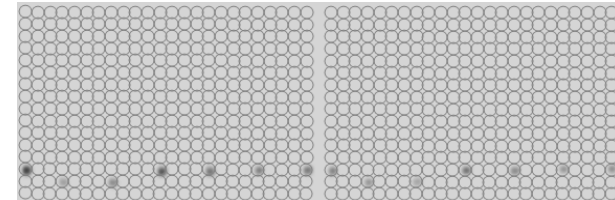
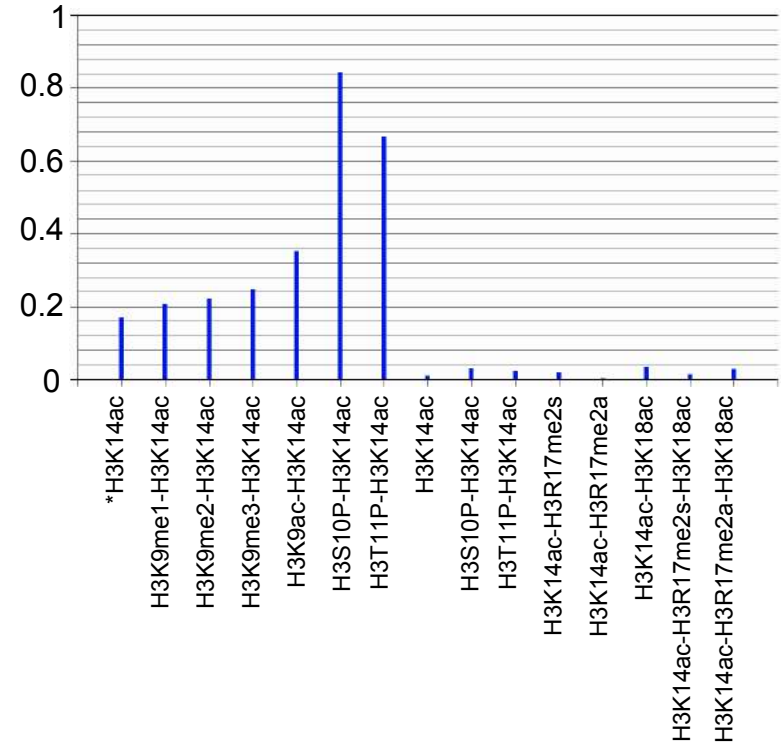
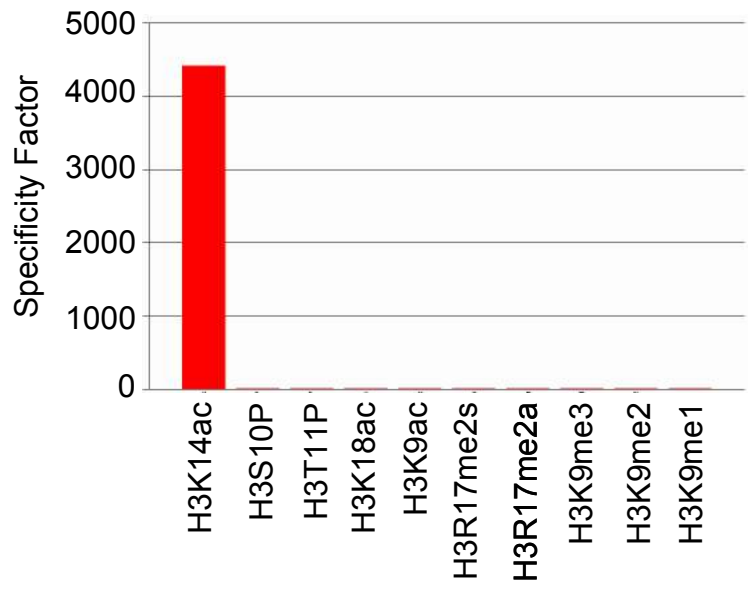
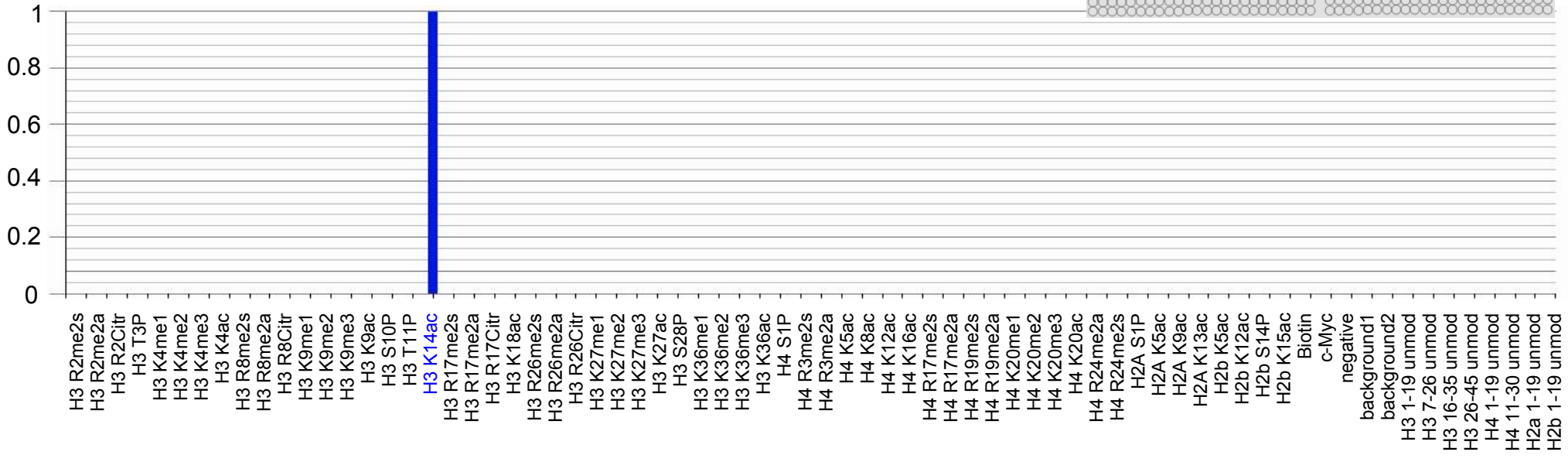
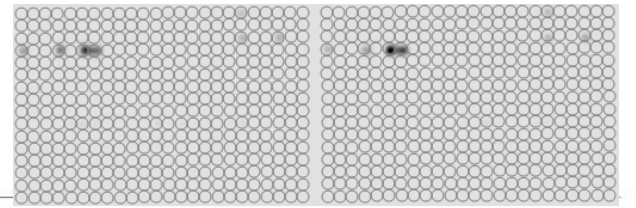


Figure 10 – figure supplement 1

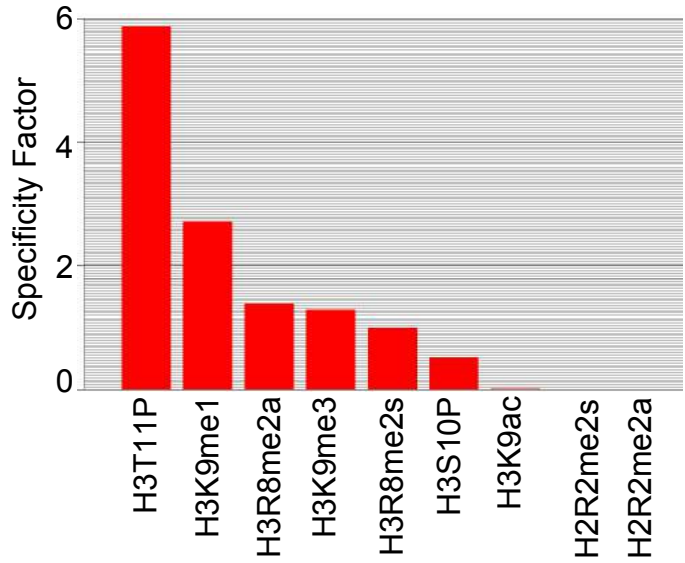
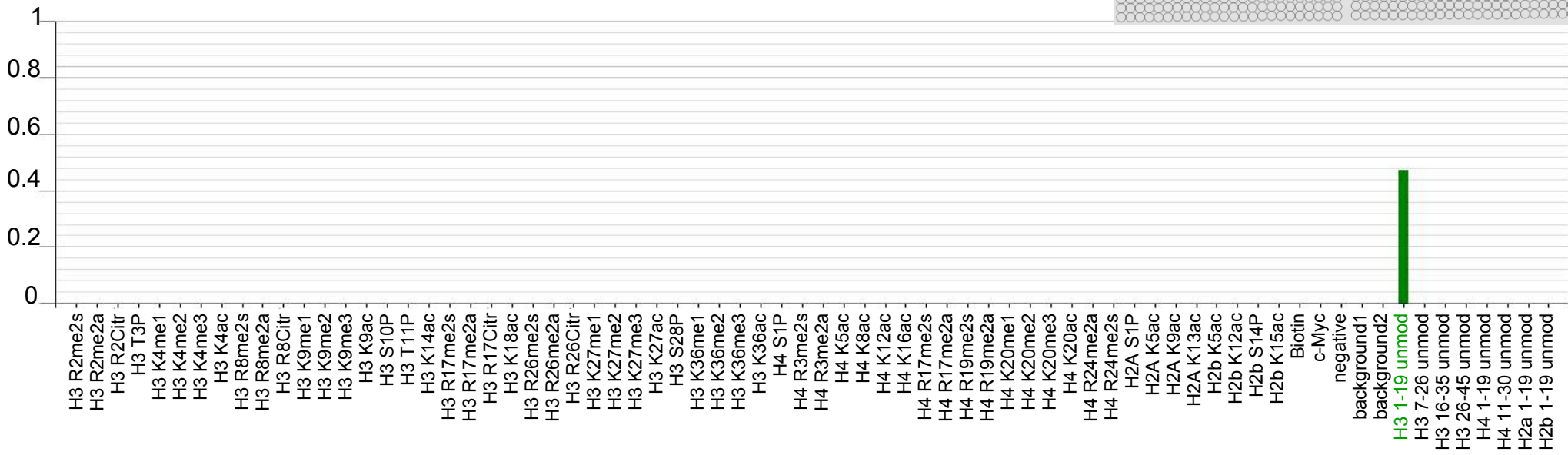
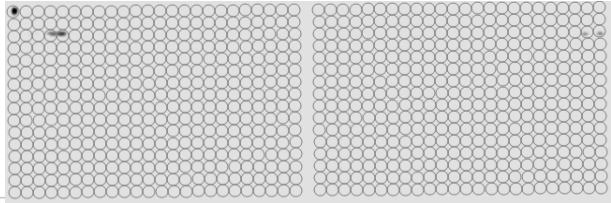
H4K20me1 (Sautel et al., 2007)



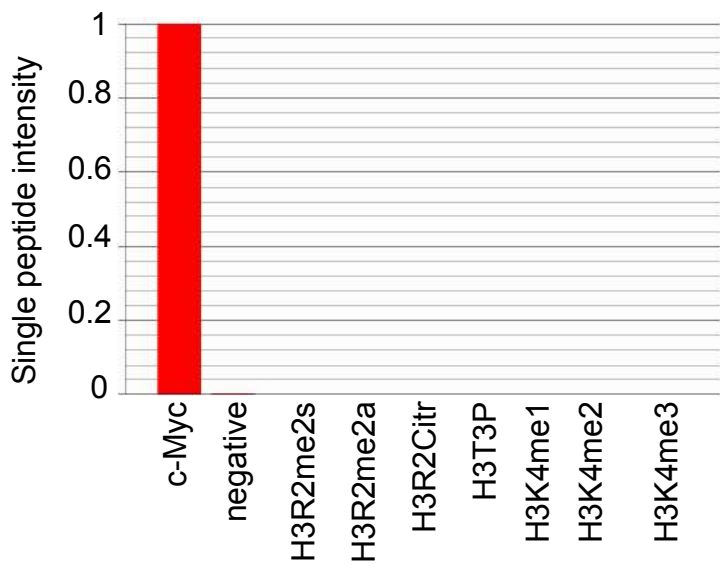
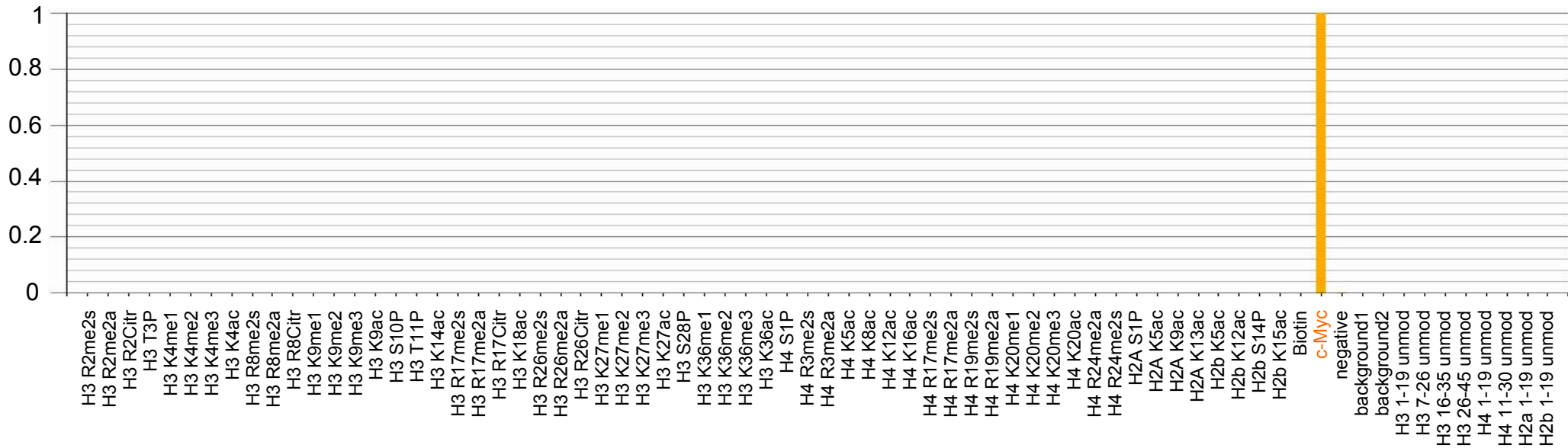
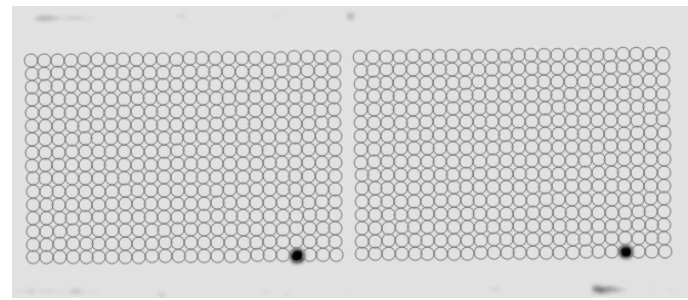
H3K14ac (Millipore # 04-1044)



H4K31ac



H4K31me1 / c-Myc control



Part III. Discussion

Developmental switching in *Toxoplasma gondii*, from the virulent tachyzoite to the relatively quiescent bradyzoite stage, is responsible for disease propagation and reactivation. To date, although the expression of many genes and their products have been identified as subject to regulation during parasite differentiation, the literature remains poor in identifying regulators of this process in *Toxoplasma gondii* and, likely, in other *Apicomplexa*. When I started my PhD, the team had strong evidence that the alteration in the rate of histone acetylation in the vicinity of stage-specific genes is one of the molecular motors that drive parasite differentiation (Saksouk et al. 2005). However, few factors have been clearly identified as being involved in these stage conversions. The discovery by our lab of TgHDAC3 as a key player in the differentiation (Saksouk et al. 2005, Bougdour et al. 2009; Maubon et al. 2010) opens new perspectives on the mechanisms involved, although many questions about the *modus operandi* of the enzyme remain unanswered.

III- 1. CRISPR/cas9-mediated gene disruption of TgHDAC3 recapitulates FR235222-mediated phenotypes

HDACi (FR235222) stimulation triggers pleiotropic phenotypes in tachyzoites ranging from (i) the growth arrest defined by a cytokinesis defect and characterized by the lack of IMC1 delineating the new formed daughter cells to (ii) the alteration of histone H4 tail acetylation levels (Figure 12A) (Bougdour A *et al.*, 2009). The contribution of TgHDAC3 to those phenotypes was only confirmed with the isolation of drug-resistant mutant. Indeed, FR235222 treatment had no effect on DNA replication and IMC1-delineated daughter cells in the TgHDAC3^{T99A} and TgHDAC3^{T99I} mutants when compared with the WT parasites suggesting that the T99A and T99I mutations in TgHDAC3 are each sufficient to circumvent the FR235222-mediated cytokinesis defect (Figure 12A). In the same line, the levels of AcH4 signals under FR235222 treatment were increased approximately eight-fold in the WT but were only increased approximately three-fold and remained unchanged in the TgHDAC3^{T99A} and TgHDAC3^{T99I} resistant lines, respectively (Figure 12B).

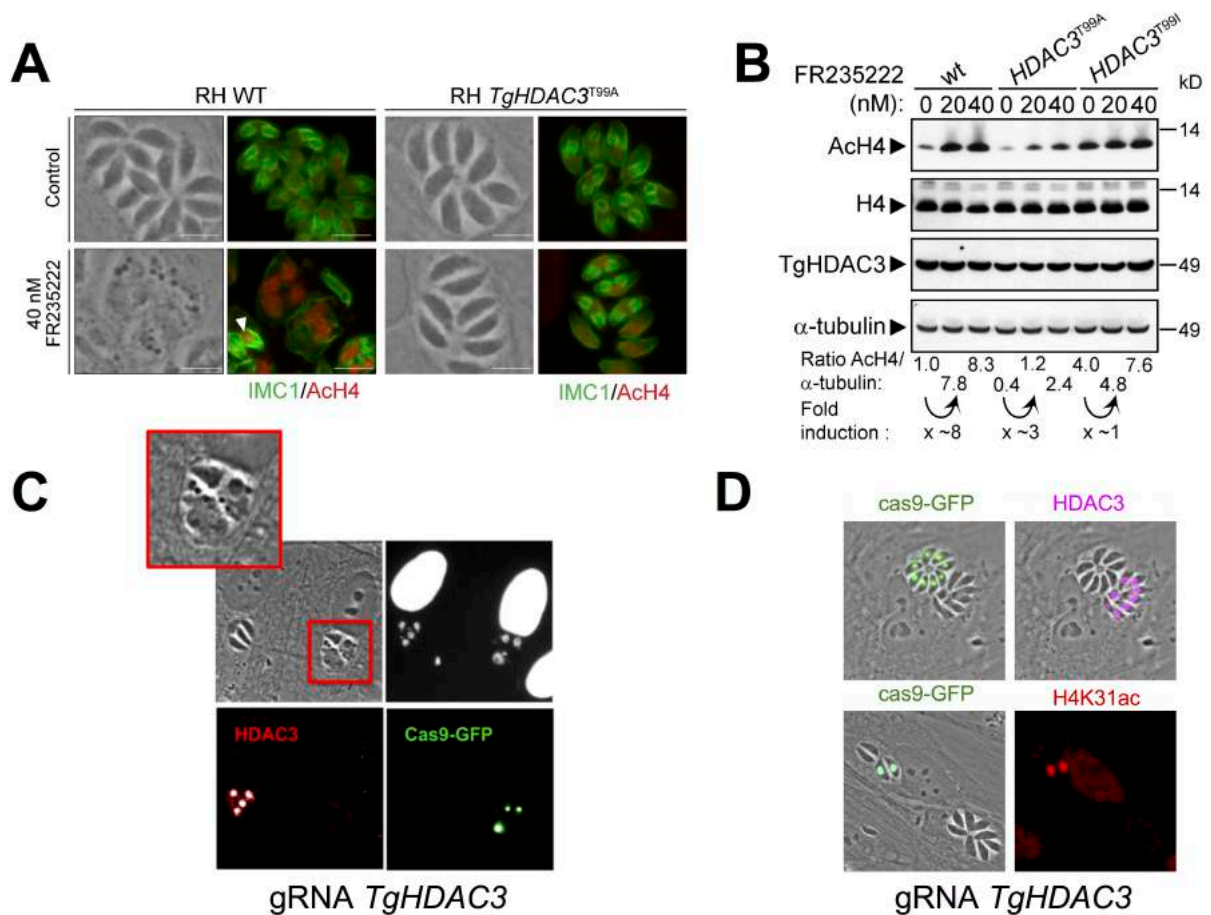


Figure 12. Chemical and genetic inactivation of TgHDAC3 cause cytokinesis defects and histone H4 hyperacetylation. (A) Images from Bougdour et al. (2009) : Effects of FR235222 on histone H4 acetylation in intracellular *T. gondii* parasites. Confluent monolayers of HFF cells were infected with *T. gondii* RH WT and R20D9 (*TgHDAC3*^{T99A}) strains in the presence of 40 nM FR235222 and 0.1% DMSO as a control. After 24 h of growth, cells were fixed and stained for AcH4 (red) and IMC1 (green). The arrowhead indicates aberrant progeny. Bars, 5 μ m. (B) Extracellular *T. gondii* parasites (RH WT, *TgHDAC3*^{T99A}, and *TgHDAC3*^{T99I}) were treated with the indicated concentrations of FR235222 for 4 h and lysed. Total cell lysates were analyzed by immunoblot with anti-AcH4, anti-H4, anti-TgHDAC3, and anti-tubulin antibodies as indicated. (C) Intracellular tachyzoites in which the *TgHDAC3* gene was disrupted by transient transfection of CRISPR/Cas9 (in green). Aberrant progeny was framed. (D) Representative micrographs showing intracellular tachyzoites in which the *TgHDAC3* gene was disrupted by transient transfection of CRISPR/Cas9. The efficiency of *TgHDAC3* disruption in Cas9-expressing parasites was monitored by the anti-TgHDAC3 staining (in pink) and cas9-GFP expression (in green). The levels of H4K31ac (in red) were monitored in *TgHDAC3*-disrupted parasites (GFP positive) and compared to untransfected parasites (GFP negative). Scale bar, 10 μ m.

Recent advances in molecular genetics using CRISPR-CAS9 technologies have considerably enhanced the ability to understand mechanisms main-played during stage conversion. Using the Cas9-mediated genetic inactivation of *TgHDAC3*, I confirmed the aforementioned

FR235222-associated phenotypes. Indeed, the cas9-mediated gene disruption of TgHDAC3 led to vacuolated tachyzoites, lacked IMC1-delineated daughter cells or aberrant progeny (Figure 12C). Moreover, the parasites displayed massive DNA over-replication ($>1N$ DNA content per cell), indicating that TgHDAC3 interferes directly or indirectly with *T. gondii* cell-cycle progression. This also provides additional support for the conclusion that the phenotype is not caused by the effect of FR235222 on the host cell but clearly mediated by TgHDAC3. Notably, the inactivation of the TgHDAC3 gene, unlike other TgHDACs, causes hyperacetylation of H4K31 in parasites nuclei (Figure 12D), thereby mimicking the effect of the cyclic tetrapeptide HDACi on the enzyme (Fig. 3A in Sindikubwabo *et al.*, and Bougdour *et al.*, 2009).

III- 2. Linking bradyzoite development to the parasite cell cycle ?

A number of elegant global gene expression studies have analyzed changes in gene expression during the parasite's lytic cycle *in vitro* in human fibroblasts (HFF cells) (Radke JR, 2005 *BMC biology* 3: 26; Gaji RY *et al.*, *Molecular microbiology* 79: 192–204; Behnke MS, *et al.* 2010 *PLoS one* 5: e12354). These have shown that *T. gondii*'s transcriptome is highly cell-cycle regulated with many genes transcribed maximally just prior to their use and then downregulated in a pattern described as “just in time” (Gaji RY *et al.*, *Molecular microbiology* 79: 192–204; Behnke MS, *et al.* 2010 *PLoS one* 5: e12354). Virtually all of the stress conditions that promote bradyzoite differentiation (alkaline pH, interferon- γ) reduce the proliferation of tachyzoites. Slowing of the parasite cell cycle has been linked to the initiation of the bradyzoite developmental program from a late-S/G2 subpopulation containing 1.8–2N DNA content (Jerome *et al.*, 1998; Radke *et al.*, 2003). During bradyzoite differentiation, these parasites proceed through M phase and then arrest in G1/G0 with uniform 1N DNA content. The unique late-S/G2 stage represents a premitotic cell cycle checkpoint for the commitment to bradyzoite formation and growth arrest following mitosis. The identification of cyclic expression of several bradyzoite-specific mRNAs, which exhibit peak expression in the late mitotic period, lends support to this model (Behnke *et al.*, 2010).

Our data indicate that FR235222-mediated TgHDAC3 inhibition affects negatively and more specifically cytokinesis without impacting karyokinesis process. The explanation could be that the treatment of parasites with FR235222 induces the activation/inhibition of protein factors involved in cytokinesis regulation. These factors may be direct substrate of

TgHDAC3, yet no obvious candidates were found while analyzing our FR235222-treated acetylome.

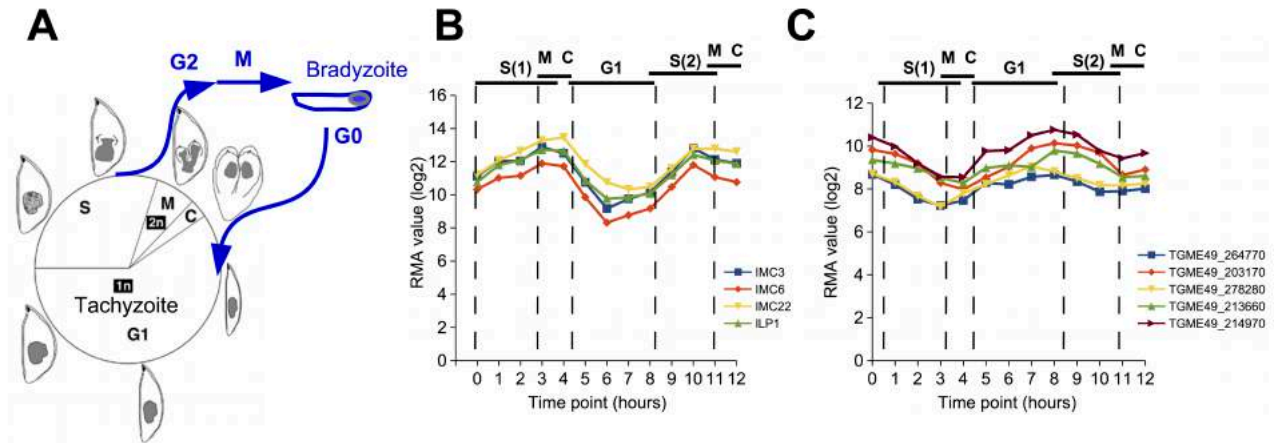


Figure 13. The FR235222-regulated proteins expression across the *T. gondii* cell-cycle. (A) Cell cycle model of *Toxoplasma gondii* development in the intermediate host. (B) and (C) Cell cycle expression mRNA profiles were assessed for proteins that were down regulated upon FR235222 treatment. Spline model curves for selected cell cycle mRNAs encoding IMC-associated (B) and unknown (C) proteins demonstrate relative abundance and time shifts indicating that the expression of these factors follows a serial order with peaks at different cell cycle stages.

Alternatively, the genes encoding cell cycle regulator may be regulated at the transcriptional levels by TgHDAC3. In fact, we identified a group of 66 proteins whose levels were decreasing upon FR235222 stimulation (Figure 8B). While most of them have unknown function, we identified 25 proteins whose genes display a typical cell cycle-regulated expression pattern (example in Figure 13), when they are not themselves involved in cell cycle progression. Four proteins belong to the inner membrane complex (IMC), that lines the interior of the plasma membrane and contains proteins important for replication (Figure 13A). The IMC localizing protein 1 (ILP1) was one of them. Notably, *Δilp1* vacuoles were shown to highly disordered and contained morphologically abnormal parasites that appeared enlarged and/or malformed (Chen AL et al., 2015 - PMID: 25691595), defects that were reminiscent of parasites treated with FR235222 or in which *TgHDAC3* was disrupted (Figure 12). Similarly, ablation of the ILP1 homolog in *Plasmodium berghei* produced enlarged, swollen parasites with microtubule and motility defects (Trempe AZ et al., 2013 - PMID: 23773015). FR235222 treatment also significantly lower downs the levels of a cyclin2-related protein (TGME49_267580) and a DNA replication licensing factor (TGME49_214970). We also identified new cell cycle-regulated genes, including one encoding a putative histone lysine methyltransferase that was shown to localize exclusively at the nascent IMC of the daughter cells (Hakimi, personal communication).

Interestingly, while looking at the proteome of tachyzoites treated with alkaline pH we did not find any cell cycle-regulated proteins whose levels were altered compared to the pH 7 mock control (data not shown). These results are in line with the observation that alkaline pH while triggering H4 acetylation did not affect cytokinesis (Figure 3) and overall question the relevance of the hypothesis that alkaline pH by reducing parasite growth stimulates stage conversion. The same negative result and resulting conclusions can be drawn concerning interferon- γ (data not shown).

III- 3. FR235222-mediated inhibition of TgHDAC3 re-programs stage-specific gene expression in tachyzoites

The importance of histone acetylation for the control of differentiation is underscored by the finding that chemical inhibition of TgHDAC3 with low doses of the compound FR235222 was able to up regulate the bradyzoite marker p36/SRS9 (Bougdour *et al.*, 2009). In the same study, analysis of the genome-wide hyperacetylation pattern induced by FR235222 led to the observation that the compound influences the expression of >370 genes, a third of which are bradyzoite-specifically expressed (Bougdour *et al.*, 2009).

In terms of studying bradyzoite differentiation, strain choice is a crucial factor. While Type I strains, particularly RH, grow faster and are easier to genetically manipulate, they are limiting in the study of cyst formation. However, the first *T. gondii* acetylome mapping (Jeffers and Sullivan, 2012 ; Xue B *et al.*, 2013) as well as our previous studies on TgHDAC3 (Saksouk *et al.*, 2005 ; Bougdour *et al.*, 2009) were made using RH strain. To understand more accurately the mechanisms involved in the process of differentiation, the choice was made to use Type II isolates that form mature cysts readily, although they grow slowly and are more difficult to manipulate. We had to develop the appropriate tools to study their acetylome and proteome in the context of TgHDAC3 inhibition.

When comparing protein levels of tachyzoites pre-stimulated with interferon- γ , alkaline pH or FR235222, the deacetylase inhibitor promotes a drastic alteration on the levels of hundreds of proteins while the other two stimuli merely change the proteome pattern when compared to their respective mock control. While it is well documented that conversion to the latent stage is a stress-mediated response, coupled with a slowing of the parasite cell cycle to which alkaline pH clearly contributes, our proteomic data in the context of a relevant cystogenic

strain tend to contradict most of the study so far published. Whether these stress treatments act on the parasite directly (while they are extracellular), and/or if they act indirectly on intracellular parasites by stressing the host cell, is unclear. Assuming that pH influences differentiation via the host cell, the cell type and the development stage of infected cells may explain this discrepancy. Indeed, the proclivity toward spontaneous differentiation for instance is influenced by the type of parasite strain but also the host cell background (Ferreira da Silva et al., 2008), and cysts are more frequently detected in differentiated host cells that are long-lived (Dubey et al., 1998).

While we expected that the trigger(s) to differentiate are complex and multifactorial, consisting of both endogenous and exogenous factors, our data brought for the first time the evidence that one unique compound can manipulate the genetic programs that govern *T. gondii* stage conversion. Unexpectedly, FR235222 stimulation triggers in tachyzoite the expression of 56 proteins whose expression is restricted to cat enteroepithelial stages (merozoite, sexual stages and oocysts) and more than 300 bradyzoite-specific proteins, one third of which being exclusively detected when isolated from brain of chronically infected intermediate hosts (e.g. mice). Although FR235222 promotes the induction of bradyzoite proteins as *compound 1* whose molecular target is still unknown (Behnke MS et al., 2008 PMID: 18433450), the HDACi is quite unique in its ability to change the genetic program in a manner to favor the expression of sexual stage-specific proteins.

Intriguingly, FR235222 induces the expression of genes that are dispersed across all parasite chromosomes. This *modus operandi* is well supported by recent studies that demonstrate primary developmental transitions leading to formation of the *T. gondii* tissue cyst are accompanied by a temporally ordered set of transcriptional events (Cleary et al., 2002; Singh et al., 2002; Radke et al., 2005). Similar observations have been made in *Plasmodium*, which has led to the ‘just in time’ hypothesis for those selected genes regulated during development (Bozdech et al., 2003; Llinas and DeRisi, 2004). It is enticing to speculate that TgHDAC3 is a master regulator that coordinates changes in the transcriptome that lead to stage conversion. This model can be extended beyond the *Apicomplexa* phylum. Indeed, Sonda et al. found that exposure of parasites to the HDACi FR235222 increased the levels of histone acetylation, altered gene transcription and inhibited *Giardia lamblia* encystation, thus providing evidence that epigenetic mechanisms are involved in stage differentiation in other parasitic protozoa (Sonda et al., 2010 - PMID: 20132448). However this model lacks elements of specificity that

would be needed to explain the strict temporal patterns of gene expression that unfold during parasite development.

III- 4. Targeting of TgHDAC3 to chromatin in a DNA-specific manner : a role for ApiAP2 transcription factors ?

In vivo, HDAC enzymes act in megadalton complexes containing corepressors, transcription factors, and linking proteins. Most often these enzymes are embedded into a network of proteins, which modify or directly support their enzymatic activity (Hakimi et al., Nature, 2002). Usually, histone deacetylases are unable to access their histone substrates, unless DNA-bound activators or repressors target them there. A major breakthrough in the understanding of the biological role of TgHDAC3 in *T. gondii* came with the observation that the enzyme forms distinct complexes.

The first TgHDAC3-interacting partners were purified by a combination of conventional and Flag affinity chromatography from a stable parasite (RH) cell line expressing ectopically a second copy of TgHDAC3 dually tagged with HA and FLAG at its C-terminus (Saksouk et al., 2005). Analysis by silver staining of Superose 6 gel filtration fractions reveals that TgHDAC3 is embedded in a high-molecular-mass complex (1 Mda) that is remarkably robust at it remains stable under stringent washing conditions (Saksouk et al., 2005). These partnerships were subsequently resolved by a combination of mass spectrometry-based proteomics that identified TgHDAC3 in a core complex gathering TBL1 and two unknown proteins, namely CRC230 (TGME49_305340) and CRC350 (TGME49_272710) whose identity was subsequently disclosed as MORC-related protein and AP2VIII-4, respectively. Several years later, while purifying the associated partners of *T. gondii* argonaute, we uncover a second TgHDAC3-containing complex, gathering TBL1 and CRC230 as well as AP2XI-2 (TGME49_310900), that led us to suggest that the deacetylase might play a role in small RNA-induced transcriptional silencing (RITS) complex (Braun et al. 2010).

Surprisingly, we did not originally identify AP2XI-2 when TgHDAC3 was used as bait (Saksouk et al., 2005). Two explanations were then considered: i) TgHDAC3 partners are substoichiometric in tachyzoite and when adding a second copy of TgHDAC3 we are eliciting a competition with the endogenous enzyme that favor partners binding to the latter. ii)

Alternatively, the C-terminal tagging of the second copy of TgHDAC3 may create steric hindrance that likely prevents a proper binding to AP2XI-2. Previous work in human cells have reported that the C-terminal region of HsHDAC3 is required for both deacetylase and transcriptional repression activity (Yang WM et al., 2002).

Therefore, we decided to reassess the interactome of TgHDAC3 by taking the opportunity to use as baits TBL1 and CRC230, two core components regularly seen associated. While we failed to epitope-tagged TBL1, purification of CRC230-associated proteins enabled us to map thoroughly a family of complexes typified by the presence of TgHDAC3, several scaffolding proteins (e.g. ELM2-containing proteins) and 14 ApiAP2 transcription factors, 10 of which were found in both type I (RH) and type II (Pru) strains (Table 1).

It has been difficult so far to determine the exact number of TgHDAC3/CRC230-containing complexes and their composition but this work is under way. Anyhow, considerable care will be taken to determine whether proteins that we have so far identified are *bona fide* components of the TgHDAC3-containing holoenzyme. Reciprocal co-immunoprecipitation and co-localization studies should be performed to validate the associations between TgHDAC3 and its potential direct partners. While we will focus our attention primarily on ApiAP2 factors, these proteins have sizes ranging between 150 and 350 kDa, making their characterization highly challenging.

Interactions of TgHDAC3 with ApiAP2 transcription factors was unexpected and is quite not common to the class I HDAC-associated repressive complexes. Usually, HDACs use co-repressor to mediate their interaction to repressors in a combination that varies from one cell type to another. For instance, the repressor REST which is responsible for modulating neural gene expression was shown to bind directly to the co-repressor subunit CoREST and thereby to recruit its associated repressive core complex gathering HDAC1,2 (Hakimi et al., PNAS 2002).

What was true for TgHDAC3 was also substantiated by the purification of other chromatin-modifying enzymes-containing complexes. For instance, four ApiAP2 factors (AP2IX-7, AP2X-8, AP2XI-2, and AP2XII-4) were identified in association with TgGCN5b, all of which are clearly distinct from those that interact with TgHDAC3 (Wang J et al., 2014). Moreover, our team identified two novel *T. gondii* ATP-dependent remodelers that are associated to distinct ApiAP2 factors (Data not shown). A question then arose: are these interactions with

ApiAP2s direct or mediated by scaffolding proteins? A high-throughput yeast two-hybrid screen using a *P. falciparum* library revealed that ApiAP2 were embedded in a complex network in which they interact directly with histone acetylase (PF08_0034), demethylase (MAL8P1.111) and PHD-containing protein (PF14_0315) (Figure 14) (Lindner et al., 2009). The screen also brought strong evidence that ApiAP2 were able to form homo- and heterodimerize in various combinations, increasing drastically the absolute number of putative TFs beyond the primary number of ApiAP2 encoding genes (Figure 14). Dimerization is expected to create diversity in recognition site specificity or influences the regulation of gene expression.

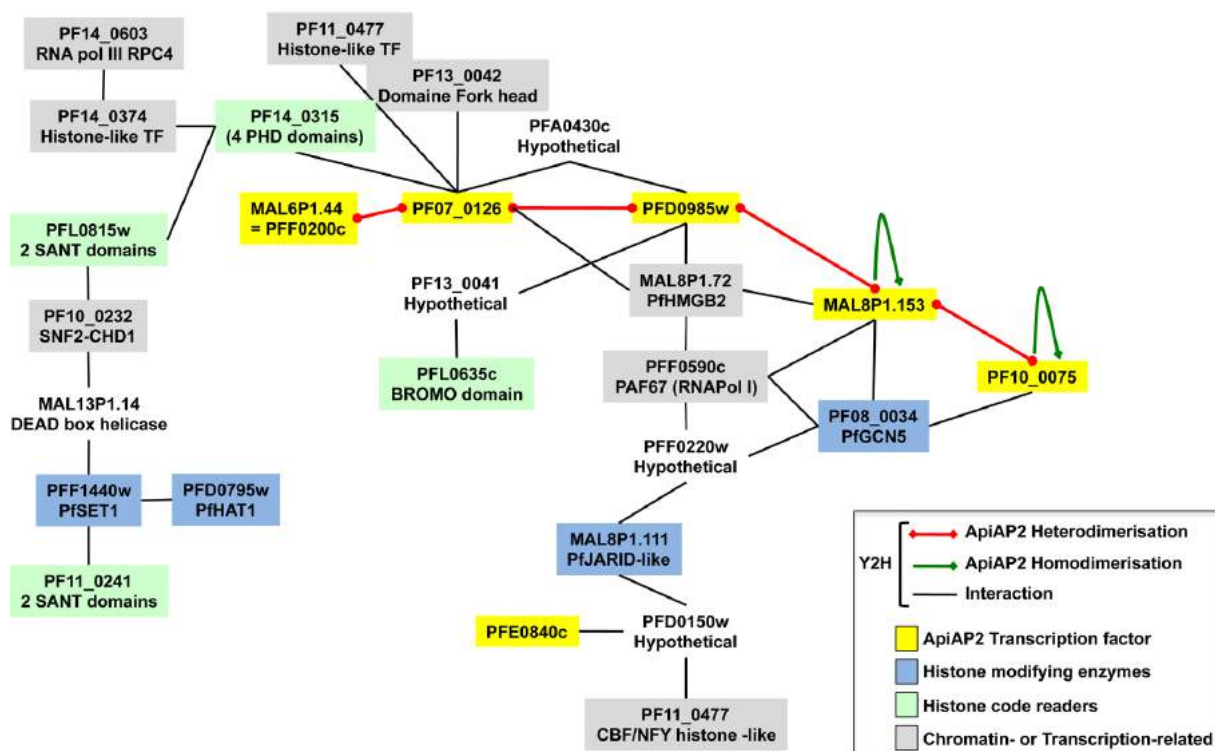


Figure 14. Protein interaction network with six *P. falciparum* ApiAP2 transcription factors. The yeast two-hybrid data are available at <http://www.plasmodb.org>. Figure from Bougdour *et al.*, 2010.

The physiological function(s) of the TgHDAC3-associated ApiAP2 can be also investigated by knocking down the genes. Based on results from a CRISPR genome-wide screen in *Toxoplasma* (Figure 15A, Sidik *et al.*, 2016) we monitored the phenotype, i.e., fitness during infection of human fibroblasts, that was associated to Cas9-mediated gene inactivation of TgHDAC3 and partners (Figure 15B). While we confirmed that CRC230 and HDAC3 are essential we uncovered that 40% of ApiAP2-encoding genes were predicted to be dispensable (Figure 15B). These results show great heterogeneity but clearly indicate the presence of

ApiAP2s that are essential to growth at least in fibroblast that make good candidate as primary regulators of the switch to a different genetic program. But it still remains to be shown.

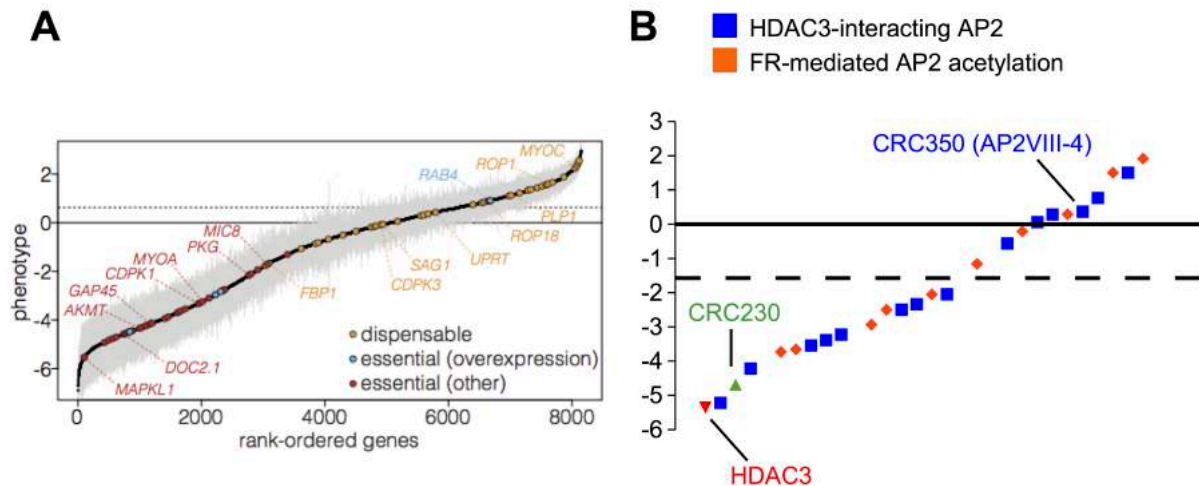


Figure 15. Monitoring TgHDAC3 and partners fitness using a genome-wide CRISPR screen in *Toxoplasma* (A) *T. gondii* genes rank-ordered based on their phenotype. Genes previously reported are highlighted, indicating whether they are dispensable (yellow), or indispensable as inferred from overexpression (blue) or another method (red). Dotted line represents the median phenotype score for the dispensable genes. The two groups are compared in a box plot where whiskers indicate the most extreme data within 1.5 times the interquartile range from the boxed quartiles (right). Image from Sidik et al., 2016, Cell 167, 1–13. (B) TgHDAC3 and partners genes rank-ordered based on their phenotype. Dotted line represents the median phenotype score for the dispensable genes.

Originally, transcription factors (TFs) and their cognate cis-acting binding sites have been difficult to identify in the phylum. Thus far, the Apicomplexan AP2 (ApiAP2) family of DNA-binding proteins are the sole family of proteins to have surfaced as candidate transcription factors in all apicomplexan species (Balaji et al., 2005; Iyer et al., 2008). Work from several laboratories begins to shed light on how the ApiAP2 proteins from *Plasmodium spp.* and *T. gondii* contribute to the regulation of gene expression at various stages of parasite development.

Recombinantly expressed *P. falciparum* AP2 domains bind specifically to a large variety of DNA sequences. Putative binding motifs were identified upstream of the majority of genes, making ApiAP2s the main candidates for generating stage-specific patterns of gene expression (Campbell et al., 2010). PfSIP2 (PF3D7_0604100) was shown to associate uniquely with the SPE2 motifs found at the chromosome ends in the telomere-associated repetitive elements (TAREs) and upstream of *var* genes (Flueck et al., 2010) while

PF3D7_1007700, a protein harboring three AP2 domains, binds a GTGCA motif resembling the rhoptry motif (Campbell et al., 2010). PfAP2Tel, harbouring a non-canonical DNA-binding AP2 domain was shown to bind to *P. falciparum* telomeres (Sierra-Miranda M et al., 2017).

Several ApiAP2 were shown to be active at different stage of parasite development. Thus, PfAP2-I (ApiAP2 involved in invasion) was identified as a key regulator of RBC invasion by the malaria parasite (Santos JM et al., CHM, 2017 - PMID: 28618269). Five ApiAP2 were shown to play key roles in parasite progression through the life cycle and were crucial for gametocytogenesis (ap2-g and ap2-g2; Kafsack et al., 2014; Sinha et al., 2014; Yuda et al., 2015), ookinete development (ap2-o; Yuda et al., 2009), sporozoite formation (ap2-sp; Yuda et al., 2010), and liver stage maturation (ap2-l; Iwanaga et al., 2012). A systematic knockout screen in rodent malaria *Plasmodium berghei* identified ten ApiAP2 genes that were essential for mosquito transmission: four were critical for the formation of infectious ookinetes, and three were required for sporogony. The authors describe non-essential functions for AP2-O and AP2-SP proteins in blood stages, and identify AP2-G2 as a repressor active in both asexual and sexual stages (Modrzynska K et al., 2017).

In *Toxoplasma gondii*, AP2IX-9 restricts *Toxoplasma* commitment to develop the mature bradyzoite tissue cyst (Radke JB et al., 2013). It was recently described a cross-talk between that two alkaline-stress-induced ApiAP2 transcription factors, i.e. AP2IX-9 and AP2IV-3. These factors were expressed in two overlapping waves during bradyzoite development, with AP2IX-9 increasing expression earlier than AP2IV-3, which peaked as AP2IX-9 expression was declining. Disruption of the AP2IX-9 gene enhanced, while deletion of AP2IV-3 gene decreased, tissue cyst formation, demonstrating that these factors have opposite functions in bradyzoite development (Hong DP et al., mSphere, 2017). Finally, AP2IX-4 (TGME49_288950) was shown to display reduced frequencies of tissue cyst formation in culture and in a mouse model of infection (Huang S et al., mSphere 2017). However, the functions of most members of the family remain unknown, including those identified in TgHDAC3 partnership.

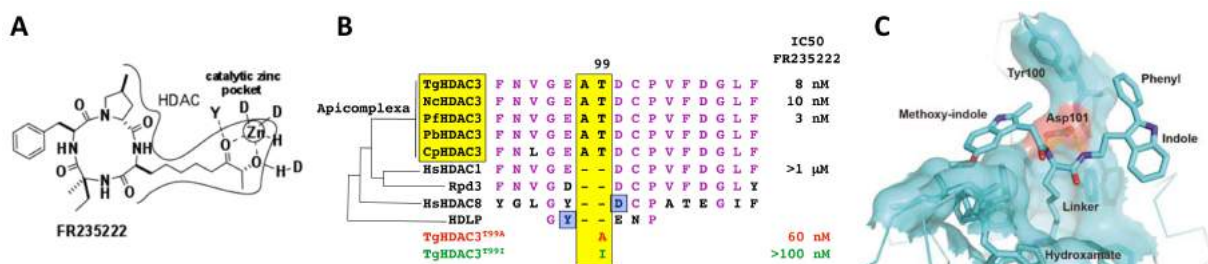
Apparently, ApiAP2 genes act together to create complex patterns of gene expression. This is reminiscent of model eukaryotes, where complex gene expression patterns that determined cell identity are the result of hierarchical networks, within which TFs function in combination and occasionally synergistically (Gerstein et al., 2012; Levo and Segal, 2014). It appears that

the apicomplexan transcriptomes are regulated in a similar way and that ApiAP2 proteins are key factors involved in this process that associate/recruit histone modifying enzymes.

III- 5. Deacetylation by TgHDAC3 of ApiAP2 transcription factors: the substrate hypothesis?

It is now understood that protein acetylation represents an additional level of regulation for multiple enzymes and, considering the prevalence of this modification on many other proteins, could dictate the flux of many other cellular processes. *In Toxoplasma*, a number of HATs have been observed to localize predominantly in the parasite cytosol, suggesting that they may also target non-histone substrates (Smith AT, et al., Eukaryot Cell, 2005). Sullivan's lab report the first 'acetylome' of tachyzoites and therefore they identified over 400 novel acetylation sites on a wide variety of proteins throughout the parasite cell, including those with roles in transcription, translation, metabolism and stress responses (Jeffers and Sullivan, 2012 ; Xue B et al., 2013). We initially thought that HDACi may have further mechanisms of action beyond the dysregulation of gene expression and this would explain the FR235222-mediated pleiotropic phenotypes aforementioned. Intriguingly, we did not find any transcriptional regulators whose expression at the protein level was affected by FR235222 treatment. However, while looking at the corresponding acetylome we uncover an extensive set of ApiAP2 and chromatin-modifying enzymes that were acetylated at specific residue upon drug treatment, expanding drastically the complexity of the TgHDAC3-regulated network.

TgHDAC3 is therefore able to accommodate different substrates apart from PTMs in histone tails or core (e.g. H4K31ac). This characteristic may lie in the particular structure of its catalytic site of the enzyme. Indeed, the residue T99 along with the amino acid A98 creates an insertion within the catalytic site of the enzyme that is exclusively conserved in apicomplexan HDAC3 family of proteins and absent in any other eukaryotic HDAC (Figure 16).



Based on sequence homology, mutations in TgHDAC3 (T99A and T99I) conferring resistance to FR235222 localize to the L2 loop of HDLP, where the residue Y91 localized at the rim of the active site contacts the cognate HDACi, TSA (Figure 16B) (Finnin, M.S. et al. 1999). Structural data of HsHDAC8 pointed out the role of the residue D101 in both substrate and HDACi recognition (Figure 16C); HsHDAC8D101A mutated enzyme was inactive on protein substrates and binding efficiency to hydroxamate inhibitor was decreased (Vannini, AC et al., EMBO Rep. 8:879–884). Given that D101 is localized in the vicinity of T99 of TgHDAC3, these data further strengthen the hypothesis of a direct inhibition of TgHDAC3 by FR235222 and are consistent with a role of T99 in the interactions with cyclopeptide inhibitors. T99A and T99I change amino acid polarity; it is therefore tempting to speculate that polar interactions at the rim of the active site support the binding to HDACi's but also expand the range of the substrates that can be handle by the enzyme.

The first evidence that the AT insertion may contribute to change the specificity of the TgHDAC3 were brought earlier by our team while monitoring P36 labeling in the TgHDAC3^{T99A} mutant background (Bougdoor et al., 2009). Indeed, they showed that the bradyzoite marker became constitutively expressed in the TgHDAC3^{T99A} mutant compared to wild-type parasites (Figure 5A). This suggests that a simple point mutation in this insertion is able to affect the basal activity of the enzyme and that was confirmed on histone acetylation (Figure 5B, compare the yield of H4 enrichment in the absence of FR235222 in WT versus mutant background). While these observations are interesting, they do not provide conclusive evidence on how TgHDAC3 selectively discriminates its substrates.

III- 6. Conclusion

The ability to perform genetic crosses, to engineer transgenic parasites lines, and the wealth of information made available through recent genome sequencing projects have made the laboratory study of these parasites important not only for understanding the diseases that they cause, but also for gaining insights into basic biological processes. During my PhD I brought new evidence that the HDAC3 family plays a role in gene expression, differentiation, and cell-cycle control. Drug inhibition of TgHDAC3 prevents the formation of the daughter cells. Moreover, parasites treated with low doses of FR235222 are committed to differentiate into bradyzoites. The emerging picture is that epigenetic changes are linked to stage conversion and virulence, although more data are needed to fully grasp the versatility and the complexity of the mechanism involved.

Part IV. Materials and methods

IV- 1. Parasites and host cells

HFF and MEF primary cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Invitrogen) supplemented with 10% heat inactivated Fetal Bovine Serum (FBS) (Invitrogen), 10mM (4-(2-hydroxyethyl)-1-piperazine ethanesulphonic acid) (HEPES) buffer pH 7.2, 2 mM L-glutamine and 50 µg ml of penicillin and streptomycin (Invitrogen). Cells were incubated at 37°C in 5% CO₂. Type I (RH wild type and RH $\Delta ku80$) and type II strains (Pru $\Delta ku80$) of *T. gondii* were maintained *in vitro* by serial passage on monolayers of HFFs. *Plasmodium falciparum* 3D7 strain was grown in RPMI 1640 media supplemented with 0.5% Albumax II, 0.1mM Hypoxanthine and Gentamicin 10 mcg/ml. The culture was maintained at 2% hematocrit and 5% parasitemia. The parasites were kept at 37°C and at 1% O₂, 5% CO₂ and 94% N₂ gas mixture concentration.

IV- 2.HDACi treatments

The final concentration of histone deacetylase inhibitors dissolved in DMSO was, as described by Bougdour A *et al.*, 2013, FR-235222 (90nM), apicidin (100nM), HC-toxin (100nM), trichostatin A (100nM), scriptaid (100nM), APHA (100mM) and sodium butyrate (5mM). They were added to infected HFF cells for 18 hours. Halofuginone (10 nM) was shown to inhibit prolyl-tRNA synthetase (Jain V *et al.*, 2015) and was used as a control.

IV- 3. Plasmid constructs

For endogenous gene tagging, we used different genes of interest (GOI) such as TGME49_268860 (ENO1), TGME49_291040 (LDH2), TGME49_207210, TGME49_216140, TGME49_262110, TGME49_316130, TGME49_20m.00351 and TGME49_305340 (CRC230) to generate stable transgenic parasites. To construct vector pLIC-TgGOI-HAFlag, the coding sequence of TgGOI was amplified using Pru $\Delta ku80$ and RH $\Delta ku80$ genomics DNA as templates and primers pLIC-GOI_F and pLIC- GOI_R. The resulting PCR product was cloned into the pLIC-HAFlag-*dhfr* vector (Bougdour A *et al.*, 2013) using the LIC cloning method as reported previously (Huynh MH and Carruthers VB, 2009) yielding pLIC-TgGOI-HA-DHFR. Cloning primers used in this study:

LIC-20.m00351_Fwd: 5'-TACTTCCAATCCAATTTAGCacggcagcatgctctggaacggtctgta-3'

LIC-20.m00351_Rev: 5'-TCCTCCACTTCCAATTTTAGCagctgtgtgagaatgtgcccgtcggtta-3'
 LIC-2072010_Fwd: 5'-TACTTCCAATCCAATTTAGCactgtaaatagaggaagacactgcactg-3'
 LIC-2072010_Rev: 5'-TCCTCCACTTCCAATTTTAGCtggtcttcttccccctccctttctagc-3'
 LIC-216140_Fwd: 5'-TACTTCCAATCCAATTTAGCtccgaagaccatccatgaatattcatgg-3'
 LIC-216140_Rev: 5'-TCCTCCACTTCCAATTTTAGCgccgtttatctcgaccacggatggcgg-3'
 LIC-268860_Fwd: 5'-TACTTCCAATCCAATTTAGCgaacatgcaggcaatggcttggtcttc-3'
 LIC-268860_Rev: 5'-TCCTCCACTTCCAATTTTAGCtttgggtgtcgaagctctctcccgcg-3'
 LIC-291040_Fwd: 5'-TACTTCCAATCCAATTTAGCgttggatgattcacaggcgacaagcattg-3'
 LIC-291040_Rev: 5'-TCCTCCACTTCCAATTTTAGCaccagcgccgctaaactcttattcaattc-3'
 LIC-262110_Fwd: 5'-TACTTCCAATCCAATTTAGCagcgtctctcaccccaagtcggcagttc-3'
 LIC-262110_Rev: 5'-TCCTCCACTTCCAATTTTAGCctgggtctggccatcttggcttggctc-3'
 LIC-316130_Fwd: 5'-TACTTCCAATCCAATTTAGCgctttgatggacaagcaaatgcagttc-3'
 LIC-316130_Rev: 5'-TCCTCCACTTCCAATTTTAGCacttcgaatgtcgtctgggacaatttc-3'
 LIC-305340_Fwd:
 5'-TACTTCCAATCCAATTTAGCACGACGCCTTCGGGAGTACCACGAGGAG-3'
 LIC-305340_Rev:
 5'-TCCTCCACTTCCAATTTTAGCCACAATCTTCGCTTCTCCATCAACCTCTG-3'

IV- 4. Cas9-mediated gene disruption

The plasmid pTOXO_Cas9-CRISPR was described in (Sangaré O et al, 2016). For gene disruption using CRISPR/Cas9 system, the GOI were: GCN5A (*TGGT1_254555*), GCN5B (*TGGT1_243440*), MYST-A (*TGGT1_318330*), MYST-B (*TGGT1_207080*), HAT1 (*TGGT1_293380*), HDAC1 (*TGGT1_281420*), HDAC2 (*TGGT1_249620*), HDAC3 (*TGGT1_227290*), HDAC4 (*TGGT1_257790*) and HDAC5 (*TGGT1_202230*). Twenty mers-oligonucleotides corresponding to specific GOI were cloned using Golden Gate strategy. Briefly, primers TgGOI-CRISP_FWD and TgGOI-CRISP_REV containing the sgRNA targeting TgGOI genomic sequence were phosphorylated, annealed and ligated into the linearized pTOXO_Cas9-CRISPR plasmid with BsaI, leading to pTOXO_Cas9-CRISPR::sgTgGOI. *Toxoplasma* tachyzoites were then transfected with the plasmid and grown on HFF cells for 18-36 hours. Cloning oligonucleotides used in this study:

TgHDAC1-CRISP-FWD: 5'- AAGTTGCGTCGCCGTTCTCTCACGCG -3'

TgHDAC1-CRISP-REV: 5'- AAAACGCGTGAGAGAACGGCGACGCA -3'

TgHDAC2-CRISP-FWD: 5'- AAGTTGCGCCCGTCGCCTCCCCCGCG -3'
TgHDAC2-CRISP-REV: 5'- AAAACGCGGGGGAGGCGACGGGGCGCA -3'
TgHDAC3-CRISP-FWD: 5'- AAGTTGATATCGGAAGTTACTACTAG -3'
TgHDAC3-CRISP-REV: 5'- AAAACTAGTAGTAACTTCCGATATCA -3'
TgHDAC4-CRISP-FWD: 5'- AAGTTGCTGTTGCTGAAGCCCAGGCG -3'
TgHDAC4-CRISP-REV: 5'- AAAACGCCTGGGCTTCAGCAACAGCA -3'
TgHDAC5-CRISP-FWD: 5'- AAGTTGGCGAGACCGGGGCAGCCGCG -3'
TgHDAC5-CRISP-REV: 5'- AAAACGCGGCTGCCCCGGTCTCGCCA -3'
TgGCN5A-CRISP-FWD: 5'- AAGTTGCGTGACGAACGACAGGCAAG -3'
TgGCN5A-CRISP-REV: 5'- AAAACTTGCCTGTCGTTTCGTCACGCA -3'
TgGCN5B-CRISP-FWD: 5'- AAGTTGGGTTTCTGTGTCGAGACCG -3'
TgGCN5B-CRISP-REV: 5'- AAAACGGTCTCGACACAGGAAACCCA -3'
TgMYSTA-CRISP-FWD: 5'- AAGTTGGCTGCTCCGCGACTCAGCGG -3'
TgMYSTA-CRISP-REV: 5'- AAAACCGCTGAGTCGCGGAGCAGCCA -3'
TgMYSTB-CRISP-FWD: 5'- AAGTTGCGCGAAGAAGGGAGAGAGCG -3'
TgMYSTB-CRISP-REV: 5'- AAAACGCTCTCTCCCTTCTTCGCGCA -3'
TgHAT1-CRISP-FWD: 5'- AAGTTGCCGACGGGTCACGGAGACTG -3'
TgHAT1-CRISP-REV: 5'- AAAACAGTCTCCGTGACCCGTCGGCA -3'

IV- 5. Antibodies

Primary antibodies : rabbit home-made anti-TgHDAC3 described in (Bougdour A *et al.*, 2009), mouse anti-HA (3F10, Roche), rabbit anti-H4K8ac (Upstate 06-760), rabbit anti-H4K12ac (Upstate 06-761), rabbit anti-H3K4ac (Diagenode C15410165), rabbit anti-H3K9ac (Diagenode C15410004), rabbit anti-H3K14ac (C15210005), rabbit anti-H3K18ac (Diagenode C15410193), rabbit anti-H3K27ac (Millipore, 04-1044-S) and mouse anti-H3K27ac (Diagenode C15200184), H4K20me3 (C15410207), H3K9me3 (Millipore, 17-625), H3K4me1 (C15410194) and H3K4me3 (C15410003-50). Western blot secondary antibodies were conjugated to alkaline phosphatase (Promega), while immunofluorescence secondary antibodies were coupled with Alexa Fluor 488 or Alexa Fluor 594 (Thermo Fisher Scientific). We also raised homemade H4K31acetylation and H4K31monomethylation-specific antibodies in rabbit against linear peptides corresponding to amino acid residues 23/24-35 of histone H4 and carrying modified residue K31: C-DNIQGITKme1PAIR; C-

DNIQGITK_{ac}PAIR and C-RDNIQGITK_{ac}PAIR. They were produced by Eurogentec and used for immunofluorescence, immunoblotting and chromatin immunoprecipitation.

IV- 6. *Toxoplasma gondii* transfection

Toxoplasma gondii RH, RH $\Delta ku80$ and Pru $\Delta ku80$ were electroporated with vectors in cytomix buffer (120mM KCl, 0.15mM CaCl₂, 10mM K₂HPO₄/ KH₂PO₄ pH7.6, 25mM HEPES pH7.6, 2mM EGTA, 5mM MgCl₂) using a BTX ECM 630 machine (Harvard Apparatus). Electroporation was performed in a 2mm cuvette at 1.100V, 25 Ω and 25 μ F. Stable transgenic parasites were selected with 1 μ M pyrimethamine, single-cloned in 96 well plates by limiting dilution and verified by immunofluorescence assay.

IV- 7. Immunofluorescence microscopy

Toxoplasma infecting HFF cells grown on coverslips were fixed in 3% formaldehyde for 20 min at room temperature, permeabilized with 0.1% (v/v) Triton X-100 for 15 min and blocked in Phosphate buffered saline (PBS) containing 3% (w/v) Bovine Serum Albumine (BSA). The cells were then incubated for 1 hour with primary antibodies followed by the addition of secondary antibodies conjugated to Alexa Fluor 488 or 594 (Molecular Probes). Nuclei were stained for 10 min at room temperature with Hoechst 33258. Coverslips were mounted on a glass slide with Mowiol mounting medium, and images were acquired with a fluorescence ZEISS ApoTome.2 microscope and images were processed by ZEN software (Zeiss). *Plasmodium* asexual stages infecting red blood cells were washed with phosphate-buffered saline (PBS) and fixed in solution with 4% paraformaldehyde and 0.0075% glutaraldehyde in PBS for 30 min. After one wash with PBS, cells were permeabilized with 0.1% Triton X-100 in PBS for 10 min. Cells were washed twice with PBS, blocked with 3% BSA in PBS for 1 hour. The cells were then incubated for 1 hour with primary antibodies followed by the addition of secondary antibodies conjugated to Alexa Fluor 488 or 594 (Molecular Probes). Nuclei were stained for 30 min at room temperature with Hoechst 33258. The parasites were finally washed 2-3 times before loading on to glass slides mixed with fluoro-gel. Images were acquired with a fluorescence ZEISS ApoTome.2 microscope and images were processed by ZEN software (Zeiss).

IV- 8. Protein extraction and Trypsin Digestion

HFF cells were grown to confluence, infected with Pru $\Delta ku80$ parasites and treated with 90nM FR235222 for 18 hours. As a control, we used 0.1% DMSO treatment. Harvested intracellular parasites were washed in PBS, lysed in ice-cold urea lysis buffer (8 M urea, 20mM HEPES-KOH pH7.5) and centrifuged for 10 minutes at 1000g at 4°C. Supernatants were collected and protein concentration was measured followed by reduction and alkylation of cysteines. Proteins were digested subsequently with endo-LysC and trypsin. Peptides were then desalted over C18 columns. Digested peptides were used for FR235222-responsive global proteome and acetylome.

IV- 9. Affinity enrichment of lysine acetylated peptides

We used a pan acetyl lysine antibody. Tryptic peptides were dissolved in immunoaffinity purification IAP buffer (50mM morpholinepropane sulfonic acid MOPS, 10mM sodium phosphate, 50mM NaCl pH 7.2) and incubated with pre-washed antibody beads (with IAP buffer and with water) at 4°C overnight. Enriched peptides were eluted from the beads with 0.15% TFA, vacuum-dried and analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS).

IV- 10. Affinity purification of Flag-tagged proteins

HFF cells were grown to confluence and infected with Pru $\Delta ku80$ and RH $\Delta ku80$ parasites expressing the endogenous protein in fusion with the HA-Flag tags. Intracellular and extracellular tachyzoites were collected by centrifugation at 4°C for 10 min at 1000xg. The cell pellets were washed in PBS1X by centrifuging at 4°C for 10 min at 1000xg and the supernatant was discarded. The whole-cell extract containing Flag-tagged protein was lysed in lysis buffer (20% glycerol, 20mM Tris-HCl pH8, 500mM KCl, 1.5mM MgCl₂, 0.5% NP-40, 100mM PMSF, 0.5mM DTT, plus complete EDTA-free protease inhibitor cocktail) and incubated at 4°C for 45 min with 5rpm rotation speed. After centrifugation at 20 000xg for 45 min at 4°C, the proteins were incubated with 500 μ l anti-Flag M2 affinity gel (Sigma- Aldrich) at 4°C for 1 hour at 5 rpm rotation speed. Beads were washed with 10- column volumes of BC500 buffer (20% glycerol, 20mM Tris-HCl pH8, 500mM KCl, 1.5mM MgCl₂, 0.5% NP-

40, 100mM PMSF, and 0.5mM DTT). Bound polypeptides were eluted with 250µg/ml Flag peptide diluted in BC100 buffer. For size-exclusion chromatography, protein eluates were loaded onto Superose 6 HR 10/30 column equilibrated with BC500. Flow rate was fixed at 0.35ml/min, and 0.5ml fractions were collected.

IV- 11. Mass spectrometry and peptide sequencing

Protein bands were excised from colloidal blue-stained gels (Invitrogen), treated with DTT and iodoacetamide to alkylate the cysteines, and then immediately subjected to in-gel tryptic digestion. Peptides were extracted with 5% vol/vol formic acid solution and acetonitrile, and injected into an Ultimate 3000 (Dionex) nanoLC system that was directly coupled to a LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific). MS and MS/MS data were acquired using Xcalibur (Thermo Fisher Scientific) and processed automatically using Mascot Daemon software (Matrix Science). Tandem mass spectra were searched against a compiled *T. gondii* database using the MASCOT program (Matrix Sciences).

IV- 12. Histones purification, Immunoblotting and mass spectrometry analysis

For histone purification, HFF cells were grown to confluence and infected with PruΔ*ku80* parasites. Intracellular tachyzoites were treated with histone deacetylase HDAC3 inhibitor, 90nM FR235222 for 18 hours. As appropriate control, we treated tachyzoites with 0.1% DMSO. Histones were extracted and purified using histone purification kit (Active motif) according to manufacturer's protocol. For western blotting, histone proteins were run on a NuPAGE 4-12% Bis-Tris polyacrylamide gels in MES-SDS running buffer (Invitrogen) and transferred to a polyvinylidene fluoride PVDF membrane (Immobilon-P; Millipore) using NuPAGE transfer buffer (Invitrogen). The blots were probed using primary antibodies: pan acetyl H4, H4K31ac and H4K31me1, followed by phosphatase-conjugated goat secondary antibodies (*Promega*). The expected band of histones were detected using NBT-BCIP (*Amresco*). Nucleosomes from *Toxoplasma gondii*-infected cells were purified and proteins separated by SDS-PAGE. The band corresponding to H4 was excised and submitted to mass spectrometry-based proteomic analysis. The presented MS/MS spectrum was identified using

Mascot search engine and shows the acetylation of K31 residue in the DNIQGITKPAIR peptide.

IV- 13. Chromatin Immunoprecipitation and Next Generation Sequencing in *T.gondii*

HFF cells were grown to confluence and infected with Pru $\Delta ku80$ strain. Harvested intracellular parasites were crosslinked with formaldehyde (final concentration 1%) for 8 mins at room temperature and the crosslinking was stopped by addition of glycine (final concentration 0.125M) for 5 min at room temperature. Crosslinked chromatin was lysed in ice-cold lysis buffer (50mM HEPES KOH pH7.5, 140mM NaCl, 1mM EDTA, 10% glycerol, 0.5%NP-40, 0.125% triton X-100, protease inhibitor cocktail) and sheared in shearing buffer (1mM EDTA pH8.0, 0.5mM EGTA pH8.0, 10mM Tris pH8.0, protease inhibitor cocktail) by sonication using a Diagenode Biorupter. Samples were sonicated, for 16 cycles (30 seconds ON and 30 seconds OFF), to 200-500 base-pair average size. Immunoprecipitation was carried out using sheared chromatin, 5% BSA, protease inhibitor cocktail, 10% triton X-100, 10% deoxycholate, DiaMag Protein A-coated magnetic beads (Diagenode) and antibodies (H4K31ac, H4K31me1, pan acetyl H4, H4K20me3, H3K9me3, H3K4me3, H3K4me1, H3K14ac). A rabbit IgG antiserum was used as a control mock. After overnight incubation at 4°C on rotating wheel, chromatin-antibody complexes were washed and eluted from beads by using iDeal ChIP-seq kit for Histones (Diagenode) according to the manufacturer's protocol. Samples were decrosslinked by heating for 4 hours at 65°C. DNA was purified by using IPure kit (Diagenode) and quantified by using Qubit Assays (*Thermo Fisher Scientific*) according to the manufacturer's protocol. For ChIP-seq, purified DNA was used to prepare libraries and then sequenced by Arraystar (USA, <http://www.arraystar.com/>).

IV- 14. Library Preparation, Sequencing and Data analysis (Arraystar)

ChIP-Sequencing library preparation was performed according to Illumina's protocol Preparing Samples for ChIP Sequencing of DNA. **Library Preparation:** 10 ng DNA of each sample was converted to phosphorylated blunt-ended with T4 DNA polymerase, Klenow polymerase and T4 polymerase (NEB); An 'A' base was added to the 3' end of the blunt phosphorylated DNA fragments using the polymerase activity of Klenow (exo minus)

polymerase (NEB); Illumina's genomic adapters were ligated to the A tailed DNA fragments; PCR amplification was performed to enrich ligated fragments using Phusion High Fidelity PCR Master Mix with HF Buffer (Finnzymes Oy). The enriched product of ~200-700 bp was cut out from gel and purified. **Sequencing:** The library was denatured with 0.1M NaOH to generate single-stranded DNA molecules, and loaded onto channels of the flow cell at 8pM concentration, amplified in situ using TruSeq Rapid SR cluster kit (#GD-402-4001, Illumina). Sequencing was carried out by running 100 cycles on Illumina HiSeq 4000 according to the manufacturer's instructions. **Data analysis:** After the sequencing platform generated the sequencing images, the stages of image analysis and base calling were performed using Off-Line Basecaller software (OLB V1.8). After passing Solexa CHASTITY quality filter, the clean reads were aligned to *Toxoplasma gondii* reference genome (Tgo) using BOWTIE (V2.1.0). Aligned reads were used for peak calling of the ChIP regions using MACS V1.4.0. Statistically significant ChIP-enriched regions (peaks) were identified by comparison of two samples, using a p-value threshold of 10^{-5} . Then the peaks in each sample were annotated by the overlapped gene using the newest *T.gondii* database. The EXCEL/BED format file containing the ChIP-enriched regions was generated for each sample. **Data visualization:** The mapped 100 bp reads represent enriched DNA fragments by ChIP experiment. Any region of interest in the raw ChIP-seq signal profile can be directly visualized in genome browser. We use 10-bp resolution intervals (10-bp bins) to partition the stacked reads region, and count the number of reads in each bin. All the 10 bp resolution ChIP-seq profiles of each sample are saved as UCSC wig format files, which can be visualized in *T.gondii* Genome Browser (http://protists.ensembl.org/Toxoplasma_gondii/Info/Index). All these raw and processed files can be found at GEO DATASET number.

Part V. References

- Ahringer, J. (2000). NuRD and SIN3 histone deacetylase complexes in development. *Trends Genet.* *16*,351-356.
- Altschul, S.F., Wootton, J.C., Zaslavsky, E., Yu, Y.K. (2010). The construction and use of log-odds substitution scores for multiple sequence alignment. *PLOS Comput Biol.* *6*, e1000852.
- Arisue, N., Hashimoto, T. (2015). Phylogeny and evolution of apicoplasts and apicomplexan parasites. *Parasitol Int.* *64*,254-259.
- Balaji, S., Babu, M.M., Iyer, L.M., Aravind, L. (2005). Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.* *33*, 3994-4006.
- Behnke, M.S., Radke, J.B., Smith, A.T., Sullivan, WJ. Jr., White, M.W. (2008). The transcription of bradyzoite genes in *Toxoplasma gondii* is controlled by autonomous promoter elements. *Mol Microbiol.* *68*, 1502-1518.
- Behnke, M.S., Wootton, J.C., Lehmann, M.M., Radke, J.B., Lucas, O., Nawas, J., Sibley, L.D., White, M.W. (2010). Coordinated progression through two subtranscriptomes underlies the tachyzoite cycle of *Toxoplasma gondii*. *PLoS One.* *5*, e12354.
- Bertrand, P. (2010). Inside HDAC with HDAC inhibitors. *Eur J Med Chem.* *45*, 2095-2116.
- Bhatti, M.M., Livingston, M., Mullapudi, N., Sullivan, WJ. Jr. (2006). Pair of unusual GCN5 histone acetyltransferases and ADA2 homologues in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell.* *5*, 62-76.
- Blackman, M.J., Bannister, L.H. (2001). Apical organelles of Apicomplexa: biology and isolation by subcellular fractionation. *Molecular & Biochemical Parasitology.* *117*, 11-25.
- Black, M.W., Boothroyd, J.C. (2000). Lytic Cycle of *Toxoplasma gondii*. *Microbiology and Molecular Biology Reviews.* *64*, 607-623.
- Bogado, S.S., Dalmasso, M.C., Ganuza, A., Kim, K., Sullivan, WJ. Jr., Angel, S.O., Vanagas, L. (2014). Canonical histone H2Ba and H2A.X dimerize in an opposite genomic localization to H2A.Z/H2B.Z dimers in *Toxoplasma gondii*. *Mol Biochem Parasitol.* *197*, 36-42
- Bohne, W., Heesemann, J., Gross, U. (1993). Induction of bradyzoite-specific *Toxoplasma gondii* antigens in gamma interferon-treated mouse macrophages. *Infect Immun.* *61*, 1141-1145.
- Bohne, W., Roos, D.S. (1997). Stage-specific expression of a selectable marker in *Toxoplasma gondii* permits selective inhibition of either tachyzoites or bradyzoites. *Mol Biochem Parasitol.* *88*,115-126.
- Boothroyd, J.C. (2009). *Toxoplasma gondii*: 25 years and 25 major advances for the field. *Int J Parasitol.* *39*, 935-46.

- Bouchut, A., Chawla, A.R., Jeffers, V., Hudmon, A., Sullivan, WJ. Jr. (2015). Proteome-wide lysine acetylation in cortical astrocytes and alterations that occur during infection with brain parasite *Toxoplasma gondii*. PLoS One. *10*, e0117966.
- Bougdour, A., Braun, L., Cannella, D., Hakimi, M.-A. (2010). Chromatin modifications: implications in the regulation of gene expression in *Toxoplasma gondii*. Cell Microbiol. *12*,413-423.
- Bougdour, A., Durandau, E., Brenier-Pinchart, M.-P., Ortet, P., Barakat, M., Kieffer, S., Curt Varesano, A., Curt-Bertini, R.-L., Bastien, O., Coute, Y., Pelloux, H., Hakimi, M.-A. (2013). Host cell subversion by *Toxoplasma* GRA16, an exported dense granule protein that targets the host cell nucleus and alters gene expression. Cell Host Microbe. *13*, 489-500.
- Bougdour, A., Maubon, D., Baldacci, P., Ortet, P., Bastien, O., Bouillon, A., Barale, J.C., Pelloux, H., Ménard, R., Hakimi, M.-A. (2009). Drug inhibition of HDAC3 and epigenetic control of differentiation in Apicomplexa parasites. J Exp Med. *206*, 953-966.
- Bozdech, Z., Llinás, M., Pulliam, B.L., Wong, E.D., Zhu, J., DeRisi, J.L. (2003). The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. PLoS Biol. *1*, e5.
- Braun, L., Cannella, D., Ortet, P., Barakat, M., Sautel, C.F., Kieffer, S., Garin, J., Bastien, O., Voinnet, O., Hakimi, M.-A. (2010). A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite *Toxoplasma gondii*. PLoS Pathog. *6*:e1000920.
- Brooks, C.F., Francia, M.E., Gissot, M., Croken, M.M., Kim, K., Striepen, B. (2011). *Toxoplasma gondii* sequesters centromeres to a specific nuclear region throughout the cell cycle. Proc Natl Acad Sci U S A. *108*, 3767-3772.
- Bunnik, E.M., Polishko, A., Prudhomme, J., Ponts, N., Gill, S.S., Lonardi, S., Le Roch, K.G. (2014). DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite *Plasmodium falciparum*. BMC Genomics. *15*, 347.
- Campbell, T.L., De Silva, E.K., Olszewski, K.L., Elemento, O., Llinás, M. (2010). Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. PLoS Pathog. *6*, e1001165.
- Chen A, L., Kim, E.W., Toh, J.Y., Vashisht, A.A., Rashoff, A.Q., Van, C., Huang, A.S., Moon, A.S., Bell, H.N., Bentolila, L.A., Wohlschlegel, J.A., Bradley, P.J. (2015). Novel components of the *Toxoplasma* inner membrane complex revealed by BioID. MBio. *6*, e02357-14.
- Chen, C.C., Carson, J.J., Feser, J., Tamburini, B., Zabaronick, S., Linger, J., Tyler, J.K. (2008). Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair. Cell. *134*, 231-243.

- Choudhary, C., Kumar, C., Gnad, F., Nielsen, M.L., Rehman, M., Walther, T.C., Olsen, J.V., Mann, M. (2009). Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science*. 325, 834-840.
- Cleary, M.D., Singh, U., Blader, I.J., Brewer, J.L., Boothroyd, J.C. (2002). *Toxoplasma gondii* asexual development: identification of developmentally regulated genes and distinct patterns of gene expression. *Eukaryot. Cell*. 1, 329-340.
- Close, P., Creppe, C., Gillard, M., Ladang, A., Chapelle, J.P., Nguyen, L., Chariot, A. (2010). The emerging role of lysine acetylation of non-nuclear proteins. *Cell Mol Life Sci*. 267, 1255-64.
- Coetzee, N., Sidoli, S., van Biljon, R., Painter, H., Llinás, M., Garcia, B.A., Birkholtz, L.M. (2017). Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual *Plasmodium falciparum* parasites. *Sci Rep*. 7, 607.
- Coleman, B.I., Skillman, K.M., Jiang, R.H.Y., Childs, L.M., Altenhofen, L.M, Ganter, M., Leung, Y., Goldowitz, I., Kafsack, B.F.C., Marti, M., Llinás, M., Buckee, C.O., Duraisingh, M.T. (2014). A *Plasmodium falciparum* histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host Microbe*. 16,177-186.
- Dalmaso, M.C., Echeverria, P.C., Zappia, M.P., Hellman, U., Dubremetz, J.F., Angel, S.O. (2006). *Toxoplasma gondii* has two lineages of histones 2b (H2B) with different expression profiles. *Mol Biochem Parasitol*. 148, 103-107.
- Dalmaso, M.C., Onyango, D.O., Naguleswaran, A., Sullivan, WJ. Jr., Angel, S.O. (2009). *Toxoplasma* H2A variants reveal novel insights into nucleosome composition and functions for this histone family. *J Mol Biol*. 392, 33-47.
- Dalmaso, M.C., Sullivan, WJ. Jr., Angel, S.O. (2011). Canonical and variant histones of protozoan parasites. *Front Biosci (Landmark Ed)*. 16, 2086-2105.
- Das, C., Lucia, M.S., Hansen, K.C., Tyler, J.K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature*. 459, 113-117.
- Daujat, S., Weiss, T., Mohn, F., Lange, U.C., Ziegler-Birling, C., Zeissler, U., Lappe, M., Schübeler, D., Torres-Padilla, M.E., Schneider, R. (2009). H3K64 trimethylation marks heterochromatin and is dynamically remodeled during developmental reprogramming. *Nat Struct Mol Biol*. 16, 777-781.
- de Ruijter, A.J., van Gennip, A.H., Caron, H.N., Kemp, S., van Kuilenburg, A.B. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem J*. 370, 737-749.
- De Silva, E.K., Gehrke, A.R., Olszewski, K., León, I., Chahal, J.S., Bulyk, M.L., Llinás, M. (2008). Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc Natl Acad Sci USA*. 105, 8393-8398.

- Di Cerbo, V., Mohn, F., Ryan, D.P., Montellier, E., Kacem, S., Tropberger, P., Kallis, E., Holzner, M., Hoerner, L., Feldmann, A., Richter, F.M., Bannister, A.J., Mittler, G., Michaelis, J., Khochbin, S., Feil, R., Schuebeler, D., Owen-Hughes, T., Daujat, S., Schneider, R. (2014). Acetylation of histone H3 at lysine 64 regulates nucleosome dynamics and facilitates transcription. *Elife*. 3, e01632.
- Di Cristina, M., Marocco, D., Galizi, R., Proietti, C., Spaccapelo, R., Crisanti, A. (2008). Temporal and Spatial Distribution of *Toxoplasma gondii* Differentiation into Bradyzoites and Tissue Cyst Formation *in vivo*. *Infection and Immunity*. 76, 3491-3501.
- Ding, Z., Gillespie, L.L., Paterno, G.D. (2003). Human MI-ER1 alpha and beta function as transcriptional repressors by recruitment of histone deacetylase 1 to their conserved ELM2 domain. *Mol Cell Biol*. 23, 250-258.
- Dixon, S.E., Stilger, K. L., Elias, E.V., Naguleswaran, A., Sullivan, WJ. Jr. (2010). A decade of epigenetic research in *Toxoplasma gondii*. *Mol Biochem Parasitol*. 173, 1-9.
- Doerig, C., Rayner, J.C., Scherf, A., Tobin, A.B. (2015). Post-translational protein modifications in malaria parasites. *Nat Rev Microbiol*. 13, 160-172.
- Donahoe, S.L., Lindsay, S.A., Krockenberger, M., Phalen, D., Šlapeta, J. (2015). A review of neosporosis and pathologic findings of *Neospora caninum* infection in wildlife. *Int J Parasitol Parasites Wildl*. 4, 216-238.
- Dubey, J.P. (2009). *International Journal for Parasitology*. 39, 877-882.
- Dubey, J.P. (2004). Toxoplasmosis- a waterborne zoonosis. *Vet Parasitol*. 126, 57-72.
- Dubey, J.P., Lindsay, D.S., Speer, C.A. (1998). Structures of *Toxoplasma gondii* tachyzoites, bradyzoites, and sporozoites and biology and development of tissue cysts. *Clin Microbiol Rev*. 11, 267-299.
- Dubey, J.P., Miller, N.L., Frenkel, J.K. (1970) The *Toxoplasma gondii* oocyst from cat feces. *J Exp Med*. 132, 636-662.
- Dzierszinski, F., Nishi, M., Ouko, L., Roos, D.S. (2004). Dynamics of *Toxoplasma gondii* differentiation. *Eukaryot Cell*. 3, 992-1003.
- Ferreira da Silva, Mda F., Barbosa, H.S., Gross, U., Lüder, C.G. (2008). Stress-related and spontaneous stage differentiation of *Toxoplasma gondii*. *Mol Biosyst*. 4, 824-834.
- Finnin, M.S., Donigian, J.R., Cohen, A., Richon, V.M., Rifkind, R.A., Marks, P.A., Breslow, R., Pavletich, N.P. (1999). Structures of a histone deacetylase homologue bound to the TSA and SAHA inhibitors. *Nature*. 401, 188-93.
- Fischle, W., Dequiedt, F., Hendzel, M.J., Guenther, M.G., Lazar, M.A., Voelter, W., Verdin, E. (2002). Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. *Mol Cell*. 9, 45-57.

Flueck, C., Bartfai, R., Niederwieser, I., Witmer, K., Alako, B.T., Moes, S., Bozdech, Z., Jenoe, P., Stunnenberg, H.G., Voss, T.S. (2010). A major role for the *Plasmodium falciparum* ApiAP2 protein PfSIP2 in chromosome end biology. *PLoS Pathog.* 6, e1000784.

Fox, B.A., Gigley, J.P., Bzik, D.J. (2004). *Toxoplasma gondii* lacks the enzymes required for de novo arginine biosynthesis and arginine starvation triggers cyst formation. *Int J Parasitol.* 34, 323-331.

Frenkel, J.K., Dubey, J.P., Miller, N.L. (1970). *Toxoplasma gondii* in cats: fecal stages identified as coccidian oocysts. *Science.* 167, 893-896.

Gaji, R.Y., Behnke, M.S., Lehmann, M.M., White, M.W., Carruthers, V.B. (2011). Cell cycle-dependent, intercellular transmission of *Toxoplasma gondii* is accompanied by marked changes in parasite gene expression. *Mol Microbiol.* 79:192-204.

Garcia, B.A., Hake, S.B., Diaz, R.L., Kauer, M., Morris, S.A., Recht, J., Shabanowitz, J., Mishra, N., Strahl, B.D., Allis, C.D., Hunt, D.F. (2007). Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem.* 282, 7641-7655.

Gay, G., Braun, L., Brenier-Pinchart, M.P., Vollaire, J., Josserand, V., Bertini, R.L., Varesano, A., Touquet, B., De Bock, P.J., Coute, Y., Tardieux, I., Bougdour, A., Hakimi, M.-A. (2016). *Toxoplasma gondii* TgIST co-opts host chromatin repressors dampening STAT1-dependent gene regulation and IFN- γ -mediated host defenses. *J Exp Med.* 213, 1779-1798.

Gerstein, M.B., Kundaje, A., Hariharan, M., Landt, S.G., Yan, K.K., Cheng, C., Mu, X.J., Khurana, E., Rozowsky, J., Alexander, R., Min, R., Alves, P., Abyzov, A., Addleman, N., Bhardwaj, N., Boyle, A.P., Cayting, P., Charos, A., Chen, D.Z., Cheng, Y., Clarke, D., Eastman, C., Euskirchen, G., Fietze, S., Fu, Y., Gertz, J., Grubert, F., Harmanci, A., Jain, P., Kasowski, M., Lacroute, P., Leng, J.J., Lian, J., Monahan, H., O'Geen, H., Ouyang, Z., Partridge EC, Patacsil D, Pauli F, Raha D, Ramirez L, Reddy TE, Reed B, Shi M, Slifer T, Wang, J., Wu, L., Yang, X., Yip, K.Y., Zilberman-Schapira, G., Batzoglou, S., Sidow, A., Farnham, P.J., Myers, R.M., Weissman, S.M., Snyder, M. (2012). Architecture of the human regulatory network derived from ENCODE data. *Nature.* 489, 91-100.

Gissot, M., Choi, S.W., Thompson, R.F., Grealley, J.M., Kim, K. (2008). *Toxoplasma gondii* and *Cryptosporidium parvum* lack detectable DNA cytosine methylation. *Eukaryot Cell.* 7, 537-540.

Gissot, M., Kelly, K.A., Ajioka, J.W., Grealley, J.M., Kim, K. (2007). Epigenomic modifications predict active promoters and gene structure in *Toxoplasma gondii*. *PLoS Pathog.* 3, e77.

Goldberg, A.D., Allis, C.D., Bernstein, E. (2007). Epigenetics: a landscape takes shape. *Cell.* 128, 635-638.

Guenther, M.G., Barak, O., Mitchell, A., Lazar, M.A. (2001). The SMRT and N-CoR Corepressors Are Activating Cofactors for Histone Deacetylase 3. *Mol Cell Biol.* 21, 6091-6101.

- Guenther, M.G., Lane, W.S., Fischle, W., Verdin, E., Lazar, M.A., Shiekhattar, R. (2000). A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness. *Genes Dev.* *14*, 1048-1057.
- Haberland, M., Montgomery, R.L., Olson, E.N. (2009). The many roles of histone deacetylases in development and physiology: implications for disease and therapy. *Nat Rev Genet.* *10*, 32-42.
- Hakimi, M.-A., Bochar, D.A., Chenoweth, J., Lane, W.S., Mandel, G., Shiekhattar, R. (2002). A core-BRAF35 complex containing histone deacetylase mediates repression of neuronal-specific genes. *Proc Natl Acad Sci U S A.* *99*, 7420-7425.
- Hakimi, M.-A., Bochar, D.A., Schmiesing, J.A., Dong, Y., Barak, O.G., Speicher, D.W., Yokomori, K., Shiekhattar, R. (2002). A chromatin remodelling complex that loads cohesin onto human chromosomes. *Nature.* *418*, 994-998.
- Hakimi, M.-A., Deitsch, K.W. (2007). Epigenetics in Apicomplexa: control of gene expression during cell cycle progression, differentiation and antigenic variation. *Curr Opin Microbiol.* *10*, 357-362.
- Hall, N., Karras, M., Raine, J.D., Carlton, J.M., Kooij, T.W., Berriman, M., Florens, L., Janssen, C.S., Pain, A., Christophides, G.K., James, K., Rutherford, K., Harris, B., Harris, D., Churcher, C., Quail, M.A., Ormond, D., Doggett, J., Trueman, H.E., Mendoza, J., Bidwell, S.L., Rajandream, M.A., Carucci, D.J., Yates, J.R.3rd., Kafatos, F.C., Janse, C.J., Barrell, B., Turner, C.M., Waters, A.P., Sinden, R.E. (2005). A comprehensive survey of the *Plasmodium* life cycle by genomic, transcriptomic, and proteomic analyses. *Science.* *307*, 82-86.
- Hong, D.P., Radke, J.B., White, M.W. (2017). Opposing Transcriptional Mechanisms Regulate *Toxoplasma* Development. *mSphere.* *2*, e00347-16.
- Howe, D.K., Sibley, L.D. (1995). *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. *J Infect Dis.* *172*, 1561-1566.
- Huang, S., Holmes, M.J., Radke, J.B., Hong, D.P., Liu, T.K., White, M.W., Sullivan, W.J. Jr. (2017). *Toxoplasma gondii* AP2IX-4 Regulates Gene Expression during Bradyzoite Development. *mSphere.* *2*, e00054-17.
- Hunter, C.A., Sibley, L.D. (2012). Modulation of innate immunity by *Toxoplasma gondii* virulence effectors. *Nat Rev Microbiol.* *10*, 766-778.
- Huynh, M.H., Carruthers, V.B. (2009). Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. *Eukaryot Cell.* *8*, 530-539.
- Innes, E.A. (2010). A Brief History and Overview of *Toxoplasma gondii*. *Zoonoses Public Health.* *57*, 1-7.

Inoue, N., Hess, K.D., Moreadith, R.W., Richardson, L.L., Handel, M.A., Watson, M.L., Zinn, A.R. (1999). New Gene family defined by MORC, a nuclear protein required for mouse spermatogenesis. *Hum Mol Genet.* 8, 1201-1207.

Iwanaga, S., Kaneko, I., Kato, T., Yuda, M. (2012). Identification of an AP2-family protein that is critical for malaria liver stage development. *PLoS One.* 7:e47557.

Iyer, L.M., Anantharaman, V., Wolf, M.Y., Aravind, L. (2008). Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. *Int J Parasitol.* 38, 1-31.

Jain, V., Yogavel, M., Oshima, Y., Kikuchi, H., Touquet, B., Hakimi, M.-A., and Sharma, A. (2015). Structure of Prolyl-tRNA Synthetase-Halofuginone Complex Provides Basis for Development of Drugs against Malaria and Toxoplasmosis. *Structure.* 23, 819-829.

Jaskelioff, M., Peterson, C.L. (2003). Chromatin and transcription: histones continue to make their marks. *Nat Cell Biol.* 5, 395-399.

Jayne, S., Zwartjes, C.G., van Schaik, F.M., Timmers, H.T. (2006). Involvement of the SMRT/NCoR-HDAC3 complex in transcriptional repression by the CNOT2 subunit of the human Ccr4-Not complex. *Biochem J.* 398, 461-467.

Jeffers, V., Sullivan, WJ. Jr. (2012). Lysine acetylation is widespread on proteins of diverse function and localization in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell.* 11, 735-742.

Jenuwein, T., Laible, G., Dorn, R., Reuter, G. (1998). SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci.* 54, 80-93.

Jerome, M.E., Radke, J.R., Bohne, W., Roos, D.S., White, M.W. (1998). *Toxoplasma gondii* bradyzoites form spontaneously during sporozoite-initiated development. *Infect Immun.* 66, 4838-4844.

Jones, J., Lopez, A., Wilson, M. (2003). Congenital toxoplasmosis. *Am Fam Physician.* 67, 2131-2138.

Jones, T.C., Bienz, K.A., Erb, P. (1986). In vitro cultivation of *Toxoplasma gondii* cysts in astrocytes in the presence of gamma interferon. *Infect Immun.* 51, 147-156.

Josling, G.A., Llinás, M. (2015). Sexual development in *Plasmodium* parasites: knowing when it's time to commit. *Nat Rev Microbiol.* 13, 573-87.

Josling, G.A., Petter, M., Oehring, S.C., Gupta, A.P., Dietz, O., Wilson, D.W., Schubert, T., Längst, G., Gilson, P.R., Crabb, B.S., Moes, S., Jenoe, P., Lim, S.W., Brown, G.V., Bozdech, Z., Voss, T.S., Duffy, M.F. (2015). A *Plasmodium Falciparum* Bromodomain Protein Regulates Invasion Gene Expression. *Cell Host Microbe.* 17, 741-751.

- Kafsack, B.F.C., Rovira-Graells, N., Clark, T.G., Bancells, C., Crowley, V.M., Campino, S.G., Williams, A.E., Drought, L.G., Kwiatkowski, D.P., Baker, D.A., Cortés, A., Llinás, M. (2014). A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature*. 507, 248-252.
- Kim, K., Weiss, L.M. (2004). *Toxoplasma gondii*: the model apicomplexan. *Int J Parasitol*. 34, 423-432.
- Kim, S.C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L., Grishin, N.V., White, M., Yang, X.J., Zhao, Y. (2006). Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell*. 23, 607-618.
- Konrad, C., Queener, S.F., Wek, R.C., Sullivan, WJ. Jr. (2013). Inhibitors of eIF2 α dephosphorylation slow replication and stabilize latency in *Toxoplasma gondii*. *Antimicrob Agents Chemother*. 57, 1815-1822.
- Kouzarides, T. (2000). Acetylation: a regulatory modification to rival phosphorylation? *EMBO J*. 19, 1176-1179.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell*. 128, 693-705.
- Kurdistani, S.K., Grunstein, M. (2003). Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol*. 4, 276-284.
- Lawrence, M., Daujat, S., Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet*. 32, 42-56.
- Lazar, M.A. (2003). Nuclear receptor corepressors. *Nuclear receptor signaling*. 1:e001.
- Li, B., Carey, M., Workman, J.L. (2007). The role of chromatin during transcription. *Cell*. 128, 707-719.
- Licausi, F., Ohme-Takagi, M., Perata, P. (2013). APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. *New Phytol*. 199, 639-649.
- Li, D.Q., Nair, S.S., Kumar, R. (2013). The MORC family: new epigenetic regulators of transcription and DNA damage response. *Epigenetics*. 8, 685-693.
- Li, D.Q., Nair, S.S., Ohshiro, K., Kumar, A., Nair, V.S., Pakala, S.B., Reddy, S.D., Gajula, R.P., Eswaran, J., Aravind, L., Kumar, R. (2012). MORC2 signaling integrates phosphorylation-dependent, ATPase-coupled chromatin remodeling during the DNA damage response. *Cell Rep*. 2, 1657-1669.
- Li, J., Wang, J., Wang, J., Nawaz, Z., Liu, J.M., Qin, J., Wong, J. (2000). Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J*. 19, 4342-4350.

- Lindner, S.E., De Silva, E.K., Keck, J.L., Llinás, M. (2010). Structural determinants of DNA binding by a *P. falciparum* ApiAP2 transcriptional regulator. *J Mol Biol.* 395, 558-567.
- Lindsay, D.S., Dubey, J.P. (2009). Long-term Survival of *Toxoplasma gondii* Sporulated Oocysts in Seawater. *Journal of Parasitology.* 95, 1019-1020.
- Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J., Winzeler, E.A. (2003). Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science.* 301, 1503-1508.
- Levo, M., Segal, E. (2014). In pursuit of design principles of regulatory sequences. *Nat Rev Genet.* 15, 453-468.
- Llinás, M., DeRisi, J.L. (2004) Pernicious plans revealed: *Plasmodium falciparum* genome wide expression analysis. *Curr Opin Microbiol.* 7, 382-387.
- Lomberk, G., Bensi, D., Fernandez-Zapico, M.E., Urrutia, R. (2006). Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat Cell Biol.* 8, 407-415.
- Lopez-Rubio, J.-J., Mancio-Silva, L., Scherf, A. (2009). Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe.* 5, 179-190.
- Lorković, Z.J. (2012). MORC proteins and epigenetic regulation. *Plant Signal Behav.* 7, 1561-1565.
- Lyons, R.E., McLeod, R., Roberts, C.W. (2002). *Toxoplasma gondii* tachyzoite-bradyzoite interconversion. *Trends Parasitol.* 18, 198-201.
- Mair, G.R., Braks, J.A., Garver, L.S., Wiegant, J.C., Hall, N., Dirks, R.W., Khan, S.M., Dimopoulos, G., Janse, C.J., Waters, A.P. (2006). Regulation of sexual development of *Plasmodium* by translational repression. *Science.* 313, 667-669.
- Maubon, D., Bougdour, A., Wong, Y.-S., Brenier-Pinchart, M.-P., Curt, A., Hakimi, M.-A., Pelloux, H. (2010). Activity of the Histone Deacetylase Inhibitor FR235222 on *Toxoplasma gondii*: Inhibition of Stage Conversion of the Parasite Cyst Form and Study of New Derivative Compounds. *Antimicrob Agents Chemother.* 54, 4843-4850.
- Meira, C.S., Pereira-Chioccola, V.L., Vidal, J.E., de Mattos, C.C., Motoie, G., Costa-Silva, T.A., Gava, R., Frederico, F.B., de Mattos, L.C. *Toxoplasma* Groups. (2014). Cerebral and ocular toxoplasmosis related with IFN- γ , TNF- α , and IL-10 levels. *Front Microbiol.* 5, 492.
- Ménard, R., Tavares, J., Cockburn, I., Markus, M., Zavala, F., Amino, R. (2013). Looking under the skin: the first steps in malarial infection and immunity. *Nat Rev Microbiol.* 11, 701-712.
- Millard, C.J., Watson, P.J., Celardo, I., Gordiyenko, Y., Cowley, S.M., Robinson, C.V., Fairall, L., Schwabe, J.W. (2013). Class I HDACs share a common mechanism of regulation by inositol phosphates. *Mol Cell.* 51, 57-67.

- Modrzynska, K., Pfander, C., Chappell, L., Yu, L., Suarez, C., Dundas, K., Gomes, A.R., Goulding, D., Rayner, J.C., Choudhary, J., Billker, O. (2017). A Knockout Screen of ApiAP2 Genes Reveals Networks of Interacting Transcriptional Regulators Controlling the *Plasmodium* Life Cycle. *Cell Host Microbe*. 21, 11-22.
- Moissiard, G., Cokus, S.J., Cary, J., Feng, S., Billi, A.C., Stroud, H., Husmann, D., Zhan, Y., Lajoie, B.R., McCord, R.P., Hale, C.J., Feng, W., Michaels, S.D., Frand, A.R., Pellegrini, M., Dekker, J., Kim, J.K., Jacobsen, S.E. (2012). MORC family ATPases required for heterochromatin condensation and gene silencing. *Science*. 336, 1448-1451.
- Montoya, J.G., Liesenfeld, O. (2004). Toxoplasmosis. *Lancet*. 363, 1965-1976.
- Mordue, D.G., Håkansson, S., Niesman, I., Sibley, L.D. (1999). *Toxoplasma gondii* resides in a vacuole that avoids fusion with host cell endocytic and exocytic vesicular trafficking pathways. *Exp Parasitol*. 92, 87-99.
- Narasimhan, J., Joyce, B.R., Naguleswaran, A., Smith, A.T., Livingston, M.R., Dixon, S.E., Coppens, I., Wek, R.C., Sullivan, WJ. Jr. (2008). Translation regulation by eukaryotic initiation factor-2 kinases in the development of latent cysts in *Toxoplasma gondii*. *J Biol Chem*. 283, 16591-16601.
- Nardelli, S.C., Che, F.Y., Silmon de Monerri, N.C., Xiao, H., Nieves, E., Madrid-Aliste, C., Angel, S.O., Sullivan, WJ. Jr., Angeletti, R.H., Kim, K., Weiss, L.M. (2013). The histone code of *Toxoplasma gondii* comprises conserved and unique posttranslational modifications. *MBio*. 4, e00922-13.
- Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., Van Noort, J., Rhodes, D., Chin, J.W. (2009). A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3K56 acetylation. *Molecular Cell*. 36, 153-163.
- Oberstaller, J., Pumpalova, Y., Schieler, A., Llinás, M., Kissinger, J.C. (2014). The *Cryptosporidium parvum* ApiAP2 gene family: insights into the evolution of apicomplexan AP2 regulatory systems. *Nucleic Acids Res*. 42, 8271-8284.
- Pereira-Chioccola, V.L., Vidal, J.E., Su, C. (2009). *Toxoplasma gondii* infection and cerebral toxoplasmosis in HIV-infected patients. *Future Microbiol*. 4:1363-79.
- Pittman, K.J., Knoll, L.J. (2015). Long-Term Relationships: the Complicated Interplay between the Host and the Developmental Stages of *Toxoplasma gondii* during Acute and Chronic Infections. *Microbiol Mol Biol Rev*. 79, 387-401.
- Ponts, N., Harris, E.Y., Prudhomme, J., Wick, I., Eckhardt-Ludka, C., Hicks, G.R., Hardiman, G., Lonardi, S., Le Roch, K.G. (2010). Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Res*. 20, 228-238.

- Pradeepa, M.M., Grimes, G.R., Kumar, Y., Olley, G., Taylor, G.C., Schneider, R., Bickmore, W.A. (2016). Histone H3 globular domain acetylation identifies a new class of enhancers. *Nat Genet.* *48*, 681-686.
- Prigent, C., Dimitrov, S. (2003). Phosphorylation of serine 10 in histone H3, what for? *J Cell Sci.* *116*, 3677-3685.
- Radke, J.B., Lucas, O., De Silva, E.K., Ma, Y., Sullivan, WJ. Jr., Weiss, L.M., Llinas, M., White, M.W. (2013). ApiAP2 transcription factor restricts development of the *Toxoplasma* tissue cyst. *Proc Natl Acad Sci U S A.* *110*, 6871-6876.
- Radke, J.R., Behnke, M.S., Mackey, A.J., Radke, J.B., Roos, D.S., White, M.W. (2005). The transcriptome of *Toxoplasma gondii*. *BMC Biol.* *3*:26.
- Radke, J.R., Donald, R.G., Eibs, A., Jerome, M.E., Behnke, M.S., Liberator, P., White, M.W. (2006). Changes in the expression of human cell division autoantigen-1 influence *Toxoplasma gondii* growth and development. *PLoS Pathog.* *2*, e105.
- Radke, J.R., Guerini, M.N., Jerome, M., White, M.W. (2003). A change in the premitotic period of the cell cycle is associated with bradyzoite differentiation in *Toxoplasma gondii*. *Mol Biochem Parasitol.* *131*, 119-127.
- Riechmann, J.L., Meyerowitz, E.M. (1998). The AP2/EREBP family of plant transcription factors. *Biol Chem.* *379*, 633-646.
- Robert-Gangneux, F., Dardé, M.L. (2012). Epidemiology of and Diagnostic Strategies for Toxoplasmosis. *Clinical Microbiology Reviews.* *25*, 264-296.
- Roth, S.Y., Allis, C.D. (1992). Chromatin condensation: does histone H1 dephosphorylation play a role? *Trends Biochem Sci.* *17*, 93-98.
- Saeij, J.P., Boyle, J.P., Boothroyd, J.C. (2005). Differences among the three major strains of *Toxoplasma gondii* and their specific interactions with the infected host. *Trends Parasitol.* *21*, 476-481.
- Saksouk, N., Bhatti, M.M., Kieffer, S., Smith, A.T., Musset, K., Garin, J., Sullivan, WJ. Jr., Cesbron-Delauw, M.F., Hakimi, M.-A. (2005). Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite *Toxoplasma gondii*. *Mol Cell Biol.* *25*, 10301-10314.
- Salcedo-Amaya, A.M., van Driel, M.A., Alako, B.T., Trelle, M.B., van den Elzen, A.M., Cohen, A.M., Janssen-Megens, E.M., van de Vegte-Bolmer, M., Selzer, R.R., Iniguez, A.L., Green, R.D., Sauerwein, R.W., Jensen, O.N., Stunnenberg, H.G. (2009). Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A.* *106*, 9655-9660.
- Sangaré, L.O., Alayi, T.D., Westermann, B., Hovasse, A., Sindikubwabo, F., Callebaut, I., Werkmeister, E., Lafont, F., Slomianny, C., Hakimi, M.-A., Van Dorsselaer, A., Schaeffer-

- Reiss, C., Tomavo, S. (2016). Unconventional endosome-like compartment and retromer complex in *Toxoplasma gondii* govern parasite integrity and host infection. *Nat. Commun.* *7*, 11191.
- Santos, J.M., Josling, G., Ross, P., Joshi, P., Orchard, L., Campbell, T., Schieler, A., Cristea, I.M., Llinás, M. (2017). Red Blood Cell Invasion by the Malaria Parasite Is Coordinated by the PfAP2-I Transcription Factor. *Cell Host Microbe.* *21*, 731-741.e10
- Sautel, C.F., Ortet, P., Saksouk, N., Kieffer, S., Garin, J., Bastien, O., Hakimi, M.-A. (2009). The histone methylase KMTox interacts with the redox-sensor peroxiredoxin-1 and targets genes involved in *Toxoplasma gondii* antioxidant defences. *Mol Microbiol.* *71*, 212-226.
- Sautel, C.F., Cannella, D., Bastien, O., Kieffer, S., Aldebert, D., Garin, J., Tardieux, I., Belrhali, H., Hakimi, M.-A. (2007). SET8-mediated methylations of histone H4 lysine 20 mark silent heterochromatic domains in apicomplexan genomes. *Mol Cell Biol.* *27*, 5711-5724.
- Scherf, A., Lopez-Rubio, J.-J., Riviere, L. (2008). Antigenic variation in *Plasmodium falciparum*. *Annu Rev Microbiol.* *62*, 445-470.
- Schlüter, D., Däubener, W., Schares, G., Groß, U., Pleyer, U., Lüderg, C. (2014). Animals are key to human toxoplasmosis. *Int J Med Microbiol.* *304*, 917-929.
- Schreiber, S.L., Bernstein, B.E. (2002). Signaling network model of chromatin. *Cell.* *111*, 771-778.
- Shao, Y., Li, Y., Zhang, J., Liu, D., Liu, F., Zhao, Y., Shen, T., Li, F. (2010). Involvement of histone deacetylation in MORC2-mediated down-regulation of carbonic anhydrase IX. *Nucleic Acids Res.* *38*, 2813-2824.
- Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science.* *311*, 844-847.
- Sibley, L.D., Boothroyd, J.C. (1992). Virulent strains of *Toxoplasma gondii* comprise a single clonal lineage. *Nature.* *359*, 82-85.
- Sidik, S.M., Huet, D., Ganesan, S.M., Huynh, M.H., Wang, T., Nasamu, A.S., Thiru, P., Saeij, J.P., Carruthers, V.B., Niles, J.C., Lourido, S. (2016). A Genome-wide CRISPR Screen in *Toxoplasma* Identifies Essential Apicomplexan Genes. *Cell.* *166*, 1423-1435.e12.
- Sierra-Miranda, M., Vembar, S.S., Delgadillo, D.M., Ávila-López, P.A., Herrera-Solorio, A.M., Lozano Amado, D., Vargas, M., Hernandez-Rivas, R. (2017). PfAP2Tel, harbouring a non-canonical DNA-binding AP2 domain, binds to *Plasmodium falciparum* telomeres. *Cell Microbiol.* *2017*, e12742.
- Singh, U., Brewer, J.L., Boothroyd, J.C. (2002). Genetic analysis of tachyzoite to bradyzoite differentiation mutants in *Toxoplasma gondii* reveals a hierarchy of gene induction. *Mol Microbiol.* *44*, 721-33.

- Sinha, A., Hughes, K.R., Modrzynska, K.K., Otto, T.D., Pfander, C., Dickens, N.J., Religa, A.A., Bushell, E., Graham, A.L., Cameron, R., Kafsack, B.F., Williams, A.E., Llinás, M., Berriman, M., Billker, O., Waters, A.P. (2014). A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. *Nature*. *507*, 253-257.
- Smith, A.T., Tucker-Samaras, S.D., Fairlamb, A.H., Sullivan, WJ. Jr. (2005). MYST family histone acetyltransferases in the protozoan parasite *Toxoplasma gondii*. *Eukaryot Cell*. *4*, 2057-2065.
- Smith, C.M., Gafken, P.R., Zhang, Z., Gottschling, D.E., Smith, J.B., Smith, D.L. (2003) Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Anal Biochem*. *316*, 23-33.
- Soète, M., Camus, D., Dubremetz, J.F. (1994). Experimental induction of bradyzoite-specific antigen expression and cyst formation by the RH strain of *Toxoplasma gondii* *in vitro*. *Exp Parasitol*. *78*, 361-370.
- Sonda, S., Morf, L., Bottova, I., Baetschmann, H., Rehrauer, H., Caflisch, A., Hakimi, M.-A., Hehl, A.B. (2010). Epigenetic mechanisms regulate stage differentiation in the minimized protozoan *Giardia lamblia*. *Mol Microbiol*. *76*, 48-67.
- Strahl, B.D., Allis, C.D. (2000). The language of covalent histone modifications. *Nature*. *403*, 41-45.
- Sullivan, WJ. Jr., Hakimi, M.-A. (2006). Histone mediated gene activation in *Toxoplasma gondii*. *Mol Biochem Parasitol*. *148*, 109-116.
- Sullivan, WJ. Jr., Smith, A.T., Joyce, B.R. (2009). Understanding mechanisms and the role of differentiation in pathogenesis of *Toxoplasma gondii*: a review. *Mem Inst Oswaldo Cruz*. *104*, 155-161.
- Telles, E., Seto, E. (2012). Modulation of cell cycle regulators by HDACs. *Front Biosci (Schol Ed)*. *4*, 831-839.
- Tenter, A.M., Heckeroth, A.R., Weiss, L.M. (2000). *Toxoplasma gondii*: from animals to humans. *Int J Parasitol*. *30*, 1217-1258.
- Tonkin, C.J., Carret, C.K., Duraisingh, M.T., Voss, T.S., Ralph, S.A., Hommel, M., Duffy, M.F., Silva, L.M., Scherf, A., Ivens, A., Speed, T.P., Beeson, J.G., Cowman, A.F. (2009). Sir2 paralogs cooperate to regulate virulence genes and antigenic variation in *Plasmodium falciparum*. *PLoS Biol*. *7*, e84.
- Trelle, M.B., Salcedo-Amaya, A.M., Cohen, A.M., Stunnenberg, H.G., Jensen, O.N. (2009). Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite *Plasmodium falciparum*. *J Proteome Res*. *8*, 3439-3450.
- Tremp, A.Z., Carter, V., Saeed, S., Dessens, J.T. (2013). Morphogenesis of *Plasmodium* zoites is uncoupled from tensile strength. *Mol Microbiol*. *89*, 552-564.

- Tropberger, P., Pott, S., Keller, C., Kamieniarz-Gdula, K., Caron, M., Richter, F., Li, G., Mittler, G., Liu, E.T., Bühler, M., Margueron, R., Schneider, R. (2013). Regulation of transcription through acetylation of H3K122 on the lateral surface of the histone octamer. *Cell*. *152*, 859-872.
- Tropberger, P., Schneider, R. (2010). Going global: novel histone modifications in the globular domain of H3. *Epigenetics*. *5*, 112-117.
- Tropberger, P., Schneider, R. (2013). Scratching the (lateral) surface of chromatin regulation by histone modifications. *Nat Struct Mol Biol*. *20*, 657-661.
- Turner, B.M. (2000). Histone acetylation and an epigenetic code. *Bioessays*. *22*, 836-845.
- Vanagas, L., Jeffers, V., Bogado, S.S., Dalmaso, M.C., Sullivan, WJ. Jr., Angel, S.O. (2012). *Toxoplasma* histone acetylation remodelers as novel drug targets. *Expert Rev Anti Infect Ther*. *10*, 1189-1201.
- Vannini, A., Volpari, C., Gallinari, P., Jones, P., Mattu, M., Carfi, A., De Francesco, R., Steinkühler, C., Di Marco, S. (2007). Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex. *EMBO Rep*. *8*, 879-84.
- Walker, R., Gissot, M., Croken, M.M., Huot, L., Hot, D., Kim, K., Tomavo, S. (2013). The *Toxoplasma* nuclear factor TgAP2XI-4 controls bradyzoite gene expression and cyst formation. *Mol Microbiol*. *87*, 641-655.
- Wang, J., Dixon, S.E., Ting, L.M., Liu, T.K., Jeffers, V., Croken, M.M., Calloway, M., Cannella, D., Hakimi, M.-A., Kim, K., Sullivan, WJ. Jr. (2014). Lysine acetyltransferase GCN5b interacts with AP2 factors and is required for *Toxoplasma gondii* proliferation. *PLoS Pathog*. *10*, e1003830.
- Wang, L., Rajan, H., Pitman, J.L., McKeown, M., Tsai, C.C. (2006). Histone deacetylase-associating Atrophin proteins are nuclear receptor corepressors. *Genes Dev*. *20*, 525-530.
- Watson, P.J., Fairall, L., Santos, G.M., Schwabe, J.W. (2012). Structure of HDAC3 bound to co-repressor and inositol tetrakisphosphate. *Nature*. *481*, 335-340.
- Weinert, B.T., Wagner, S.A., Horn, H., Henriksen, P., Liu, W.R., Olsen, J.V., Jensen, L.J., Choudhary, C. (2011). Proteome-wide mapping of the *Drosophila* acetylome demonstrates a high degree of conservation of lysine acetylation. *Sci Signal*. *4*, ra48.
- Weiss, L.M., Kim, K. (2000). The development and biology of bradyzoites of *Toxoplasma gondii*. *Front Biosci*. *5*, 391-405.
- Weiss, L.M., Ma, Y.F., Takvorian, P.M., Tanowitz, H.B., Wittner, M. (1998). Bradyzoite development in *Toxoplasma gondii* and the hsp70 stress response. *Infect Immun*. *66*:3295-3302.

- Whitfield, M.L., Sherlock, G., Saldanha, A.J., Murray, J.I., Ball, C.A., Alexander, K.E., Matese, J.C., Perou, C.M., Hurt, M.M., Brown, P.O., Botstein, D. (2002). Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol Biol Cell*. *13*, 1977-2000.
- Williams, S. K., Truong D., Tyler, J. K. (2008). Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc. Natl. Acad. Sci. USA*. *105*, 9000-9005.
- Yabsley, M.J., Shock, B.C. (2012). Natural history of Zoonotic Babesia: role of wildlife reservoirs. *Int J Parasitol Parasites Wildl*. *2*, 18-31.
- Yahiaoui, B., Dzierszynski, F., Bernigaud, A., Slomianny, C., Camus, D., Tomavo, S. (1999). Isolation and characterization of a subtractive library enriched for developmentally regulated transcripts expressed during encystation of *Toxoplasma gondii*. *Mol Biochem Parasitol*. *99*, 223-235.
- Yang, W.M., Tsai, S.C., Wen, Y.D., Fejer, G., Seto, E. (2002). Functional domains of histone deacetylase-3. *J Biol Chem*. *277*, 9447-9454.
- Yang, X.J., Seto, E. (2008). The Rpd3/Hda1 family of lysine deacetylases: from bacteria and yeast to mice and men. *Nat Rev Mol Cell Biol*. *9*, 206-218.
- Yap, G.S., Sher, A. (1999). Cell-mediated immunity to *Toxoplasma gondii*: initiation, regulation and effector function. *Immunobiology*. *201*, 240-247.
- Young, R.A. (2011). Control of Embryonic Stem Cell State. *Cell*. *144*, 940-954.
- Yuda, M., Iwanaga, S., Kaneko, I., Kato, T. (2015). Global transcriptional repression: An initial and essential step for *Plasmodium* sexual development. *Proc Natl Acad Sci U S A*. *112*, 12824-12829.
- Yuda, M., Iwanaga, S., Shigenobu, S., Kato, T., Kaneko I. (2010). Transcription factor AP2-Sp and its target genes in malarial sporozoites. *Mol Microbiol*. *75*, 854-863.
- Yuda, M., Iwanaga, S., Shigenobu, S., Mair, G.R., Janse, C.J., Waters, A.P., Kato, T., Kaneko, I. (2009). Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. *Mol Microbiol*. *71*, 1402-1414.
- Zhang, J., Kalkum, M., Chait, B.T., Roeder, R.G. (2002). The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Mol Cell*. *9*, 611-623.
- Zhang, J., Sprung, R., Pei, J., Tan, X., Kim, S., Zhu, H., Liu, C.F., Grishin, N.V., Zhao, Y. (2009). Lysine acetylation is a highly abundant and evolutionarily conserved modification in *Escherichia coli*. *Mol Cell Proteomics*. *8*, 215-225.

Zhang, K., Williams, K.E, Huang, L., Yau, P., Siino, J.S., Bradbury, E.M., Jones, P.R., Minch, M.J., Burlingame, A.L. (2002). Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. *Mol Cell Proteomics*. 1,500-508.

Zhang, M., Fennell, C., Ranford-Cartwright, L., Sakthivel, R., Gueirard, P., Meister, S., Caspi, A., Doerig, C., Nussenzweig, R.S., Tuteja, R., Sullivan, WJ. Jr., Roos, D.S., Fontoura, B.M., Ménard, R., Winzeler, E.A., Nussenzweig, V. (2010). The *Plasmodium* eukaryotic initiation factor-2alpha kinase IK2 controls the latency of sporozoites in the mosquito salivary glands. *J Exp Med*. 207, 1465-1474.

Xue, B., Jeffers, V., Sullivan, WJ. Jr., Uversky, V.N. (2013). Protein intrinsic disorder in the acetylome of intracellular and extracellular *Toxoplasma gondii*. *Mol Biosyst*. 9; 9645-9657.

Xu, K., Dai, X.L., Huang, H.C., Jiang, Z.F. (2011). Targeting HDACs: a promising therapy for Alzheimer's disease. *Oxid Med Cell Longev*. 2011, 143269.

Annexes

ARTICLE

Received 11 Nov 2015 | Accepted 26 Feb 2016 | Published 11 Apr 2016

DOI: 10.1038/ncomms11191

OPEN

Unconventional endosome-like compartment and retromer complex in *Toxoplasma gondii* govern parasite integrity and host infection

Lamba Omar Sangaré¹, Tchilabalo Dilezitoko Alayi^{2,3}, Benoit Westermann², Agnes Hovasse², Fabien Sindikubwabo⁴, Isabelle Callebaut⁵, Elisabeth Werkmeister⁶, Frank Lafont⁶, Christian Slomianny⁷, Mohamed-Ali Hakimi⁴, Alain Van Dorselaer², Christine Schaeffer-Reiss² & Stanislas Tomavo^{1,3}

Membrane trafficking pathways play critical roles in Apicomplexa, a phylum of protozoan parasites that cause life-threatening diseases worldwide. Here we report the first retromer-trafficking interactome in *Toxoplasma gondii*. This retromer complex includes a trimer Vps35-Vps26-Vps29 core complex that serves as a hub for the endosome-like compartment and parasite-specific proteins. Conditional ablation of *TgVps35* reveals that the retromer complex is crucial for the biogenesis of secretory organelles and for maintaining parasite morphology. We identify *TgHP12* as a parasite-specific and retromer-associated protein with functions unrelated to secretory organelle formation. Furthermore, the major facilitator superfamily homologue named *TgHPO3*, which is a multiple spanning and ligand transmembrane transporter, is maintained at the parasite membrane by retromer-mediated endocytic recycling. Thus, our findings highlight that both evolutionarily conserved and unconventional proteins act in concert in *T. gondii* by controlling retrograde transport that is essential for parasite integrity and host infection.

¹Center for Infection and Immunity of Lille, INSERM U 1019, CNRS UMR 8204, Institut Pasteur de Lille, Université de Lille, 59000 Lille, France. ²Laboratory of Bio-Organic Mass Spectrometry, IPHC, CNRS UMR 7178, Université de Strasbourg, 67087 Strasbourg, France. ³Plateforme de Protéomique et des Peptides Modifiés (P3M), Institut Pasteur de Lille, CNRS, Université de Lille, 59000 Lille, France. ⁴CNRS UMR5163, LAPM, Université Joseph Fourier, Grenoble 38000, France. ⁵CNRS UMR7590, Sorbonne Universités, Université Pierre et Marie Curie-Paris 6, MNHN, IRD-IUC, Paris 75005, France. ⁶Bioimaging Platform, IBL, CNRS, Université de Lille, 59000 Lille, France. ⁷Laboratory of Cell Physiology, INSERM U 1003, Université de Lille, 59655 Villeneuve d'Ascq, France. Correspondence and requests for materials should be addressed to S.T. (email: stanislas.tomavo@gmail.com).

The phylum Apicomplexa comprises an ancient group of early divergent eukaryotes, including some of the most deadly pathogens of medical and veterinary importance. *Plasmodium* species are responsible for malaria, which causes as many as 700,000 deaths per year, while *Toxoplasma gondii* chronically infects up to 30% of the human population, with immunocompromised patients and pregnant women at risk for adverse outcomes, such as toxoplasmic encephalitis and spontaneous abortion, respectively¹. *T. gondii* is considered a model system not only for its pathogenic relatives but also for intracellular parasitism and infection biology in general. *T. gondii* has common eukaryotic organelles, including the nucleus, endoplasmic reticulum and a single Golgi stack, but also specific secretory organelles named dense granules, micronemes and rhoptries that contain parasite-derived factors required for host infection. Rhoptries and micronemes are formed *de novo* during parasite replication, and this process requires significant protein and lipid trafficking through the secretory pathway.

The trafficking mechanisms employed by *T. gondii* retain several typical eukaryote components as well as evolving divergent features. Protein trafficking of this parasite is mediated by entry into a canonical endoplasmic reticulum followed by vesicle packaging through a single Golgi complex^{2,3}. Post-Golgi protein sorting to specific organelles requires the function of dynamin-related protein B, which is involved in fission events⁴. Downstream Rab-GTPases function throughout the parasite secretory pathway⁵. *T. gondii* soluble *N*-ethylmaleimide-sensitive-factor attachment protein receptor (SNARE) proteins in docking and fusion at target membranes have also been described^{6,7}. However, unlike in mammalian cells, *T. gondii* endoplasmic reticulum is reduced so that the nuclear envelope itself contributes to a substantial proportion of its total volume². Whereas in mammalian cells hundreds of Golgi stacks occupy the perinuclear area⁸, the Golgi apparatus is limited to a single discrete structure in *T. gondii*⁹. The post-Golgi system, also named the endosome-like compartment (ELC), is involved in the trafficking of microneme proteins^{10,11}. The ELC is decorated by the small GTPases, Rab5 and Rab7, which are typically associated with the endosomal system. Nevertheless, classical endocytosis has not yet been validated in *T. gondii*. This parasite has no lysosomes; rather the parasite harbours acidic vesicles that were thought to be precursors of the rhoptry organelles¹². The parasite lacks most components of endosomal sorting complexes, which are known for their roles in forming multivesicular bodies that deliver ubiquitinated membrane proteins and lipids to lysosomes for degradation^{3,13}. The machinery required for caveogenesis and caveola-dependent invaginations have not yet been identified in the parasite¹⁴. Furthermore, while evidence of conventional clathrin-dependent endocytosis by *T. gondii* is lacking, clathrin is present exclusively in post-Golgi compartments where its function is restricted to post-Golgi trafficking¹⁵, and the uptake of cytosol proteins by the tachyzoites of *T. gondii* has recently been described using an endocytosis assay¹⁶. However, the mechanisms underlying the events of this unconventional endocytosis in the parasite remain to be determined. Clearly, the secretory pathway of *T. gondii* can be considered a stripped-down version of the more complex trafficking machinery that characterizes higher eukaryotes. Despite this minimal trafficking machinery, the parasites actively rely on a membrane vesicle formation and transport during its intracellular lifecycle; however, to date, comparatively little is known about the mechanisms involved in trafficking pathways in *T. gondii*.

We previously reported a *T. gondii* sortilin-like receptor (TgSORTLR) that regulates protein transport and is essential for apical secretory organelle biogenesis and host infection¹⁷.

Moreover, the C-terminal tail of TgSORTLR was shown to be involved in recruiting many cytosolic cargo proteins including two homologues of the core retromer components, Vps26 and Vps35 (ref. 17), which are known to regulate retrograde transport from endosomes to the *trans*-Golgi network (TGN) in yeast and mammals^{18,19}.

Here, we report that a singular architecture with a trimer Vps35–Vps26–Vps29 core complex acts as the major endosomal cargo recycling machinery and is required for parasite integrity and more specifically for secretory organelle biogenesis and maintenance of a multiple ligand-binding transporter at the *T. gondii* membrane. Our findings provide strong evidence that the unconventional TgSORTLR-containing ELC is involved in distinct mechanisms for the delivery of major retromer-dependent cargo. They also demonstrate a role for the endocytic recycling pathway in *T. gondii* pathogenesis.

Results

Features of the retromer interactome of *T. gondii*. To identify proteins that interact with the *T. gondii* retromer complex, we chromosomally appended an encoded hemagglutinin (HA) epitope to TgVsp35 and TgVps26. This knock-in strategy allows steady-state levels of epitope-tagged protein expression via homologous promoters. We also tagged TgVps29 identified in the parasite genome (TGME49_252490, www.toxodb.org) with a cMyc epitope as above. We performed a series of immunoprecipitation experiments under native conditions; revealing that TgVps35-HA, TgVps26-HA and TgVps29-cMyc were specifically pulled down (Fig. 1a, lanes 2–4; and Supplementary Fig. 1) using HA or cMyc-tagged protein extracts and antibodies specific to HA and cMyc, respectively. No protein signals were detected in the negative controls using naïve sera and the same protein extracts, as expected (Fig. 1a, lanes 5 and 6). In addition, immunoprecipitation of TgVps35-HA and TgVps26-HA also revealed a faint protein band corresponding to TgSORTLR protein using rat antibodies anti-TgSORTLR (Fig. 1b, lane 3 (E) and blue stars in left and middle panels) while immunoprecipitation of TgVps29 did not, most likely due to its low-expression level (Fig. 1b, lane 3 (E), right panel). Mass spectrometry analysis of the eluates corroborates the presence of TgVps35, TgVps26 and TgVps29 in each immunoprecipitation sample (Supplementary Data 1). Consistent with the immunoblots shown in Fig. 1b, the presence of TgSORTLR was only confirmed in immunoprecipitates of TgVps35 and TgVps26 by mass spectrometry (Supplementary Data 1). To gain unbiased insight into the genuine retromer composition in *T. gondii*, we developed a quantitative approach using micro liquid chromatography-selected reaction monitoring (microLC-SRM) and stable isotope-labelled standard peptides. The absolute quantification of TgVps35, TgVps26 and TgVps29 was carried out using three proteotypic peptides per protein (Supplementary Data 2). This approach yielded a stoichiometry of ~1:1 for TgVps35 relative to TgVps26 and 3:1 between TgVsp35 and TgVps29 (Supplementary Table 1). This stoichiometry between TgVps35 and TgVps29 is in contrast to the formation of a functional core retromer complex at a ratio of 1:1:1, as in mammalian and yeast cells²⁰. However, this discrepancy may also be explained by the fact that TgVps29 may associate with TgVps35 at a much lower affinity than TgVps26, thus leading to the reduced levels of TgVps29 identified by co-immunoprecipitations. In addition, 17 retromer-interacting proteins were identified in the interactome (Fig. 1c) and ranked according to the following filtering criteria: protein common to at least two co-immunoprecipitations, absent in the control and identified with at least two unique peptides (Supplementary Data 1). Most interactors (12 out of 17) were immunoprecipitated

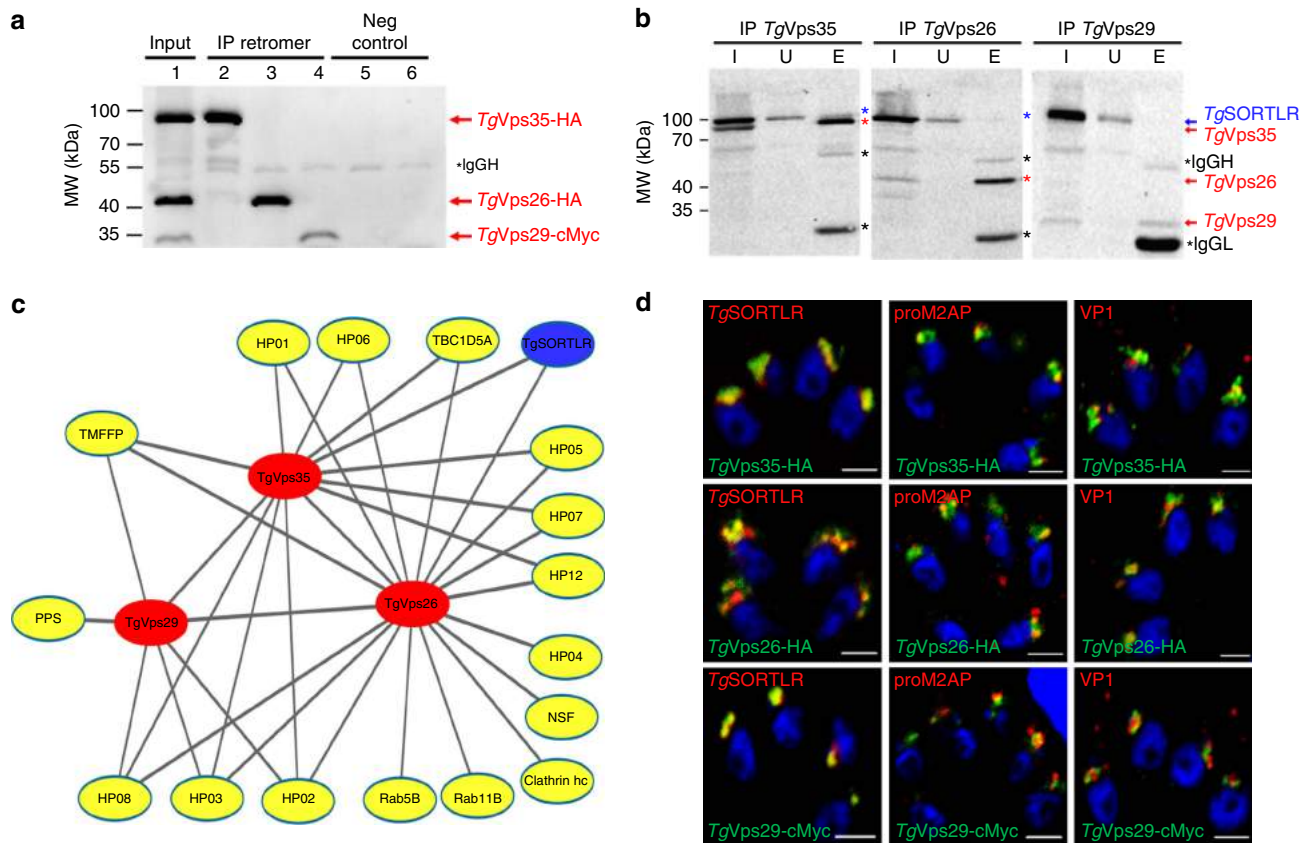


Figure 1 | Interactome reveals cargo-selective complex and other interactors with the retromer of *T. gondii*. (a) Immunoblots of co-immunoprecipitates of *TgVps35*-HA and *TgVps26*-HA (designated IP retromer) probed with rabbit polyclonal specific anti-HA antibodies (lanes 2 and 3) using total detergent protein extracts from the knock-in *TgVps35*-HA and *TgVps26*-HA parasites, respectively, and anti-HA beads. Immunoblot of co-immunoprecipitate of *TgVps29*-cMyc (IP retromer) probed with rabbit polyclonal anti-cMyc antibodies (lane 4) using total detergent protein extract from the knock-in *TgVps29*-cMyc parasites and anti-cMyc beads. Negative controls (Neg control) using total detergent protein extracts from untagged parental RH TaTi parasites incubated with anti-HA (lane 5) and anti-cMyc (lane 6) beads. Lane 1 (designated input) corresponds to equally mixed sample of all three detergent extracts containing *TgVps35*-HA, *TgVps26*-HA and *TgVps29*-cMyc proteins and revealed by a mixed probe containing both anti-HA and anti-cMyc antibodies. Molecular weights (kDa) of protein markers are shown on left. IgG_H means heavy chain of IgG. (b) Immunoblots of *TgVps35*-HA, *TgVps26*-HA and *TgVps29*-cMyc as described in a probed with rat specific anti-*TgSORTLR* antibodies. (I, input) corresponds to total detergent protein extracts from *TgVps35*-HA, *TgVps26*-HA and *TgVps29*-cMyc knock-in parasites, respectively; (U) unbound lysates to the anti-HA or anti-cMyc beads and (E) eluates corresponding to co-immunoprecipitates. The blots were simultaneously incubated with rat anti-*TgSORTLR* and rabbit anti-HA or rat anti-*TgSORTLR* and rabbit anti-cMyc antibodies. Protein markers (kDa) are also shown on left. IgG_H means heavy chain of IgG, IgG_L means light chain of IgG. (c) Retromer interactome was constructed by analysing the co-immunoprecipitates of *TgVps35*-HA, *TgVps26*-HA and *TgVps29*-cMyc validated by immunoblotting in a and mass spectrometry (Supplementary Data 1). The interactome identified *TgVps35*, *TgVps29* and *TgVps26* (red) and *TgSORTLR* (blue) in addition to the putative phosphatidylinositol synthase (PPS), transporter major facilitator family protein (TMFFP), putative N-ethylmaleimide sensitive fusion protein (NSF), multi-pass transmembrane protein (MTP), Rab5, Rab11B, Rab7-GTPase-activating protein (GAP) regulator TBC1D5A homologue and nine parasite-specific HP. (d) Confocal imaging of *TgVps35*, *TgVps26* and *TgVps29* that co-localize with *TgSORTLR*, proM2AP and vacuolar protein 1 (VP1) using intracellular tachyzoites of the respective knock-in parasites stained with anti-HA or anti-cMyc antibodies followed by probing with anti-*TgSORTLR*, anti-proM2AP and anti-VP1 antibodies, respectively. Bar, 2 μm.

with both *TgVps35* and *TgVps26* (Fig. 1c), confirming the potential of predominant *TgVps35*-*TgVps26* complexes in which only a fraction of *TgVps29* is bound to generate a functional retromer complex, as determined by the quantitative proteomics described in Supplementary Table 1. Functional classification by gene ontology analysis revealed that some of these interactors played roles in cell trafficking; Rab5B, an endosome marker; Rab11B, a factor essential for inner membrane complex recycling²¹; the TBC1D5A homologue, a Rab7-GTPase-activating protein that negatively regulates the core retromer function²²; the aforementioned *TgSORTLR* receptor¹⁷; and N-ethylmaleimide-sensitive protein, a factor involved in SNARE-dependent membrane fusion. In addition to the established binding partners, 9 out of 17 proteins are new parasite-specific proteins (that is, hypothetical proteins (HP);

Fig. 1c). Confocal imaging revealed that *TgVps35*-HA, *TgVps26*-HA and *TgVps29*-cMyc co-localize with the ELC markers pro-microneme 2-associated protein (proM2AP), vacuolar protein 1 (VP1) and *TgSORTLR* (Fig. 1d). We therefore conclude that the ELC defines the sub-cellular compartment where retromer-mediated vesicle recycling or retrograde trafficking operates via an endolysosomal-like system in *T. gondii*.

***TgVps35* silencing abrogates host infection by *T. gondii*.** To establish the functional roles of *TgVps35* in *T. gondii* infection, we generated conditional anhydrotetracyclin (ATc)-inducible knockout mutants (iKo*TgVps35*) using the strategy described in Fig. 2a. We selected three positive clones from the emerging stable parasite population and the genome editing of these clones

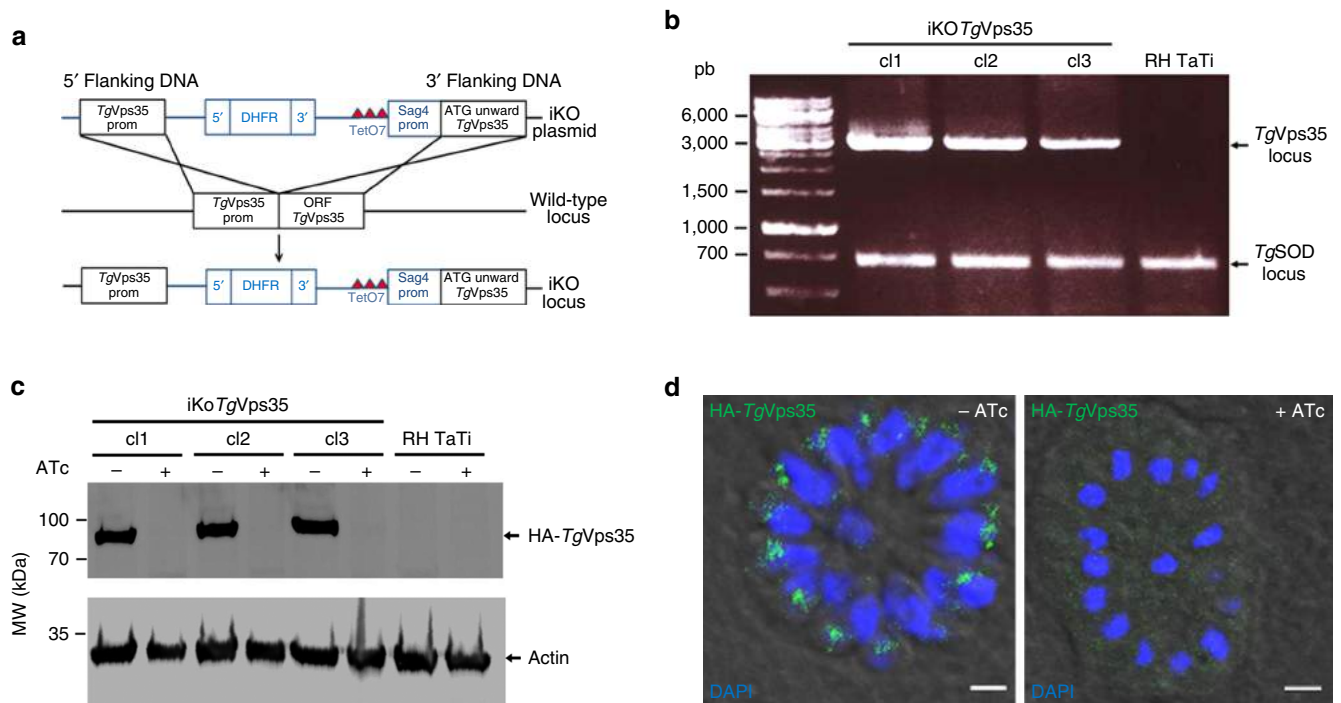


Figure 2 | Conditional ablation of *TgVps35* gene. (a) Schematic of the vector and experimental approach used for the conditional ablation of the *TgVps35* gene. (b) PCR analysis of three clones with conditional disruption of *TgVps35* and the parental line. Superoxide dismutase (SOD) served as the positive control. Also see the primers used for these PCR in Supplementary Table 3. (c) Immunoblots of the three conditional iKo*TgVps35* mutants and RH TaTi parasites, which were grown in the presence or absence of ATc for 48 h, harvested and purified. Each lane refers to a total SDS-protein extract corresponding to the equivalent of 2×10^6 parasites. Immunoblots were probed with anti-HA antibodies. Actin probed with specific monoclonal antibodies served as a loading control. Molecular weights (kDa) of protein markers are indicated on left. (d) Intracellular vacuoles containing 16-daughter iKo*TgVps35* mutants corresponding to one of the three clones analysed by PCR and western blots and PCR confirmed the conditional depletion *TgVps35* protein (right panel) by confocal imaging after 48 h post-infection in the presence of ATc. The left panel showed the same mutant in which *TgVps35* protein was detected in the endosome-like compartment (ELC) closely located to the nuclei in the absence of ATc, as expected. 4', 6-diamidino-2-phenylindole (DAPI) was used to stain nuclei. Rabbit specific anti-HA antibodies was also used. Bar, 2 μ m.

was verified by PCR, demonstrating the perfect integration of the knockout vector at the *TgVps35* locus (Fig. 2b). Following ATc treatment, we subsequently observed the disappearance of HA-*TgVps35* protein by western blotting (Fig. 2c) and confocal imaging (Fig. 2d). To ascertain *bona fide* morphological phenotypes that stem from the inducible targeted disruption of the *TgVps35* gene, we complemented this mutant with full-length cMyc-tagged *TgVps35* (Comp-iKo*TgVps35*), which was introduced in the uracil phosphoribosyl transferase locus, as this gene is known to be non-essential for parasite survival²³. The iKo*TgVps35* mutants were severely impaired in their ability to invade host cells (Fig. 3a) and did not form plaques after multiple rounds of host cell invasion and lysis (Fig. 3b). Complementation of the iKo*TgVps35* mutant that allows obtaining Comp-iKo*TgVps35* parasite lines restored the ability of these complemented mutants to efficiently reinvade host cells (Fig. 3a), yielding normal plaque sizes in the presence of ATc similar to those of parental RH TaTi parasites (Fig. 3b). These later observations demonstrate that the lack of host cell invasion and the subsequent inability of the iKO mutants to establish several rounds of cell lysis and reinvasion are directly linked to the depletion of *TgVps35* and the absence of functional retromer complex in these mutants, thus excluding pleiotropic and non-specific phenotypes.

To examine the role of *TgVps35* in *Toxoplasma* infection *in vivo*, mice were infected with lethal doses of iKo*TgVps35*, Comp-iKo*TgVps35* or parental parasites followed by *TgVps35* suppression *in vivo* by providing ATc in the drinking water. Strikingly, the ATc-treated mice inoculated with iKo*TgVps35*

survived, whereas animals inoculated with iKo*TgVps35* but not treated with ATc succumbed to the infection by day 9 (Fig. 3c). Mice infected with Comp-iKo*TgVps35* mutants and the parental strains succumbed to the infection regardless of the initiation of ATc treatment (Fig. 3c). It should be mentioned that RH TaTi background was genetically attenuated in virulence compared with the parental and wild-type RH strain, thus allowing challenging mice with sub-lethal parasite doses. When mice were inoculated with sub-lethal doses of these mutants or parental parasites and re-challenged with lethal doses of the wild-type parental RH strain, all iKo*TgVps35*-infected mice succumbed in a manner similar to the naïve primo-infected animals, whereas those infected with Comp-iKo*TgVps35* and the parental strains survived (Fig. 3d). Thus, the conditional ablation of *TgVps35* transformed a *T. gondii* into a complete non-lethal strain of parasites, and furthermore, infection with iKo*TgVps35* parasites does not confer sterile immunity to reinfection, which is also consistent with phenotypic traits previously described for iKo*TgSORTLR* mutants¹⁷.

Retromer is essential for parasite integrity. We observed that the disappearance of rooptries peaks at 24 h of ATc treatment while micronemes were mostly affected 48 h after ATc pressure, which also corresponds to the time necessary for the complete depletion of *TgVps35*, as shown by western blots and confocal microscopy in Fig. 2. Following 48 h of ATc treatment, we also found a complete disorganized morphology with the marked absence of the typical banana-shaped bodies in *TgVps35*-depleted mutants using electron microscopy (Fig. 4b), whereas untreated

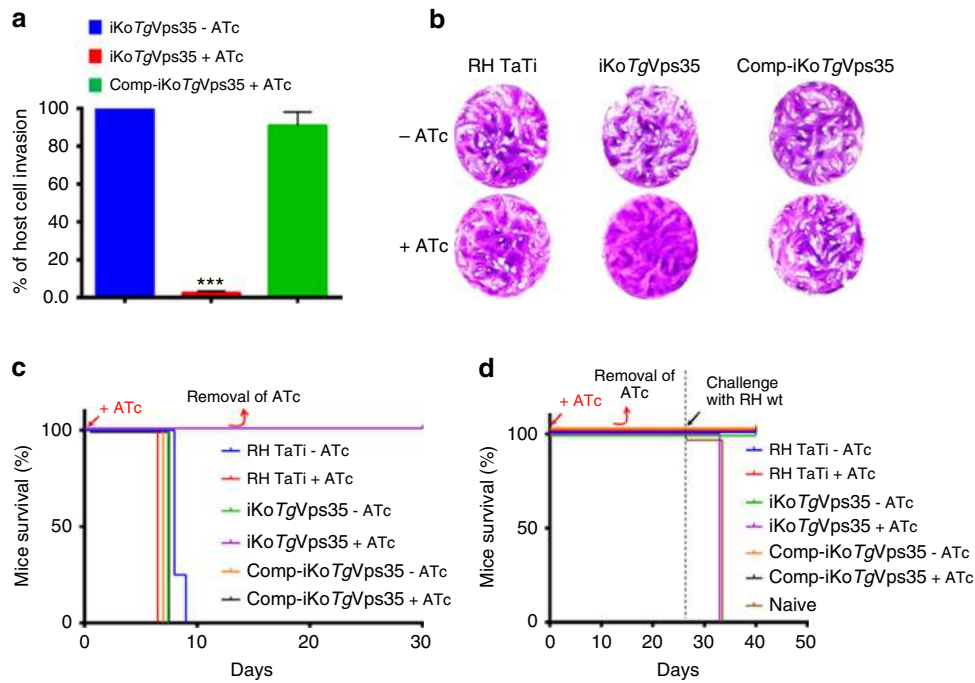


Figure 3 | Resultant phenotypic traits of conditional disruption of *TgVps35* gene. (a) Host cell invasion was assayed in *iKoTgVps35* and complemented mutants (Comp-*iKoTgVps35*) in the presence and absence of ATc. Bars indicate mean \pm s.d. ($n = 3$, $P < 0.001$ by Student's test). (b) Host cell lytic plaques were examined in *TgVps35*-deficient mutants, Comp-*iKoTgVps35* and parental RH TaTi parasites in the presence or absence of ATc. (c) Survival of mice infected with lethal doses of *TgVps35*-deficient mutants, Comp-*iKoTgVps35* and parental RH TaTi parasites in the presence and absence of ATc in the drinking water. The ATc was removed after 12 days and mice survival was monitored for 30 days post-infection. (d) Avirulent *TgVps35*-depleted parasites do not confer protective immunity to reinfection with lethal doses of RH wild-type parasites. ATc treatment and mice survival were monitored as above, except that sub-lethal doses of parasites were used during the primo-infection.

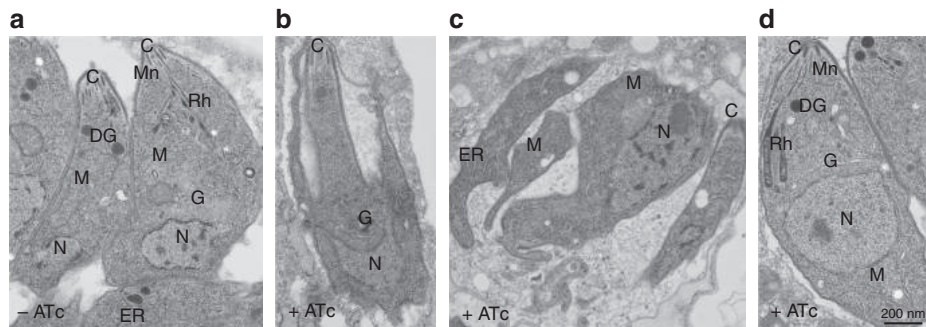


Figure 4 | The *TgVps35* gene is essential for secretory organelle biogenesis and parasite morphology. (a) Transmission electron microscopy showing normal banana-shape morphology of *iKoTgVps35* parasites in the absence of ATc. (b) An *iKoTgVps35* mutant showing an aberrant morphology. (c) Four replicated daughter *iKoTgVps35* mutants with disorganized body shapes without rhoptries, micronemes and dense granules. (d) Complementation of *iKoTgVps35* mutants restored normal parasite morphology with the presence of secretory organelles. C, conoid; DG, dense granules; G, Golgi; M, mitochondria; Mn, micronemes; N, nucleus; Rh, rhoptries. Bar, 200 nm.

iKoTgVps35 parasites appeared structurally normal with all secretory organelles (Fig. 4a). It should be mentioned that the membrane localization of the major glycosyl-phosphatidyl inositol-anchored surface antigens²⁴, SAG1 and SAG3 of *T. gondii* was not impaired by the suppression of *TgVps35* (Fig. 9f), indicating that the traffic to and the integrity of the parasite pellicle were not affected. This aberrant parasite morphology was confirmed in intravacuolar dividing mutants that were also devoid of rhoptries, micronemes and dense granules (Fig. 4c). Complementation of *iKoTgVps35* mutants restored the ability of the parasites to form the secretory organelles *de novo* even in the ATc pressure (Fig. 4d). Using confocal microscopy, we showed that in the absence of rhoptry and microneme organelles, the ROP and MIC proteins were all

mis-localized in the cytoplasm as well as in the parasitophorous vacuole of *TgVps35*-depleted mutants, leading to the loss of the typical apical end staining of these proteins (Fig. 5a, lower, left and middle panels, respectively). As expected, the parental parasites normally contained rhoptries and micronemes (Fig. 5a, upper, left and middle panels). In addition, the dense granule GRA1 protein staining that typically surrounds the parasitophorous vacuole (Fig. 5a, upper and right panel) was also altered in *TgVps35*-depleted mutants (Fig. 5a, lower and right panel). This later observation is in sharp contrast to the phenotypic traits of *iKoTgSORTLR* mutants in which dense granule biogenesis and secretion were not affected¹⁷. We also confirmed that the secretory organelles were correctly localized in complemented *iKoTgVps35* parasites and that the mis-sorting of

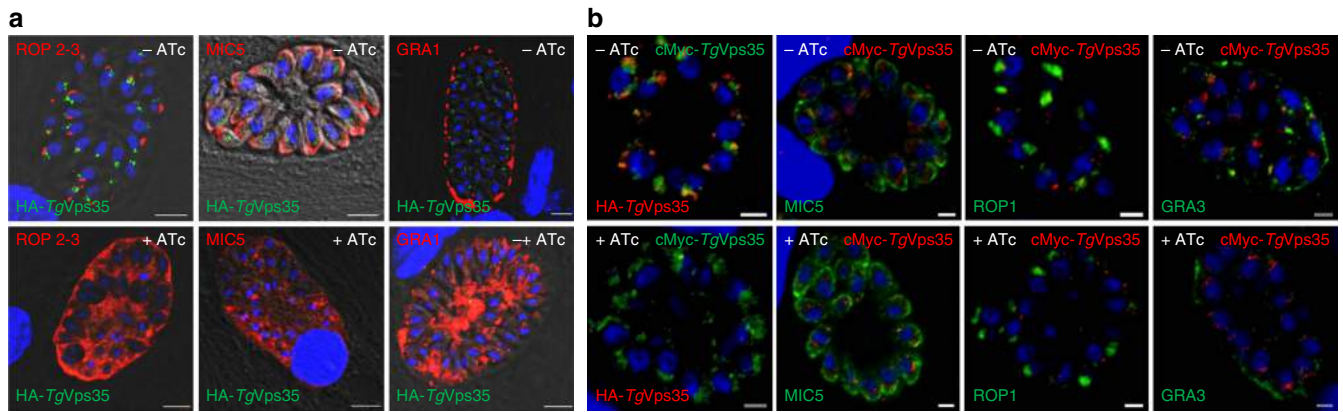


Figure 5 | Conditional ablation of *TgVps35* results in mis-sorting of ROP, MIC and GRA proteins. (a) Confocal immunofluorescence microscopy of ROP 2-3, MIC5 and GRA1 proteins in iKo*TgVps35* mutants in the presence (lower panels) or absence of ATc (upper panels) for 48 h using specific antibodies to ROP2-3, MIC5 and GRA1 proteins (see the complete list of antibodies in Supplementary Table 5). Bar, 2 μ m. (b) Confocal immunofluorescence microscopy of MIC5, ROP1 and GRA3 proteins in complemented iKo*TgVps35* mutants in the presence (lower panels) or absence of ATc (upper panels) as above. Bar, 2 μ m.

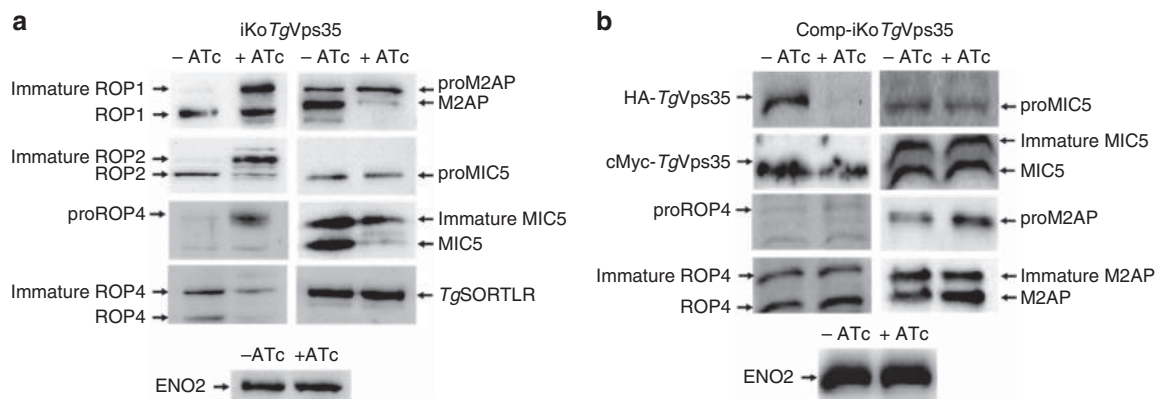


Figure 6 | Disruption of *TgVps35* accumulates unprocessed and immature ROP and MIC proteins. (a) Immunoblots of iKo*TgVps35* mutants probed with specific anti-ROP1, ROP2, proROP4, ROP4, M2AP, proMIC5 and MIC5 antibodies after 48 h post-infection in the presence or absence of ATc. *TgSORTLR* and ENO2 protein levels were identical in both ATc-treated and untreated mutants. (b) Immunoblots of complemented iKo*TgVps35* mutants probed with specific anti-ROP1, ROP2, proROP4, ROP4, M2AP, proMIC5 and MIC5 antibodies after 48 h post-infection in the presence or absence of ATc. ENO2 protein levels were identical in both ATc-treated and untreated mutants.

ROP, MIC and GRA proteins was rescued in the presence of ATc using confocal microscopy (Fig. 5b).

In wild-type parasites, formation of rhoptries and micronemes is correlated with proteolytic maturation of ROP and MIC proteins^{3,25}. Likewise, this proteolytic maturation was defective in the *TgVps35*-depleted mutants compared with the parental strain, leading to the accumulation of unprocessed ROP1, ROP2, ROP4, M2AP and MIC5 (Fig. 6a and Supplementary Fig. 2). Notably, iKo*TgVps35* parasites that were not treated with ATc displayed typical proteolytic maturation of the aforementioned ROP and MIC proteins (Fig. 6a and Supplementary Fig. 2). Next, we probed these blots with specific antibodies that recognized the N-terminal pro-peptides of ROP4 and MIC5 and found a significant accumulation of both pro-protein and immature forms in the *TgVps35*-deficient mutants (Fig. 6a and Supplementary Fig. 2). In contrast, processing of the receptor *TgSORTLR* was unchanged in these mutant parasites, suggesting that neither the processing of pre-protein in the endoplasmic reticulum is impaired nor this receptor is subjected to the typical lysosomal-like degradation in the *TgVps35*-deficient parasites (Fig. 6a). This later behaviour of *T. gondii* lacking *TgVsp35* and retromer functions differs greatly from what is normally observed in the absence of functional retromer complex in other eukaryotes

in which the cargo sorting receptors such as sortilin and mannose-6-phosphate receptor are targeted to lysosomes for degradation^{26–29}. Since the levels of the control protein, the glycolytic enzyme enolase ENO2, were similar between mutant and parental parasites, we speculated that both pro-ROP and pro-MIC specifically accumulated in iKo*TgVps35* mutants as a consequence of conditional disruption of *TgVps35* functions, which are clearly recovered in the presence of iKo*TgVps35* complementation as cMyc-*TgVps35* is able to fully restore proteolytic processing and maturation of ROP and MIC proteins under ATc pressure (Fig. 6b). Taken together, these results suggest that retromer-mediated recycling is likely required to deliver and maintain one or more proteases that process pro-ROP and pro-MIC proteins, a proteolytic processing that is a key parameter for secretory organelle formation and host infectivity by *T. gondii*.

Secretory organelle biogenesis depends on retromer. The cellular location of *TgVps35*, *TgVps26* and *TgVsp29* of the retromer complex with the ELC markers proM2AP and vacuolar protein 1 (Fig. 1d), as for Rab5 or Rab7, prompted us to investigate the outcome of *TgVps35* depletion on *TgSORTLR*, which also co-distributes in the parasite with these two

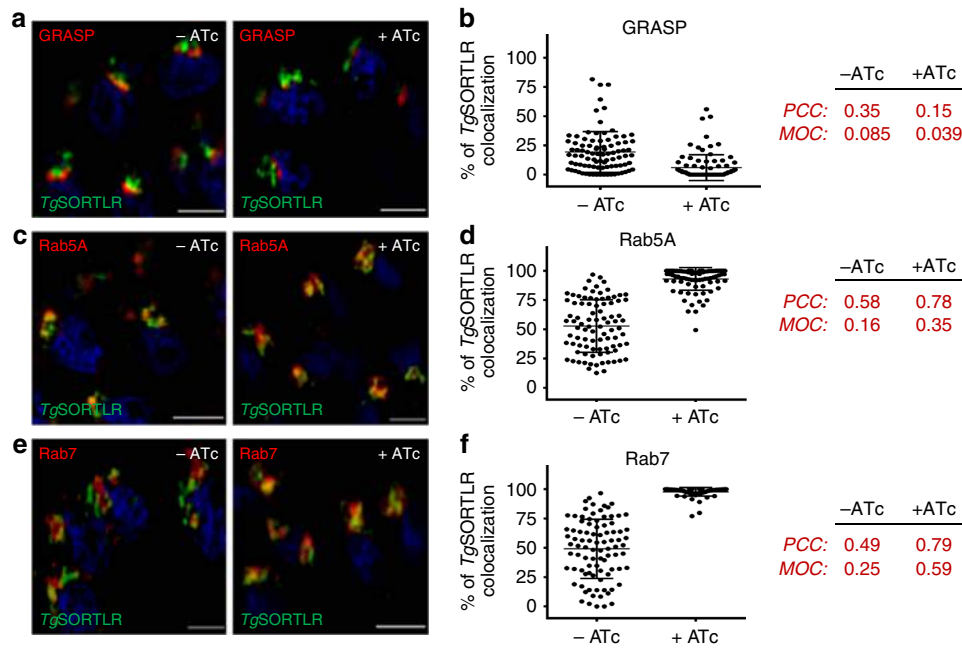


Figure 7 | Recycling of *TgSORTLR* is blocked in Rab5A- and Rab7-decorated ELCs. (a) SIM imaging of the co-distribution of *TgSORTLR* with Golgi marker GRASP in iKo*TgVps35* mutants in the absence or presence of ATc. Bar, 2 μ m. (b) Quantification of the co-distribution of *TgSORTLR* with Golgi marker GRASP in iKo*TgVps35* mutants by SIM in the absence or presence of ATc. Bars indicated the mean $n = 90$ parasites from 7 to 8 vacuoles \pm s.d., $P < 0.0001$. (c) Co-distribution of *TgSORTLR* with the marker of early endosome Rab5A in iKo*TgVps35* mutants grown in the presence or absence of ATc and treated for SIM imaging as above. Bar, 2 μ m. (d) Quantification of *TgSORTLR* co-distributing with the marker of early endosome Rab5A in iKo*TgVps35* mutants by SIM in the absence or presence of ATc. Bars indicated the mean $n = 90$ parasites from 7 to 8 vacuoles \pm s.d., $P < 0.0001$. (e) Co-distribution of *TgSORTLR* with the marker of late endosome marker Rab7 in iKo*TgVps35* mutants by SIM imaging as above. Bar, 2 μ m. (f) Quantification of *TgSORTLR* co-distributing with the marker of late endosome marker Rab7 in iKo*TgVps35* mutants by SIM in the absence or presence of ATc. Bars indicated the mean $n = 90$ parasites \pm s.d., $P < 0.0001$. The Pearson correlation coefficient (PCC) and the Mander's overlap coefficient (MOC) used to quantify the degree of colocalization between the red and green fluorophores were shown on the right.

small GTPases¹⁷. Using high-resolution structured illumination microscopy (SIM), we monitored the discrete compartments that contained *TgSORTLR* (Fig. 7) by co-labelling Golgi apparatus marker with the Golgi reassembly stacking protein (GRASP)-RFP, the early endosome with HA-Rab5A, and the late endosome with HA-Rab7. While endogenous *TgSORTLR* (up to 20%) colocalized with GRASP-RFP in the Golgi compartment (Fig. 7a, left panel) in the parental strain, the co-distribution drastically decreased to a marginal level in *TgVps35*-deficient mutants (Fig. 7a, right panel and Fig. 7b). Consequently, we observed significantly increased colocalization of endogenous *TgSORTLR* with Rab5A and Rab7-positive ELC in *TgVps35*-depleted mutants (Fig. 7c,e; right panels). Quantitation of this co-distribution in *TgVps35*-deficient mutants indicated that 90% of *TgSORTLR* localized with Rab5A (Fig. 7d) and that up to 99% of endogenous *TgSORTLR* was colocalized with Rab7-positive ELC (Fig. 7f). In comparison, only 60% (Rab5A) and 50% (Rab7) of *TgSORTLR* was colocalized with the Rab GTPase-labelled ELC in parasites that were not treated with ATc (Fig. 7d,f). Biochemical data also supported the notion that only *TgVps26* interacts with Rab7 in a GTP-dependent manner, while none of these three subunits of the *T. gondii* core retromer partnered with Rab5B or Rab11B in the presence of GTP (Supplementary Fig. 3). Clearly, these data demonstrate that the retromer complex drives *TgSORTLR* recycling from Rab7-positive ELC to the TGN, thus sustaining another round of protein transport for proper secretory organelle biogenesis.

***TgHP12* is a parasite-specific retromer-associated partner.** The conditional disruption of the *TgVps35* gene strongly suggests that the retromer complex is likely involved in other functions, such as

controlling parasite shape in addition to secretory organelle biogenesis. In our quest to discover other roles for the retromer complex in *T. gondii*, we sought to decipher the functions of some parasite-specific proteins also named HP that were identified in the interactome (Fig. 1c and Supplementary Data 1). Towards this goal, we searched by bioinformatics for striking sequence features that could potentially define retromer-associating proteins among these parasite-specific HP proteins. We identified a typical type I transmembrane *TgHP12* protein that harbours a putative coiled-coil region downstream of the transmembrane segment (Supplementary Fig. 4a). These structural features are conserved in *TgHP12* homologues in all tested parasites across the *Apicomplexa* phylum (Fig. 8a). Potential relationships were identified between the *TgHP12* helical region and two helical structures present in rabenosyn-5 and FIP2 (Supplementary Fig. 4b), which are known to be involved in the interaction with the Rab GTPases^{30,31}, and also with the coiled-coil region of syntaxin, which shares typical heptad repeats with *TgHP12* (Supplementary Fig. 4b).

Because these similarities represent only short segments, statistical analysis was unable to establish a significant relationship with any of the proteins. As a result, we have not further investigated the significance of these possible structural features, but instead, we determined the molecular relationships between *TgHP12* and the retromer complex by knock-in *TgHP12*-cMyc into iKo*TgVps35* mutants. These experiments revealed that *TgHP12* co-localizes with *TgVps35*-HA and *TgSORTLR* (Fig. 8b, upper panels) but not with cathepsin L or M2AP, two markers of the ELC (Fig. 8b, lower panels), suggesting that *TgHP12* is a resident protein of both the Golgi and TGN compartments. Likewise, mass spectrometry was used to

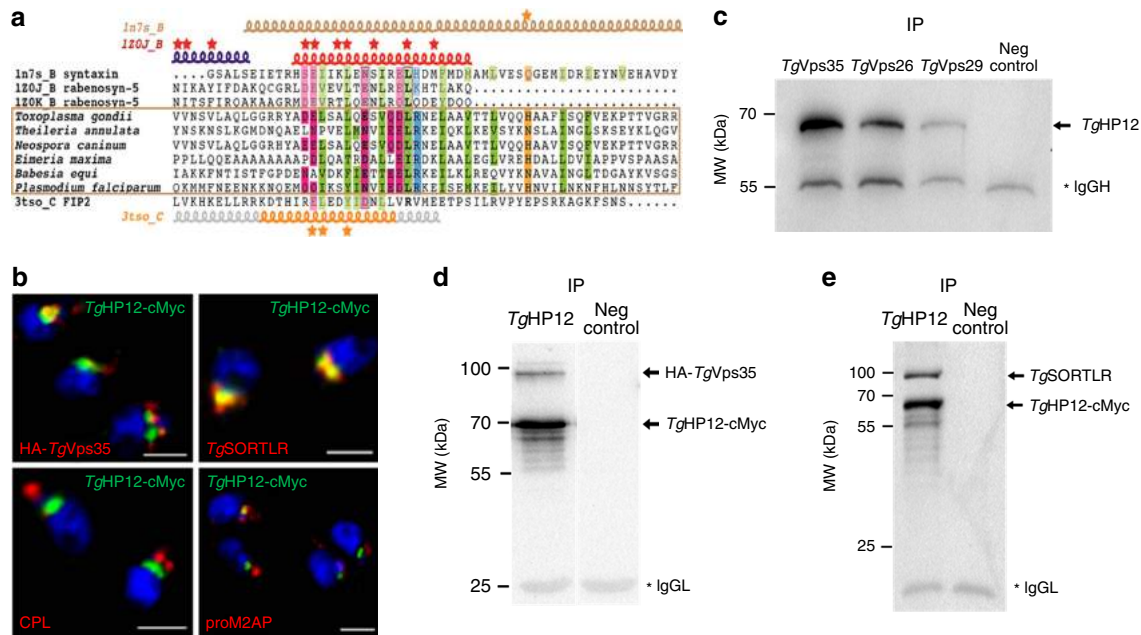


Figure 8 | *TgHP12* is a novel parasite-specific retromer-associating partner. (a) Alignment of the helical region in *TgHP12* protein and in orthologues from apicomplexan parasites with human syntaxin (pdb 1n7s_B, syntaxin), rabenosyn-5 (120j_B and 120k_B), and FIP2 (pdb 3tso_C). Secondary structures are reported above and below the alignment. Stars indicate amino acids involved in Rab binding (rabenosyn-5 and FIP2) and the conserved glutamine of syntaxin that participates in the ionic central layer of SNARE complexes. Genbank identifier (gi) and N- and C-terminal limits are as follows: *Theileria annulata* (85000999, aa 239–297), *Neospora caninum* (401395596, aa 42–100), *Eimeria maxima* (557188226, aa 283–342), *Babesia equi* (510902511, aa 253–311) and *Plasmodium falciparum* (583212139, aa 375–433). Also see Supplementary Fig. 4. (b) Confocal immunofluorescences showing co-distribution of *TgHP12* protein with *TgVps35* protein, and *TgSORTLR* protein (upper panel) while no colocalization was detected with cathepsin L (CPL) or proM2AP protein. Mouse or rabbit polyclonal anti-cMyc antibodies were used in addition to rat anti-*TgSORTLR*, rabbit anti-HA, anti-CPL and proM2AP antibodies. Bar, 2 μ m. (c) Reverse immunoprecipitation using total detergent extract proteins from the knock-in *TgVps35*-HA, *TgVps29*-cMyc or *TgVps26*-HA parasites demonstrated that all three components of the core retromer complex can be pulled down by *TgHP12*. IgG_H, IgG heavy chain; IP, immunoprecipitations; Neg control, negative control using naïve sera; molecular weights (kDa) were shown on left. (d) *TgHP12*-cMyc and *TgVps35*-HA proteins were concomitantly co-immunoprecipitated from total detergent protein extracts from iKo*TgVps35* parasites in which *TgHP12*-cMyc protein was expressed by knock-in strategy. The blots were probed with rabbit anti-cMyc and anti-HA. IgG_L, IgG light chain; IP, immunoprecipitations; Neg control, negative control using naïve sera; molecular weights (kDa) were shown on left. (e) *TgHP12*-cMyc and *TgSORTLR* proteins were concomitantly co-immunoprecipitated from parasites total detergent protein extracts from iKo*TgVps35* parasites in which *TgHP12*-cMyc protein was expressed by knock-in strategy. The blots were stained with rabbit anti-cMyc and rat anti-*TgSORTLR*. IgG_L, IgG light chain; IP, immunoprecipitations; Neg control, negative control using naïve sera; molecular weights (kDa) were shown on left.

demonstrate that *TgHP12* specifically pulled down *TgVps35*, *TgVps29* and *TgVps26* in addition to *TgSORTLR* (Supplementary Table 2). These interactions were confirmed by reverse immunoprecipitations in which *TgVps35*, *TgVps29* and *TgVps26* also specifically pulled down *TgHP12* protein (Fig. 8c). Furthermore, we confirmed that the eluates of *TgHP12* also contained HA-*TgVps35* (Fig. 8d) and *TgSORTLR* (Fig. 8e) by western blots. To obtain additional insight into the functions of *TgHP12*, we disrupted this gene using the CRISPR-Cas9 strategy (Supplementary Fig. 4c). We confirmed the efficient disruption of *TgHP12* gene, as no *TgHP12* protein was detected in these mutants using rat polyclonal antibodies that we specifically raised and purified against the recombinant *TgHP12* protein (Supplementary Fig. 4d, first and left of upper panels). As expected, this protein was normally expressed in wild-type parasites (Supplementary Fig. 4d, red, first and left of lower panels). Since no deleterious effects in rhoptries, micronemes or dense granules occurred following CRISPR-Cas9 disruption of *TgHP12*, we conclude that this type I transmembrane protein is likely involved in functions distinct from those described for *TgSORTLR*. In line with this hypothesis, proteomic analysis indicated that *TgHP12* binds to other partners (Supplementary Data 3), suggesting a possible role in alternative trafficking pathways or in the regulation of other functions as a retromer-associated partner.

Retromer maintains a parasite transporter at the membrane. The generation of iKo*TgVps35* mutants that are deficient in retromer-mediated transport allows us to investigate the recycling mechanisms that deliver and maintain transmembrane proteins at the parasite membrane. We also used bioinformatics to search for candidate multi-spanning transmembrane proteins among the HP identified in the retromer interactome. We discovered that *TgHP03* exhibits the topology and positions of 12 transmembrane helices that were predicted by Phyre2 alignments³² to align with known three-dimensional structures of several members of the major facilitator superfamily (MFS), suggesting that *TgHP03* may belong to this superfamily (Fig. 9a). These membrane transporters facilitate movement of a wide range of small substrates such as metabolites, oligosaccharides, amino acids, oxyanions and drugs that were all transported by MFS across the cell membranes³³. Even though a significant relationship with the MFS can be established, no direct link with one specific MFS member was determined by bioinformatics. To assess the functional links between *TgHP03* and the retromer complex, we first examined its sub-cellular localization by detecting a cMyc epitope-tagged version in the iKo*TgVps35* mutants. In the absence of ATc, *TgHP03*-cMyc displayed homogenous membrane staining on both extracellular (Fig. 9b,c, left upper panel) and intracellular (Fig. 9c, right upper panel) parasites,

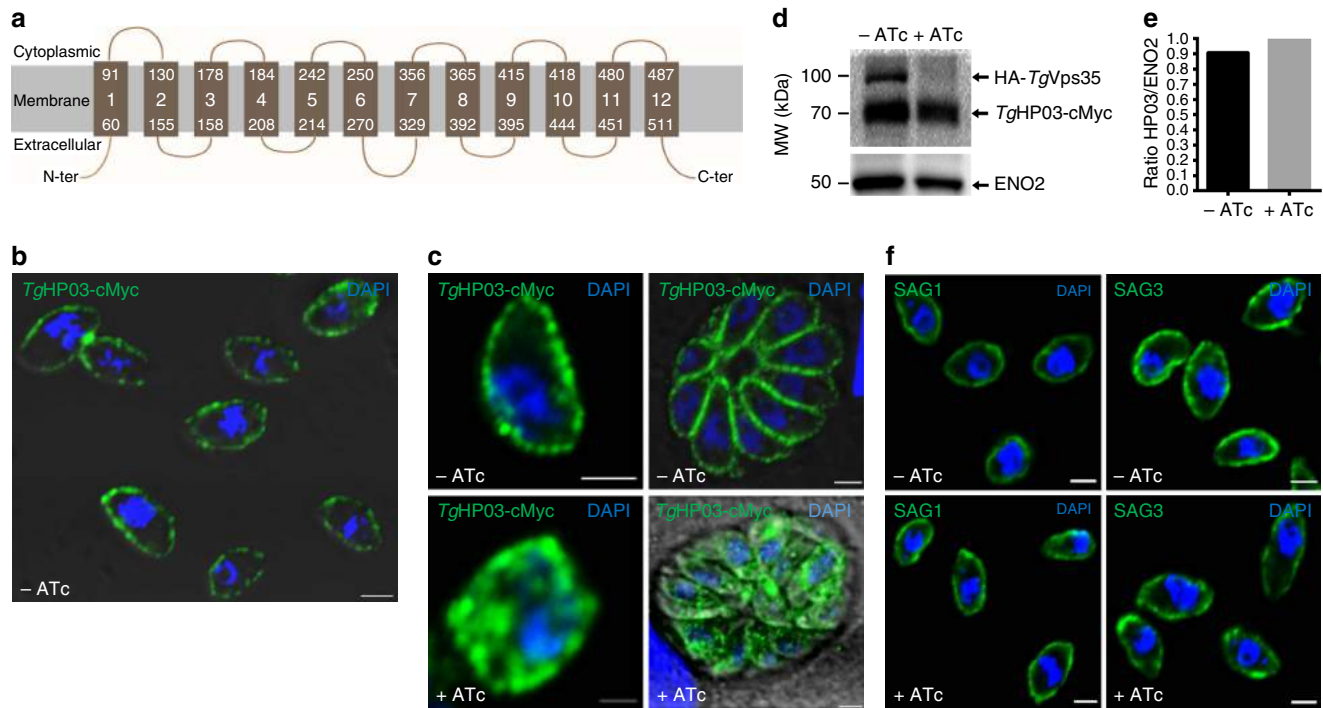


Figure 9 | Retromer is required to maintain TgHP03 in the parasite membrane. (a) Schematic representation of TgHP03 topology inside the parasite membrane. (b) Confocal imaging of extracellular parasites with knock-in TgHP03-cMyc in the iKoTgVps35 mutants. Immunofluorescence assay (IFA) was performed in the absence of ATc and detergent permeabilization. Nuclei of the parasites as stained with 4', 6-diamidino-2-phenylindole (DAPI; blue). Bar, 2 μ m. (c) Magnified image of an extracellular parasite expressing TgHP03-cMyc protein in the iKoTgVps35 mutants grown in the absence of ATc and detergent permeabilization (upper, left panel); intracellular parasites expressing TgHP03-cMyc protein in the iKoTgVps35 genetic background. IFA was performed in the absence of ATc and in the presence detergent permeabilization (upper, right panel); magnified image of an intracellular parasites expressing TgHP03-cMyc protein in the iKoTgVps35 genetic background in the presence of ATc and detergent permeabilization (lower, left panel) and intracellular parasites expressing TgHP03-cMyc protein in the iKoTgVps35 genetic background in the presence of ATc and detergent permeabilization (lower, right panel). Bar, 2 μ m. (d) Immunoblots of parasites expressing TgHP03-cMyc protein in the knock-in iKoTgVps35 mutants grown in the presence or absence of ATc. ENO2 was used as a loading control. Molecular weights (kDa) were shown on left. (e) Quantification of TgHP03 levels in these parasites expressing TgHP03 protein in the knock-in iKoTgVps35 mutants that were grown in the absence or presence of ATc. (f) The surface localization of glycosylphosphatidylinositol (GPI)-anchored SAG1 and SAG3 were determined in iKoTgVps35 mutants in the presence or absence of ATc using monoclonal antibodies specific to SAG1 and SAG3. Bar, 2 μ m.

as expected for transmembrane transporters. TgVps35 suppression by ATc led to a mis-localization and enhanced accumulation of TgHP03 in intra-cytoplasmic vesicular structures in both extracellular (Fig. 9c, left lower panel) and intracellular mutants (Fig. 9c, right lower panel), thus leading to a decrease of TgHP03 at the parasite surface. Since ATc treatment did not affect the levels of TgHP03-cMyc protein (Fig. 9d,e), it is likely that no TgHP03 protein degradation may occur on suppression of TgVps35-mediated endocytic recycling. This observation is also consistent with the absence of TgSORTLR degradation following suppression of TgVps35 (Fig. 6a). Again, the accumulation of mis-sorted cargo in intra-cytoplasmic vesicles in *T. gondii* is also in sharp contrast to transmembrane transporter degradation occurring via lysosomes in mammalian cells³⁴. Interestingly, we found that the surface localization of the glycosylphosphatidylinositol-anchored major surface antigens, SAG1 and SAG3 of *T. gondii* was not influenced by the suppression of TgVps35 (Fig. 9f), suggesting that only the sorting of multi-spanning transmembrane proteins may be influenced by retromer-dependent endocytic recycling. Together, our data identify a role for TgVps35 in the endosome-to-plasma membrane sorting of multi-spanning transmembrane transporter cargo and provide the first evidence for the mechanistic role of TgVps35 in the process linked to retromer-dependent endocytic recycling.

Discussion

Nascent apical organelles require proficient cell trafficking to fulfil their critical role during invasion and intracellular development of *T. gondii*. Here, we show that retromer-mediated recycling is essential for secretory organelle biogenesis, parasite morphology and maintenance of a transmembrane transporter at the parasite membrane. The retromer complex was first identified in yeast and mammals as a heteropentameric complex typified by a cargo-selective complex that was built around the Vps35–Vps29–Vps26 trimer and a dimer of different sorting nexins^{28,35–38}. Our previous findings that TgSORTLR receptor is involved in the recruitment of Vps26 and Vps35 homologues prompted us to investigate how the retromer complex regulates retrograde transport as well as other functions in *T. gondii*. Towards this goal, we characterized the retromer interactome that includes TgVps35, TgVps29 and TgVps26 proteins. We failed to identify homologues of nexins (SNX) in the immunoprecipitates of *T. gondii* using all three retromer components, confirming that no SNX proteins can be pulled down in association with the retromer cargo-selective trimer³⁹. However, we also noticed the paucity of nexin-like proteins in the genome databases of *T. gondii*, and this may also explain our inability to pull down recognizable nexin-like proteins in the parasite. Rather, we identified a singular retromer complex typified by the TgVps35–TgVps29–TgVps29 trimer. This complex interacts

with *TgSORTLR*, *N*-ethylmaleimide-sensitive protein, Rab5B, Rab11B and the TBC1D5A homologue, a member of the Rab7-GTPase-activating proteins known to negatively regulate the retromer complex through Rab7 dissociation from endosomal membranes²². However, we cannot rule out the possibility that some parasite-specific proteins may act as SNX and Bin/Amphiphysin/Rvs domain-containing proteins that are required for recruitment of the retromer complex to endosome membranes, membrane curvature and tubulation^{27,35}. Furthermore, we demonstrated that *TgSORTLR* recycling from the ELC to the TGN is severely compromised in *TgVps35* mutants, suggesting that the retromer complex functions in endocytic recycling in *T. gondii*. In contrast to yeast and human cells in which Rab7-Vps35 interaction is necessary for binding to endosome membranes^{37,38}, we found an unconventional *TgRab7*-*TgVps26* interaction, a singular feature previously described in *Entamoeba histolytica*⁴⁰.

In addition, we describe additional evidence for the phenotypic links between an absence of proteolytic processing of ROP/MIC proteins, mis-sorting of their unprocessed forms, and loss of rhoptry and microneme organelles. Our hypothesis is that *TgSORTLR*, in addition to acting as a sorting receptor for ROP/MIC proteins, delivers proteases, which are involved in the processing of pro-proteins, to secretory organelles.

The retromer interactome also includes several parasite-specific proteins. We first described *TgHP12* as a novel type I transmembrane protein with helical regions that share potential similarities with Rab-binding or SNARE-like domains. We showed that *TgHP12* localizes to the TGN and physically interacts with the core retromer complex *TgVps35*-*TgVps29*-*TgVps26* as well as with *TgSORTLR*. Genetic ablation of *TgHP12* using CRISPR-Cas9 indicates that its function is neither essential nor related to secretory organelle biogenesis. Instead, proteomic analyses revealed other unrelated secretory organelle proteins whose functions remain to be determined. Our current study also describes for the first time in apicomplexan parasites, the endocytic recycling of the multi-pass transmembrane transporter *TgHP03*, which belongs to the MFS, a class of membrane

transport proteins that facilitate movement of small solutes across cell membranes such as drugs, metabolites, oligosaccharides and amino acids in response to chemiosmotic gradients⁴¹. *TgHP03* accumulates in endocytic vesicles that are detected in the cytoplasm of *TgVps35*-deficient mutants, indicating that maintenance of *TgHP03* at the surface of *T. gondii* requires endocytic recycling from endosomes to the plasma membrane. Clearly, our study underscores the wide range of possible cargo molecules that are recycled by the retromer complex in light of the numerous identified transmembrane proteins that require future examination. A global analysis of retromer-mediated cargo *per se* will aid the delineation of the diverse metabolite and metal ion transporters required to maintain parasite nutrient homeostasis and intracellular replication.

It is also worth noting that the absence of *TgVps35* has negatively impacted the biogenesis of dense granules and body morphology in addition to that of rhoptries and micronemes, whereas *TgSORTLR* functions are restricted to rhoptry and microneme formation¹⁷. This disorganization of parasite body shape in the absence of *TgVps35* suggests a possible additional role of the retromer in regulating cytoskeleton and endosome functions as previously described in mammalian cells^{42,43}. However, the sorting nexins that are involved in this process in mammalian cells are not presently identified in *T. gondii*. Alternatively, we cannot rule out the possibility that the parasite-specific HP or other retromer-associated proteins could play this novel function that controls body morphology in *T. gondii*.

In conclusion, we demonstrated that the retromer complex of *T. gondii* is a major endosomal recycling hub required for sorting different cargo proteins that regulate diverse functions vital for parasite survival, as indicated in the our model presented in Fig. 10. This model also supports the notion that the retromer complex drives *TgSORTLR* recycling from Rab7-positive ELC to the TGN, thus sustaining another round of protein transport for proper secretory organelle biogenesis (Fig. 10a). In contrast to higher eukaryotes, in which an increase of lysosome degradation of the glucose transporter and other cargo sorting proteins,

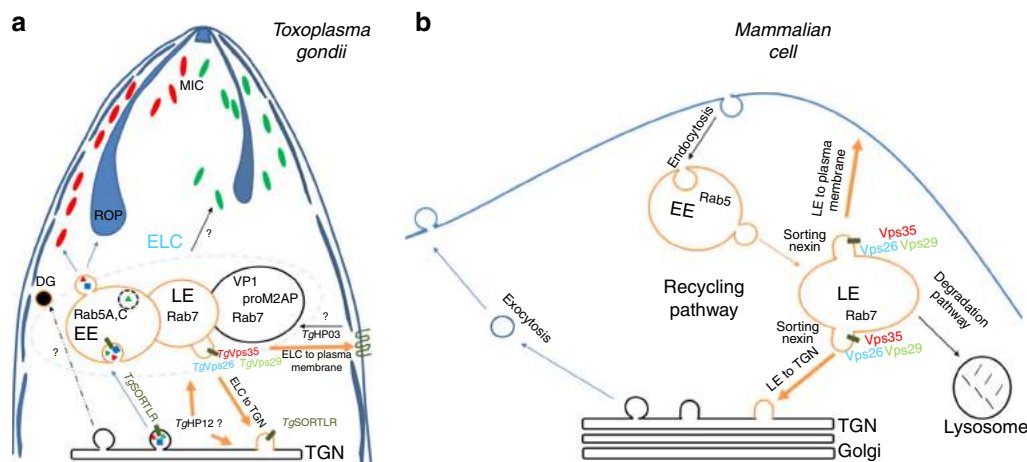


Figure 10 | Model of retromer-mediated recycling of diverse sorting cargoes in *T. gondii* versus mammalian cells. No recognizable sorting nexins, Bin/Amphiphysin/Rvs (BAR)-containing domain proteins or lysosomal-like protein degradation have been identified in *T. gondii* (a), in contrast to the situation in mammalian cells (b). Instead, the retromer-dependent recycling is essential for secretory organelle formation and parasite shape. We propose a model suggesting that Rab7 is the key small GTPase, which is involved in the endocytic recycling of *TgSORTLR* from endosomes to TGN for another round of ROP and MIC transport and secretory organelle biogenesis. We further raised the possibility that in contrast to mammalian cells, *T. gondii* lysosomal-like organelles only promote proteolytic maturation of proteins destined to secretion and that the endosomal system is adapted for organellar discharge of virulence-like factors required for the intracellular lifestyle of the parasite. Moreover, we provide the first evidence that a multiple ligand transmembrane transporter *TgHP03* is maintained at the surface of *T. gondii* through endocytic recycling from endosomes to the plasma membrane. BAR, Bin/Amphiphysin/Rvs; MIC, microneme; ROP, rhoptry; *TgHP03*, *T. gondii* hypothetical protein O3.

sortilin and mannose-6-phosphate receptor occurred in the absence of Vps35 (refs 28,34), the disruption of *TgVsp35* did not affect *TgSORTLR* and *TgHP03* degradation. This later observation also supports our model that in contrast to mammalian cells (Fig. 10b), *T. gondii* lysosomal-like organelles^{11,44,45}, only promote proteolytic maturation of proteins destined to secretion and that the endosomal system is adapted for organelle biogenesis and discharge of factors required for the intracellular lifestyle of the parasite (Fig. 10a). Our findings are expected to advance our understanding of endocytic recycling by *T. gondii*, highlighting the identity of unconventional endosomal factors, and this new knowledge may ultimately reveal new targets for managing toxoplasmosis.

Methods

Growth of host cells and parasite strains. *T. gondii* RH wild-type strain and its derivative RHΔku80 (ref. 46) or RH TaTi (ref. 47) strains were grown in human foreskin fibroblasts (HFF; from ATCC, USA) as described¹⁷. Plaque size assays were performed using HFF cells infected with 250 parasites in the presence or absence of $1.5 \mu\text{g ml}^{-1}$ of ATc for 6 days followed by crystal violet staining. All stained coverslips were imaged with an Axioplan microscope (Zeiss).

Generation of knockout and knock-in *T. gondii* strains. All primers used for PCR are listed in Supplementary Table 3. The conditional iKo*TgVps35* mutants were generated in the RH TaTi strain using the pG13-D-T7S4 plasmid, which contains 2 kb of 5' and 3' genomic DNA and the dihydrofolate reductase (DHFR) gene for pyrimethamine selection. After transfection of 5×10^6 parasites with 50 μg of linear plasmid, a stable line was cloned by limiting dilution. For complementation, iKo*TgVps35* mutants were transfected with 50 μg of plasmid containing the full-length cMyc-tagged *TgVps35* gene inserted into the uracil phosphoribosyl transferase locus. Stable transgenic and cloned parasites were selected with 5 μM 5-fluoro-2'-deoxyuridine. Transgenic *TgVps35*-HA, *TgVps26*-HA and *TgVps29*-cMyc parasites were generated with a knock-in strategy with DNA fragments of 2.1, 2.1 and 1.8 kb, respectively, cloned upstream of the stop codon from TGGT1_242660, TGGT1_263500 and TGGT1_252490 genes. DNA sequences were cloned into the pLIC-HA-DHFR and pLIC-cMyc-DHFR plasmids⁴⁶. Tachyzoites (5×10^6 parasites) of the RHΔku80 strain were transfected with 25 μg of linearized plasmids. Transgenic parasites expressing cMyc-tagged *TgHP03* (TGGT1_240810) and *TgHP12* (TGGT1_294220) were also generated as described above. All plasmids used in this study are listed in Supplementary Table 4.

Transient transformation and Cas9-mediated gene disruption. The plasmid pTOXO_Cas9-CRISPR (see map in Supplementary Fig. 4, panel c) corresponds to pUC57 carrying the C-terminally HA/GFP tagged *S. pyogenes Cas9* gene⁴⁸ fused to 2 nuclear localization sequences expressed under the control of the TUB8 promoter as well as the *TgU6* promoter driving the gRNA. Twenty mers-oligonucleotides corresponding to specific gene of interest were cloned using Golden Gate strategy⁴⁹. The ccdB positive-selection marker acts by killing the background of cells with no cloned DNA. The plasmid was synthesized and fully sequenced by GenScript (Singapore).

Confocal microscopy. Extracellular or intracellular parasites were fixed with 4% paraformaldehyde in phosphate-buffered saline for 20 min and processed as described^{17,50} using the indicated antibodies. The sources, origins and dilutions of all antibodies used for immunofluorescence assays were listed in Supplementary Table 5. Samples were observed with a Zeiss Confocal or Axiophot microscope, and images were processed using ZEN software (Zeiss).

Structured illumination microscopy. SIM was used to obtain high-resolution images using an ElyraPS1 microscope system (Zeiss) with a $\times 100$ oil immersion lens (alpha Plan Apochromat $\times 100$, NA 1.46, oil immersion). The resolution was measured using beads with a diameter of 100 nm (Tetraspeck multicolor). The point spread function was calculated using the metroloj plugin (ImageJ, NIH), which gave an x - y resolution of 125 nm and a z resolution of 500 nm. A voxel size of $0.050 \times 0.050 \times 0.150 \mu\text{m}^3$ was used for the measurement. Three lasers (405, 488 and 561 nm) were used for excitation. SIM images were acquired with an EMCCD camera (Andor Technology Ltd, UK) and processed with ZEN software, exposure times varied between 100 and 120 ms. Three-dimensional images were generated using a z -step of 150 nm (total thickness $\sim 5 \mu\text{m}$), while reconstructions and co-distributions were determined with IMARIS software (Bitplane). For co-distribution analysis, we used the colocalization module implemented in the IMARIS software. To process the images, we first applied a threshold for each channel (threshold adapted according to the labelling). For the Alexa488 channel, we used a threshold fixed at 15,000. For the red channel, the threshold varied between 8,000 and 25,000, depending on the dye used. The percentage (%) of

volume B (green channel) above the threshold that colocalized (colocalized volume B above threshold/total volume B above threshold) and thresholded Mander's coefficient B and Pearson coefficient were calculated. The acquisition was performed sequentially using 43HE, 38HE and BP 420–480 Zeiss filter sets, and 15 frames were acquired to reconstruct one image (five rotations \times three phases, with a SIM grating period of 51 μm for the blue channel, 42 μm for the green channel and 34 μm for the red channel). Beads with a diameter of 100 nm were imaged to measure the chromatic misalignment of our system (fit procedure by the ZEN software), and these parameters enabled further correction of the alignment on each acquired multi-channel stack. Image reconstruction was achieved using ZEN software with the following parameters: Noise Level -4, Sectioning 99-83-83 and Frequency Weighting 1.

Specifically, iKo*TgVps35* mutants transiently transfected were treated with ATc or left untreated for 24 h or 48 h, fixed and stained with primary specific antibodies and either Dylight594- or Alexa488-conjugated secondary antibodies. Forty eight hours of ATc treatment corresponds to the time point where optimal effects on rhoptry and microneme biogenesis were observed and expression of *TgVps35* protein was completely ablated as demonstrated by western blots and IFA. Therefore, all other experiments were then performed at this time point in this study. For GRASP-RFP expressing parasites, the fluorescence signal was directly visualized by SIM. Colocalization was quantified as overlap between green and red images using the IMARIS software for 90 intracellular parasites from 7 to 8 vacuoles.

Electron microscopy. For transmission electron microscopy, intracellular tachyzoites of iKo*TgVps35*, complemented mutants, or parental parasites were fixed overnight at 4 °C with 2.5% glutaraldehyde/4% paraformaldehyde prepared 0.1 M cacodylate buffer, cells were fixed in 2.5% glutaraldehyde prepared in 0.1 M cacodylate buffer and post-fixed in 1% osmium tetroxide in the same buffer. After acetonitrile dehydration, the pellet was embedded in Epon. Ultrathin sections (90 nm) were cut using a Leica UC7 ultramicrotome and collected on 150 mesh hexagonal barred copper grids. After staining with 2% uranyl acetate prepared in 50% ethanol and incubation with a lead citrate solution, sections were observed on a Hitachi H-600 transmission electron microscope at 75 kV accelerating voltage.

Host cell invasion assays. The conditional iKo*TgVps35* mutants were treated with ATc or left untreated for 48 h and then used to infect HFF cells for 1 h at 37 °C. Twenty-four hours post-infection, coverslips were fixed and processed for immunofluorescence assays. The intracellular parasites were labelled with anti-SAG1 antibodies, and the nucleus was stained with 4', 6-diamidino-2-phenylindole. For each condition, 400 microscopic fields ($\times 40$) were observed. The number of SAG1-positive parasites was compared with the number of cells stained with 4', 6-diamidino-2-phenylindole. The ratio parasites/cells are presented as mean values \pm s.d. from three independent experiments.

Mouse infectivity studies. A group of 4 female BALB/C mice (6–8-week-old, from Janvier Labs, France) were intraperitoneally injected with 2×10^6 tachyzoites of RH TaTi, iKo*TgVps35* or Comp-iKo*TgVps35* parasites. To suppress *TgVps35* expression, the drinking water was supplemented with 0.2 mg ml⁻¹ of ATc and 5% of sucrose. As controls, a group of four mice was also infected with the same parasite lines without ATc. Survival was monitored over 4 weeks. For vaccination assays, BALB/C mice were infected by intraperitoneal injection with 3×10^2 tachyzoites of RH TaTi, iKo*TgVps35* or Comp-iKo*TgVps35* parasites followed by ATc treatment or not. After 4 weeks, the survived mice were re-challenged with 1×10^3 tachyzoites of wild-type RH strain and survival monitored for 30 days. All animal experiments were performed following the guidelines of the Pasteur Institute of Lille animal study board, which conforms to the Amsterdam Protocol on animal protection and welfare, and Directive 86/609/EEC on the Protection of Animals Used for Experimental and Other Scientific Purposes, updated in the Council of Europe's Appendix A (<http://conventions.coe.int/Treaty/EN/Treaties/PDF/123-Arev.pdf>). The animal work also complied with the French law (no. 87-848 dated 19 october 1987) and the European Communities Amendment of Cruelty to Animals Act 1976. All animals were fed with regular diet and all procedures were in accordance with national regulations on animal experimentation and welfare authorized by the French Ministry of Agriculture and Veterinary committee (permit number: 59-009145).

Co-immunoprecipitation and western blots. Tachyzoites (1×10^9 parasites) from *TgVps35*-HA, *TgVps26*-HA, *TgVps29*-cMyc and RHΔku80 strains were lysed with 10 mM HEPES, pH 7.9; 1.5 mM MgCl₂; 10 mM KCl; 0.5 mM dithiothreitol (DTT); 0.1 mM EDTA; 0.65% NP40; 0.5 mM phenylmethanesulfonylfluoride (PMSF); and a cocktail of protease inhibitors (Sigma Aldrich). After 30 min on ice, the lysates were centrifuged at 21, 693g for 30 min at 4 °C, and the supernatants were then incubated with anti-HA or anti-cMyc agarose beads (Thermo Pierce) overnight at 4 °C under rotating shaker. After five washings with 10 mM Tris, pH 7.5; 150 mM NaCl; 0.2% Triton X-100; 0.5 mM PMSF, and a final washing with 62.5 mM Tris, pH 6.8, immunoprecipitates were eluted with Laemmli buffer (0.2% SDS, 100 mM DTT, and 10% sucrose). These samples were analysed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) followed by silver staining before proteomics analyses. For western blots, parasites were lysed with Laemmli

buffer, and lysate proteins were separated on 12% acrylamide gels and processed for immunoblotting using primary specific antibodies and then secondary antibodies conjugated to alkaline phosphatase (Thermo Pierce). All primary antibodies used for western blots were listed in Supplementary Table 5. The blots were imaged using ChemiDoc XRS⁺ (Bio-Rad).

Liquid chromatography coupled to mass spectrometry. For liquid chromatography and mass spectrometry performed using Q-Exactive mass spectrometer (Thermo Scientific, Bremen, Germany), each sample (30 μ l) was denatured with Laemmli sample buffer and loaded onto one-dimensional SDS-PAGE (12%), stained with colloidal Coomassie G-250 (Bio-Rad, Hercules, CA) and six slices per sample were excised for reduction and alkylation. For this step, gel slices were cut in small pieces (1 mm³) and the staining of gel pieces were removed thrice with 120 μ l of a mixture of 50/50 (v/v), 25 mM ammonium bicarbonate (NH₄CO₃)/acetonitrile for 10 min. In-gel reduction and alkylation of protein disulfide bonds were performed, respectively, with 100 μ l of 10 mM of DTT in 50 min at 57 °C and 100 μ l of 50 mM of iodoacetamide (IAM) for 30 min at room temperature. After a washing step with 120 μ l of 25 mM NH₄CO₃ and the dehydration step with 100 μ l acetonitrile for 5 min, an in-gel digestion was performed on each sample with 0.07 μ g of sequencing grade porcine trypsin (Promega, Madison, WI) for 16 h at 37 °C. The peptide were extracted thrice from gel with a mixture of 60/40/0.1 (v/v/v), acetonitrile/25 mM of NH₄CO₃ (v/v) and 0.1% formic acid. The extracted solution were then dried with vacuum centrifuge (Uniequip GmbH, Munich, Germany) and resuspended in 10 μ l of water containing 0.1% formic acid.

Each extracted solution of 3 μ l was injected into the Ultimate 3,000 RSLC nano-System (Dionex, Thermo Scientific) through a trap column (Acclaim PepMap, 5 mm \times 300 μ m inner diameter, C18, 5 μ m, 100 Å; Dionex) at 5 μ l min⁻¹ with water containing 0.1% formic acid and 2% acetonitrile. After 5 min, the trap column was set on-line with analytical column (Acclaim PepMap RSLC, 15 cm \times 75 μ m inner diameter, C18, 2 μ m, 100 Å; Dionex, Sunnyvale, CA). The elution was carried out by applying mixture of solvent A (HPLC grade water with 0.1% formic acid) and solvent B (HPLC grade acetonitrile with 0.1% formic acid) at the flow rate of 300 nl min⁻¹. The separations were performed by applying a linear gradient of 2–40% solvent B over 38 min followed by a washing step (5 min at 70% solvent B) and an equilibration step (11 min at 2% solvent B).

The eluted peptides were analysed by a Q-Exactive mass spectrometer. For ionization, a nanospray Flex Ion Source was used with a voltage of 1.9 kV, and the capillary temperature was 275 °C. Full MS scans were acquired in the Orbitrap mass analyser over m/z 300–3,500 range with resolution of 70,000 at m/z 200. The target automatic gain control value of 1×10^6 was used with a maximum allowed injection time (Maximum IT) of 250 ms. For MS/MS, an isolation window of 2 m/z was utilized. Ten most intense peaks (TopN) with charge state between 2 and 6 were selected for fragmentation in the higher-energy collisional induced dissociation cell with normalized collision energy of 35. The tandem mass spectra were acquired over m/z 200–2,000 range in the Orbitrap mass analyser with resolution 35,000 at m/z 200 and an automatic gain control of 2×10^5 . The ion intensity selection threshold was 6.7×10^4 , and the maximum injection time was 150 ms. The dynamic exclusion time was 10 s and the total run time was 60 min. All these systems were fully controlled by Thermo Xcalibur 3.0 (Thermo Fisher Scientific).

All data files (*.raw) collected were processed with a specific workflow designed in Proteome Discoverer 1.4 (Thermo Fisher Scientific). MS/MS data was interpreted using two search engine Mascot (version 2.4.1, Matrix Science, London, UK) and Sequest HT (Thermo Fisher Scientific). Searches were performed against *T. gondii* (TGVEG, TGME49 and TGGT1 stain) protein sequences downloaded from www.toxodb.org at the 11th December 2014 (18,954 entries). The Mascot ion score were > 20 and Sequest HT XCorr > 1.5. The target-decoy database search allowed us to control and to estimate the false positive identification rate⁵¹.

Nanoscale liquid chromatography and tandem mass spectrometry.

For nanoscale liquid chromatography coupled to tandem mass spectrometry (NanoLC-MS/MS), samples were electrophoresed onto 12% SDS-PAGE and stained overnight with colloidal Coomassie Brilliant Blue. Gel bands were manually excised, reduced 1 h at 60 °C by adding DTT to a final concentration of 10 mM and alkylated by adding iodoacetamide to a final concentration of 40 mM. An overnight digestion was performed by adding trypsin (Promega). Tryptic peptides were extracted (60% acetonitrile, 0.1% HCOOH) before mass spectrometry analyses. NanoLC-MS/MS analyses were performed on three different systems: nano-ACQUITY Ultra-Performance-LC system (UPLC; Waters, Milford, MA, USA) hyphenated to either Q-TOF Impact HD or MaXis 4G (Bruker Daltonics, Bremen, Germany) and a nanoLC-Chip/Cube (Agilent Technologies, Palo Alto, CA, USA) hyphenated to an ion trap amaZon (Bruker Daltonics). For MaXis 4G and amaZon analysis, methods used were previously described⁵² with slight modifications. For Impact HD analysis, peptides were first trapped on a 0.18 \times 20 mm², 5 μ m Symmetry C18 pre-column (Waters), and then separated on an ACQUITY UPLC BEH130 C18 column (Waters), 75 mm \times 250 mm with 1.7 μ m particle size. The solvent system consisted of 0.1% HCOOH in water (solvent A) and 0.1% HCOOH in acetonitrile (solvent B). Trapping was performed for 3 min at 5 μ l min⁻¹ with 99% A and 1% B. Elution was performed at a flow rate of 450 nl min⁻¹, using a 1–35% gradient (solvent B) over 30 min at 60 °C.

The mass spectrometer was equipped with a Captive Spray source and a nano-Booster operating in positive mode, with the following settings: source temperature was set at 150 °C while drying gas flow was at 3 l min⁻¹. The nano-electrospray voltage was optimized to $-1,300$ V. External mass calibration of the time-of-flight (TOF) was achieved before each set of analyses using Tuning Mix (Agilent Technologies) in the mass range of 322–2,722 m/z . Mass correction was achieved by recalibration of acquired spectra to the applied lock masses hexakis (2,2,3,3-tetrafluoropropoxy) phosphazine ([M + H]⁺ = 922.0098 m/z).

For tandem MS experiments (CID), the system was operated with fixed cycle time of 3 s in the range of 150–2,200 m/z . MS/MS scan speed was monitored in function of precursor intensity from 4 to 25 Hz. Ions were excluded after acquisition of one MS/MS spectra and the exclusion was released after 1 min. The complete system was fully controlled by Hystar 3.2 (Bruker Daltonics).

Bioinformatics and protein identification. Mass data collected during nanoLC-MS/MS analyses were processed, converted into *.mgf files with Data Analysis 4.0 (Bruker Daltonics) and interpreted using MASCOT 2.5.1 algorithm (Matrix Science, London, UK) running on a local server. Searches were performed without any molecular weight or isoelectric point restrictions against an in-house generated protein database composed of protein sequences of *T. gondii* (ToxoDB database, October 2014) and known contaminant proteins such as human keratins and trypsin. All proteins were concatenated with reversed copies of all sequences (49,328 entries) with an in-house database generation toolbox <https://msda.unistra.fr>⁵³. Trypsin was selected as enzyme, carbamidomethylation of cysteine (+ 57 Da) was set as fixed modification, oxidation of methionine (+ 16 Da) were set as variable modification and both precursor and fragment mass tolerances were adapted according to instrumental mass accuracy. Mascot results were loaded into the Proline software (Proline Studio Release 1.0) and filtered to obtain a false discovery rate of < 1%.

Absolute quantitation using LC-SRM. For microLC-SRM assay, three proteotypic peptides per targeted protein (TgVps35, TgVps26 and TgVps29) were selected. A total of nine high-purity isotopically labelled equivalent peptides were purchased (AQUA peptides, Thermo Fischer Scientific, Bremen; Germany). Previous nanoLC-MS/MS analyses afforded a representative MS/MS spectrum for each peptide. Four to six transitions were monitored for both endogenous and heavy-labelled peptides. Thus, a total of 78 transitions corresponding to 20 precursors and 3 proteins were measured. For the lower limits of quantification and detection determination, a dilution series of the labelled peptides was realized at different concentrations in a mixture containing all tagged proteins and injected in triplicate on a QQQ-6490 triple quadrupole mass spectrometer (Agilent Technologies). The area under curve of the three best transitions per peptide were summed and drawn versus the peptide concentration. Two calibration curves were drawn *per* peptide: high calibration curve (15 fmol μ l⁻¹–238 fmol μ l⁻¹) and low-calibration curve (2 fmol μ l⁻¹–30 fmol μ l⁻¹). We evaluated the lower limits of quantification and the lower limits of detection determination by applying recognized definitions⁵⁴.

For the SRM analyses, samples were electrophoresed onto 4% SDS-PAGE and stained for 45 min with colloidal Coomassie Brilliant Blue. The stacking gel bands were predigested and digested as previously described and 1 μ l of a mixture of heavy-labelled peptides was added to each sample before LC-SRM analyses.

All separations were carried out on an Agilent 1100 Series HPLC system (Agilent Technologies). For each analysis, the sample was loaded into a trapping column ZORBAX 300SB-C18 MicroBore Guard 5 μ m, 1.0 \times 17 mm² (Agilent Technologies) at 50 μ l min⁻¹ with aqueous solution containing 0.1% (v/v) formic acid and 2% acetonitrile. After 3 min trapping, the column was put on-line with a ZORBAX 300SB-C18 3.5 μ m, 0.3 \times 150 mm² column (Agilent Technologies). Peptide elution was performed at 5 μ l min⁻¹ by applying a linear gradient of solvent A (water with 2% acetonitrile and 0.1% (v/v) formic acid) and B (acetonitrile with 0.1% (v/v) formic acid), from 8 to 42% solvent B over 30 min followed by a washing step (2 min at 90% solvent B) and an equilibration step (13 min at 8% solvent B). The isolation width for both Q1 and Q3 was set to 0.7 m/z unit. The collision energy was experimentally optimized by testing nine values centred on the calculated value from the one given by the supplier. Time-scheduled SRM method targeted the pairs of isotopically labelled peptides/endogenous peptides in ± 5 min retention time windows within a cycle time of 3 s. Mass data collected during LC-SRM were processed with the Skyline open-source software package 3.1.1 (ref. 55). Area intensity ratios of the heavy and the light forms of each peptide were manually checked. The endogenous peptide amount calculation was performed by using the most suitable calibration curve. The mass spectrometry and LC-SRM data were deposited to the ProteomeXchange⁵⁶ Consortium via PRIDE⁵⁷ partner repository, and the peptide atlas SRM experiment library (PASSEL), respectively.

Production of recombinant glutathione S-transferase (GST)-TgHP12 and specific antibodies. The DNA corresponding to C-terminal sequence of 254 amino acids long from 1,051 to 1,812 nucleic acid was amplified by PCR using the following primers: sense 5'-CCGGGGATCCGTAGAAAAAGCCTACAACGGTG GGG-3', and antisense, 5'-CCGGGGCGCCGCTACAATCTGTCAAGTCTTCTCT

CCAGTC-3'. The amplified DNA was purified and cloned in frame into pGEX6P3. After verification by sequencing, the plasmid was used to transform *E. coli* BL21 for recombinant protein expression. Protein was purified by GST column and 100 µg of protein was used to immunize one Wistar (RjHan:WI) rat (Janvier Labs, France) using complete Freund adjuvant. The rat was challenged three times with 50 µg of protein and incomplete Freund adjuvant before bleeding 10 days after the last boost and the serum was purified.

GST-pull down experiments. Recombinant Rab5B, Rab7 and Rab11B proteins were fused to GST using pGEX6P3. After transformation of BL21 *E. coli*, lysates were incubated to Glutathione-beads and washed four times with buffer A: 50 mM Tris-HCl pH 7.5, 0.5M NaCl, 270 mM sucrose 1 mM EGTA, 1 mM EDTA, 1% Triton X-100 and 0.5 mM PMSF) and washed six times with buffer A without Triton X-100. Total TgVps35-HA and TgVps26-HA knock-in parasite extracts were prepared from 10⁹ tachyzoites that were lysed with buffer B: 10 mM HEPES pH 7.9 1.5 mM MgCl₂, 10 mM KCl, 0.5 mM DTT, 0.1 mM EDTA, 0.65% NP40 and 0.5 mM PMSF. The parasite lysate (equivalent to 2.0 × 10⁸ tachyzoites) were added in pull down buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl and 0.5 mM PMSF and incubated with beads containing 2 µg GST-Rab5, GST-Rab7, GST-Rab11B or GST alone in the presence of 1 mM GTP or GDP overnight at 4 °C. Precipitants were washed three times with the pull down buffer containing 0.1% Triton X-100 and eluted by Laemmli buffer and analysed by western blots, which were probed with anti-HA antibodies and anti-GST.

Statistical analysis. All data were analysed with Graph Pad Prism software. A Student's *t*-test was used for statistical analysis. The Mann-Whitney test was also used for analysis of mice survival curves.

References

- Pappas, G., Roussos, N. & Falagas, M. E. Toxoplasmosis snapshots: Global status of *Toxoplasma gondii* seroprevalence and implications for pregnancy and congenital toxoplasmosis. *Int. J. Parasitol.* **39**, 1385–1394 (2009).
- Hager, K. M., Striepen, B., Tilney, L. G. & Roos, D. S. The nuclear envelope serves as an intermediary between the ER and Golgi complex in the intracellular parasite *Toxoplasma gondii*. *J. Cell Sci.* **112**, 2631–2638 (1999).
- Tomavo, S., Slomianny, C., Meissner, M. & Carruthers, V. B. Protein trafficking through the endosomal system prepares intracellular parasites for a home invasion. *PLoS Pathog.* **9**, e1003629 (2013).
- Breinich, M. S. *et al.* A dynamin is required for the biogenesis of secretory organelles in *Toxoplasma gondii*. *Curr. Biol.* **19**, 277–286 (2009).
- Kremer, K. *et al.* An overexpression screen of *Toxoplasma gondii* Rab-GTPases reveals distinct transport routes to the micronemes. *PLoS Pathog.* **9**, e1003213 (2013).
- Jackson, A. J., Clucas, C., Mamczur, N. J., Ferguson, D. J. & Meissner, M. *Toxoplasma gondii* syntaxin 6 is required for vesicular transport between endosomal-like compartments and the Golgi complex. *Traffic* **14**, 1166–1181 (2013).
- Morion-Guyot, J., Pastore, S., Berry, L., Lebrun, M. & Daher, W. *Toxoplasma gondii* Vps11, a subunit of HOPS and CORVET tethering complexes is essential for the biogenesis of secretory organelles. *Cell. Microbiol.* **17**, 1157–1178 (2015).
- Mogelvang, S., Marsh, B. J., Ladinsky, M. S. & Howell, K. E. Predicting function from structure: 3D structure studies of the mammalian Golgi complex. *Traffic* **5**, 338–345 (2004).
- He, C. Y. Golgi biogenesis in simple eukaryotes. *Cell. Microbiol.* **9**, 566–572 (2007).
- Brydges, S. D., Harper, J. M., Parussini, F., Coppens, I. & Carruthers, V. B. A transient forward-targeting element for microneme-regulated secretion in *Toxoplasma gondii*. *Biol. Cell* **100**, 253–264 (2008).
- Parussini, F., Coppens, I., Shah, P. P., Diamond, S. L. & Carruthers, V. B. Cathepsin L occupies a vacuolar compartment and is a protein maturase within the endo/exocytic system of *Toxoplasma gondii*. *Mol. Microbiol.* **76**, 1340–1357 (2010).
- Dubremetz, J.-F. Rhoptries are major players in *Toxoplasma gondii* invasion and host cell interaction. *Cell Microbiol.* **9**, 841–848 (2007).
- Schmidt, O. & Teiss, D. The ESCRT machinery. *Curr. Biol.* **22**, R116–R120 (2012).
- Lige, B. *et al.* Introduction of caveolae structural proteins into the protozoan *Toxoplasma* results in the formation of heterologous caveolae but not caveolar endocytosis. *PLoS ONE* **12**, e51773 (2012).
- Pieperhoff, M. S., Schmitt, M., Ferguson, D. J. & Meissner, M. The role of clathrin in post-Golgi trafficking in *Toxoplasma gondii*. *PLoS ONE* **8**, e77620 (2013).
- Dou, Z., McGovern, O. L., Di Cristina, M. & Carruthers, V. B. *Toxoplasma gondii* ingests and digests host cytosolic proteins. *MBio* **5**, e01188-14 (2014).
- Sloves, P. J. *et al.* *Toxoplasma* sortilin-like receptor regulates protein transport and is essential for apical secretory organelle biogenesis and host infection. *Cell Host Microbe* **11**, 515–527 (2012).
- Bonifacino, J. S. & Hurley, J. H. Retromer. *Curr. Opin. Cell Biol.* **20**, 427–436 (2008).
- Gallon, M. & Cullen, P. J. Retromer and sorting nexins in endosomal sorting. *Biochem. Soc. Trans.* **43**, 33–47 (2015).
- Norwood, S. J. *et al.* Assembly and solution structure of the core retromer protein complex. *Traffic* **12**, 56–71 (2011).
- Agop-Neresian, C. *et al.* Biogenesis of the inner membrane complex is dependent on vesicular transport by the alveolate specific GTPase Rab11B. *PLoS Pathog.* **6**, e1001029 (2010).
- Seaman, M. N. J., Harbour, M. E., Tattersall, D., Read, E. & Bright, N. Membrane recruitment of the cargo-selective retromer subcomplex is catalyzed by the small GTPase Rab7 and inhibited by the Rab-GAP TBC1D5. *J. Cell Sci.* **122**, 2371–2382 (2009).
- Fox, B. A. & Bzik, D. J. De novo pyrimidine biosynthesis is required for virulence of *Toxoplasma gondii*. *Nature* **415**, 926–929 (2002).
- Tomavo, S., Schwarz, R. T. & Dubremetz, J.-F. Evidence for glycosyl-phosphatidylinositol anchoring of *Toxoplasma gondii* major surface antigens. *Mol. Cell. Biol.* **9**, 4576–4580 (1989).
- Di Crisanti, M., Spaccapelo, R., Soldati, D., Bistoni, F. & Crisanti, A. Two conserved amino acid motifs mediate protein targeting to the micronemes of the apicomplexan parasite *Toxoplasma gondii*. *Mol. Cell. Biol.* **20**, 7332–7341 (2000).
- Arighi, C. N. *et al.* Role of the mammalian retromer in sorting of the cation-independent mannose 6-phosphate receptor. *J. Cell Biol.* **165**, 123–133 (2004).
- Seaman, M. N. J. The retromer complex-endosomal protein recycling and beyond. *J. Cell Sci.* **125**, 4693–4702 (2012).
- Steinberg, F. *et al.* A global analysis of SNX27-retromer assembly and cargo specificity reveals a function in glucose and metal ion transport. *Nat. Cell Biol.* **15**, 461–471 (2013).
- Maruzs, T. *et al.* Retromer ensures the degradation of autophagic cargo by maintaining lysosome function in *Drosophila*. *Traffic* **16**, 1088–1107 (2015).
- Eathiraj, S., Pan, X., Ritacco, C. & Lambright, D. G. Structural basis of family-wide Rab GTPase recognition by rabenosyn-5. *Nature* **436**, 415–419 (2005).
- Lall, P. *et al.* Structural and functional analysis of FIP2 binding to the endosome-localised Rab25 GTPase. *Biochim. Biophys. Acta* **1834**, 2679–2690 (2013).
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
- Law, C. J., Maloney, P. C. & Wang, D. N. Ins and outs of major facilitator superfamily antiporters. *Annu. Rev. Microbiol.* **62**, 289–305 (2008).
- McGough, I. J. *et al.* Identification of molecular heterogeneity in SNX27-retromer-mediated endosome-to-plasma membrane recycling. *J. Cell Sci.* **127**, 4940–4953 (2014).
- Cullen, P. J. Endosomal sorting and signalling: an emerging role for sorting nexins. *Nat. Rev. Mol. Cell Biol.* **9**, 574–582 (2008).
- Harrison, M. S. *et al.* A mechanism for retromer endosomal coat complex assembly with cargo. *Proc. Natl Acad. Sci. USA* **111**, 267–272 (2014).
- Rojas, R. *et al.* Regulation of retromer recruitment to endosomes by sequential action of Rab5 and Rab7. *J. Cell Biol.* **183**, 513–526 (2008).
- Liu, T. T., Gomez, T. S., Sackey, B. K., Billadeau, D. D. & Burd, C. G. Rab GTPase regulation of retromer-mediated cargo export during endosome maturation. *Mol. Biol. Cell.* **23**, 2505–2515 (2013).
- Swarbrick, J. D. *et al.* VPS29 is not an active metallo-phosphatase but is a rigid scaffold required for retromer interaction with accessory proteins. *PLoS ONE* **6**, e20420 (2011).
- Nakada-Tsukui, K., Saito-Nakano, Y., Ali, V. & Nozaki, T. A retromer-like complex is a novel Rab7 effector that is involved in the transport of the virulence factor cysteine protease in the enteric protozoan parasite *Entamoeba histolytica*. *Mol. Biol. Cell* **16**, 5294–5303 (2005).
- Pao, S. S., Paulsen, I. T. & Saier, M. H. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **62**, 1–34 (1998).
- Cullen, P. J. & Korswagen, H. C. Sorting nexins provide diversity for retromer-dependent trafficking events. *Nat. Cell Biol.* **14**, 29–37 (2011).
- Gautreau, A., Oguievetskaia, K. & Ungermann, C. Function and regulation of the endosomal fusion and fission machineries. *Cold Spring Harb. Perspect. Biol.* **1**, 6 (2014).
- Miranda, K. *et al.* Characterization of a novel organelle in *Toxoplasma gondii* with similar composition and function to the plant vacuole. *Mol. Microbiol.* **76**, 1358–1375 (2010).
- Tomavo, S. Evolutionary repurposing of endosomal systems for apical organelle biogenesis in *Toxoplasma gondii*. *Int. J. Parasitol.* **44**, 133–138 (2014).
- Huynh, M. H. & Carruthers, V. B. Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. *Eukaryot. Cell* **8**, 530–539 (2009).
- Sheiner, L. *et al.* A systematic screen to discover and analyze apicoplast proteins identifies a conserved and essential protein import factor. *PLoS Pathog.* **7**, e1002392 (2011).
- Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
- Weber, E. *et al.* Assembly of designer TAL effectors by Golden Gate cloning. *PLoS ONE* **6**, e19722 (2011).

50. Sloves, P. J. *et al.* Apical organelle secretion by *Toxoplasma* controls innate and adaptive immunity and mediates long-term protection. *J. Infect. Dis.* **212**, 1449–1458 (2015).
51. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207–214 (2007).
52. Gissot, M. *et al.* *Toxoplasma gondii* Alba proteins are involved in translational control of gene expression. *J. Mol. Biol.* **425**, 1287–1301 (2013).
53. Carapito, C. *et al.* MSData, a proteomics software suite for in-depth Mass Spectrometry Data Analysis using grid computing. *Proteomics* **14**, 1014–1019 (2014).
54. Domanski, D. *et al.* MRM-based multiplexed quantitation of 67 putative cardiovascular disease biomarkers in human plasma. *Proteomics* **12**, 1222–1243 (2012).
55. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968 (2010).
56. Vizcaíno, J. A. *et al.* ProteomeXchange provides globally co-ordinated proteomics data submission and dissemination. *Nat. Biotechnol.* **30**, 223–226 (2014).
57. Vizcaíno, J. A. *et al.* 2016 update of the PRIDE database and related tools. *Nucleic Acids Res.* **44**, 447–456 (2016).

Acknowledgements

L.O.S. would like to thank Thomas Mouveaux, Pasteur Institute of Lille; Etienne Dewailly, University of Lille; Joelle Duflot, Pasteur Institute of Lille for excellent technical assistance and Pierre Julien Sloves, Pasteur Institute of Lille for his early guidance. This work was supported by the following grants: Laboratoire d'Excellence (LabEx) ParaFrap from the National Agency for Research ANR-11-LABX-0024 and the ANR-14-CE14-0002-01. We also acknowledge additional financial support from the INSERM, Pasteur Institute of Lille and CNRS. PhD and Post-doc fellowships were from 'Métropole Européenne de Lille' (MEL) to L.O.S. and 'FEDER-Région' for T.D.A. M.A.H is supported by the European Research Council (ERC Consolidator grant no. 614880).

Author contributions

L.S.O., T.D.A., B.W., A.H., I.C., C.S.R. and A.V.D. conceived and designed the experiments, analysed data and wrote paper; F.S. and M.A.H.: engineering CRISPR-Cas9 vectors; E.W. and F.L.: SIM imaging; C.S.: electron microscopy; S.T. designed, supervised and wrote this study.

Additional information

Accession codes: The mass spectrometry proteomics data collected during nanoLC-MS/MS analyses have been deposited to the ProteomeXchange Consortium via PRIDE partner repository with the data set identifier PXD003603. The LC-SRM data have been deposited in the Peptide Atlas SRM Experiment Library (PASSEL) with the data set identifier PASS00824.

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interest.


Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Sangaré, L. O. *et al.* Unconventional endosome-like compartment and retromer complex in *Toxoplasma gondii* govern parasite integrity and host infection. *Nat. Commun.* **7**:11191 doi: 10.1038/ncomms11191 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

SCIENTIFIC REPORTS



OPEN

A Critical Role for *Toxoplasma gondii* Vacuolar Protein Sorting VPS9 in Secretory Organelle Biogenesis and Host Infection

Takaya Sakura¹, Fabien Sindikubwabo², Lena K. Oesterlin³, Hugo Bousquet³, Christian Slomianny⁴, Mohamed-Ali Hakimi², Gordon Langsley⁵ & Stanislas Tomavo¹

Accurate sorting of proteins to the three types of parasite-specific secretory organelles namely rhoptry, microneme and dense granule in *Toxoplasma gondii* is crucial for successful host cell invasion by this obligate intracellular parasite. Despite its tiny body architecture and limited trafficking machinery, *T. gondii* relies heavily on transport of vesicles containing proteins, lipids and important virulence-like factors that are delivered to these secretory organelles. However, our understanding on how trafficking of vesicles operates in the parasite is still limited. Here, we show that the *T. gondii* vacuolar protein sorting 9 (*TgVps9*), has guanine nucleotide exchange factor (GEF) activity towards Rab5a and is crucial for sorting of proteins destined to secretory organelles. Our results illuminate features of *TgVps9* protein as a key trafficking facilitator that regulates protein maturation, secretory organelle formation and secretion, thereby ensuring a primary role in host infection by *T. gondii*.

Toxoplasma gondii is an important food and waterborne pathogen causing toxoplasmosis, a usually mild disease in immunocompetent humans that can turn into a major threat in immunocompromised patients and during primary infection of pregnant woman. *T. gondii* is a member of the *Apicomplexa*, a phylum of numerous medically important parasites causing life-threatening diseases in human and animals worldwide. The phylum is typified by specific secretory organelles called rhoptries, micronemes and dense granules that are essential for host cell invasion and host pathway modulation. In *Toxoplasma*, rhoptries contain two groups, termed rhoptry (ROP) and rhoptry neck (RON), of effector proteins some of which are virulence factors; whereas micronemes secrete MIC proteins that are involved in parasite gliding, host cell attachment and invasion^{1,2}. After invasion, dense granules discharge GRA proteins involved in parasitophorous vacuole (PV) formation and in hijacking host cell gene expression and metabolism³.

Despite having a single cell architecture, the parasite relies on active and abundant vesicle and protein trafficking. *T. gondii* and likely all *Apicomplexa* have reutilized classical endosomal and endocytic trafficking pathways more typical of higher eukaryotes towards building specialized secretory organelles that release parasite effectors to interplay with host cell signaling pathways as a way to take control over host immunity and ultimately to promote long-term parasitism^{4–8}. It is now well established that apicomplexan parasites operate an unconventional endosome-like system (ELC) to traffic proteins from the Golgi apparatus to rhoptries and micronemes^{6–8}. However, the mechanisms involved in endosome-like vesicle formation and delivery to the aforementioned organelles in general remain elusive. In mammalian cells, the endosomal system is used for the uptake of plasma membrane-associated components, which after passage through Rab5-positive early endosomes (EE) enter either Rab11A-positive recycling endosomes to return to the plasma membrane, or Rab7-positive late endosomes to be delivered to lysosomes (LE)⁹.

¹Laboratory of Cellular and Molecular Biology of Toxoplasma, Université de Lille, Institut Pasteur de Lille, Center for Infection and Immunity of Lille, INSERM U 1019, CNRS UMR 8204, 59000, Lille, France. ²Institute for Advanced Biosciences (IAB), INSERM U1209, CNRS UMR5309, Université Grenoble Alpes, 38700, Grenoble, France. ³Institut Curie, PSL Research University, CNRS UMR144, Molecular Mechanisms of Intracellular Transport Group, 75248, Paris, France. ⁴Laboratory of Cell Physiology, INSERM U 1003, Université de Lille, 59655, Villeneuve d'Ascq, France. ⁵Laboratoire de Biologie Cellulaire Comparative des Apicomplexes, INSERM U1016, CNRS UMR8104, Institut Cochin, 75014, Paris, France. Correspondence and requests for materials should be addressed to S.T. (email: stanislas.tomavo@gmail.com)

Clearly, regulated vesicular traffic allows different cargoes to correctly reach their specific organelle destinations at the right time^{4–8}, and this is essential for successful parasite infection of its host^{7,10}. For example, dynamin-related protein B (DrpB) and clathrin, which reside in the post-Golgi network (TGN) and the endosomal-like compartment (ELC) contribute to the formation of transport vesicles that are essential for secretory organelle biogenesis^{11,12}. Vacuolar protein sorting 11 (Vps11) that is the subunit of CORVET (class C core vacuole/endosome tethering) and HOPS (homotypic fusion and vacuolar protein sorting) complexes are required for transport of MIC and ROP proteins to micronemes and rhoptries of *T. gondii*¹³. Thus, the parasite intra-vesicular trafficking of the endolysosome pathway involves functions of the CORVET and HOPS tethering complex. In addition, *T. gondii* Rab5+ and Rab7+ effector complexes likely interact with CORVET and HOPS in a manner similar to mammalian cells to induce membrane fusion within the endolysosome pathway of the parasite^{14,15}.

We have described that transport of MIC and ROP proteins to microneme and rhoptry organelles, respectively, also required an essential sortilin-like receptor named *TgSORTLR*¹⁰ and traffic through a non-conventional ELC⁷. The C-terminal tail of *TgSORTLR* interacts with clathrin, three components of the AP1 adapter complex, Sec23/24 and three vacuolar protein sorting namely Vps26, Vps35 and Vps9¹⁰. Furthermore, the retromer composed of Vps35-Vps29-Vps26 that recycles *TgSORTLR* from *TgRab5*- to *TgRab7*-dependent ELC before delivery to Golgi, is also essential for secretory organelle biogenesis and parasite shape⁷.

Here, we report that the *T. gondii* counterpart of Vps9 (herein named *TgVps9*) is a *bona fide* Rab5 GTP-Exchange Factor (GEF) that is crucial for ROP protein maturation and processing, and its loss leads to a reduced number of rhoptries. Absence of *TgVps9* also impairs peripheral microneme biogenesis and disturbs dense granule secretion resulting in an accumulation of novel vesicles present both within and outside the parasite. Together with the rhoptry defect, absence of peripheral microneme formation and dense granule secretion severely affects parasite invasion of host cells. Collectively, these observations support the notion that *TgVps9*-mediated loading of GTP to *TgRab5* is crucial for fine-tuning vesicle sorting to secretory organelles, the latter being essential for *T. gondii* host cell infection.

Results

***T. gondii* vacuolar protein sorting 9 is a bona fide Rab5 guanine nucleotide exchange factor.** In eukaryotic cells, Vps9 domain-containing proteins are known as guanine nucleotide exchange factors (GEF) that stimulate the release of monomeric guanosine diphosphate (GDP)-bound to Rab5, allowing guanosine triphosphate (GTP) to bind and activate Rab5 that, in turn, regulates endosome vesicle trafficking^{16–18}. Previous work led us to identify an association with the C-terminus of *TgSORTLR*¹⁰, a protein with a predicted molecular mass of 140 kDa typified by a Vps9-like domain localized between amino acid (aa) 945 and 1117. This putative parasite Vps9 homologue harbors a region of 1326 amino acids extended at the N-terminal end and in this respect differs from its yeast and human counterparts that contain a shorter N-terminal end (Supplementary Fig. S1A). We first demonstrated that the predicted *TgVps9* domain operates *in vitro* as a GEF towards Rab5, by testing activity of a bacterial expressed *TgVps9* recombinant protein towards human Rab5, as previously described^{19,20}. Based on its homology with the catalytic core of mammalian Rabex5²¹; a truncated recombinant version (aa₈₄₉ to aa₁₁₃₄) of *TgVps9* was purified from *E. coli*. Human recombinant Rab5A was purified in its GDP bound form and nucleotide exchange to GppNHp, a non-hydrolysable GTP analogue, was monitored by tryptophan fluorescence measurements. A dose dependent GEF activity of recombinant *TgVps9* towards human recombinant Rab5A was detected and compared to human recombinant Rabex5 (Fig. 1A). As expected, no tryptophan fluorescence change was observed in the presence of excess GDP as no conformational change was induced during nucleotide exchange from GDP to GTP (Fig. 1A).

***TgVps9* localizes to the endosome-like compartment of *T. gondii*.** Having established that recombinant *TgVps9* possesses GEF activity towards recombinant Rab5, we then sought to determine in *T. gondii* whether the protein resides in the same subcellular compartment as Rab5. To this end, we chromosomally appended the hemagglutinin (HA) epitope at the 3'-end of the endogenous *TgVps9* gene and validated by western blot that the tagged *TgVps9* was readily expressed by transgenic tachyzoites (Supplementary Fig. S1B). HA-tagged *TgVps9* migrated with an apparent molecular mass of 170 kDa, which is higher than the predicted 140.0 kDa, likely due to the observation that *TgVps9* is heavily phosphorylated with 30 different phospho-sites indicated in Supplementary Figure S1A and collated at ToxoDB (www.toxodb.org).

Consistent with its *in vitro* GEF activity towards Rab5^{22,23}, *TgVps9*-HA *in vivo* co-localized with *TgRab5A* (Fig. 1B, top and left panel). As expected, we confirmed a co-localization between *TgVps9* and *TgSORTLR* (Fig. 1B, top panel), the endosomal-like compartment (ELC) receptor that has been used to efficiently pull down *TgVps9*¹⁰. The unprocessed precursor pro-ROP4 (Fig. 1B, top panel) known to be present in the ELC also co-distributes with *TgVps9*, while surprisingly proM2AP, a microneme marker, does not (bottom panel). In addition, VP1 a marker of the plant-like vacuole that is present in close vicinity to ELC co-distributes with *TgVps9* (Fig. 1B, top panel). In contrast, CPL (lysosomal-related compartment marker), GRASP (Golgi reassembly stacking protein), M2AP (MIC2-associated protein) or ROP4 (rhoptry marker) do not co-distribute with *TgVps9* (bottom panel). Taken together, these co-localization studies indicate that *TgVps9* is embedded in the endosomal-like compartment (ELC) together with Rab5, a compartment with an established role in the formation of secretory organelles of *T. gondii*^{4–8,10}.

Conditional ablation of *TgVps9* affects secretory organelle biogenesis to generate large novel vesicular-like structures. To examine *TgVps9* function in a clonal homogenous parasite population, we generated conditional anhydrotetracyclin (ATc)-inducible knockout *TgVps9* mutants (named *iKOTgVps9*) using the strategy depicted in Fig. 2A. We selected several positive clones from the emerging stable parasite population

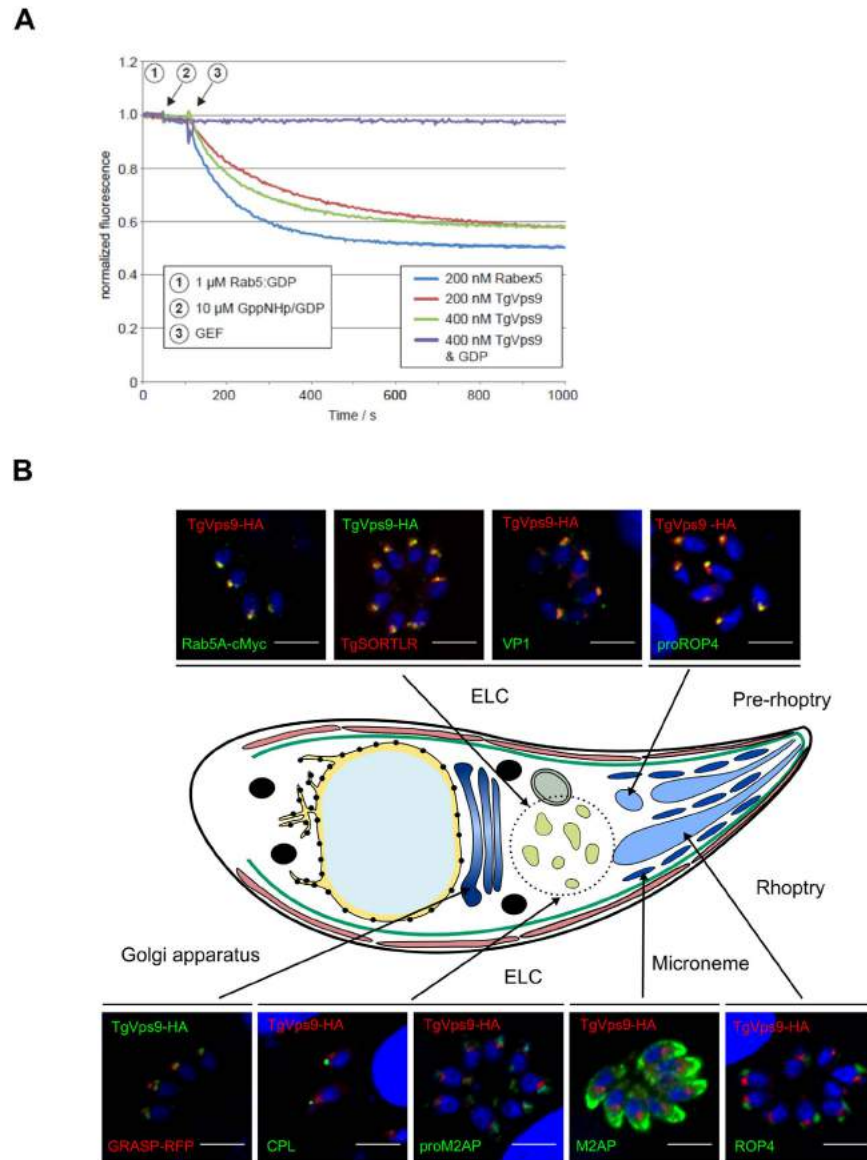


Figure 1. *T. gondii* Vps9 is a bona fide Rab5 GEF that localizes to endosome-like compartment.

(A) *In vitro* GEF assay of recombinant TgVps9. The conformation changes of Rab5 upon GDP/GppNHp exchange were measured by monitoring tryptophan auto-fluorescence in the presence of 200 and 400 nM of TgVps9, respectively. As a positive control, the human Rabex5 known for its GEF activity to Rab5 was also tested. In the presence of GDP, no change in auto-fluorescence was observed either for TgVps9 or Rabex5 as no conformational change should be induced when exchanging GDP for GDP. (B) Subcellular location of TgVps9-HA in the knocked in parasites was compared to that of several other organelle markers such as TgRab5A-cMyc (early endosome marker), TgSORTLR (Golgi and ELC marker), VP1 (Plant-like vacuole marker), pro-ROP4 (pre-rhoptry marker), GRASP-RFP (Golgi marker), CPL (Lysosomal-related compartment marker), pro-M2AP (immature microneme marker), and ROP4 (Rhoptry marker). Plasmids expressing GRASP-RFP and Rab5-cMyc were transiently transfected and RFP-positive parasites were directly visualized while specific anti-cMyc antibodies were used to stain Rab5-cMyc positive parasites. Specific antibodies were used to detect all other proteins. Scale bar indicates 5 μ m.

and the genome editing of one expanded clone *in vitro* was verified by PCR using the two specific primers (see nucleotide sequences in Methods) shown in blue arrow (Fig. 2A), thus demonstrating the perfect integration of the knock-out vector at the TgVps9 locus (Fig. 2B). Following ATc treatment, while a significant reduction of HA-TgVps35 protein was observed 24 h post-treatment, 48 h or 72 h of ATc-treatments were required for a complete and reproducible disappearance of HA-TgVps35 protein by western blotting (Fig. 2C). We confirmed these latter observations by confocal imaging (Fig. 2D) and further investigated all phenotypic consequences of this ATc-inducible TgVps9 knock out mutant at least at 48 h post-treatment. Next, we examined iKO TgVps9 mutants by electron microscopy and observed several striking ultrastructural changes associated with the loss of TgVps9. Both apical (yellow arrows) and peripheral micronemes (white arrows) were observed in cytoplasm anterior to

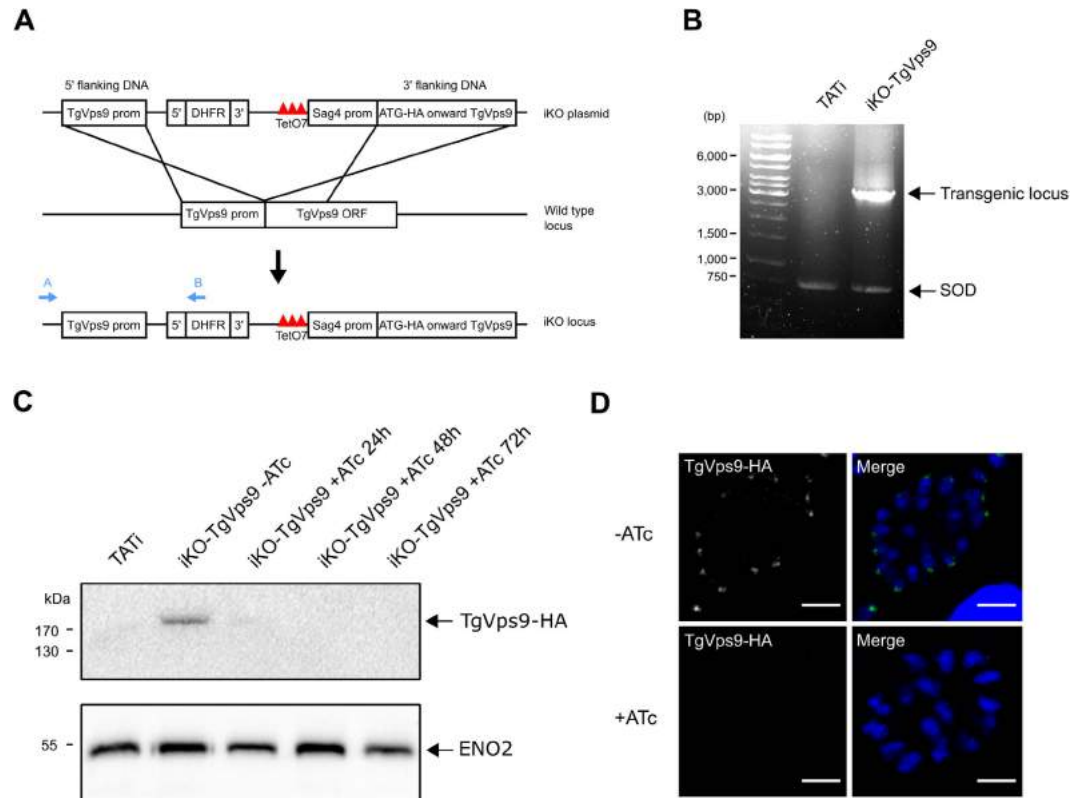


Figure 2. Conditional disruption of *TgVps9* gene. (A) Schematic representation of the ATc-inducible strategy to generate a knock out *TgVps9* mutant after a double homologous recombination in *T. gondii* genome. (B) PCR analysis to demonstrate the perfect integration of the knock-out vector into *TgVps9* locus using two specific primers A and B indicated in blue color. PCR of superoxide dismutase (SOD) gene corresponds to a positive control indicating that equal quantity of genomic DNA was used for all parasite strains tested. (C) Western blots of total SDS protein extracts from wild type and iKOTgVps9 mutants, which were incubated without or with ATc for 24, 48, and 72 hours. Blots were probed with anti-HA antibodies but also with antibodies specific to the glycolytic enzyme ENO2 used as protein loading control. (D) Confocal images of ATc-untreated or ATc-treated iKOTgVps9 mutants that were stained with anti-HA antibodies at 48 h post-treatment and infection. Scale bar indicates 5 μ m.

the nucleus of ATc-untreated tachyzoites (Fig. 3A,B) whereas ATc-treated iKOTgVps9 parasites, micronemes were only observed in the apical tip close to the conoid (Fig. 3C–E). We counted a total number of 30 micronemes located at the extreme apical end of 43 ATc-treated iKO-TgVps9 mutants using electron microscopy while a total number of 199 apical and peripheral micronemes were seen in 29 ATc-untreated parasites, indicating that there were about 5-fold fewer micronemes in *TgVps9*-deficient mutants *versus* the parental strain. Clearly, these data indicate the absence of peripheral micronemes in ATc-induced iKOTgVps9 mutants and the presence in the cytoplasm at the proximity of the nucleus of a novel large vesicular structure of approximately 500-nm diameter (panel D, black arrows). Elevated numbers of novel vesicles of variable size and morphology were also observed in the PV space delimited by the PVM (panel E, shown with *). We estimated that about 17% of iKOTgVps9 mutants examined by electron microscopy contained these aforementioned 500-nm intra-parasite vesicles while approximately 20% of mutants had novel vesicles of variable size and morphology in their PV space. In addition, we observed a significant reduction of the overall number of rhoptries per mutants using electron microscopy. Specifically, ATc-treated iKOTgVps9 mutants contained less than two thirds the relative number of rhoptries per mutant compared to the parental line. Also, we observed a disorganized ultrastructural morphology with the marked absence of the typical banana-shaped bodies in several *TgVps9*-depleted mutants (Fig. 3D), in a manner similar to the retromer iKOTgVsp35 mutants⁷ whereas untreated iKOTgVps9 parasites appeared structurally normal with all secretory organelles (Fig. 3A,B). These latter observations suggest that the cytoskeleton of parasite bodies may also be affected in these mutants. It should be mentioned that electron microscopy was used to show that other organelles including the mitochondrion, the nucleus, the Golgi apparatus, the inner complex membrane (IMC) and the plasma membrane appeared morphologically normal in these iKOTgVps9 mutants treated with ATc for 48 h (Supplementary Fig. S2, see panel A–E). In addition, iKOTgVps9-deficient mutants appear to undergo normal endodyogeny with two daughters forming within the mother cell (Supplementary Fig. S2F, see stars indicating the nucleus of two dividing daughter tachyzoites). Collectively, these data suggest that the traffic to and the integrity of the other parasite organelles were not altered by the loss of *TgVps9*. Furthermore, rhoptries, micronemes and dense granules were not completely absent either or not morphologically affected *per*

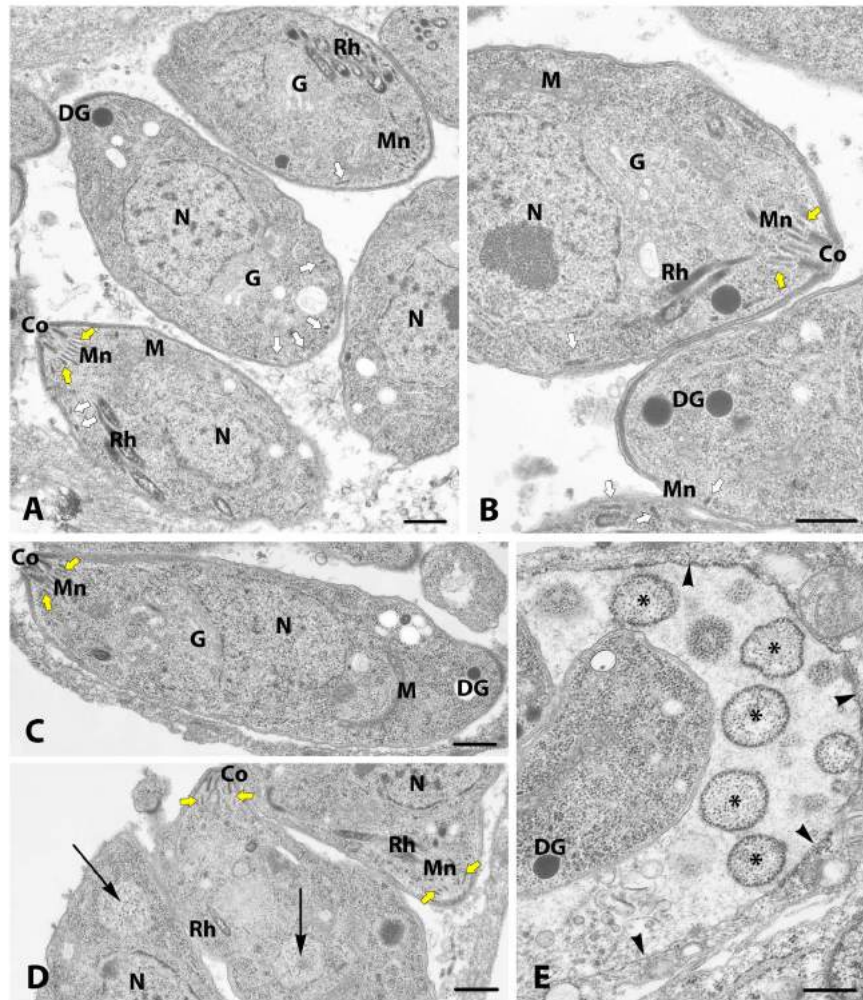


Figure 3. Electron microscopy of *TgVps9* mutants. (A,B) Ultrastructural sections of typical parental parasites were shown. Secretory organelles such as rhoptries (Rh), both peripheral (Mn, white arrows) and apical micronemes (Mn, yellow arrows) and dense granules (DG) can easily be visualized. (C–E) Ultrastructural images of sections representing *TgVps9*-depleted mutants after 48 h of ATc treatment were shown. Only apical micronemes, few rhoptries and dense granules can be seen in *TgVps9*-depleted mutants. Examples of novel 500 nm-diameter vesicles present in the cytoplasm above the nuclei of these mutants were indicated by arrows in panel D. A region of the parasitophorous vacuole (PV) space that also containing several novel vesicles of different sizes and morphology (see asterisks) were shown along with the delimited PV membrane (PVM) indicated by black arrowheads. The nucleus (N), mitochondrion (M), Golgi apparatus (G) and conoid (Co) were labelled. Scale bar is 500 nm.

se in *TgVps9*-deficient mutants (Fig. 3 and Supplementary Fig. S2), only organelle number was reduced in these mutants. Taken together, these data indicate that *TgVps9* likely regulates the turnover of vesicle precursors and pre-organelles destined to become fully mature secretory organelles.

The loss of *TgVps9* causes aberrant organelle secretion. We also examined the phenotypic consequences of *TgVps9* loss on the subcellular localizations of different secretory organelle markers. In ATc-induced *iKOTgVps9* mutants, ROP2-3 and ROP4 proteins were abnormally sorted into the host cytoplasm and decorated the host cell nuclear envelope (Fig. 4A,B, right panels, see white arrows). In the absence of ATc, *iKOTgVps9* mutants displayed normal apical localization of ROP proteins (Fig. 4A,B, left panels). In ATc-induced *iKOTgVps9* parasites, pro-ROP4 was also profoundly mis-sorted (Fig. 4C, right panels) with diffuse and weak labeling in both the PV space and the host cell cytoplasm (white arrow) unlike non-ATc-induced parasites that showed the typical apical pre-rhoptry localization of pro-ROP4, i.e. proximal to parasite nuclei (Fig. 4C, left panels). The imaging data suggest that the *iKOTgVps9* mutant-hosting PV may also be leaky following ATc induction for 48 h, thus resulting in the diffusion of pro-protein and mature protein in the vacuolar space and beyond the PV. Nevertheless, these data indicate that *TgVps9* loss results in an accumulation of ROP precursor proteins, their

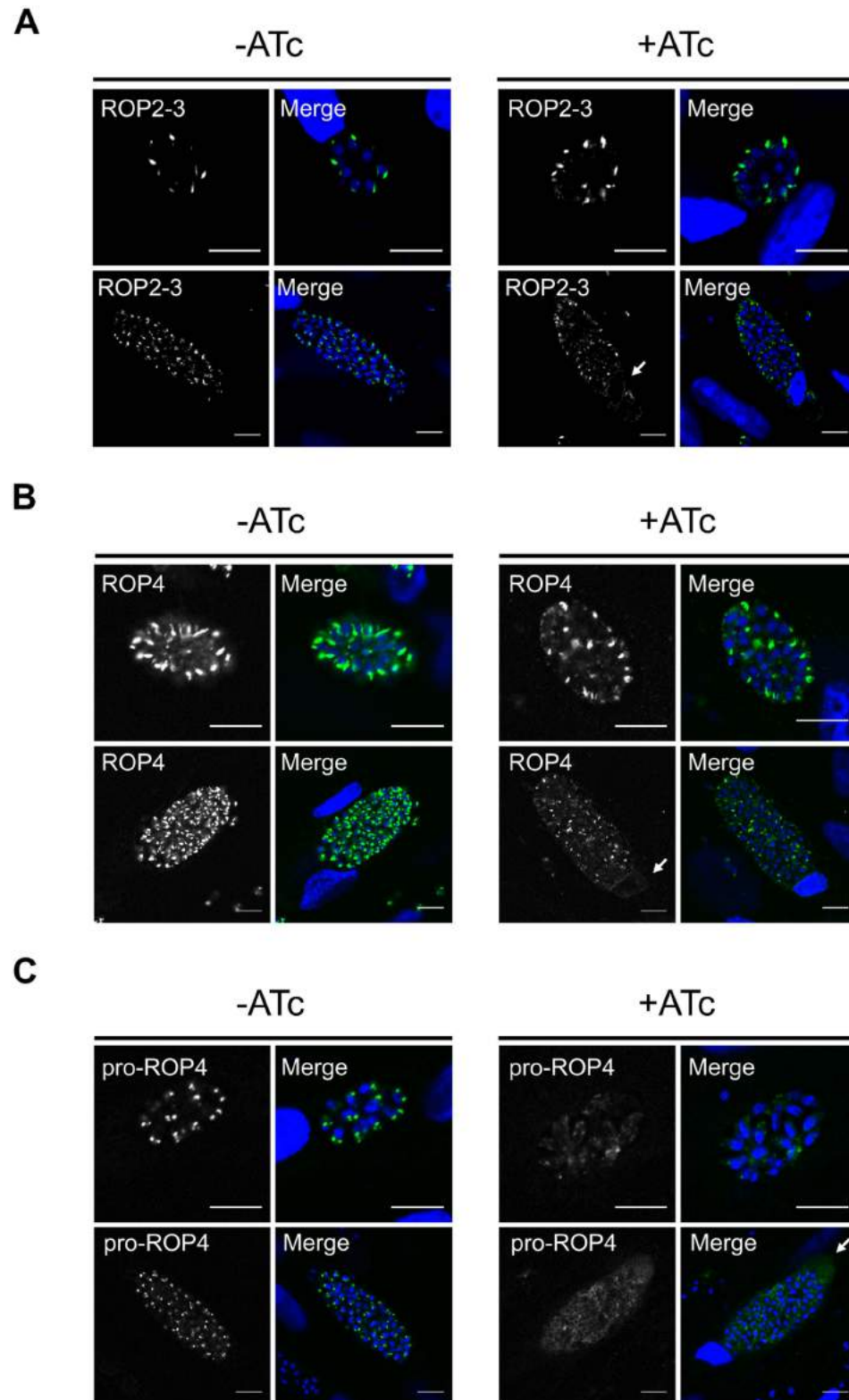


Figure 4. Conditional ablation of *Tgyps9* results in mis-sorting of ROP proteins. (A) Confocal immunofluorescences of ROP2-3 proteins in *iKOTgVps9* mutants in the presence (right panels) or absence of ATc (left panels) using specific antibodies to ROP2-3. (B) Confocal immunofluorescences of ROP4 proteins in *iKOTgVps9* mutants in the presence (right panels) or absence of ATc (left panels) using specific antibodies to ROP4. (C) Confocal immunofluorescences of proROP4 proteins in *iKOTgVps9* mutants in the presence (right panels) or absence of ATc (left panels) using specific antibodies to proROP4. Upper panels in (A–C) images correspond to small vacuoles containing 16 or less daughter parasites. Lower panels represent large vacuoles containing 32 or more daughter parasites. Scale bar on all images correspond to 10 μ m.

mis-sorting to other subcellular compartments and a significant reduction of the number of rhoptries per parasites as observed by electron microscopy.

We confirmed in the conditional *TgVps9* mutants that the typical conical microneme M2AP and MIC3 signals (Fig. 5A, left panels, red arrows) were completely changed to fluorescence signals at the extreme tip of each *TgVps9*-depleted mutant (Fig. 5A, right panels, yellow arrows). The most impressive and marked mis-sorting affects the dense granule GRA3 that was exclusively retained within the PV space (Fig. 5B, right panels) in the *iKOTgVps9* mutants whereas this protein decorated the PV membrane of vacuoles containing parental parasites, as expected (Fig. 5B, left panels, white arrows). The location of *TgSORTLR* was unchanged (Fig. 5C), confirming that not all proteins in the secretory ER-Golgi and ELC pathways are mis-sorted in *TgVps9*-deficient parasites. Altogether, these data clearly indicate that *TgVps9* is required for correct protein trafficking, sorting and delivery to the three main secretory organelles: rhoptry, microneme and dense granules.

Conditional *TgVps9* silencing dysregulates ROP protein maturation. In *T. gondii*, formation of rhoptries and micronemes is correlated with proteolytic processing followed by maturation of ROP and MIC proteins, respectively^{24–28}. Therefore, we investigated the role of *TgVps9* in the processing and maturation of representative ROP and MIC proteins. Specific antibodies that exclusively recognized the N-terminal pro-peptides of ROP4 revealed enhanced accumulation of pro-protein ROP4 in *TgVps9*-deficient mutants (Fig. 6A, left panel, single star). By calculating the ratio of the pro-protein to the mature ROP4 protein, using the housekeeping ENO2 loading control, we estimated that pro-ROP4 protein level was increased to almost 2-fold in ATc-treated *iKOTgVps9* mutants compared to untreated parental parasites (Fig. 6F). Using specific antibodies to the mature ROP4 protein, we detected a rise of proROP4 protein level to 4-fold (Fig. 6B, single stars), suggesting that pro-ROP4 protein accumulated between 2- to 4-fold higher in *TgVps9*-deficient mutants *versus* parental parasites (compare Fig. 6B,E, single star). As a consequence, the amount of processed mature ROP4 diminished (Fig. 6B,E, right lane, double stars). Similarly, we observed an increase level of unprocessed proROP2 (Fig. 6C,F, single star) that was estimated to be approximately 3-fold with a reduced amount of the mature ROP2 protein. We observed no significant changes for pro-M2AP protein (Fig. 6D–F, single star) or processed M2AP protein (Fig. 6E,F, double stars). We therefore concluded that pro-ROP proteins specifically accumulated in *iKOTgVps9* mutants, suggesting that *TgVps9* is likely involved in the delivery of both proteases and pro-ROP proteins to the same subcellular compartment and that their proteolytic processing is important for rhoptry organelle formation.

Conditional *TgVps9* silencing abrogates host cell invasion by *T. gondii*. More importantly, homogeneous clonal populations of *iKOTgVps9* mutants allowed us to address the role of *TgVps9* in *T. gondii* infection. The *iKOTgVps9* mutants were severely impaired in their ability to invade host cells with an 80% decrease after 48 h of ATc-treatment (Fig. 7A). In addition, at 7 day post-infection, *TgVps9*-depleted mutants did not form plaques associated with multiple rounds of host cell invasion in the presence of ATc (Fig. 7B), indicating that *TgVps9* is essential for ensuring proper formation of secretory organelles that are necessary for parasite propagation through multiple cycles of invasion, lysis and reinvasion of host cells. We have not observed any obvious deficiency in parasite egress from the host cell as the ATc-treated *iKOTgVps9* mutants spontaneously lysed out and freshly egressed tachyzoites can be recovered at 72 hours post-infection. Thus, we suggest that the reduction in number of secretory organelles *per* parasite and their default in organelle secretion represent the critical functions of *TgVps9* that are required for proper host infection by *T. gondii*.

Discussion

In this study, we established *in vitro* that *TgVps9* is a *bona fide* Rab5 GEF and locates *in vivo* in the same compartment as *TgRab5A*. We are suggesting that the GEF activity of *TgVps9* towards one of the *TgRab5* isoforms is essential for host cell invasion by *T. gondii* and its intracellular propagation. Genetic ablation of *TgVps9* by ATc-inducible knock out system, led to impairment of microneme biogenesis and default in proper dense granule secretion. Overall, we observed a significant mis-sorting of rhoptry, microneme and dense granule proteins and we argue that this underpins the loss in host cell invasion, vacuole formation, PV leaking and intracellular propagation. We found an accumulation of unprocessed ROP proteins such as ROP2 and ROP4 following the loss of *TgVps9*. However, we did not observe pro-ROP4 in the endosomal-like compartment and the pre-organelles of *TgVps9*-deficient mutants. Instead, pro-ROP4 is released into the parasitophorous vacuole and the host cell cytoplasm, where together with other ROP proteins, it appears perinuclear. The common phenotypic traits of *TgVps9* null mutants were reduced numbers of both rhoptries and micronemes *per* parasites, as seen by confocal imaging and electron microscopy. In contrast, we did not observe mis-sorting of proteins destined to mitochondrion, apicoplast, inner complex membrane and nucleus after *TgVps9* loss using confocal imaging. No other morphological changes of the aforementioned compartments have been detected by electron microscopy. This suggests that the default in secretory organelles such as rhoptry, microneme and dense granule is specifically restricted to the function of *TgVps9*, as expected for a genuine partner of *TgSORTLR*¹⁰, the Golgi and endosomal-like receptor that has been previously shown to be involved in protein transport and biogenesis of these parasite-specific secretory organelles.

Based on these observations, we propose a model in which *TgVps9* contributes to the regulation of ROP protein processing/maturation and the proper protein sorting to pre-rhoptries (Fig. 7C). This model is fully supported by the observation that the processing and maturation of ROP proteins have been shown to take place in the pre-organelles¹ and pro-ROP4 co-localizes with *TgVps9*. In contrast, neither the pro-MIC2 associated protein (pro-M2AP), nor the pro-M2AP processing enzyme CPL^{24–28} colocalize with *TgVps9*. This suggests that both processing and maturation of MIC proteins likely occurs in the distal sub-compartment of the ELC^{29,30} by means of CPL and in a *TgVps9*-independent fashion. It is also tempting to speculate that pre-micronemes and pre-rhoptries may bud from the novel and yet-uncharacterized vesicles that are visible above the nuclei of

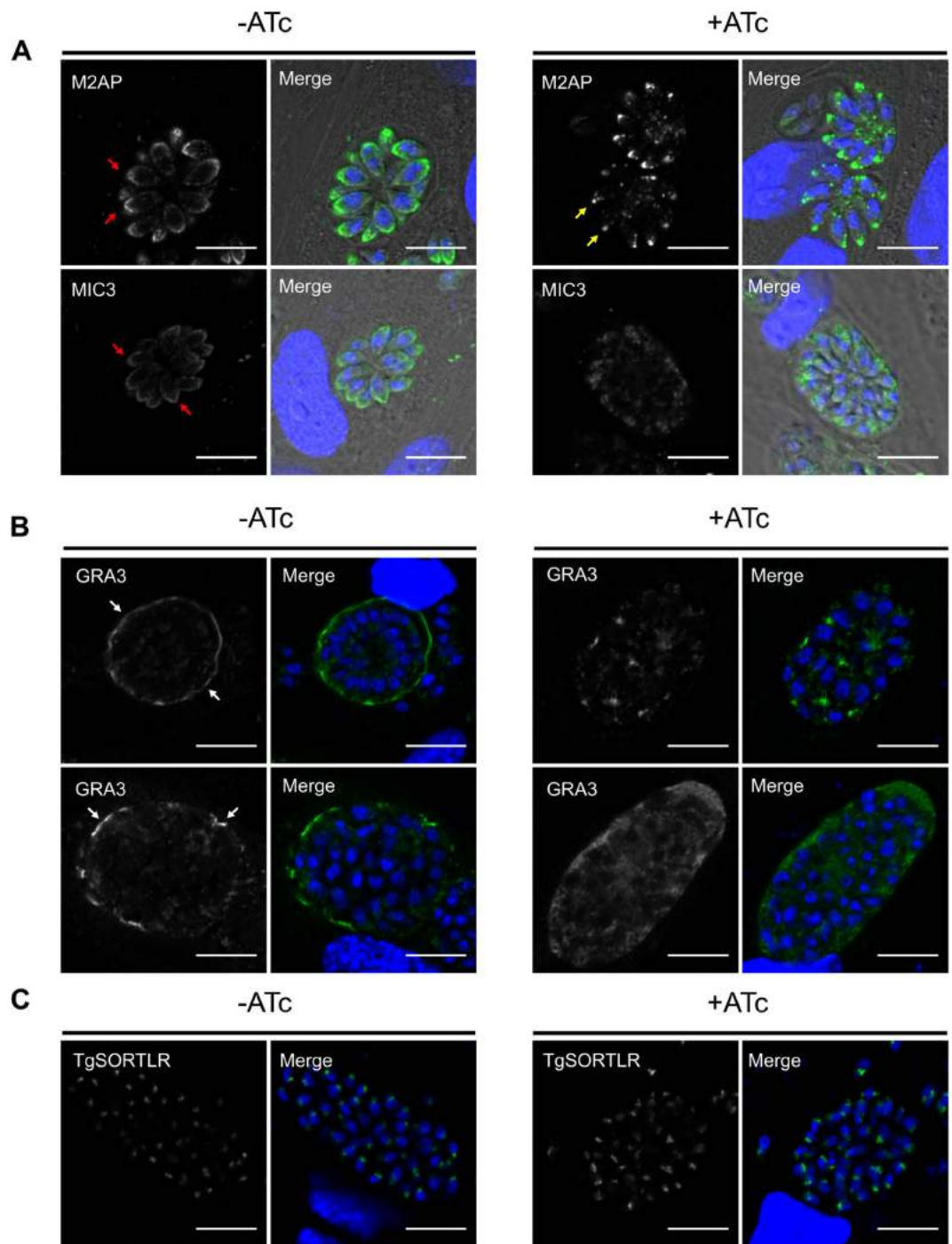


Figure 5. Conditional *TgVps9* silencing affected microneme biogenesis and dense granule secretion. (A) Confocal immunofluorescences of M2AP and MIC3 proteins in iKO *TgVps9* mutants in the presence (right panels) or absence of ATc (left panels) using antibodies specific to each protein, respectively. Note that the typical and conical signal of M2AP and MIC3 proteins (red arrows) in the parental strain disappears. Instead, only residual punctuated signal was seen in *TgVps9*-depleted mutants (yellow arrows), indicating the absence of peripheral micronemes in these mutants. The whole bodies of intracellular parasites were shown by phase contrast in order to indicate fluorescence signals corresponding to micronemes located at the extreme apical end of Cas9-GFP positive parasites (yellow arrows). (B) Confocal immunofluorescences of GRA3 protein in iKO *TgVps9* mutants in the presence (right panels) or absence of ATc (left panels) using specific antibodies to GRA3. Note the absence of GRA3 protein delivery to the parasitophorous vacuole membrane (PVM) that contrasts to the situation in the parental parasites (white arrows). (C) The presence of *TgSORTLR* in the Golgi-ELC region was unchanged in parental parasites and mutants regardless of treatment with ATc or not. Scale bar is 10 μ m.

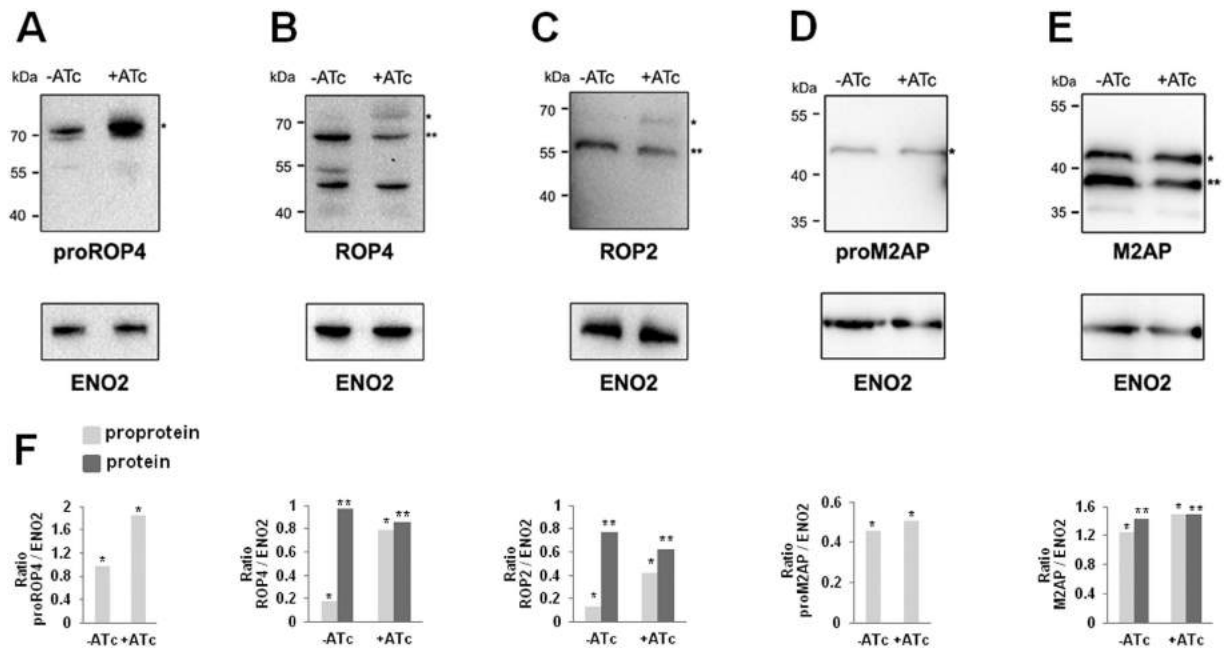


Figure 6. Upon conditional disruption of *TgVps9* unprocessed and immature ROP proteins accumulate.

(A) Immunoblots of intracellular *iKOTgVps9* mutants probed with specific anti-proROP4 at 48 h post-infection in the absence or presence of ATc. (B) Immunoblots of intracellular *iKOTgVps9* mutants probed with specific anti-ROP4 antibodies at 48 h post-infection in the absence or presence of ATc. (C) Immunoblots of intracellular *iKOTgVps9* mutants probed with specific anti-ROP2 antibodies at 48 h post-infection in the absence or presence of ATc. (D) Immunoblots of intracellular *iKOTgVps9* mutants probed with specific anti-proM2AP antibodies at 48 h post-infection in the absence or presence of ATc. (E) Immunoblots of intracellular *iKOTgVps9* mutants probed with specific anti-M2AP antibodies at 48 h post-infection in the absence or presence of ATc. (F) Quantification of protein intensity by densitometry that shows ratio between pro-protein or mature protein and ENO2 levels. The housekeeping glycolytic enzyme ENO2 was used as negative and loading control. A single star (*) indicates the unprocessed precursor protein while double stars (**) correspond to processed and mature protein. Molecular weights of protein markers (kDa) were indicated on the left of each panel.

TgVps9-depleted mutants. As discussed above, *TgVps9* displays *in vitro* GEF towards human Rab5A, comparable to that of Rabex5^{18–23}. It follows then that some phenotypic traits of *TgVps9*-deficient mutants resemble those previously reported for *TgRab5A* protein after its encoding gene has been disrupted³¹. As both *TgRab5A* and *Vps11* probably interact through the CORVET-tethering complexes¹³, this could explain their phenotypic similarities with those observed for *iKOTgVps9* mutants described here.

Protein trafficking that relies on VPS9 has been described in several other eukaryotic cells such as yeast, which has three Vps9-domain containing proteins, Vps9, Muk1 and Vrl1, all exhibiting GEF activity towards Rab5 paralogs³². Mammalian cells contain at least nine Vps9 domain-containing proteins fulfilling diverse functions including regulation of protein transport, endocytosis and signaling pathways²². Additionally, it has been reported that Vps9 domains also interact with retromer complex and phosphatidylinositol 3-phosphate (PI3P) to promote the enrichment of PI3P lipids at the endosomes³³. Knowing that *TgSORTLR*¹⁰ and the retromer machinery⁷ in *T. gondii* share similarities with those of *TgVps9*, the latter may participate in regulation of retromer and endosomal lipid content. However, *TgVps9* is not associated with the Golgi apparatus like *TgSORTLR*, suggesting that this parasite Rab5 GEF is likely involved in anterograde transport and secretory organelle formation, rather than protein recycling in *T. gondii*.

In conclusion, loss of *TgVps9* inhibits rhoptry protein processing/maturation, impairs secretory organelle biogenesis and secretion, leading to an inability of *TgVps9*-deficient mutants to invade host cells and to achieve multiple rounds of invasion, proliferation and reinvasion of host cells. However, there is still a missing link between protein processing, maturation, vesicular traffic and secretory organelle formation. Further investigation of *TgVps9* functions that will define how different parasite *TgRab* proteins are precisely regulated by its GEF activity could provide this missing link.

Methods

Parasite culture. We used *T. gondii* tachyzoites of RH strain for CRISPR/Cas9 knockout experiment, RHΔ*Ku80*³⁴ for the knock in of *TgVps9* gene (TGME49_230140) and RHΔ*Ku80*TATi for inducible knockout (iKO) strain^{35,36} that were grown using Human Foreskin Fibroblast (HFF) cells from ATCC (USA) as described⁷. The *iKOTgVps9* mutants were cultured in the presence of 1.5 μg/ml anhydrotetracycline (ATc).

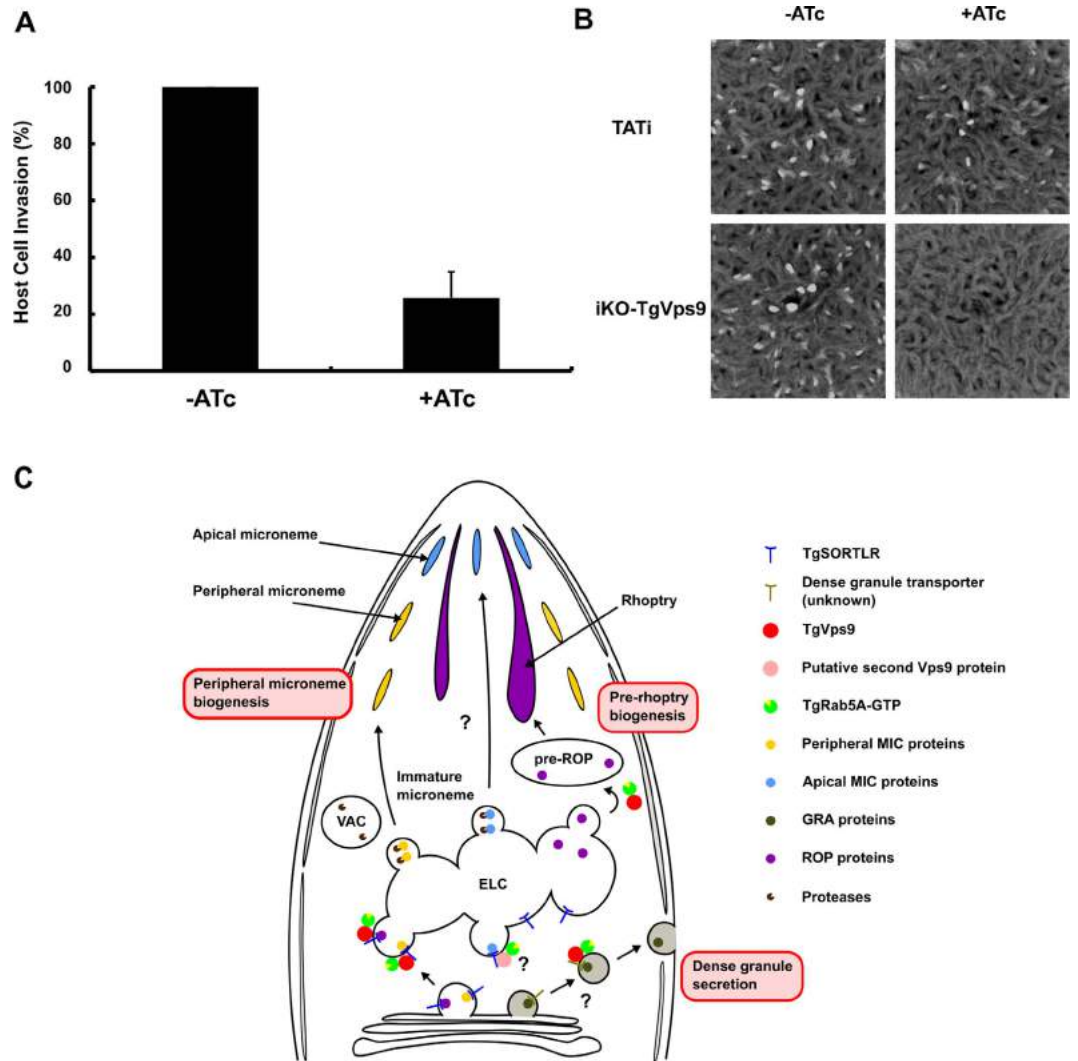


Figure 7. *TgVps9* is essential for host cell invasion by *T. gondii*. (A) Host cell invasion was assayed in *iKOTgVps9* mutants in the presence or absence of ATc. Bars indicate mean \pm SD ($n = 3$, $P < 0.001$ by Student's test). (B) Host cell lytic plaques were examined in *TgVps9*-deficient mutants and parental RH Δ *Ku80*TATI parasites in the presence or absence of ATc. (C) Model of *TgVps9*-mediated trafficking and regulation of diverse sorting cargoes required for secretory organelle biogenesis in *T. gondii*. Based on the processing inhibition of precursor ROP proteins in *TgVps9*-depleted mutants and the mis-sorting of pro-ROP4, we conclude that the site of action of *TgVps9* starts from the early endosome to late endosome until pre-rhoptry organelles. In addition, this also suggests that *TgVps9* is required for the trafficking of the proteolytic enzymes to the pre-rhoptries where the pro-peptides of pro-ROP proteins were processed during protein maturation. In contrast, *TgVps9* protein was not required to transport the cathepsin protease-like (CPL) enzyme³², a type II transmembrane protein involved in the processing and maturation of MIC proteins within the ELC. Indeed, *TgVps9* function was limited to the biogenesis of peripheral micronemes. We further evidenced that *TgVps9* protein was also important for the proper discharge of GRA protein inside the PVM. However, the molecular mechanisms underlying how GRA were transported to the dense granules and thereafter released to the PVM remain to be elucidated. We also propose that Rab5 is the key small GTPase, which was regulated by *TgVps9* according to our biochemical GEF assays and the previous data of Vps11 of CORVET tethering complexes¹³ and those of Rab5A³³. We suggest that function of *TgVps9* is important for the transport to and secretion of organelles required for the successful intracellular lifestyle of the parasite.

Production of recombinant *TgVps9*, Rabex5, and Rab5 proteins. Recombinant protein of the catalytic core of *TgVps9* from aa₈₄₉ to aa₁₁₃₄ was generated using a modified pET19 plasmid that expresses His-tag protein with a TEV cleavage site using the following primers: forward (Recomb-Vps9a.d4_F) CCGGC ATATGGCGTCTTCTGCCTCTTTTCTGCC and reverse (Recomb-Vps9a.d4_R) CCGGGGATC CTTAGCGTTCGCGTTCGCGTTCGATTC. Human recombinant Rabex5 from aa₁₃₂ to aa₃₉₇ and full length Rab5a with a C-terminal CVIL mutation were prepared as previously described³⁷. The recombinant *TgVps9* was expressed in BL21 Codon Plus (DE3)-RIPL, and cell pellet was resuspended in buffer A (50 mM HEPES pH 8, 300 mM NaCl, 1 mM Tris (2-carboxyethyl) phosphine (TCEP), 20 mM imidazole pH 8), sonicated, and centrifuged

at 20,000 rpm for 1 hour at 4 °C. Cell lysate was incubated with 2 ml agarose Ni-NTA beads, washed with buffer A and subsequently buffer B (50 mM HEPES pH 8, 300 mM NaCl, 1 mM TCEP, 30 mM imidazole pH 8). Bound protein was eluted using buffer C (50 mM HEPES pH 8, 300 mM NaCl, 1 mM TCEP, 500 mM imidazole pH 8). His-tags were removed by TEV protease cleavage during overnight dialysis in Buffer A without TCEP and imidazole. In order to remove the His-tagged TEV protease, dialyzed protein solution was incubated with Ni-NTA beads again before the GEF assays were performed.

GEF assay. GEF activities were analyzed by intrinsic tryptophan fluorescence measurements showing fluorescence changes due to the conformational change from GDP to GTP state^{19,20}. Rab5 fluorescence was excited at 297 nm and emission signals were detected at 340 nm. The fluorescence was recorded using a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies). 10 μM of 5'-Guanylyl imidodiphosphate (GppNHp), non-hydrolysable analog of GTP, or 10 μM GDP were added to 1 μM of Rab5 GDP in GEF assay buffer (25 mM HEPES pH 7.5, 200 mM NaCl, 1 mM MgCl₂, 2 mM dithioerythritol (DTE)), subsequently 200 nM or 400 nM of TgVps9 or 200 nM of Rabex5 were added.

Generation of stable transgenic strains. Endogenous gene tagging methodology using pLIC-HA-DHFR plasmid³⁴ was used to generate TgVps9-HA knock in parasites. DNA of TgVps9 was cloned in this plasmid using the following primers: forward (F-KL_Vps9) TACTTCCAATCCAATTTAATGCCCTGCTTGCCCCTCGCCT and the reverse (R-KL_Vps9) TCCTCCACTTCCAATTTTAGCTTTCCTGTCACATATGTTTCGCGTCCG. To obtain iKOTgVps9 mutants, we used pG13-D-T7S4 plasmid³⁶ in which a 2-kb DNA containing the promoter sequence was cloned using the following primers: forward (iKOTgVps9-5' F) CCGGCATATGCTTCTAACGGCACCCTTAAGGTGC and reverse (iKOTgVps9-5' R) CCGGCATA TGTGCGCCTTCTCGTGTCTGTTG; and another 2-kb DNA containing the coding sequence of TgVps9 gene using the following primers: forward (iKOTgVps9-3' F) CCGGTGATCAATGTACCCAT ACGATGTTCCGATTACGCTCGTCACGGGAAGAAGACCAGCAGTC and reverse (iKOTgVps9-3' R) CCGGCCTAGGGGGAGAAGAGGAGACAGAAACATCTCGACTACGACC with the HA-tag sequence underlined in the forward primers inserted at the N-terminus of TgVps9 protein, right after the initiation ATG codon. 1 × 10⁷ parasites were transfected with 50 μg of linearized plasmid and selected with 2 μM of pyrimethamine. The emerging pyrimethamine-resistant population was cloned by limiting dilution. The clones were checked for plasmid integration by PCR using genomic DNA and two primers: forward (named A or Test_iKOVPS9.F2) ATTACAGCCAGCAGTGGCCAACCGAAT and reverse (named B, DHFR-int. R) GGCGTTGAATCTCTTGCCGACTGTGGAGAGGGGAAGTCC.

Immunofluorescence microscopy. Confocal microscopy was performed as described previously^{10,38}. Briefly, intracellular parasites within HFF cells on 24-well coverslips were fixed by 4% paraformaldehyde for 10–15 min at room temperature. Fixed cells were permeabilized with 0.2% Triton X-100 and blocked with 5% fetal bovine serum (FBS). These cells were incubated with primary antibodies for 1 h at 37 °C and sequentially stained with secondary antibodies conjugated with Alexa 488, 594 or 647 in addition to DAPI for 30–45 min at 37 °C. Stained cells were mounted with Mowiol. All images were captured by a confocal microscope LSM780 or 880 (Carl Zeiss). Image processing was performed by open-source Fiji software.

Western blots. Intracellular iKOTgVps9 mutants or parental parasites incubated with ATc or not for 48 h were scrapped and washed twice with PBS. The intracellular parasites were pelleted before suspension by Laemmli buffer (62.5 mM Tris-HCl pH 6.8; 2% SDS; 100 mM DTT; 10% sucrose) and boiled for SDS-PAGE. 2 × 10⁶ parasites were fractionated on 10% acrylamide gels, which were transferred to nitrocellulose membranes as previously described³⁹. Immunoblot was performed using several anti-MIC and ROP antibodies in TNT buffer (100 mM Tris-HCl pH 7.6; 150 mM NaCl; 0.1% Tween20). All membranes were stained with antibodies specific to the glycolytic enzyme anti-ENO2⁴⁰ as a loading control after stripping antibody.

Electron microscopy. For transmission electron microscopy, cells were fixed in 2.5% glutaraldehyde prepared in 0.1 M cacodylate buffer and post-fixed in 1% osmium tetroxide in the same buffer. After acetonitrile dehydration, the pellet was embedded in Epon. Ultrathin sections (90 nm) were cut using a Leica UC7 ultramicrotome and collected on 150 mesh hexagonal barred copper grids. After staining with 2% uranyl acetate prepared in 50% ethanol and incubation with a lead citrate solution, sections were observed on a Hitachi H-600 transmission electron microscope at 75 kV accelerating voltage.

Invasion and plaque assays. Wild type RHΔ*Ku80*TATi strain and iKOTgVps9 parasites were incubated under ATc condition for 48 h and mechanically lysed by passage through a syringe. 1 × 10⁵ parasites were inoculated to HFF cells and incubated for 1 h at 37 °C. After infection for 1 hour, extracellular parasites were washed out with PBS and used to infect HFF cells before growing for 24 h at 37 °C. These infected cells were fixed by PFA and sequentially stained by GAP45 antibody with DAPI and counted by Axioimager Z1 (Carl Zeiss). Host cell invasion values were calculated using the ratio of intracellular parasite/host nucleus numbers as described⁴¹. For plaque assays, 400 freshly lysed parasites were used to infect HFF cells followed by incubation for 7 days with or without ATc. These cells were fixed by ethanol and stained by crystal violet.

References

- Boothroyd, J. C. & Dubremetz, J.-F. Kiss and spit: the dual roles of *Toxoplasma* rhoptries. *Nat. Rev. Microbiol.* **6**, 79–88 (2008).
- Carruthers, V. B. & Tomley, F. M. Microneme proteins in apicomplexans. *Subcell. Biochem.* **47**, 33–45 (2008).
- Bougdour, A., Tardieux, I. & Hakimi, M. A. *Toxoplasma* exports dense granule proteins beyond the vacuole to the host cell nucleus and rewires the host genome expression. *Cell Microbiol.* **16**, 334–343 (2014).

4. Tomavo, S., Slomianny, C., Meissner, M. & Carruthers, V. B. Protein trafficking through the endosomal system prepares intracellular parasites for a home invasion. *PLoS Pathog.* **9**, e1003629 (2013).
5. Tomavo, S. Evolutionarily repurposing of endosomal systems for apical organelle biogenesis in *Toxoplasma gondii*. *Intern. J. Parasitol.* **44**, 133–138 (2014).
6. Jimenez-Ruiz, E., Morlon-Guyot, J., Daher, W. & Meissner, M. Vacuolar protein sorting mechanisms in apicomplexan parasites. *Mol. Biochem. Parasitol.* doi: 10.1016/j.molbiopara.2016.01.007 (2016).
7. Sangaré, L. O. *et al.* Unconventional endosome-like compartment and retromer complex in *Toxoplasma gondii* govern parasite integrity and host infection. *Nat. Commun.* **7**, 11191, doi: 10.1038/ncomms11191 (2016).
8. McGovern, O. L. & Carruthers, V. B. Toxoplasma Retromer Is Here to Stay. *Trends Parasitol.* doi: 10.1016/j.pt.2016.05.007 (2016).
9. Luzio, J. P., Hackmann, Y., Dieckmann, N. M. G. & Griffiths, G. M. The biogenesis of lysosomes and lysosome-related organelles. *Cold Spring Harb Perspect. Biol.* **6**, a016840 (2014).
10. Sloves, P.-J. *et al.* *Toxoplasma* sortilin-like receptor regulates protein transport and is essential for apical secretory organelle biogenesis and host infection. *Cell Host Microbe.* **11**, 515–527 (2012).
11. Breinich, M. S. *et al.* A dynamin is required for the biogenesis of secretory organelles in *Toxoplasma gondii*. *Curr. Biol.* **19**, 277–286 (2009).
12. Pieperhoff, M. S., Schmitt, M., Ferguson, D. J. P. & Meissner, M. The role of clathrin in post-golgi trafficking in *Toxoplasma gondii*. *PLoS One.* **8**, 1–16 (2013).
13. Morlon-Guyot, J., Pastore, S., Berry, L., Lebrun, M. & Daher, W. *Toxoplasma gondii* Vps11, a subunit of HOPS and CORVET tethering complexes, is essential for the biogenesis of secretory organelles. *Cell Microbiol.* **17**, 1157–1178 (2015).
14. Balderhaar, H. J. K. & Ungermann, C. CORVET and HOPS tethering complexes- coordinators of endosome and lysosome fusion. *J. Cell Sci.* **126**, 1307–1316 (2013).
15. Rink, J., Ghigo, E., Kalaidzidis, Y. & Zerial, M. Rab conversion as a mechanism of progression from early to late endosomes. *Cell.* **122**, 735–749 (2005).
16. Stenmark, H. Rab GTPases as coordinators of vesicle traffic. *Nat. Rev. Mol. Cell Biol.* **10**, 513–525 (2009).
17. Balderhaar, H. J. *et al.* The CORVET complex promotes tethering and fusion of Rab5/Vps21-positive membranes. *Proc. Natl. Acad. Sci. USA* **110**, 3823–3828 (2013).
18. Delprato, A., Merithew, E. & Lambright, D. G. Structure, exchange determinants, and family-wide Rab specificity of the tandem helical bundle and Vps9 domains of Rabex-5. *Cell.* **118**, 607–617 (2004).
19. Pan, J. Y., Sanford, J. C. & Wessling-Resnick, M. Effect of guanine nucleotide binding on the intrinsic tryptophan fluorescence properties of Rab5. *J. Biol. Chem.* **270**, 24204–24208 (1995).
20. Simon, I., Zerial, M. & Goody, R. S. Kinetics of interaction of Rab5 and Rab7 with nucleotides and magnesium ions. *J. Biol. Chem.* **271**, 20470–20478 (1996).
21. Delprato, A. & Lambright, D. G. Structural basis for Rab GTPase activation by VPS9 domain exchange factors. *Nat. Struct. Mol. Biol.* **14**, 406–412 (2007).
22. Carney, D. S., Davies, B. & Horazdovsky, B. F. Vps9 domain-containing proteins: activators of Rab5 GTPases from yeast to neurons. *Trends Cell Biol.* **16**, 27–35 (2006).
23. Hama, H., Tall, G. G. & Horazdovsky, B. F. Vps9p is a guanine nucleotide exchange factor involved in vesicle-mediated vacuolar protein transport. *J. Biol. Chem.* **274**, 15284–15291 (1999).
24. Di Crisanti, M., Spaccapelo, R., Soldati, D., Bistoni, F. & Crisanti, A. Two conserved amino acid motifs mediate protein targeting to the micronemes of the apicomplexan parasite *Toxoplasma gondii*. *Mol. Cell Biol.* **20**, 7332–7341 (2000).
25. Que, X. *et al.* The cathepsin B of *Toxoplasma gondii*, toxopain-1, is critical for parasite invasion and rhoptry protein processing. *J. Biol. Chem.* **277**, 25791–25797 (2002).
26. Harper, J. M. *et al.* cleavable propeptide influences *Toxoplasma* infection by facilitating the trafficking and secretion of the TgMIC2-M2AP invasion complex. *Mol. Biol. Cell.* **17**, 4551–4563 (2006).
27. Brydges, S. D., Harper, J. M., Parussini, F., Coppens, I. & Carruthers, V. B. A transient forward-targeting element for microneme-regulated secretion in *Toxoplasma gondii*. *Biol. Cell.* **100**, 253–264 (2008).
28. Lagal, V. *et al.* *Toxoplasma gondii* protease TgSUB1 is required for cell surface processing of micronemal adhesive complexes and efficient adhesion of tachyzoites. *Cell Microbiol.* **12**, 1792–1808 (2010).
29. Miranda, K. *et al.* Characterization of a novel organelle in *Toxoplasma gondii* with similar composition and function to the plant vacuole. *Mol. Microbiol.* **76**, 1358–1375 (2010).
30. Parussini, F., Coppens, I., Shah, P. P., Diamond, S. L. & Carruthers, V. B. Cathepsin L occupies a vacuolar compartment and is a protein maturase within the endo/exocytic system of *Toxoplasma gondii*. *Mol. Microbiol.* **76**, 1340–1357 (2010).
31. Kremer, K. *et al.* An overexpression screen of *Toxoplasma gondii* Rab-GTPases reveals distinct transport routes to the micronemes. *PLoS Pathog.* **9**, e1003213 (2013).
32. Paulsel, A. L., Merz, A. J. & Nickerson, D. P. Vps9 family protein Muk1 is the second rab5 guanine nucleotide exchange factor in budding yeast. *J. Biol. Chem.* **288**, 18162–18171 (2013).
33. Bean, B. D. M. *et al.* Rab5-family guanine nucleotide exchange factors bind retromer and promote its recruitment to endosomes. *Mol. Biol. Cell.* **26**, 1119–1128 (2015).
34. Huynh, M.-H. & Carruthers, V. B. Tagging of endogenous genes in a *Toxoplasma gondii* strain lacking Ku80. *Eukaryot Cell.* **8**, 530–539 (2009).
35. Meissner, M., Schlüter, D. & Soldati, D. Role of *Toxoplasma gondii* myosin A in powering parasite gliding and host cell invasion. *Science.* **298**, 837–840 (2002).
36. Sheiner, L. *et al.* A systematic screen to discover and analyze apicoplast proteins identifies a conserved and essential protein import factor. *PLoS Pathog.* **7**, e1002392 (2011).
37. Oesterlin, L. K., Goody, R. S. & Itzen, A. Posttranslational modifications of Rab proteins cause effective displacement of GDP dissociation inhibitor. *Proc. Natl. Acad. Sci. USA* **109**, 5621–5626 (2012).
38. Sloves, P. J. *et al.* Apical organelle secretion by *Toxoplasma* controls innate and adaptive immunity and mediates long-term protection. *J. Infect. Dis.* **212**, 1449–1458 (2015).
39. Que, X., Engel, J. C., Ferguson, D., Wunderlich, A., Tomavo, S. & Reed, S. L. Cathepsin Cs are key for the intracellular survival of the protozoan parasite, *Toxoplasma gondii*. *J. Biol. Chem.* **282**, 4994–5003 (2007).
40. Mouveaux, T. *et al.* Nuclear glycolytic enzyme enolase of *Toxoplasma gondii* functions as a transcriptional regulator. *PLoS One.* **9**, e105820 (2014).
41. Olguin-Lamas, A. *et al.* A novel *Toxoplasma gondii* nuclear factor TgNEF3 is a dynamic chromatin-associated component, modulator of nucleolar architecture and parasite virulence. *PLoS Pathog.* **7**, e1001328 (2011).

Acknowledgements

We would like to thank Prof. David J.P. Ferguson (Oxford University, UK) for critically reading this manuscript, Dr Roger S. Goody (MPI for molecular Physiology, Dortmund) for providing plasmid DNA for the expression of Rab5 and Rabex5; and Dr Tchilabalo Dilezitoko Alayi for quantification of western blots. This work was supported by the following grants: Laboratoire d'Excellence (LabEx) ParaFrap from the National Agency for Research ANR-

11-LABX-0024 and the ANR-14-CE14-0002-01. We also acknowledge additional financial support from the INSERM, Pasteur Institute of Lille, CNRS and the DFG/ANR grant GO 284/8-1. Post-doc and PhD fellowships were from LabEx ParaFrap to T.S. and F.S., respectively. M.A.H is supported by the European Research Council (ERC Consolidator grant no. 614880).

Author Contributions

T.S. conceived and designed the experiments, analyzed data and wrote paper; F.S. and M.A.H. involved in reverse genetics approaches and wrote paper; C.S., electron microscopy; L.K.O. and H.B., involved in testing GEF activity; G. L. involved in GEF activity design; S.T. designed, supervised and wrote this study. The manuscript has been seen and approved by all authors.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Sakura, T. *et al.* A Critical Role for *Toxoplasma gondii* Vacuolar Protein Sorting VPS9 in Secretory Organelle Biogenesis and Host Infection. *Sci. Rep.* **6**, 38842; doi: 10.1038/srep38842 (2016).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016

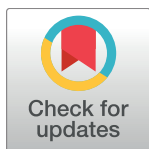
RESEARCH ARTICLE

Dual role of the *Toxoplasma gondii* clathrin adaptor AP1 in the sorting of rhoptry and microneme proteins and in parasite division

Kannan Venugopal¹, Elisabeth Werkmeister¹, Nicolas Barois¹, Jean-Michel Saliou¹, Anais Poncet¹, Ludovic Huot¹, Fabien Sindikubwabo², Mohamed Ali Hakimi², Gordon Langsley³, Frank Lafont¹, Sabrina Marion^{1*}

1 Centre d'Infection et d'Immunité de Lille, Université de Lille, Inserm U1019, CNRS UMR 8204, CHU Lille, Institut Pasteur de Lille, Lille, France, **2** IAB, Team Host-pathogen interactions & immunity to infection, Université Grenoble Alpes, Inserm U1209, CNRS UMR5309, Grenoble, France, **3** Laboratoire de Biologie Cellulaire Comparative des Apicomplexes, Faculté de Médecine, Université Paris Descartes—Sorbonne Paris Cité, France. Inserm U1016, CNRS UMR8104, Institut Cochin, Paris, France

* sabrina.marion@pasteur-lille.fr



OPEN ACCESS

Citation: Venugopal K, Werkmeister E, Barois N, Saliou J-M, Poncet A, Huot L, et al. (2017) Dual role of the *Toxoplasma gondii* clathrin adaptor AP1 in the sorting of rhoptry and microneme proteins and in parasite division. *PLoS Pathog* 13(4): e1006331. <https://doi.org/10.1371/journal.ppat.1006331>

Editor: Isabelle Coppens, Johns Hopkins School of Public Health, UNITED STATES

Received: July 12, 2016

Accepted: April 3, 2017

Published: April 21, 2017

Copyright: © 2017 Venugopal et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: KV, FS, MAH, GL, and SM are supported by the Laboratoire d'Excellence (LabEx) ParaFrap from the National Agency for Research ANR-11-LABX-0024 and ANR-14-CE14-0002-01. MAH is supported by the European Research Council (ERC Consolidator grant no. 614880). SM is supported by a Chaire d'Excellence Université Lille Nord de

Abstract

Toxoplasma gondii possesses a highly polarized secretory system, which efficiently assembles *de novo* micronemes and rhoptries during parasite replication. These apical secretory organelles release their contents into host cells promoting parasite invasion and survival. Using a CreLox-based inducible knock-out strategy and the ddFKBP over-expression system, we unraveled novel functions of the clathrin adaptor complex *TgAP1*. First, our data indicate that AP1 in *T. gondii* likely functions as a conserved heterotetrameric complex composed of the four subunits γ , β , $\mu 1$, $\sigma 1$ and interacts with known regulators of clathrin-mediated vesicular budding such as the unique ENTH-domain containing protein, which we named Epsin-like protein (*TgEpsL*). Disruption of the $\mu 1$ subunit resulted in the mis-sorting of microneme proteins at the level of the Trans-Golgi-Network (TGN). Furthermore, we demonstrated that *TgAP1* regulates rhoptry biogenesis by activating rhoptry protein exit from the TGN, but also participates in the post-Golgi maturation process of preROP compartments into apically anchored club-shaped mature organelles. For this latter activity, our data indicate a specific functional relationship between *TgAP1* and the Rab5A-positive endosome-like compartment. In addition, we unraveled an original role for *TgAP1* in the regulation of parasite division. AP $\mu 1$ -depleted parasites undergo normal daughter cell budding and basal complex assembly but fail to segregate at the end of cytokinesis.

Author summary

The phylum Apicomplexa comprises a large group of obligate intracellular parasites of wide human and agricultural significance. Most notable are *Plasmodium*, the causative agent of malaria, and *Toxoplasma gondii*, one of the most common human parasites, responsible for disease of the developing fetus and immune-compromised individuals. Apicomplexa are characterized by the presence of an apical complex consisting of

France/Centre National de la Recherche Scientifique (CNRS). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

secretory organelles named micronemes (MIC) and rhoptries (ROP). MIC and ROP proteins, released upon host cell recognition, are essential for host cell invasion and parasite survival. After invasion, these organelles are neo-synthesized at each parasite replication cycle. In our study, we demonstrate a crucial role for the *T. gondii* clathrin adaptor complex AP1 in the vesicular transport of neo-synthesized MIC and ROP proteins, thereby regulating mature apical organelle formation. In addition, we unravel an original role for TgAP1 in the late stages of the parasite division process during daughter cell segregation. Therefore, our study provides new insights into key regulatory mechanisms of the vesicular trafficking system essential for host invasion and intracellular survival of *Toxoplasma gondii*.

Introduction

Eukaryotic parasitic pathogens belonging to the phylum *Apicomplexa* are responsible for causing severe mortality in humans and great economic losses in livestock. *Toxoplasma gondii* (*T. gondii*) is of critical importance to pregnant women, as primary infections have the potential to cause neonatal malformations and even death of the developing foetus. In addition, the opportunistic nature of this obligate intracellular parasite can lead to the development of encephalitis in immunosuppressed individuals after reactivation of lifelong persistent cysts in the central nervous system [1]. As their name suggests *Apicomplexa* have a complex of unique apical secretory organelles called rhoptries, micronemes and dense granules that sequentially release their content enabling parasite invasion and intracellular survival. Microneme proteins (MIC) and Rhoptry Neck proteins (RON) are first secreted and trigger the formation of a transient structure, the moving junction (MJ) that anchors the parasite to the host cell and forms a ring through which the parasite penetrates [2] [3]. Rhoptry protein (ROP) contained in the bulb portion of these club-shaped organelles are immediately discharged after MJ formation and participate in the establishment of the intracellular parasitophorous vacuole (PV) in which the parasite intensively multiplies [4]. ROP proteins secreted into the host cell also play a crucial role in the manipulation of host innate immune responses to promote parasite survival [5]. Dense granule proteins (GRAs) are key parasite effectors exocytosed during parasite entry into the vacuolar space, where a certain sub-population contributes to the formation of a nano-tubulo-vesicular network called the intravacuolar network [6] [7]. This tubular network has been shown to be essential for nutrient import and regulation of parasite antigen exposure at the PV [8]. In addition, similar to ROP proteins, GRA proteins can be secreted beyond the PV membrane to actively modulate host gene expression and immune responses triggered upon infection [9].

The stripped-down and polarized version of the eukaryotic intracellular trafficking system has facilitated the use of *T. gondii* in studying the biogenesis of conserved organelles like the Golgi apparatus [10], and, more recently, of the apicomplexan-specific rhoptries, micronemes and dense granules [11]. These secretory organelles are formed *de novo* during each parasite replication cycle by budding and fusion of vesicles emerging from the ER and Golgi. Earlier studies have characterized sorting motifs within MIC and ROP proteins required for their trafficking from the Golgi towards their final destination [12] [13] [14] [15] [16] [17]. These studies led to the conclusion that protein processing and protein sorting were inter-dependent activities. For instance, the prodomain of soluble MIC3, MIC5 and M2AP proteins was shown to be essential for targeting the proteins to the micronemes [14] [15] [18]. Processing of ROP proteins takes place at a post-Golgi level and by contrast to MIC proteins, the presence of the

pro-region of ROP1 was not a prerequisite for its targeting to the rhoptries [19]. More recently the trafficking routes taken by MIC and ROP proteins were delineated by examining the functions of some regulators of the endocytic compartments [11] [20] [21]. Key trafficking molecules were identified, such as the sortilin-like receptor (SORTLR) [22], the dynamin-related protein B (DrpB) [23] and the HOPS/ CORVET complex subunits Vps11, Vps18, Vps39, Mon1 and Vps9, recently described as the Guanine nucleotide Exchange Factor (GEF) of Rab5A [20] [24], all involved in the anterograde pathway regulating secretory organelle biogenesis. In addition, *TgStx6*, a parasite SNARE homolog of syntaxin 6 and the retromer protein Vps35, which are involved in the retrograde transport of molecules from the endosomal-like compartment (ELC) to the Golgi, were shown to be required for the biogenesis of dense granules [25] and rhoptries/micronemes [26], respectively. These recent findings suggest that *T. gondii* has functionally repurposed evolutionarily conserved regulators of the endosomal system to the secretory pathway to form secretory organelles [11] [20]. SORTLR was identified as the unique receptor transporting both, ROP and MIC proteins from the Golgi to the ELC [22]. Depletion of SORTLR led to parasites deprived of apical secretory organelles and ROP and MIC proteins were released into the vacuolar space, or the host cell cytoplasm via the default constitutive secretion pathway. However, so far, little is known about how neo-synthesized ROP and MIC proteins loaded on the unique receptor SORTLR are differentially sorted at the level of the Trans-Golgi-Network (TGN) to reach their distinct final destinations. ROP proteins and a sub-population of MIC proteins (MIC3/MIC8 complex) were shown to be transported via a Rab5A/C-dependent and Rab7-independent pathway, while MIC2/M2AP complex trafficking was Rab5A/C- and Rab7-independent [21]. However, depletion of Mon1, the putative GEF factor for Rab7, as well as depletion of the different HOPS complex subunits involved in Rab7 endosomal compartment biogenesis and integrity, was recently shown to impair rhoptry, microneme and dense granule formation, thus leading to contradictory conclusions concerning the role of Rab7 in secretory organelle formation [20]. Interestingly, the clathrin adaptor protein 1 complex (*TgAP1*) was found associated with the C-terminal tail of SORTLR, suggesting that SORTLR-mediated transport of ROP and MIC proteins might occur via *TgAP1*- and clathrin-dependent budding from the Golgi [22]. In eukaryotes, the AP1 complex has a highly conserved regulatory function in the transport of cargos at the level of the TGN and the early / sorting endosomal compartment [27] [28] [29]. Notably, AP1 regulates the targeting of resident hydrolases to lysosomes in mammals or to the digestive vacuole in plants [30] and yeast [27] [30]. AP1 also plays an essential role in the polarized sorting of vesicles from the TGN to the plasma membrane in epithelial cells [31] and in the secretion of plasma membrane and cell wall proteins in plants [30] [32]. In addition, AP1 has also been shown to have a conserved role in the regulation of the cell division process in lower and higher eukaryotes. For the latter, AP1 is crucial for the final step of daughter cell segregation by delivering Golgi-derived vesicles at the cleavage furrow of dividing cells or the developing cell plate in plants [33] [34] [35]. Finally, AP1 is also involved in the retrograde pathway from the early/sorting endosomes to the TGN [29] and in the retrieval of membrane and other factors from immature secretory granules to promote their maturation [36]. A recent phylogenetic analysis of AP complexes in apicomplexans revealed that these parasites have undergone repeated secondary losses of adaptin complex genes [37]. While the four subunits of the AP1 complex were retained in all studied apicomplexan genomes, the entire AP3 complex was neither found in *Theileria*, nor in *Cryptosporidium parvum* and *Babesia bovis*. This study also indicated a possible degeneration of AP3 subunits in *Plasmodium*, while AP μ 2 was lost in *C. parvum*. In addition, like many other eukaryotes, the apicomplexans possess a single AP β 1/2 subunit. Therefore, *T. gondii* appears as the unique apicomplexan parasite having conserved in its genome all the genes encoding for AP1, AP2, AP3 and AP4 complexes [37]. AP1 is

composed of four subunits: two large subunits γ and β , a medium subunit μ , and a small subunit σ . Sorting motifs present in the cytoplasmic domain of cargo receptors are specifically recognized by the different sub-units of the complex. While γ and σ recognize the dileucine motif, β and μ bind to the tyrosine-based motif [27] [38]. In *T. gondii*, a mutagenesis analysis of the cytoplasmic domain (CD) of the transmembrane MIC2 protein has revealed that two conserved motifs are necessary and sufficient for targeting the protein to the micronemes [39]. One of these signals contains tyrosine residues, whereas the other one is composed of a stretch of acidic residues. These motifs are also present and conserved in the CD of MIC6 and are sufficient for microneme targeting [17]. These data suggested the existence of an AP-dependent mechanism for MIC protein sorting at the level of the TGN or ELC. Concerning ROP protein sorting mechanisms, a previous study indicated that ROP2 possesses a dileucine and a tyrosine-based motif located in the C-terminal part of the protein required for the export of ROP proteins from the TGN/ELC [40]. The authors of this study also demonstrated that these motifs were specifically recognized by the μ subunit of AP1 [40] [41]. Over-expressing AP μ 1 mutated at residues that bind the tyrosine-motif, led to accumulation of ROP2 in a post-Golgi multi-vesicular compartment resembling endosomes and immature rhoptries [41]. Similarly, perturbing the function of AP μ 1 by siRNA interference led to major defects in ROP biogenesis while, microneme and dense granule organelle biogenesis was not perturbed [41]. However, this model was challenged by the resolution of the ROP2 protein structure. This study revealed the absence of the predicted transmembrane domain and demonstrated that the association of ROP2 with the parasitophorous vacuole membrane is mediated by an amphipathic peptide enclosed in the N-terminal domain [42]. Of note, SORTLR that associates with ROP and MIC proteins also possesses a dileucine motif in its cytoplasmic tail, suggesting an additional AP-dependent sorting mechanism for ROP and MIC proteins [22]. In *P. falciparum*, AP1 localizes at the Golgi/ER compartment and in rhoptries at the schizont stages. The AP μ 1 subunit was found associated with the rhoptry-associated protein 1 (RAP1) suggesting a role in rhoptry protein trafficking [43].

In the present study, we demonstrated that TgAP1 regulates both, rhoptry and microneme formation but not dense granule biogenesis. In addition, despite a significant difference in the cell division process of *T. gondii* compared to other eukaryotes, our study revealed a conserved role for TgAP1 in the late stages of cytokinesis, by regulating the final step of daughter cell segregation.

Results

TgAP μ 1 localizes at the Trans-Golgi-Network and on secretory vesicles

In order to define the subcellular localization of the TgAP1 complex, we generated knock-in (KI) parasites expressing the μ 1 subunit (TGTT1_289770) fused to a HA tag at its C-terminus. Western blot analysis confirmed the expression of the tagged protein at the expected size (Fig 1A). A clonal parasite line was isolated and AP μ 1 localization was analyzed by an immunofluorescence assay (IFA) using confocal (Fig 1B) and super-resolution microscopy SIM (structured illumination microscopy)(Fig 1C and 1D). As expected for the AP1 complex, a localization at the Golgi area was observed and confirmed by co-localization with the TGN marker SORTLR (Fig 1B and 1C). In addition to the TGN, a faint but specific AP μ 1 signal was systematically detected in the parasite cytoplasm by confocal microscopy. After saturating the stronger Golgi-associated signal, we clearly identified this weaker signal as AP μ 1-positive vesicles spread throughout the cell cytoplasm and close to the cell periphery (Fig 1D, arrows). This pattern of distribution strongly suggested that TgAP1 is involved in additional trafficking pathways apart from the ones involved in MIC and ROP protein transport, in particular, in vesicle delivery to the plasma membrane or the inner membrane complex (IMC).

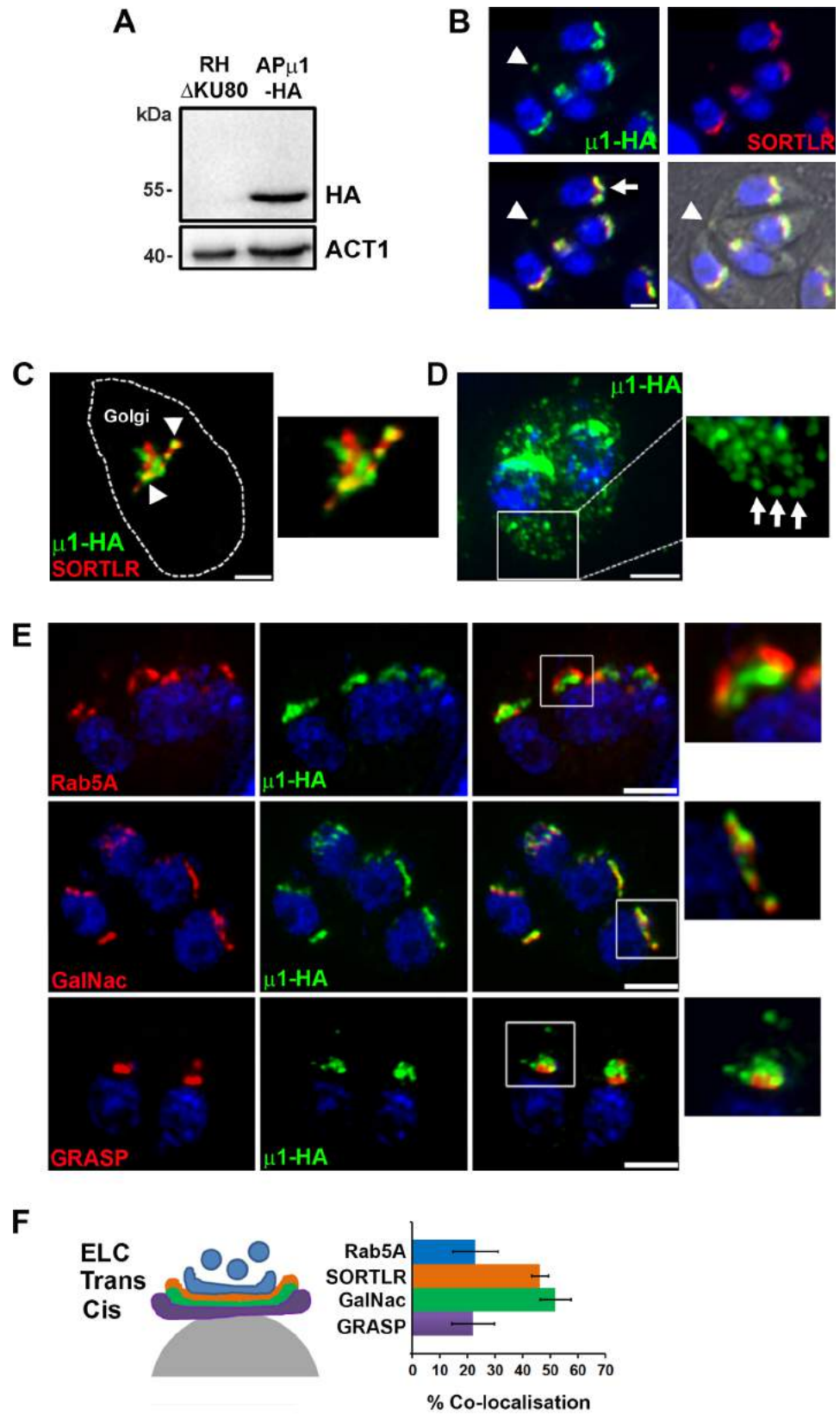


Fig 1. AP μ 1 localizes at the Trans-Golgi-Network and on secretory vesicles. **A-** Western Blot image showing the expression of the endogenous HA-tagged μ 1 subunit at the expected size of 49 kDa in knock-in parasites (RH Δ KU80: parental strain). Actin (ACT1) was used as a loading control. **B-** Confocal microscopy images showing the localization of μ 1-HA (green) with SORTLR (red) at the Golgi region (arrow). Nuclei are shown by DNA staining (blue). Note the very discrete signal of μ 1-HA (green) at a localization corresponding to the residual body (arrowhead). Bar: 2 μ m. **C-** SIM image showing the partial co-localization (indicated by arrowheads) of μ 1-HA (green) with SORTLR (red) in sub-regions of the Golgi apparatus. Bar: 1 μ m. **D-** SIM image showing the localization of μ 1-HA (green) in vesicles spread throughout the parasite cytoplasm and also present in proximity to the plasma membrane (arrows in the inset). Bar: 2 μ m. **E-** The co-localization of μ 1-HA (green) with markers of the endosomal compartment (Rab5A-YFP), trans-Golgi (GalNAc-GFP) and cis-Golgi (GRASP-RFP) (all shown in red) was examined by SIM microscopy. For each marker, a zoom of the Golgi region is shown (insets). Bars: 2 μ m. **F-** Scheme (left) illustrating the localization of the different markers associated with the endosomal-like compartment (ELC), the trans-Golgi (Trans) and the cis-Golgi (Cis). The histogram (right) indicates the percentage of co-localization between μ 1-HA and Rab5A-YFP, SORTLR, GalNAc-GFP and GRASP-RFP, in at least 30 parasites that were analyzed for each condition. μ 1-HA displays the strongest co-localization with the TGN markers SORTLR and GalNAc-GFP.

<https://doi.org/10.1371/journal.ppat.1006331.g001>

SIM acquisition and Imaris software analysis indicated a $46.4 \pm 3.8\%$ co-localization between AP μ 1 and SORTLR (Fig 1C and 1F) at the TGN. In agreement, AP μ 1 shows a similar high percentage of co-localization ($52.0 \pm 5.6\%$) with another TGN marker, GalNAc, in contrast to the cis-Golgi marker GRASP ($22.0 \pm 7.7\%$) (Fig 1E and 1F). Using a KI line expressing Rab5A-YFP under the endogenous promoter, the Rab5A-positive endosome-like compartment (ELC) was mostly detected as vesicles emerging posteriorly from the AP μ 1- and SORTLR-positive TGN (Fig 1E, upper panel, Fig 1F and S1 and S2 Movies). In agreement, a weaker co-localization of AP μ 1 with Rab5A was quantified ($22.9 \pm 8.2\%$) compared to the TGN markers SORTLR and GalNAc (Fig 1E and 1F). However, we noticed that the Rab5A-positive ELC and the TGN appeared physically connected, in particular during Golgi duplication at the G1/S transition phase of the cell cycle (S1A Fig). This notion was also supported by the observation that Brefeldin A (BFA) treatment, which disperses the Golgi apparatus by inhibiting the activity of the ARF1 GTPase required for COPII-mediated vesicular transport, led to the partial dispersion of the Rab5A-positive compartment similar to what was observed for the AP1-positive TGN (S1B Fig). As previously suggested [19], these results indicate that *T. gondii* possesses an unusual ELC, which is physically and likely functionally connected to the TGN, resembling the early endosomal/TGN hybrid compartment of plants [44].

TgAP μ 1 belongs to a conserved tetrameric complex and interacts with the epsin-like protein

To assess, whether AP μ 1 is a component of the highly conserved tetrameric complex composed of the three other subunits σ 1, β and γ , we performed an immunoprecipitation (IP) assay using anti-HA antibodies on AP μ 1-HA KI parasite lysate followed by mass spectrometry analysis. AP μ 1 was reproducibly found associated with the three other subunits (Table 1, S2 Fig) confirming the formation of a conventional tetrameric complex described in other eukaryotes. This finding was also supported by the localization of the σ 1 sub-unit at the SORTLR-positive TGN in KI parasites expressing the fusion protein σ 1-HA under the native promoter (S3A and S3B Fig). In addition, the unique ENTH domain-containing protein encoded in the *T. gondii* genome, which we named Epsin-like protein (TgEpsL) was identified, however with only one unique peptide (Table 1, S2 Fig). We decided to further investigate the possible interaction between TgAP1 and TgEpsL because epsin proteins are very well known AP2 and AP1 binding proteins involved in the activation of clathrin-mediated vesicular budding at the plasma membrane and at the TGN, respectively. Epsins bind to phospholipids via their ENTH domain and regulate clathrin coat formation by inducing curvature of the lipid

Table 1. List of proteins identified by mass spectrometry following the IP of μ 1-HA proteins in KI parasites expressing μ 1-HA under the endogenous promoter. The detailed list is included in S2 Fig. The parental strain RH Δ KU80 was used as a control for non-specific binding to the anti-HA antibody-coated beads. The table indicates the number of “unique peptides / spectra” for each identified protein in two biological independent assays (IP1 and IP2).

Protein name	Accession numbers	Molecular weight (Da)	Total number of unique peptides / spectra	
			IP-1	IP-2
			μ 1-HA	μ 1-HA
mu1 adaptin	TGGT1_289770	48 918	11/55	8/16
gamma adaptin	TGGT1_313670	107 036	20/53	24/77
beta adaptin	TGGT1_240870	101 920	18/47	19/76
sigma1 adaptin	TGGT1_270370	19 679	2/4	1/2
clathrin heavy chain	TGGT1_290950	194 480	4/6	1/1
ENTH-domain containing protein	TGGT1_214180	65 903	1/1	1/1

<https://doi.org/10.1371/journal.ppat.1006331.t001>

bilayers [38]. Sequence analysis indicated that *TgEpsL* contains conserved clathrin and phosphoinositide binding sites, such as recently described elsewhere [45] (S3C Fig). However, besides the ENTH domain, no similarities were found between *TgEpsL* and other epsins (S3C Fig), suggesting a specificity of binding partners and regulatory mechanisms of the protein activity in *T. gondii*. First, we generated single KI parasites expressing *EpsL*-cMyc under its natural promoter as well as double KI parasites expressing both, *EpsL*-cMyc and AP μ 1-HA proteins (Fig 2A). IFA analysis by confocal and SIM microscopy confirmed the co-localization of *TgEpsL* with *TgAP1* at the TGN (Fig 2B). To verify the interaction between *TgAP1* and *TgEpsL*, an IP was performed on double KI *EpsL*-cMyc / AP μ 1-HA expressing parasites, using either anti-cMyc or anti-HA antibodies. Western blot analysis confirmed the interaction between AP μ 1-HA and *EpsL*-cMyc in both IP assays (Fig 2C and 2D). In agreement with this result, IP of *EpsL*-cMyc followed by mass spectrometry identified the β , γ and μ 1 subunits of the AP1 complex and the small GTPase ARF1 as the main proteins associated with *TgEpsL* (Table 2, S4 Fig), thereby confirming the result obtained by western blot. Importantly, no subunit of the AP2 complex was identified, suggesting that *TgEpsL* might not function in AP2-mediated endocytosis. In other eukaryotes, the AP1 complex interacts with epsinR via the exposed GAE (“Gamma Appendage Ear”) domain of the γ subunit [46] [47]. Sequence alignment analysis showed a strong conservation of *T. gondii* GAE and BAE (“Beta Appendage Ear”) domains with the corresponding domains of the AP1 β and γ subunits from other species (S5 Fig, S1 Table). Therefore, the GAE and BAE domains of *TgAP1*, fused to GST were produced (Fig 2E). GST pull-down experiments in presence of a total extract of *EpsL*-cMyc/ AP μ 1-HA double KI parasites indicated that the GAE domain is sufficient to pull-down the *TgEpsL* protein, while no binding of the μ 1 subunit was monitored (Fig 2F). In contrast, a weak interaction was detected between *TgEpsL* and the BAE domain (Fig 2F), suggesting a preferential association of *TgEpsL* with the γ sub-unit, as a similar quantity of the two domains was used in the assay (Fig 2E). Interestingly, we also found that the BAE domain pulled down the μ 1 subunit. As no direct interaction between this domain and the μ 1 subunit has been described in other eukaryotes, it was likely that the BAE domain could interact with a complex of proteins that includes μ 1 and other *TgAP1* binding proteins. In agreement with this finding, SORTLR that directly interacts with the μ 1 subunit, was also found in the pull down eluate of the β -ear but not in the eluate of the γ -ear (Fig 2F).

Together these data indicate that the AP1 complex in *T. gondii* is conserved at the molecular level and likely functions similar to its mammalian homologue as a heterotetrameric complex regulating epsin-mediated vesicular transport of parasite proteins.

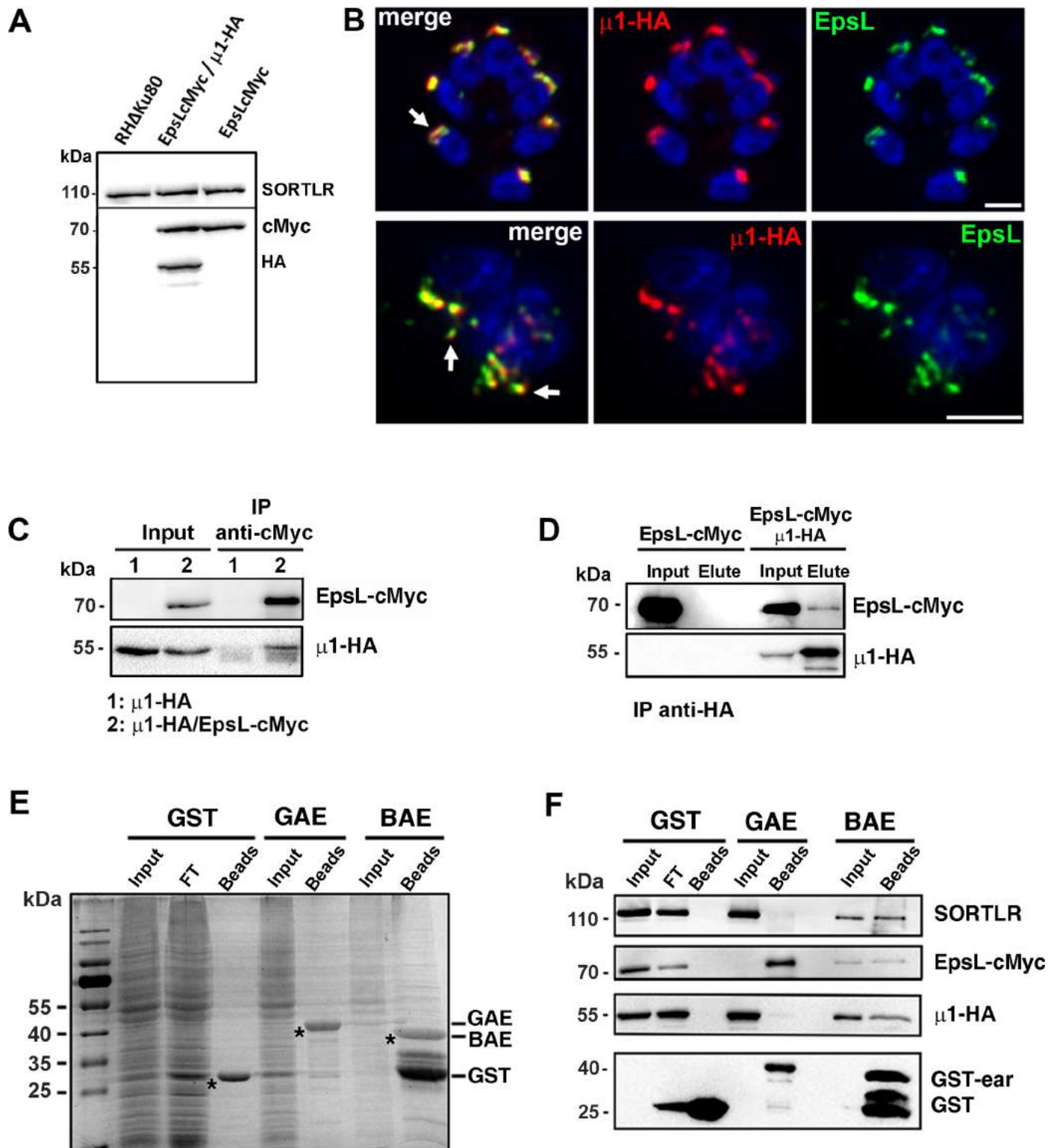


Fig 2. The unique ENTH-domain containing protein of *T. gondii* is a key partner of AP μ 1. A- Western blot showing the expression of the cMyc-tagged EpsL protein at the expected size of 66 kDa in single KI parasites (lane 2), and together with μ 1-HA (49 kDa) in double knock-in parasites (lane 3). The parental strain (RHΔKU80) is shown in lane 1. SORTLR was used as a loading control. B- Images illustrating the co-localization of EpsL-cMyc (green) with μ 1-HA (red) (arrows), acquired with confocal microscopy (upper panel) or SIM (lower panel). Bars: 2 μ m. C- Co-immunoprecipitation of μ 1-HA with EpsL-cMyc in double KI parasites (lanes 2) using anti-cMyc antibodies. No binding of the μ 1-HA protein on anti-cMyc coated beads was

detected in single KI μ 1-HA expressing parasites (lane 1). D- Reverse co-immunoprecipitation of EpsL-cMyc with μ 1-HA in double KI parasites using anti-HA antibodies. No binding of EpsL-cMyc protein with anti-HA antibody coated beads was detected in the single KI EpsL-cMyc expressing parasites. E, F- A GST-pull down experiment with the GST-tagged gamma appendage ear (GAE) and GST-tagged beta appendage ear (BAE) domains of *TgAP1* was performed in presence of a total lysate from EpsL-cMyc/ μ 1-HA double KI parasites. E: SDS-PAGE gel stained with coomassie blue showing that a similar quantity of GST, GST-GAE and GST-BAE (asterisks) was bound on the glutathion beads used in the assay shown in F. F: WB analysis demonstrated the preferential binding of EpsL-cMyc to the ear domain of the γ sub-unit (GAE). A weak interaction of SORTLR and the μ 1 sub-unit with the BAE domain was also detected. GST alone was used as a control. FT: Flow-Through.

<https://doi.org/10.1371/journal.ppat.1006331.g002>

TgAP1 regulates microneme formation

To investigate *TgAP1* functions, inducible knock-out (KO) parasites lacking the μ 1 subunit were generated using the CreLox strategy [48], where excision of the endogenous locus and subsequent expression of the YFP protein is triggered upon addition of rapamycin (Fig 3A). Integration of the LoxP- μ 1HA-LoxP cassette at the endogenous locus was validated by PCR (Fig 3B) and expression of the corresponding protein by Western Blot (Fig 3C). By IFA, integration of the LoxP- μ 1HA-LoxP cassette led to the expression of a tagged protein that co-localized with SORTLR at the TGN (Fig 3D, upper panel). Incubation of transgenic parasites with rapamycin triggered AP μ 1 depletion and YFP expression, but only in 12.0±4% of the whole population (Fig 3B and 3D, lower panel). Of note, increasing rapamycin concentration did not lead to a higher yield of KO parasites. Longer rapamycin induction periods improved the rate of YFP-positive parasites but led to random excision events over the successive cell cycles making it difficult to analyze the heterogenous phenotypes associated with AP1 depletion. Thus, we decided to induce AP μ 1-KO parasites for 6 hours after invasion in order to trigger μ 1 gene excision during the first division cycle.

TgAP1 had been previously shown to regulate rhoptry biogenesis at a post-Golgi level while microneme and dense granule formation was not perturbed [41]. Upon depletion of AP μ 1, we found that the soluble MIC3 protein was re-directed towards the vacuolar space, demonstrated by the GAP45 labeling of parasite contours, whereas MIC8 was found retained in the TGN, confirmed by its co-localization with SORTLR (Fig 4A). A weaker MIC8 staining was also detected at the plasma membrane suggesting that a part of the protein escaped by the constitutive secretory pathway to the parasite surface (Fig 4A). On the other hand, M2AP and MIC2 exit from the TGN was not impaired as the proteins were not seen retained in this compartment, nor secreted into the vacuolar space. However, both proteins appeared concentrated at the apex of the parasites, while lateral micronemes were weakly detected (Fig 4A). Similar to MIC2, the transmembrane MIC6 (Fig 4B) and AMA1 (S6A Fig) proteins, as well as MIC1 (S6A Fig), were found concentrated in apical micronemes, while the soluble protein MIC4 was found re-routed towards the vacuolar space (Fig 4B). Furthermore, in support of our data

Table 2. List of proteins identified by mass spectrometry following the IP of EpsL-cMyc protein in double KI parasites expressing EpsL-cMyc and μ 1-HA proteins. The detailed list is included in S4 Fig. The single KI parasites expressing μ 1-HA was used as a control for non-specific binding to the anti-cMyc antibody-coated beads. The table indicates the number of “unique peptides / spectra” for each identified protein.

Protein name	Accession numbers	Molecular weight (Da)	Total number of unique peptides / spectra	
			Control:	IP:
			μ 1-HA	EpsL-cmyc/ μ 1-HA
ENTH domain-containing protein	TGGT1_214180	65 903	3/4	15/76
beta adaptin	TGGT1_240870	101 920	/	19/28
gamma 1 adaptin	TGGT1_313670	107 036	2/3	17/33
mu1 adaptin	TGGT1_289770	48 918	/	9/11
ADP ribosylation factor ARF1	TGGT1_276140	61 745	2/2	6/12

<https://doi.org/10.1371/journal.ppat.1006331.t002>

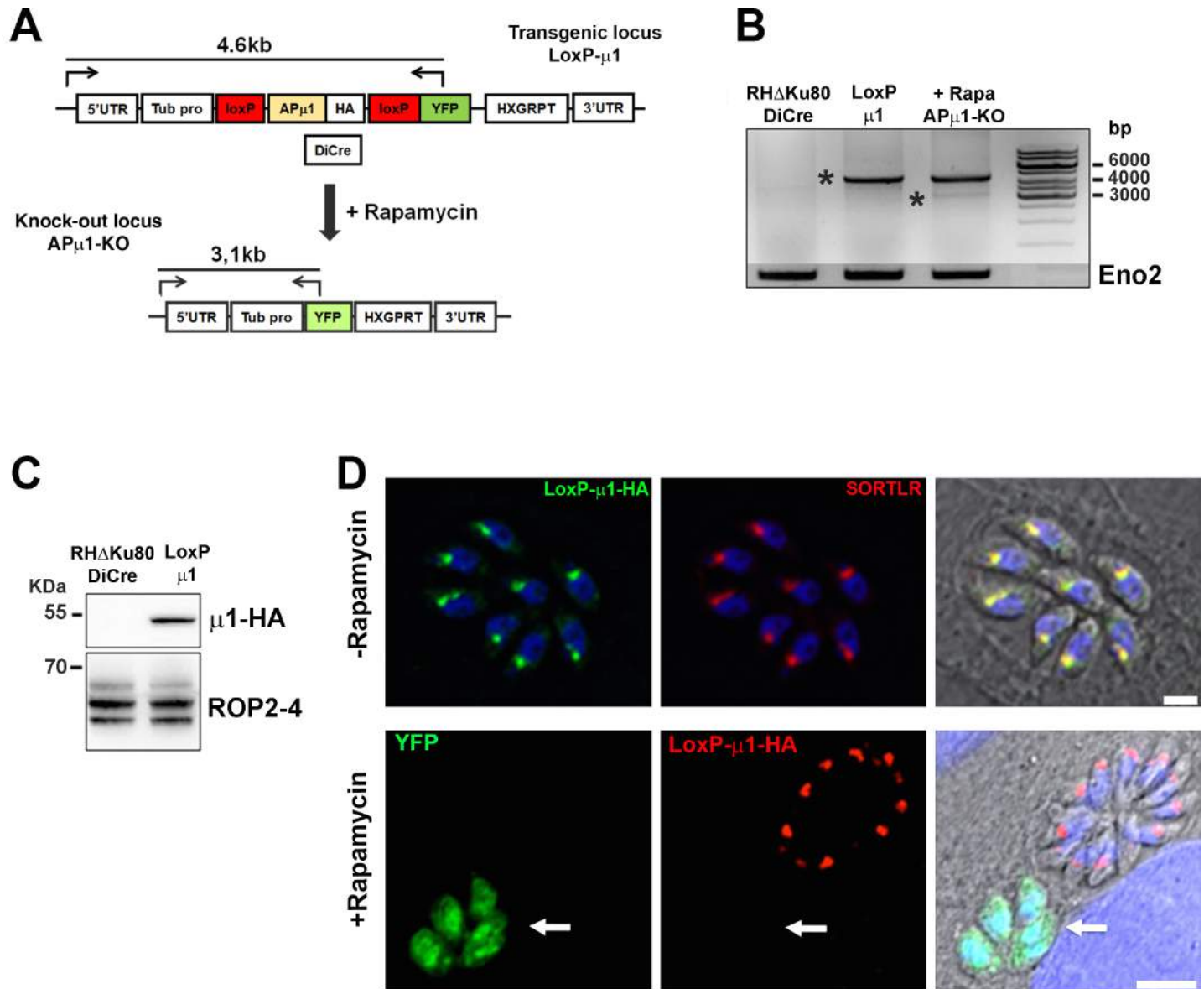


Fig 3. CreLox-based strategy used to deplete AP μ 1. **A-** Scheme depicting the cloning strategy used to replace the endogenous μ 1 locus by the LoxP- μ 1-HA-LoxP insert. Upon rapamycin induction, the DiCre recombinase excised the LoxP flanked locus leading to YFP expression. The positions of the primers used to verify the integration of the insert into the genome and its excision upon rapamycin incubation, are indicated. **B-** PCR confirming the integration of the LoxP- μ 1-HA-LoxP insert at the endogenous AP μ 1 locus resulting in the amplification of a band at 4.6 kb (asterisk). Rapamycin induction resulted in the amplification of a lower and weak band at 3.1 kb (asterisk) corresponding to the low percentage of AP μ 1-KO parasites. The primers used for the PCR are depicted in A-. Amplification of the enolase 2 (Eno2) gene was used as a control. **C-** WB showing the expression of integrated μ 1-HA protein at the expected size in a clonal population (DiCre RH Δ KU80: parental strain). ROP 2–4 was used as a loading control. **D- Upper panel:** Confocal microscopy images showing the localization of μ 1-HA (green) together with SORTLR (red) at the TGN, after integration of the sequence flanked by the LoxP sites in DiCre RH Δ KU80 parasites. Nuclei are shown by DNA staining (blue). Bar: 2 μ m. **Lower panel:** Confocal microscopy images showing the absence of μ 1-HA signal (red) in YFP positive parasites (arrow) upon rapamycin treatment. Bars: 5 μ m.

<https://doi.org/10.1371/journal.ppat.1006331.g003>

suggesting that TgAP1 might not be involved in the MIC2/M2AP complex exit from the TGN, we observed by SIM microscopy that AP μ 1 co-localized with immature proMIC3 but not with immature proM2AP (Fig 4C). In addition, SORTLR was not mis-localized in AP μ 1-KO parasites (S6B Fig), showing that TgAP1 likely functions downstream of SORTLR in the antero-grade secretory pathway. Finally, we observed that dense granule biogenesis was not affected

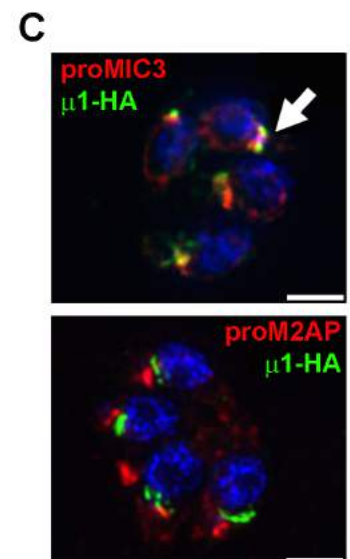
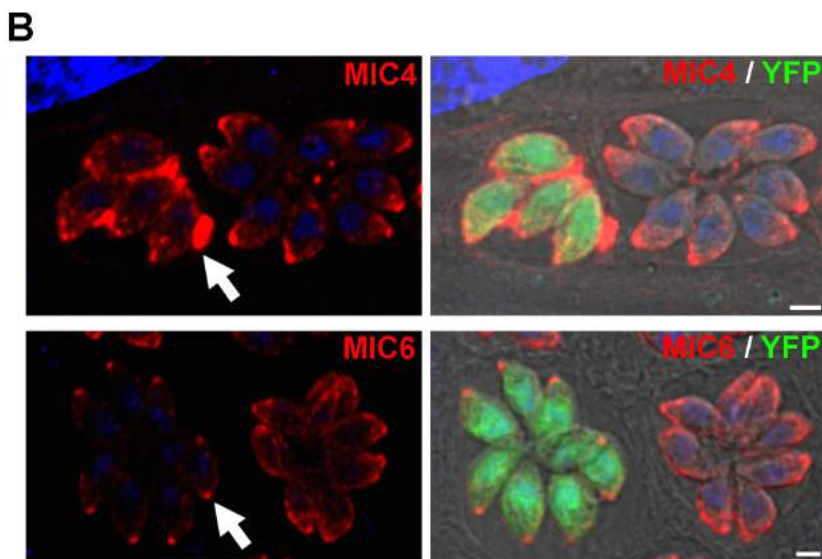
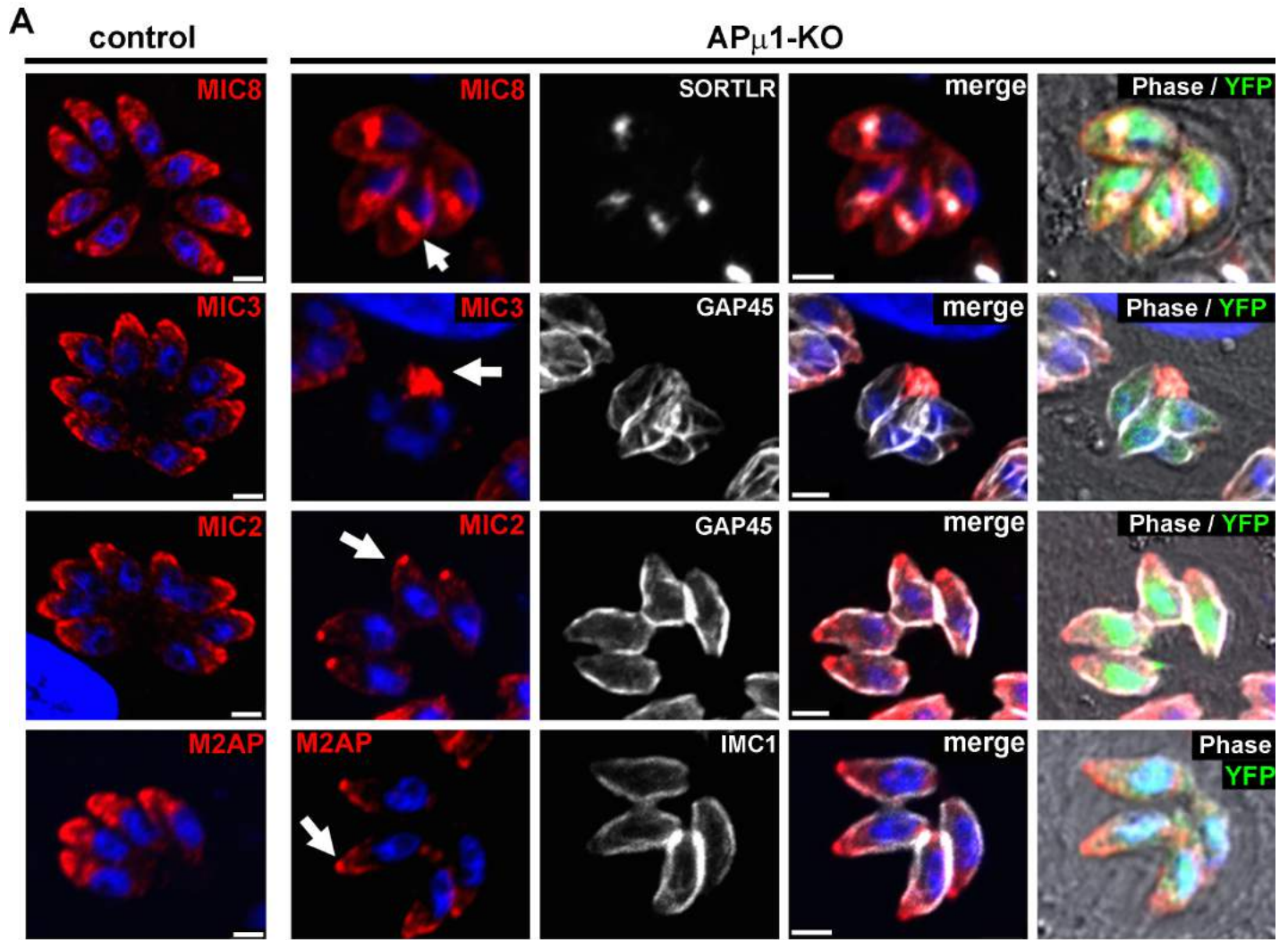


Fig 4. AP μ 1 ablation impairs microneme protein localization. **A-** Confocal images showing the localization of MIC8, MIC3, MIC2 and M2AP proteins (red) in control (YFP-negative parasites) and AP μ 1-KO parasites (YFP-positive). MIC8 accumulated in the parasite TGN, confirmed by its co-localization with SORTLR (white). MIC3 was found to be secreted into the parasitophorous vacuolar space (arrow). Parasite contours were stained with the IMC marker GAP45 (white). M2AP and MIC2 (parasite contours: IMC markers IMC1 and GAP45, respectively, both shown in white) displayed a preferential apical localization, while lateral micronemes were weakly detected. Nuclei are shown by DNA staining (blue). Bar: 2 μ m. **B-** Confocal microscopy images showing the localization of MIC4 (red, upper panel), and MIC6 proteins (red, lower panel) in control (YFP-negative parasites) and AP μ 1-KO parasites (YFP-positive). MIC4 was found to be secreted into the parasitophorous vacuolar space (arrow), while MIC6 was concentrated at the apex (arrow). **C-** SIM microscopy images showing the localization of μ 1-HA (green) and proMIC3 (top) or proM2AP (bottom) in μ 1-HA KI parasites. No co-localization between μ 1-HA and proM2AP was observed in contrast to proMIC3. Bars: 2 μ m.

<https://doi.org/10.1371/journal.ppat.1006331.g004>

upon AP μ 1 ablation (S6C Fig). Therefore, our data indicate that TgAP1 is part of the early sorting machinery that regulates distinct MIC protein complex transport from the TGN to their final destinations.

TgAP1 regulates rhoptry biogenesis

In YFP-positive parasites, we also observed that deletion of the AP μ 1 subunit drastically affected the formation of rhoptry organelles, which were detected as dispersed compartments distributed throughout the cell cytoplasm (Fig 5A, upper panel). We also found that ROP proteins were re-routed to the vacuolar space (Fig 5A, middle panel). Only $16.6 \pm 4.7\%$ of the examined vacuoles showed typical apically located club-shaped rhoptries in AP μ 1-KO parasites (Fig 5B). Next, we monitored immature pre-rhoptry compartment formation (Fig 5A and 5C). Quantification of parasites displaying a positive signal for proROP4 proteins indicated that $26.1 \pm 1.2\%$ of control parasites contained immature rhoptries, corresponding to dividing parasites in S/M phase of the cell cycle. In AP μ 1-KO parasites, an increase in the percentage of vacuoles positive for the proROP4 signal was counted ($40.5 \pm 1.2\%$), indicating no defect in ROP protein neosynthesis and suggesting a defect in ROP protein maturation. In agreement, immature proROP4 proteins were detected in the vacuolar space or at the residual body (Fig 5A, lower panel) in $71.8 \pm 5.6\%$ of the proROP4-positive vacuoles. Therefore, our data indicate that TgAP1 is involved in immature proROP protein exit from the TGN compartment. In absence of AP μ 1, the neo-synthesized ROP proteins are secreted into the vacuolar space likely via the constitutive pathway, leading to an increase of vacuoles positive for proROP4 signal over the successive cell cycles in our unsynchronized parasite population. In addition, the presence of dispersed rhoptries containing mature ROP proteins (Fig 5A) may indicate an additional role of TgAP1 in the rhoptry maturation process at a post-Golgi level. The fact that we observed both, the early proROP protein targeting to the vacuolar space and the later step of rhoptry organelle maturation defect, is likely a consequence of the unsynchronized nature of the AP μ 1 locus excision events together with the disappearance of the remaining endogenous protein, which might be effective at different stages of the rhoptry maturation process in this fast growing parasite.

Finally, we monitored host cell invasion and egress activity in AP μ 1-KO parasites. In agreement with the observed defect in rhoptry and microneme formation, we found that host cell invasion was drastically inhibited (Fig 5D). In contrast, egress was moderately impaired (Fig 5E), similar to what was observed for the Vps11-KO mutant [20].

Over-expression of TgAP μ 1 leads to defects only in rhoptry formation

It has been previously shown that over-expression of AP μ 1 or a point-mutated form of AP μ 1 caused a drastic defect in rhoptry formation at the post-Golgi level, while microneme biogenesis was not impaired [41]. To understand the distinct observed phenotypes compared to the AP μ 1-KO parasites, we applied a similar strategy and conditionally over-expressed the μ 1

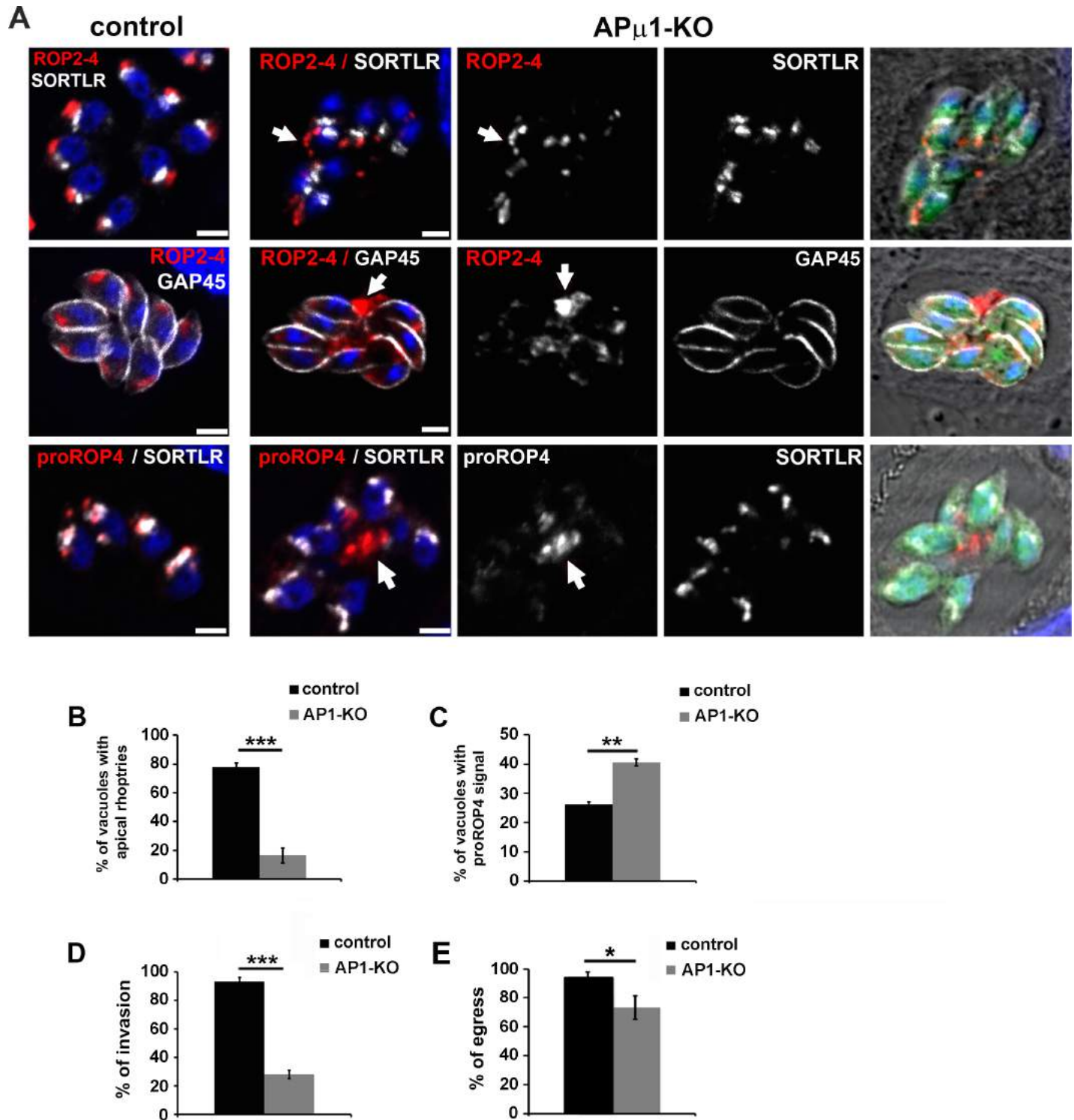


Fig 5. AP μ 1-KO parasites show defects in rhoptry formation. A-Confocal images showing the localization of ROP2-4 and proROP4 proteins (red) in control (YFP-negative vacuoles) and AP μ 1-KO parasites (YFP-positive vacuoles) together with the TGN marker SORTLR and the IMC marker GAP45 (both in white). In AP μ 1-KO parasites, mature rhoptries were found dispersed within the cytosol (upper panel, arrow) or in the vacuolar space (middle panel, arrow), while immature proROP4 proteins were found re-routed towards the vacuolar space and residual body (lower panel, arrow). Bar: 2 μ m. B- Histogram indicating the percentage of examined vacuoles displaying apically positioned rhoptries in control and AP μ 1-KO parasites. Mean values of three independent assays are shown \pm SEM, *** p <0.001 (Student's t-test). C- Histogram indicating the percentage of examined vacuoles positive for the immature protein proROP4 staining in control and AP μ 1-KO parasites. Mean values of three independent assays are shown \pm SEM, ** p <0.01 (Student's t-

test). D- Histogram depicting the percentage of invaded parasites after 45 min incubation with host cells of mechanically released parasites for control (YFP-negative) and AP μ 1-KO (YFP-positive) parasites. Mean values of three independent assays are shown \pm SEM, *** p <0.001 (Student's t-test). E- Histogram depicting the percentage of egressed vacuoles after induction with the calcium ionophore A23187 in control (YFP-negative) and AP μ 1-KO (YFP-positive) parasites. Mean values of three independent assays are shown \pm SEM, * p <0.05 (Student's t-test).

<https://doi.org/10.1371/journal.ppat.1006331.g005>

subunit (DD μ 1 parasites) using the ddFKBP system [49]. Western blot analysis confirmed the time-dependent accumulation of the cMyc-tagged μ 1 subunit upon shield-1 addition in the growth medium (Fig 6A). By IFA, the over-expressed cMyc μ 1 protein was detected at the Golgi area, but also in cytoplasmic vesicles particularly concentrated at the basal pole of the parasite (Fig 6B). We found that cMyc μ 1-overexpressing parasites displayed less pronounced defects compared to AP μ 1-KO parasites. First, MIC2/M2AP and MIC8/MIC3 complexes were both correctly targeted to micronemes (S7A Fig) as previously observed after over-expression of the mutated μ 1D176A subunit [41], and MIC protein processing was not affected (S7B Fig). In contrast to micronemes, rhoptry formation was perturbed, and ROP proteins were mostly detected in dispersed compartments throughout the parasite cytoplasm (Fig 6C, insets). Only $19.0 \pm 2.1\%$ of the population showed apically localized, mature rhoptries (Fig 6D). Furthermore, in agreement with the observed defect in mature rhoptry formation, a partial decrease in host cell invasion was monitored (Fig 6I). However, in opposite to AP μ 1-KO parasites, in DD μ 1-induced parasites, proROP4 proteins were not re-routed to the vacuolar space or into the residual body, and the percentage of proROP4 positive vacuoles was similar to control parasites (Fig 6C and 6E). Furthermore, our analysis by confocal microscopy indicated that proROP4 proteins exit normally from the SORTLR-positive TGN compartment (S7C and S7D Fig), as the percentage of co-localization between these two compartments did not vary compared to control parasites. In addition, we also monitored no change in the percentage of co-localization between immature proROP4 and mature ROP2-4 proteins (Fig 6F and 6G) suggesting no defect in ROP protein processing, which was also confirmed by WB analysis (Fig 6H). Of note, we could also observe a co-localization between proROP4-positive compartments and the over-expressed cMyc μ 1 protein (Fig 6F, insets). To verify that the over-expression of the μ 1 subunit induced defects in rhoptry formation via a TgAP1-dependent activity and not indirectly by titrating other μ 1-interacting factors, we performed an IP of the cMyc μ 1 protein. Immunoprecipitation of cMyc μ 1 using anti-cMyc antibodies identified the three other subunits of the TgAP1 complex and TgEpsL as the main protein partners (Table 3 and S8 Fig). In addition, AP σ 1-HA correctly localized at the TGN (S7E Fig) suggesting that cMyc μ 1 over-expression does not cause the mis-location of the endogenous TgAP1 complex. These data suggest that the over-expressed cMyc μ 1 protein likely integrates into a functional AP1 complex. In summary, we found that over-expressed AP μ 1 impaired the post-Golgi maturation process of rhoptry organelles without perturbing proROP protein proteolytic processing, thus acting at a similar or posterior step to this event in the secretory pathway. We therefore examined the morphology of the endosome-like compartment in DD μ 1 parasites.

Over-expression of TgAP μ 1 perturbed the Rab5A-positive ELC morphology

We transiently transfected HA-tagged Rab5A and Rab7 in DD μ 1 parasites and examined the distribution of the ELC compartment after shield-1 induction (Fig 7A and 7B). Interestingly, we found that the Rab5A compartment displayed an altered morphology, appearing as large vesicular structures scattered along the Golgi area in comparison to the more homogenous and continuous distribution observed in control parasites (Fig 7A). In contrast, the Rab7 compartment did not show detectable changes by confocal microscopy after overexpression of

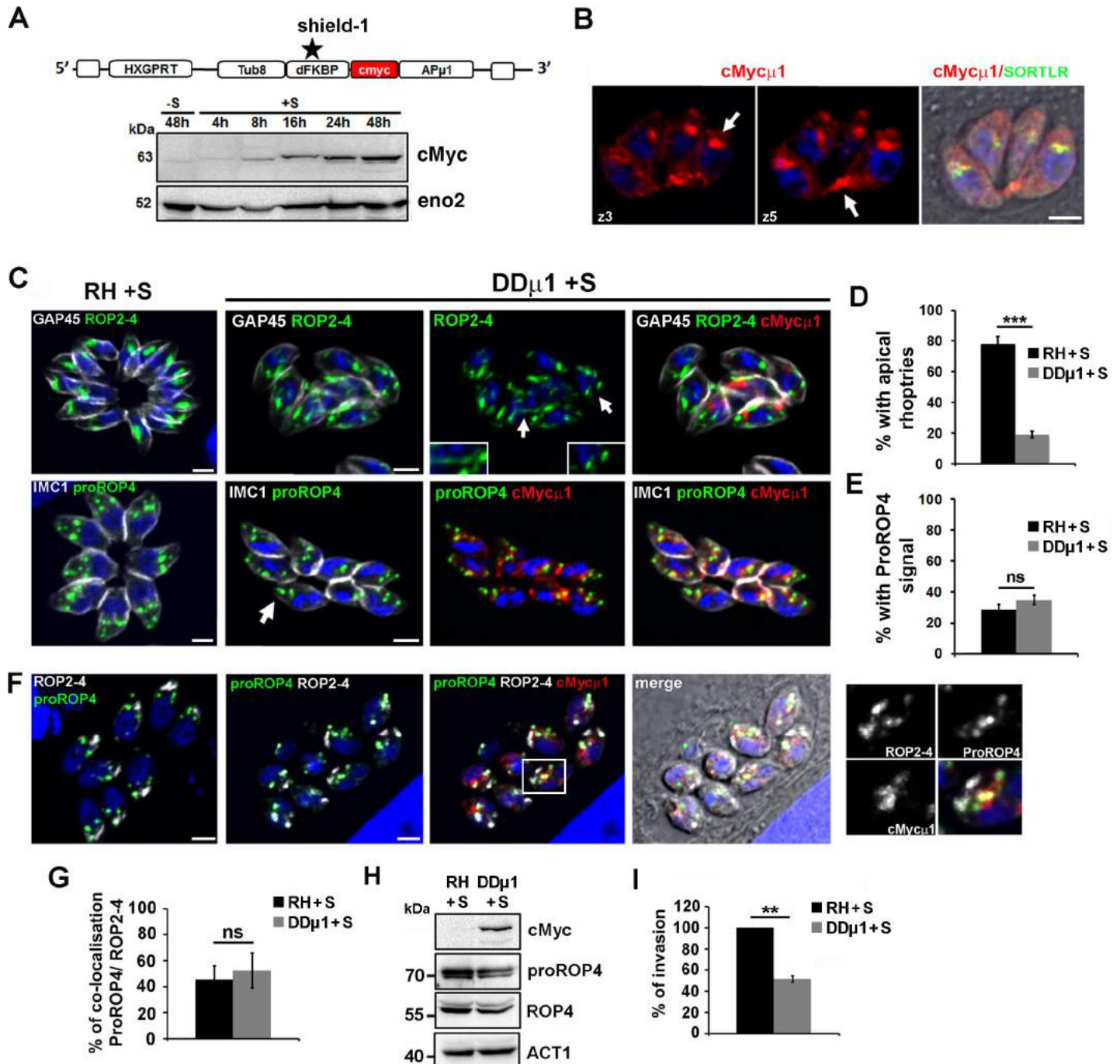


Fig 6. The inducible over-expression of AP1 only perturbed rhoptry formation. **A**— Scheme showing the cloning strategy employed to insert the cMyc-tagged μ 1 subunit under the influence of the destabilisation domain ddFKBP (DD). After addition of the synthetic ligand shield-1, the protein is no longer degraded but accumulated in the parasites (Tub8: tubulin promoter; HXGPRT: resistance cassette). The WB image shows the accumulation of the cMyc μ 1 protein upon shield-1 treatment for the indicated time periods. The protein eno2 was used as a loading control. **B**— Confocal microscopy images showing the localization of the over-expressed cMyc μ 1 protein (red) in a clonal population of DD μ 1 parasites in two confocal planes (z3 and z5). cMyc μ 1 was detected at the Golgi apparatus (co-localization with SORTLR shown in green in the merged image) and in vesicles accumulating at the basal pole of the parasites (arrows). Bar: 2 μ m. **C**— Confocal images showing the localization of ROP2-4 (upper panel) or proROP4 (lower panel) proteins (both in green) and cMyc μ 1 (red) in control RH and DD μ 1 parasites incubated with shield-1 (+S) for 24 hours. The contours of the parasites are delineated by staining of the IMC markers, GAP45 or IMC1 (white). Rhoptries are detected as dispersed atrophied compartments (arrows and insets in upper panel) in DD μ 1 parasites, while proROP4 compartments were normally formed (arrow in lower panel). **D**— Histogram indicating the percentage of examined vacuoles displaying apically positioned rhoptries in control and DD μ 1 parasites induced with shield-1 (+S) for 24 hours. Mean values of three independent assays are shown \pm SEM, *** p <0.001 (Student's t-test). **E**— Histogram indicating the percentage of vacuoles positive for the immature protein proROP4 staining in

control and DDμ1 parasites induced with shield-1 (+S) for 24 hours. Mean values of three independent assays are shown ± SEM. **F-** Confocal images showing the co-localization of ROP2-4 (white), proROP4 (green) and cMycμ1 (red) in control RH and DDμ1 parasites incubated with shield-1 for 24 hours. Bar: 2μm. On the right, a zoom of the Golgi region indicated by a white frame in the merge image is shown. **G-** The histogram indicates the percentage of co-localization between the proROP4 signal and the ROP2-4 signal after image acquisition by airyscan confocal microscopy. Data are indicated as average ± SD, n = 15 vacuoles. **H-** WB analysis of ROP4 protein proteolytic processing in control RH and DDμ1 parasites incubated with shield-1 (+S) for 24 hours. No defect was found as the immature proROP4 and mature ROP4 proteins were detected at similar amounts in both parasites lines. Actin (ACT1) was used as a loading control and the detection of the cMycμ1 protein was used as a control for the shield-1 induction. **I-** Invasion assay. Histogram depicting the percentage of invaded parasites after 45 min incubation with host cells of mechanically egressed parasites for both the parental strain and DDμ1 parasites induced with shield-1 (+S) for 16 hours. Mean values of three independent assays are shown ± SEM, **p<0.01 (Student's t-test).

<https://doi.org/10.1371/journal.ppat.1006331.g006>

cMycμ1 (Fig 7B). Therefore, we next examined in details the relationship between the Rab5A-positive compartment and the ROP maturation process using SIM microscopy. In control parasites, at the onset of ROP protein neo-synthesis during parasite division, preROP compartments seemed to emerge from the TGN as already formed large vesicular compartments co-distributing with Rab5A-positive vesicles or displaying a faint Rab5A signal at their limiting membrane (Fig 7C, upper panel). More distant preROP compartments from the TGN, likely en route to the apical pole of the parasite, were detected as Rab5A-negative compartments (Fig 7D, upper panel, arrows) suggesting that the maturation process involves a transient step of transport through the Rab5A-positive ELC. In DDμ1 induced parasites, proROP4 proteins were found retained in large vesicular compartments emerging from the Golgi/ELC, which displayed a strong Rab5A signal at their limiting membrane (Fig 7C, lower panel). We calculated that 63.6 ± 18.2% of proROP4-positive preROP compartments were also positive for Rab5A compared to only 21.9 ± 13.0% in control parasites (Fig 7E). Notably, the formation of large and dispersed Rab5A-positive compartments empty of proROP4 proteins could also be observed in DDμ1 induced parasites (Fig 7D, lower panel) and not in control cells. This observation suggests a more general process of AP1-dependent regulation of Rab5A-positive endosomal membrane dynamics apart from its contribution in the process of preROP compartment maturation.

Collectively, our data suggest that the maturing preROP compartments require a TgAP1-dependent activity to further evolve towards the mature apically anchored club-shaped organelles. Importantly, this TgAP1 activity seems to be tightly connected to the activity of the Rab5A-positive ELC. Therefore, the study of the less severe phenotype associated with over-expression of the APμ1 subunit allowed us to identify an additional step of TgAP1-mediated regulation of the rhoptry maturation process at the level of the endosome-like compartment.

Altogether, our data indicate that TgAP1 acts at different trafficking steps during secretory organelle biogenesis. First, TgAP1 regulates the anterograde transport of the studied microsome complexes from the TGN. Second, our data suggest that TgAP1 regulates rhoptry formation by acting on both, at the level of immature ROP protein exit from the TGN as well as at the ELC level to ensure the rhoptry maturation process into apically anchored, club-shaped organelles.

Table 3. List of proteins identified by mass spectrometry following the IP of cMycμ1 protein in cMycμ1-overexpressing parasites induced with shield-1 for 24 hours. The detailed list is included in S8 Fig. The parental strain RH was used as a control for non-specific binding to the antibody-coated beads (no peptides corresponding to the indicated proteins were found in the control IP). The table indicates the number of unique peptides and spectra for each identified protein.

Protein name	accession numbers	molecular weight (Da)	Number of peptides	Number of spectra
Gamma 1 adaptin	TGGT1_313670	107 036	49	189
beta adaptin	TGGT1_240870	101 920	51	160
Sigma1 adaptin	TGGT1_270370	19 679	9	63
ENTH domain-containing protein	TGGT1_214180	65 903	4	4

<https://doi.org/10.1371/journal.ppat.1006331.t003>

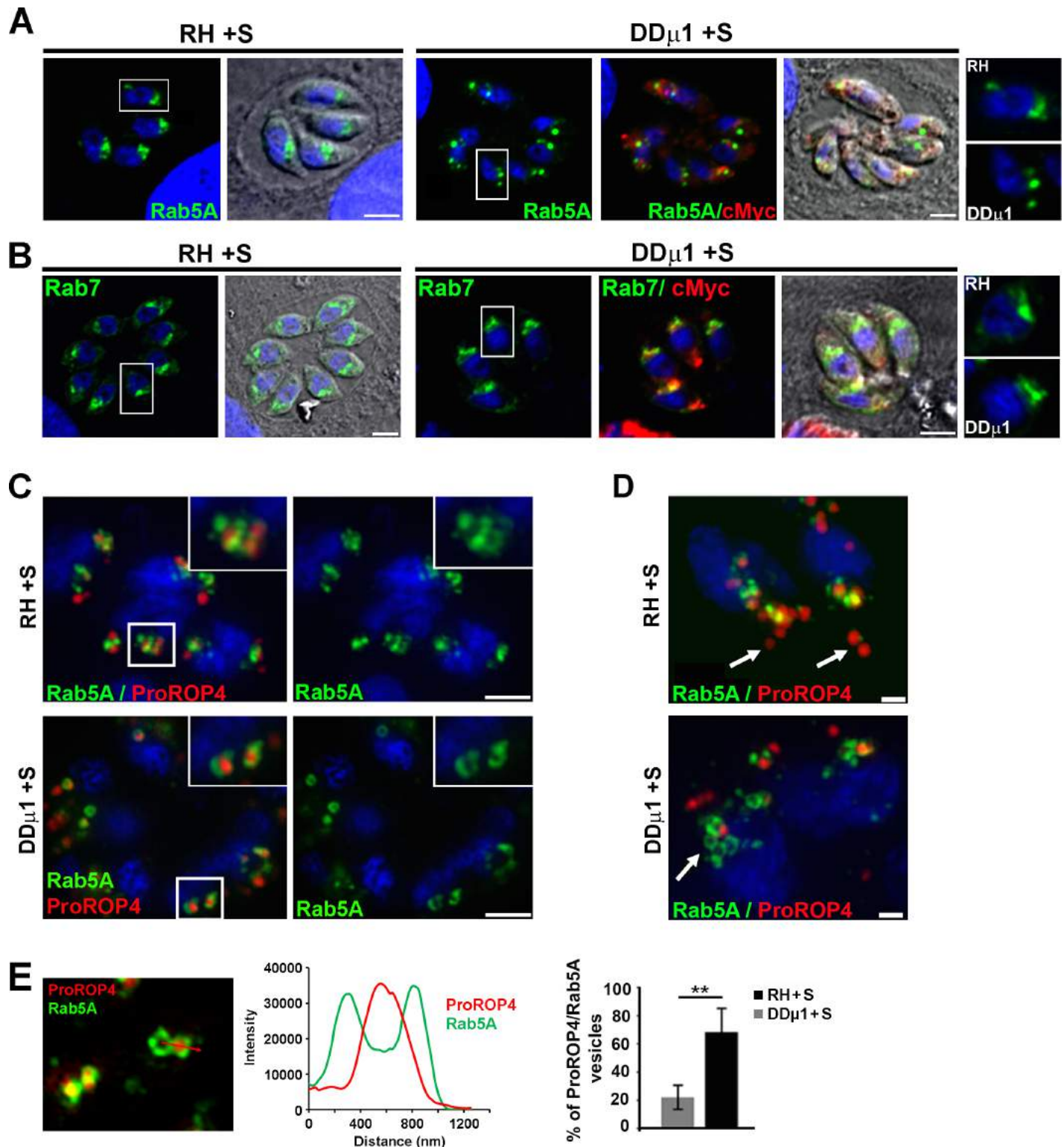


Fig 7. Over-expression of AP1 perturbed the Rab5A compartment morphology. **A, B-** Confocal images showing the localization of Rab5A (A, green), Rab7 (B, green) and cMyc1 (A, B, red) in the parental strain RH and DD μ 1 parasites treated with shield-1 (+S) for 16 hours. A zoom of the Golgi area indicated by the white frame in each image is shown as an inset on the right. Bar: 2 μ m. **C-** SIM images showing the localization of proROP4 (red) and Rab5A-HA (green) proteins in control RH parental strain (upper panel) and cMyc1 over-expressing parasites (lower panel) treated with shield-1 (+S) for 16 hours. Bars: 2 μ m. **D-** SIM images showing the co-distribution of Rab5A-positive vesicles surrounding proROP4 vesicular compartments in RH parental

strain (upper panel), while TGN-distant preROP compartments were negative for Rab5A staining (arrows). Induced DD μ 1 parasites exhibited Rab5A-positive enlarged vesicular compartments (lower panel, arrow) empty of proROP4 proteins (red). Bars: 500 nm. **E-** Left: SIM image of DD μ 1 parasites induced with shield-1 showing proROP4 proteins (red) contained in vesicles with a strong Rab5A (green) signal at their limiting membrane illustrated by the intensity profile of each signal (graph). Right: Histogram depicting the percentage of proROP4-positive vacuoles showing a Rab5A signal at their limiting membrane in RH and DD μ 1 parasites treated with shield-1 (+S). Data are presented as average \pm SD (n = 50 parasites), ***p<0.001 (Student's t-test).

<https://doi.org/10.1371/journal.ppat.1006331.g007>

TgAP1 regulates cell division

When observed by IFA, both the induced DD μ 1 and AP μ 1-KO parasites display a clear growth defect. Quantification indicated a delay in parasite growth starting from stage 4 to 8 and 8 to 16 parasites per vacuole in AP μ 1-KO and DD μ 1 parasites, respectively (Fig 8A and 8C). In particular, after 48 hours, the AP μ 1-KO parasites appeared to have stopped dividing compared to neighboring YFP-negative vacuoles (Fig 8B). The growth defective parasites seemed to be tethered by lateral contact sites or even fused together. Parasites that have broken out were also often observed (Fig 8B, arrows). Very similar but milder morphological defects were detected in DD μ 1 parasites (Fig 8D). After the stage with 8 parasites per vacuole, the parasites were no longer organized in a rosette-like structure and also appeared “tethered” by their lateral sides or basal pole (Fig 8D, i, arrows and insets). Importantly, the over-expressed cMy μ 1 protein was found enriched at sites of lateral contacts between parasites (Fig 8D, ii, region 1, insets) and at their basal pole (Fig 8D, ii, region 2, insets).

These observations suggest that the TgAP1-defective parasites suffer from a defect of division at the late stages of cytokinesis when daughter cells segregate. To verify this hypothesis, we examined in detail the different steps of the division process in these parasites. In AP μ 1-KO parasites, we found that centromeres replicate normally (Fig 8E, middle panel). New daughter buds positive for the inner membrane complex (IMC) marker IMC3 were also assembled (Fig 8E, middle and lower panel) and the AP μ 1-KO parasites displayed a similar percentage of IMC3-positive budding daughter cells compared to control parasites (Fig 8F). Similar observations were obtained in DD μ 1 parasites (S9 Fig). Interestingly, while labeling for the centromeric protein chromo1, we noticed that AP μ 1-KO parasites, which had almost completed their budding process, displayed a chromo1 localization at the basal pole in a region that seemed to connect mother parasites (Fig 8E, lower panel, arrows and insets). The dynamic nature of the chromo1 protein localization during the different stages of the cell cycle has been previously described [50]. We confirmed this localization by IFA in control parasites (Fig 8E, upper panel) and found that the protein also transiently localized at the basal pole of the mother parasite while daughter parasites terminated the budding process (Fig 8E, upper panel, image 3, arrow). Interestingly, this chromo1 dynamic localization is somehow reminiscent of the Chromosomal Passenger Complex (CPC) in higher eukaryotes, which translocates from the spindle pole in M phase to the mid-body at the end of cytokinesis [51]. This observation again suggested that the division defect in AP μ 1-KO parasites likely occurs at the late stage of parasite segregation. To investigate further this hypothesis, we examined the localization of the basal complex protein MORN1, which forms a contractile ring required to segregate nascent daughter cells [52]. We transiently expressed MORN1-cherry in AP μ 1-KO parasites and monitored its localization both in fixed cells and by live imaging (Fig 8G and 8H and S3 Movie). We were able to visualize the formation of MORN1-cherry positive contractile rings at the basal pole of nascent daughter parasites, suggesting that basal complex assembly is not impaired in AP μ 1-depleted parasites (Fig 8H and S3 Movie). However, in opposite to control parasites where MORN1-positive basal complexes were clearly seen separated one from another, in AP μ 1-KO parasites, the mother parasites appeared to remain connected by their MORN1-positive basal pole (Fig 8G and 8H, insets: region 2). When examined more carefully,

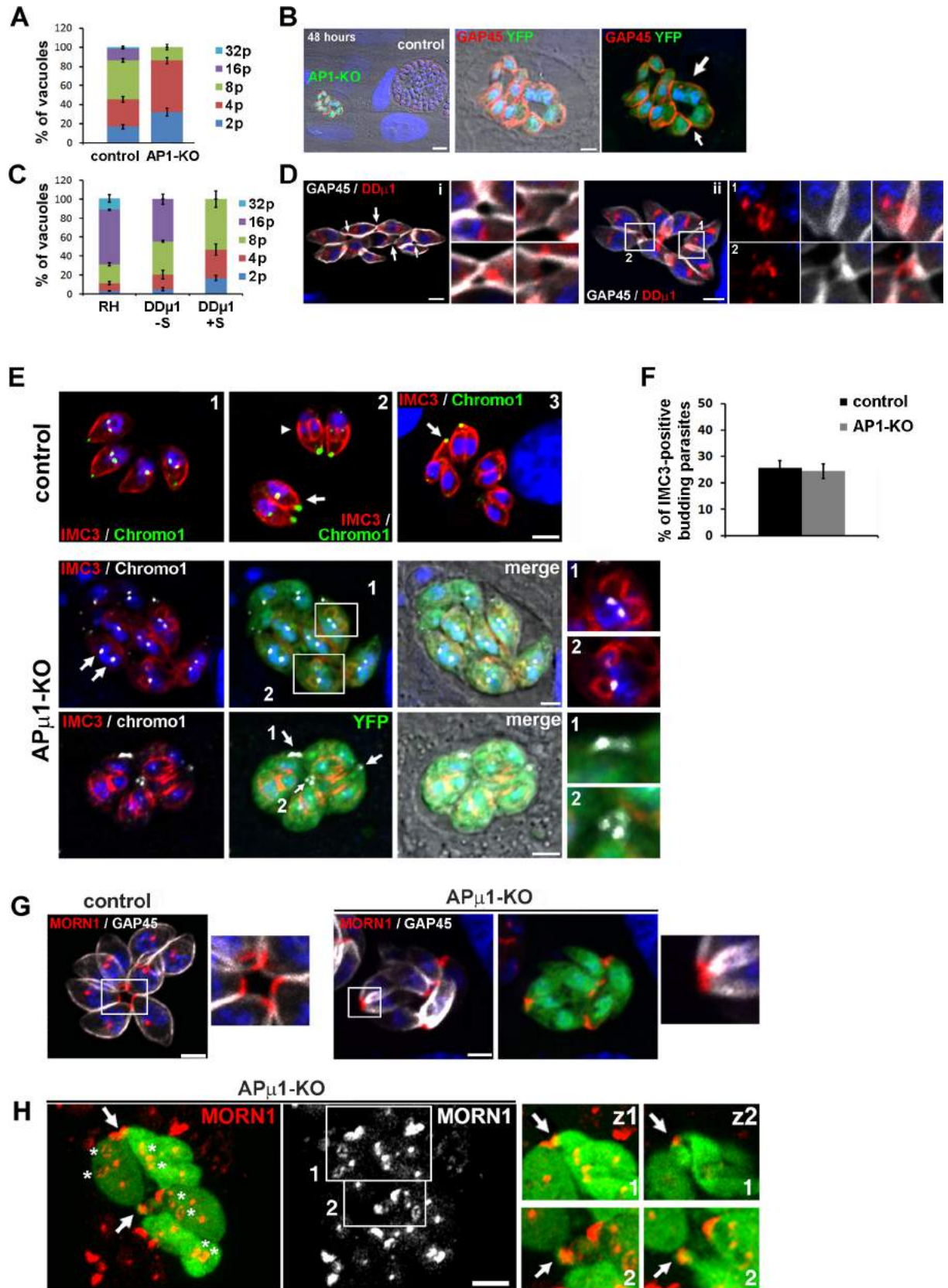


Fig 8. *TgAP1* regulates parasite growth. **A, C:** Intracellular growth assay performed in AP μ 1-KO parasites after rapamycin treatment (A) and DD μ 1 parasites after \pm shield-1 induction (C) at 24 hours post-invasion revealed defects in parasite replication. The histograms depict the percentage of vacuoles containing 2, 4, 8, 16 or 32 parasites. Mean values of three independent assays are shown \pm SEM. **B, D:** Confocal images showing the disorganized appearance of dividing AP μ 1-KO (B) and DD μ 1 parasites (D). **B-** After 48 hours of growth, the AP μ 1-KO parasites (YFP-positive, green) have stopped to grow and parasites seem to display a defect in segregation. Note the apparent rupture of the cortex revealed by the GAP45 staining (red) (arrow). Nuclei are shown by DNA staining (blue). **D-** Induced DD μ 1 parasites were labeled for GAP45 (white) and overexpressed cMyc μ 1 proteins (red). Left (i): Mis-organised vacuole showing tethered parasites with an apparent defect in lateral segregation revealed by the GAP45 staining (insets: zoom of the different areas indicated by arrows in the main image). Right (ii): Confocal images showing the enrichment of the overexpressed cMyc μ 1 protein at sites, where parasites have remained tethered to each other (insets: region 1: lateral sides, region 2: basal pole). **E-** Upper panel: Confocal microscopy images in control parasites showing the duplication of the centromeres (centromeric protein chromo1, green, image 1) and the formation of daughter buds (IMC3 protein, red, images 2, arrowhead). Note that chromo1 transiently accumulates at the basal pole of parasites at the very end of the daughter cell budding process (image 3, arrow). Lower two panels: Confocal microscopy images of AP μ 1-KO parasites (YFP-positive parasites) showing the duplication of the centromeres (chromo1, white arrows) and the formation of daughter buds (IMC3, red, middle and lower panels). Note the accumulation of chromo1 at sites connecting mother parasites, while daughter cells complete bud formation (lower panel, arrows). Bars: 2 μ m. **F-** Histogram depicting the percentage of vacuoles containing budding daughter cells (IMC3 staining) in control and AP μ 1-KO parasites. Mean values of three independent assays are shown \pm SEM. **G-** Airyscan confocal microscopy images showing the localization of MORN1-cherry (red) in control (YFP-negative, left) and AP μ 1-KO parasites (YFP-positive, right) parasites. The parasite contours were delineated by staining the IMC marker GAP45 (white). The insets show a zoom of the region indicated by a frame in the main image. **H-** Airyscan confocal images showing AP μ 1-KO parasites (YFP-positive, green) expressing the MORN1-cherry protein (red). MORN1-positive daughter rings were normally assembled (asterisks) but mother parasites seemed to be attached by their basal pole (arrows), which appeared as deformed elongated membranous structures (insets: two confocal planes z1 and z2 of the regions 1 and 2 indicated with a white frame in the main image, arrows). Bars: 2 μ m.

<https://doi.org/10.1371/journal.ppat.1006331.g008>

we observed that the basal complex constriction seemed prolonged by unusual elongated and wider membranous structures (Fig 8H, insets: region 1). Interestingly, we were also able to visualize similar deformed elongations of the parasite basal pole after performing Correlative Light Electron microscopy (CLEM) in AP μ 1-KO YFP-positive parasites (Fig 9 and S10A Fig). In control parasites, the residual body appeared as a well-organized structure, in which each parasite basal pole displayed a discrete constricted region connecting the parasites to a central mass of membrane, previously described as the remnant of the mother cell formed at the end of cytokinesis (Fig 9A–9D). In AP μ 1-KO parasites, the basal pole of each parasite appears elongated, wider and connected to the residual body in a disorganised manner (Fig 9E–9H). In addition, while control parasites displayed a typical rosette-like organization, AP μ 1-KO parasites exhibited a general deformed morphology that could indicate a defect in their cortical integrity (Fig 9E). In more drastically affected vacuoles, parasites appeared to have broken out or to remain connected by their lateral sides (S10B Fig, images e and f). In the latter case, the pellicles from neighbouring daughter cells appeared to not have been fully segregated one from another, resembling the parasite division defects observed in Rab11A defective parasites [53]. Supporting the hypothesis that AP μ 1-KO parasites suffer from a segregation blockage, we noticed that when egress was artificially triggered in AP μ 1-KO parasites by incubation with the calcium ionophore A23187, most of the monitored vacuoles initiated the process (Fig 5E) demonstrated by the visualization of some parasites escaping the vacuole, however, the other parasites remained in the vacuole and appeared attached together, unable to escape (S4 Movie). Finally, in agreement with our IFA observations (Fig 5), AP μ 1-KO parasites did not possess apically localized mature rhoptries and only apical micronemes were detected (S10B Fig, images c and d).

We also performed transmission electron microscopy (TEM) in induced DD μ 1 parasites. We found that the parasite morphology was similarly affected, together with parasite organization within the vacuole (Fig 10G–10I), in opposite to control parasites (Fig 10A and 10B), which were organized in a rosette-like structure. Similar to AP μ 1-KO parasites, the basal pole of the DD μ 1 parasites displayed a pronounced change in their morphology (Fig 10F, 10G and 10I, arrows). In contrast to control cells (Fig 10C), our images suggest that DD μ 1 induced

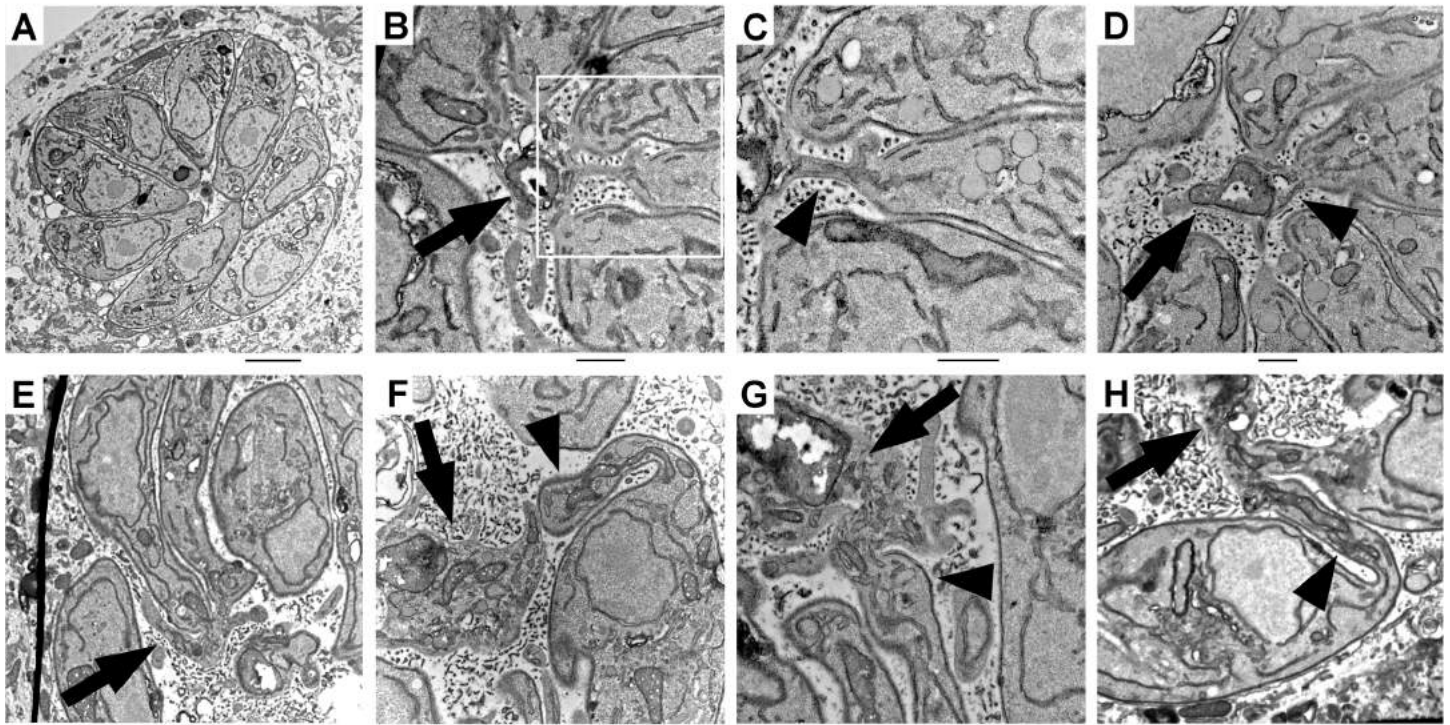


Fig 9. Correlative Light Electron Microscopy (CLEM) images illustrating the organization of the basal pole (arrow heads) and the residual body (arrows) in control and AP1-KO parasites. A-D: Non-YFP control vacuoles (C: zoom of the region indicated by a white frame in B). Note the typical organization in a rosette-like structure and the correct morphology of the parasites. **E-H:** YFP-positive AP1-KO parasites (detected in the region 3T shown in S10 Fig). Bars: 500 nm.

<https://doi.org/10.1371/journal.ppat.1006331.g009>

parasites were unable to properly constrict their basal extremity, which appeared deformed with wide elongations (Fig 10F and 10G). In addition, we also observed parasites that were attached by discrete lateral contact sites (Fig 10I, arrow). Furthermore, in agreement to our IFA images (Fig 4 and Fig 5), in DD μ 1 induced parasites, apically anchored mature rhoptries were not detected or present as dispersed compartments throughout the cell (Fig 10D and 10E), while apical and lateral micronemes were still visualized (Fig 10D). The parasites also accumulated large lucent vesicles of unknown nature in their cytoplasm (Fig 10D, 10G and 10H). We could detect membranous or vesicle-like structures within their lumen and some of our images suggest that they could be formed by internal budding of the limiting membrane (S11 Fig), as previously observed for the VAC compartment [54]. However, presently, we cannot conclude whether these vesicles represent dispersed and fragmented VAC-related compartments or enlarged Rab5-positive endosomal structures that we detected by SIM microscopy (Fig 7).

Collectively, our data suggest that TgAP1 is involved in parasite division by regulating the very late stages of cytokinesis after the budding process has been completed. Though we did not directly demonstrate it, our data converge towards the hypothesis that a TgAP1-dependent delivery of vesicles at the plasma membrane, directly from the Golgi, or indirectly via a recycling activity of the mother cell plasma membrane from the residual body, is required to complete daughter cell segregation. By this activity, apart from its role in ROP and MIC protein transport, TgAP1 could deliver lipids as well as important regulatory factors, such as regulators of cytoskeleton dynamics. These hypotheses are supported by our observations that the over-expressed cMyc μ 1 protein was found accumulated at the basal pole of dividing parasites and

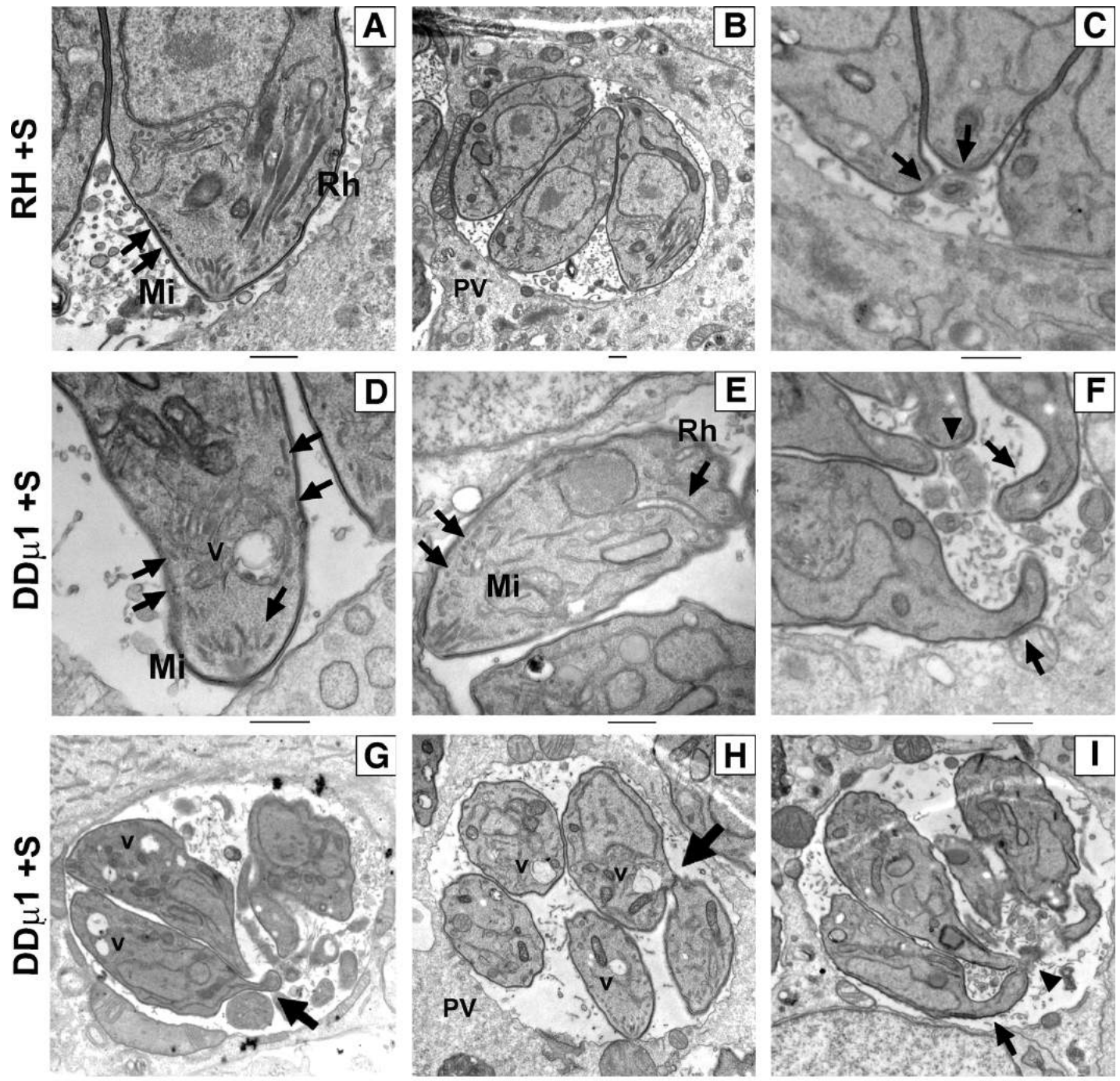


Fig 10. Transmission electron microscopy images showing the formation of mature rophtries (Rh) and micronemes (Mi) anchored at the apical pole in control parasites (A) and the normal distribution of parasites in rosette-like structures (B). Bars: 500nm. (C)- Zoom of the posterior end of the control parasites showing the tight constriction of the basal pole with a thin continuity to the residual body (arrowhead). In DD μ 1 parasites induced with shield-1 (+S) for 24 hours (D-I), apically positioned rophtries could not be detected in contrast to micronemes (D) or they were found dispersed in the cytoplasm (E, arrow). Numerous giant lucent vesicles (V) were also observed (D, G, H). In addition, the parasites were found disorganised within the vacuole with a distorted morphology (G-I) particularly at the basal pole (F, G and I, arrow), which appeared deformed and elongated (arrows in F) despite the detection of the residual body (F and I, arrowhead). In addition, some parasites seemed to remain attached by discrete lateral contact sites (H, arrow). Bars: 500nm. Mi: micronemes, Rh: rophtries, PV: parasitophorous vacuole, V: vesicles.

<https://doi.org/10.1371/journal.ppat.1006331.g010>

lateral contact sites (Figs 6B and 8D) and that endogenous *TgAP1* was also detected in numerous peripheral vesicles at the parasite cortex in AP μ 1-HA KI parasites (Fig 1D). Of note, such as illustrated in Fig 1B, we also noticed in some vacuoles a very discrete localization of endogenous AP μ 1-HA at a point connecting the basal pole of the parasites.

Discussion

Role of *TgAP1* in ROP and MIC protein trafficking

Previous studies have shown that MIC proteins navigate through the secretory pathway as complexes presumably assembled at the TGN level, which include one transmembrane escorter and one or two soluble partners. For example, MIC6 escorts MIC1/MIC4 and MIC8 associates with MIC3. All the members of the complex possess sorting signals required to address the proteins to the mature organelles, in particular the prodomain of the soluble partner and the cytoplasmic domain of the transmembrane partner. After depletion of the MIC6-CD containing sorting signals, both MIC1 and MIC4 were retained together with MIC6 Δ CD in the ER/Golgi [17]. The prodomain of MIC3 was shown to be essential for complex transport through the secretory pathway [14], however, deletion of MIC8 did not impact the targeting of MIC3 to micronemes, suggesting distinct regulatory trafficking mechanisms [16]. In AP μ 1-KO parasites, soluble MIC3 was re-directed towards the vacuolar space, whereas the transmembrane MIC8 protein was mainly retained in the TGN. Interestingly, we also found that similar to MIC3, the soluble MIC4 protein was re-routed to the vacuolar space; however the transmembrane MIC6 protein was not retained in the Golgi but localized to a sub-population of apical micronemes. Therefore, this data suggests that these two MIC complexes exhibit different trafficking mechanisms in relation to *TgAP1* function, MIC8 being the only transmembrane MIC protein that we found retained in the Golgi upon AP μ 1 ablation. We also observed that MIC2 and AMA1 were localized in apical micronemes while the typical lateral staining was weakly detected. However, in opposite to MIC4 and MIC3, the soluble M2AP was not re-routed to the vacuolar space but accumulated together with MIC2 in apical micronemes. This result shows that despite the lack of lateral MIC2/M2AP containing micronemes, both proteins of the complex likely exit the Golgi and are targeted to apical micronemes. Presently, we cannot conclude whether this loss is due to an indirect effect on parasite cortical integrity, which could impair lateral anchoring of these organelles but not anchoring to the conoid region, or caused by a direct effect on protein trafficking. In addition, the MIC1/MIC4/MIC6 complex was dissociated in absence of *TgAP1* similarly to the MIC3/MIC8 complex, supporting the hypothesis that *TgAP1* might play an important role in complex stabilization at the TGN before their export. Of note, SORTLR has been shown to interact with the soluble proteins MIC1, MIC4 and MIC5 and to be involved in the trafficking of the corresponding MIC1/4/6 complex and the MIC5 protein, while interaction with transmembrane MIC proteins has not been investigated. However, the MIC3/MIC8 and MIC2/M2AP complexes were also mis-targeted. Therefore, further studies are now required to understand better at the molecular level, whether *TgAP1* interacts with both components of the sorted MIC complex, which could include SORTLR loaded with the soluble MIC as well as the associated transmembrane MIC protein (see proposed model in Fig 11) by their respective sorting signals. To address this question, it will be interesting to investigate the role of the dileucine motif present in the cytoplasmic tail of SORTLR in the differential sorting of soluble versus transmembrane MIC proteins at the TGN.

Concerning the role of *TgAP1* in ROP protein trafficking, our data indicate that immature proROP proteins were re-directed towards the vacuolar space and the basal body in AP μ 1-KO parasites, indicating an *TgAP1*-dependent transport in the anterograde secretory pathway

before their proteolytic processing occurs, therefore most likely at the TGN level. In contrast, perturbing *TgAP1* functions by an inducible over-expression of the $\mu 1$ subunit led to less severe effects. MIC proteins were not mis-sorted but maturing rhoptries were retained in Rab5A-positive compartments and the formation of apically anchored club-shaped rhoptries was impaired. Of note, a defect in both, the anterograde and retrograde transport of proteins from the Rab5A-positive endosomes would lead to an accumulation of membrane and disorganization of these compartments. Presently, we cannot exclude a role of *TgAP1* in either of these pathways (Fig 11). Indeed, in mammalian cells, AP1 also regulates the retrograde transport of proteins from the early/sorting endosomal compartment to the TGN, while GGA (Golgi-localized, γ ear-containing, ADP-ribosylation factor-binding) is involved in anterograde transport from the TGN to the endosomes [27]. GGA molecules are not encoded in the *T. gondii* genome [11]. Thus, one can envision a similar role for *TgAP1* in the retrograde transport of molecules from the Rab5A-positive ELC to the TGN, in particular during rhoptry biogenesis. Indeed, rhoptry biogenesis could follow similar mechanisms than the ones involved in secretory granule maturation. In specialized secretory cells, immature secretory granules emerge from the TGN as preformed large vesicular compartments [36], similar to what we observed for the PreRhop compartments in *T. gondii*. Mature granule formation requires an AP1-dependent retrieving of SNAREs and proteases from the immature granule in order to form granules competent to be released upon an external stimulus. Therefore, it is possible that similarly, the defect in the preRhop maturation process that we observed is the consequence of an impairment of *TgAP1*-dependent retrieving of membrane factors, such as SNAREs. This mechanism would be crucial to pursue the rhoptry maturation process, in

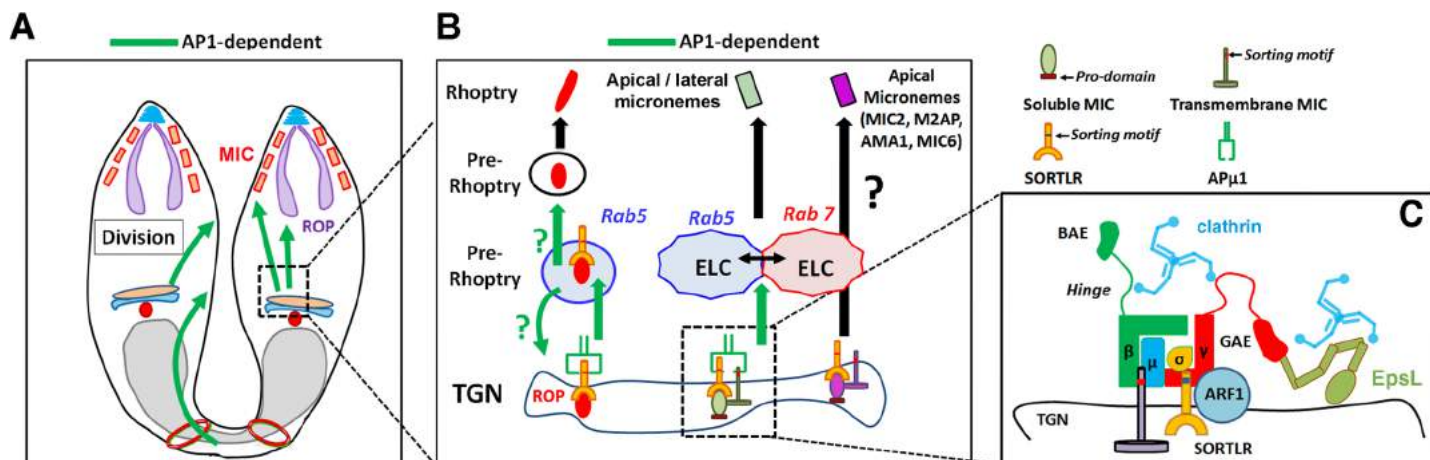


Fig 11. Model summarizing the different functions of AP1 in *T. gondii*. **A**— Our data indicate that *TgAP1* is involved in the sorting of MIC proteins from the TGN and rhoptry biogenesis, as well as, participates to daughter cell segregation. This latter activity might be regulated by a *TgAP1*-dependent recycling activity of the mother plasma membrane from the residual body or a direct transport of vesicles from the Golgi to the nascent daughter pellicles. **B**— *TgAP1* regulates the sorting and transport of all the different studied MIC protein complexes from the TGN (green arrow), including MIC3/8, MIC1/4/6 and M2AP/ MIC2, resulting in the loss of lateral micronemes containing these proteins. However, a subpopulation of apical micronemes, containing the proteins MIC2/ M2AP, AMA1 and MIC6, were still detected upon AP1μ1 ablation. At the molecular level, one can envision that *TgAP1* recognizes via its subunits γ - σ , the dileucine motif present in the cytoplasmic tail of SORTLR, which has loaded the soluble MIC partner (such as MIC3). *TgAP1* could simultaneously bind to the tyrosine motif of the transmembrane MIC partner (such as MIC8) via its subunits β - μ , thereby participating to the complex stabilization and transport into clathrin-coated vesicles from the TGN. These putative sorting mechanisms have to be confirmed. Finally, we found that *TgAP1* regulates ROP protein transport from the TGN to the Rab5A-positive ELC and also the subsequent steps of the rhoptry maturation process. *TgAP1* could regulate the latter activity either, by stimulating ROP protein exit from the ELC or by retrieving material from the preRhop compartments in a Rab5A-dependent manner to ensure the following steps of maturation into club-shaped apically anchored organelles (green arrows). The Rab7-positive ELC also likely participates in ROP and MIC trafficking, though a specific functional relationship was found between *TgAP1* and the Rab5A-positive ELC. **C**— Our data also indicate that the AP1 complex in *T. gondii* functions as a conserved heterotetrameric complex composed of the $\mu 1$, $\sigma 1$, γ and β subunits and interacts with ARF1 and clathrin. We also found that the ear appendage domain of the γ subunit associates with the unique ENTH-domain containing protein *TgEpsL*.

<https://doi.org/10.1371/journal.ppat.1006331.g011>

particular the further steps of transport, apical anchoring and final remodeling of the preROP compartments into club-shaped organelles. This hypothesis would be in line with data obtained in other eukaryotic systems showing that AP1 triggers clathrin-dependent vesicular budding from Rab5- and Rab4-positive early/sorting endosomal compartments [55] [28] [29]. We also observed that perturbing *TgAP1* function induces a more general defect in the Rab5A-positive compartment, distinct from immature pre-rhoptries, suggesting that a specific functional relationship might exist between *TgAP1* and the Rab5A-positive compartment in *T. gondii*. At the molecular level, AP1 and GGA1 proteins have been demonstrated to regulate Rab5 membrane dynamics by binding directly to the Rab5 effector Rabaptin-5 in mammalian cells [56] [57]. Notably, over-expression of Rabaptin-5 shifts the localization of GGA1- and TGN-associated cargos into enlarged Rab5 endosomes [56] [57] [58]. Therefore, further studies are needed to explore a putative role of *TgAP1* in the regulation of the Rab5 membrane dynamics in *T. gondii* and to identify the involved molecular mechanisms.

Moreover, the fact that a unique ENTH-domain containing protein is expressed in *T. gondii* raised the question of the role of clathrin-mediated endocytosis at the plasma membrane, a question that is still a matter of debate. In particular, clathrin was found mainly localized at the TGN and in cytoplasmic vesicles and perturbing its function led to defects in Golgi duplication and ROP and MIC biogenesis [59]. However, it was recently shown that parasites depleted for the cathepsin CPL internalized GFP proteins by a still unknown mechanism [60]. The role of clathrin- or *TgAP2*-mediated endocytosis in this process was not investigated. Here, we found that the unique *T. gondii* epsin-like protein mostly co-localizes with *TgAP1* at the TGN and in cytoplasmic vesicles and that *TgEpsL* interacts with the *TgAP1* complex but not the *TgAP2* complex. These findings argue against a clathrin- and epsin-mediated mechanism for protein internalization at the parasite plasma membrane, but further experiments are required to confirm this hypothesis. In particular, we are currently investigating *TgEpsL* function by the generation of inducible KO parasites. Due to the unique organisation of the parasite cortex, that comprises three lipid bilayers (the plasma membrane and the inner membrane complex), it is possible that *T. gondii* uses alternative specific pathways for the internalization of macromolecules compared to mammalian cells.

TgAP1 regulates parasite growth

Our data strongly suggest that the cell division defect we found upon AP μ 1 ablation is not linked to rhoptry and microneme biogenesis but rather to a *TgAP1*-mediated vesicular transport at the level of the basal pole and lateral sides of segregating daughter cells. Indeed, we could already observe at the stage of 4 parasites per vacuole, the ROP and MIC protein trafficking defects before the cell division process was drastically affected (Fig 4 and Fig 5). In AP μ 1-KO parasites, the budding process of nascent daughter cells was not perturbed as well as the formation of the MORN1-positive contractile rings. However, we observed both by IFA (Fig 8H) and electron microscopy (Fig 9), that the basal pole of the parasites is elongated and deformed, suggesting a later defect after the contractile rings of the basal complex have reached the posterior end of the parasite. In higher eukaryotes, mid-body abscission, which takes place last at the end of cytokinesis, is a complex process timely regulated by the sequential recruitment of different factors such as members of the ESCRT protein family, the Chromosomal Passenger Complex, Rab GTPases and kinase/phosphatases [51]. In *T. gondii*, after daughter cell budding has been completed, parasites remained attached within the vacuole by their basal pole via a highly organized structure called the residual body [61]. The role of the residual body is not clear. Apart from receiving the remnant material from the mother parasite, it probably plays an important role in the maintenance of synchronous division cycles within the

vacuole. A recent study has described the well organized and regular structure of the residual body, connecting each parasite basal pole via a thin membranous connection, which follows the constricted basal complex region [61]. This organization seems to be perturbed in *TgAP1*-defective parasites. Thus, two scenarios could be proposed to explain the division defects observed in AP μ 1-KO parasites. First, a *TgAP1*-dependent delivery of vesicles directly from the TGN/ELC to the basal pole could be required to orchestrate the final step of cytokinesis and in particular, to spatially organise the specific attachment structure that remains between parasites. These vesicles could deliver crucial regulatory factors, such as regulators of cytoskeleton components required to modulate the contractile force involved in parasite attachment at the residual body. Second, *TgAP1* could be involved in a recycling activity of the mother plasma membrane from the residual body to terminate daughter cell segregation, such as previously observed for the IMC [62] or contribute to a direct transport of *de novo* synthesized lipids from the TGN to the plasma membrane. In agreement with the latter hypothesis, a recent study has demonstrated that parasite depleted for the FAS II enzyme, which is responsible for fatty acid biosynthesis at the apicoplast, displayed drastic division defects and were unable to segregate after the budding process has completed, forming a mass of tethered cells [63]. This study revealed the requirement for *de novo* lipid synthesis and therefore, we believe, of regulated trafficking pathways for the delivery of these lipids, to complete daughter cell segregation at the end of cytokinesis. Importantly, a similar role for the AP1-dependent delivery of Golgi-derived vesicles at the cleavage furrow of dividing cells, has been previously described in different organisms, such as *S. pombe* [41], *D. discoideum* [42], *C. elegans* embryo and plants [32], suggesting a conserved function for the complex AP1 in cell division among various eukaryotic organisms including *T. gondii*. Therefore, a major perspective of our work will be to dissect the mechanisms regulating a possible *TgAP1*-mediated transport of vesicles from the parasite basal pole and/or the TGN/ELC to the cortical area and to identify cargos that could be transported via this pathway and possible regulatory factors. In particular, it will be interesting to investigate potential links with other compartments involved in constitutive secretion such as the Rab11A-positive compartment. Indeed, deregulation of Rab11A activity also results in incomplete pellicle assembly in the inner regions between daughter cells, leading to a cell separation block late in cytokinesis [53].

In conclusion, in plant cells, secretory and endocytic routes intersect at the hybrid trans-Golgi network /early endosomes, where cargos from both the anterograde and the retrograde pathways are further correctly sorted in a timely manner [44]. In *T. gondii*, we also found a tight physical and functional association between the TGN and the ELC throughout the cell cycle. BFA treatment led to the dispersion of both, the TGN and the Rab5A-positive compartment. Although tightly connected, these two compartments are functionally distinct, as recently suggested by the study of the retromer functions, where depletion of *TgVps35* led to the retention of SORTLR in the endosomal-like compartment by inhibition of its retrograde transport to the TGN [26]. In plants, late endosomes, also called multi-vesicular bodies (MVBs) or pre-vacuolar compartments (PVCs) were shown to directly emerge from the hybrid TGN/early endosomal compartment as immature large vesicular compartments containing intra-luminal membranes, such as observed for the preROP compartments in *T. gondii* [64]. Furthermore, AP1 was shown to be critical for BFA-sensitive post-Golgi trafficking events from the TGN/EE to the MVB [30] [32]. Finally, fully differentiated AP μ 1-KO plant cells contained fragmented vacuoles rather than a large central vacuole as in wild type cells, suggesting an additional role of AP1 in the later step of vacuolar fusion similar to what is observed in yeast [32] [33]. Our findings suggest that AP1 in *T. gondii* may also be involved in both steps of proROP protein exit from the TGN compartment towards the ELC but also at the later step of preROP vesicular compartments maturation into apically anchored club-

shaped rhoptries. Because of the functional and structural similarities between *T. gondii* and the plant trafficking system, one can envision that many vesicular trafficking pathways and the corresponding molecular regulatory mechanisms are conserved.

Materials and methods

Ethics statement

No study on human participants, specimens or tissue samples, or vertebrate animals, embryos or tissues have been conducted.

Cloning strategies

Genomic DNA was isolated from the Type I RH Δ KU80 strain parasites using the Promega Wizard genomic DNA purification kit and used as template for PCR. The p5RT70 loxp-AP1 μ -loxp-YFP-HXGPRT plasmid was generated by a 3 step cloning. First, a 2kb fragment from the endogenous 3'UTR of TgAP1 μ (TGGT1_289770) was amplified using primers CCGGGAGCTCAAATCAACAAGGGGGGGCGAGG and GCGCGAGCTCACGGAGAAGGAACGAGGAGCAAAG and cloned into a unique SacI site of the mother vector p5RT70loxPKillerRedloxPYFP-HXGPRT [48]. As a second step the coding sequence of the gene was amplified with a HA epitope tag added at the C terminus using primers GCGCCCTAGGATGGCGGGGCGTCTGCGGTGT and GCGCAGATCTCTAAGCGTAATCTGGAACATCGTATGGTAGGAGAGTCTCAGTTGGTACTCTCCA and inserted into the plasmid using the restriction sites AvrII and BglII respectively. As a final step, a 2.5kb fragment from the 5'UTR of TgAP1 μ was amplified with the primers GCGCGGTACCCAAGTTCCCGTTTGTCCTGG and GCGCGGGCCCTCTTGGGACTGCAAGATCGACTG cloned using the sites KpnI and ApaI respectively. The DDcMyc μ 1 parasites were obtained using the ddFKBP over-expression system [49] as follows: The TgAP1 μ gene was amplified with the following primers: GCGCATGCATATGGCGGGGGCGTCTGCG and GCGCTTAATTAAGTAGGAGAGTCTCAGTTG TACTCTCCATTTTGAGTGATG and cloned into the pG12-Tub8-DD-mCherryMyc-HXGPRT vector using the restriction sites NsiI and PacI. The plasmid was then digested by AvrII and BglII to remove the DD and mCherry fragments. The DD cassette was re-introduced into the resulting vector after amplification by PCR (F: CTTTGTAGATCTAAAATGGGAGTGCAGG, R: GCGCCCTAGGTTCCGGTTTTAGAAGCTCCAC) and ligation into the AvrII and BglII sites.

Primers used to generate 3'-terminally tagged genes integrated at the endogenous locus (knock-in parasites) and produce recombinant proteins are indicated in Table 4.

Parasite culture and transfection

Toxoplasma gondii Type I RH Δ KU80 Δ HXGPRT and DiCre Δ KU80 Δ HXGPRT parasites were grown on confluent Human Foreskin Fibroblast (HFF) cells (CCD-1112Sk (ATCC, CRL-2429TM)) which were cultured in complete DMEM (gibcoLife Technologies) supplemented with 10% Fetal Bovine Serum (GibcoLife Technologies) and 1% Pen Strep (gibcoLife Technologies). To obtain the DD μ 1 parasites, 50 μ g of the pG12-Tub8-DD-cmyc-AP1 μ -HXGPRT plasmid was transfected in RH Δ KU80 Δ HXGPRT parental strain by electroporation following standard procedures. To obtain the AP1 μ -KO parasites, 50 μ g of loxp-AP1 μ HA-loxp-YFP-HXGPRT construct was transfected in the DiCre Δ Ku80 Δ HXGPRT strain parasites. Following transfection, in both cases the parasites were subjected to Mycophenolic acid/Xanthine drug selection and verified for the transfection efficiency by immunofluorescence analysis. Subsequently the non-clonal populations of parasites were subjected to cloning by serial dilution.

Table 4.

Plasmid	Primers (F: Forward; R: reverse)	Linearization enzyme
pLic EpsL-cmyc (HXGPRT)	F: TACTTCCAATCCAATTTAATGCCCTCGTTCTCTCCTTCTCAGACGTT R: TCCTCCACTTCCAATTTTAGCGAACCCCGTCGTAGCAGGAGAT	NcoI
pLic μ 1-HA (DHFR)	F: TACTTCCAATCCAATTTAATGCGGATCTTCCCTAGTTGCGGCCAGTCAC R: TCCTCCACTTCCAATTTTAGCGGAGAGTCTCAGTTGGTACTACTCTCCATTTTGGAGT	SnaBI
pLic-Rab5A-YFP (DHFR)	F: TACTTCCAATCCAATTTAATGCACTTTTGCCTCCACATGCACACC R: TCCTCCACTTCCAATTTTAGCGTGAGTGTCTCAGAAGGGAAGAACG	Eco47III
pLic σ 1-HA (DHFR)	F: TACTTCCAATCCAATTTAATGCGTGATCCACCACCTTTGTCGAGATCTTGG R: TCCTCCACTTCCAATTTTAGCGTCATGTAAGCTTGACTCCACCTTTAGTGTGCTC	EcoRV
GST- β ear	F: GGATCCGAGAAGCTCTCTGCCGACAAGGACGTTTTTCAGA R: GAATTCTCACGACCGTGCGCTCAGCC	
GST-year	F: GGATCCTTTCCGCCGATGAATGTCTTGAACGAGGACG R: GAATTCTCACGCGAGGAGTCCCGCGG	

<https://doi.org/10.1371/journal.ppat.1006331.t004>

For the AP μ 1-KO clonal parasites, integration of the transgenic construct at the endogenous locus was verified by a genotyping PCR using a forward primer (GACGCGTTTCACTTCCTCTGCTTCCTC) located upstream of the cloned 5'UTR, and a reverse primer (GTTTACGTCGCCGTCCAGCTCGAC) located on the YFP cassette. To obtain clonal knock-in parasites, 25 μ g of plasmids were linearized over-night and transfected into the RH Δ KU80 Δ HXGPRT parental strain by electroporation followed by drug selection and cloning. Transient transfections were performed in 10^*10^6 parasites with 50 μ g of the following plasmids: HA-tagged *TgRab5A* (V. Carruthers) / cMyc-tagged *TgRab7* and cMyc-tagged *TgRab5A* (M. Meissner); GalNac-YFP (D. Roos); GRASP-RFP (K. Hager), MORN1-cherry and IMC3-cherry (M.J. Gubbels) and parasites were allowed to invade HFF cells for 24 h prior analysis.

Western blot

Parasites were lysed in lysis buffer (NaCl 150mM, TrisHCl 20mM, EDTA 1mM, 1% TritonX100, protease inhibitors) and total proteins were subjected to electrophoresis in a 10% polyacrylamide gel. The proteins were transferred onto a nitrocellulose membrane (AmershamTMProtranTM 0.45 μ NC) by a standard western blot procedure. The membrane was blocked with 5% milk (non-fat milk powder dissolved in TNT buffer: 100mM Tris pH8.0, 150mM NaCl and 0.1% Tween20) and probed with primary antibodies diluted in the blocking buffer. The primary antibodies were followed by respective species specific secondary antibodies conjugated to HRP. The antibody incubations were followed by thorough washing using the TNT buffer. The membranes were visualized using ECL Western blotting substrate (Pierce).

Immunofluorescence assays (IFA)

When indicated, infected confluent HFF monolayers were incubated for 1 h with 5 μ M of Brefeldin A (Sigma-Aldrich) before fixation with 4% paraformaldehyde (PFA) in phosphate buffered saline (PBS), for 20 minutes. After quenching with 50mM NH₄Cl, the coverslips were permeabilised with 0.2% triton dissolved in 5% FBS-PBS for 30 minutes. The coverslips were then incubated with primary antibodies in 0.1% triton dissolved in 2%FBS-PBS or 0.05% Saponin for 1 h and then washed with PBS, followed by goat anti-rabbit or goat anti-mouse secondary antibodies conjugated to Alexa Fluor 488 or Alexa Fluor 594 (Molecular Probes, Invitrogen). Images were acquired using a Zeiss LSM880 confocal microscope. Antibodies used for IFA experiments are the following: rabbit anti-HA (Cell Signaling Technology), rat

anti-cMyc (Abcam), mouse anti-SAG1 (our lab), rabbit anti-GAP45, rabbit anti-MIC8, anti-MIC4, anti-MIC6 (D. Soldati-Favre), mouse anti-MIC2, rabbit anti-M2AP, mouse anti-pro-MIC3, rabbit anti-proM2AP (V. Carruthers), mouse anti-MIC3, mouse anti-ROP 2–4, mouse anti-GRA3 (J.F. Dubremetz), rat anti-SORTLR (our lab), mouse anti-chromo1 (our lab), rabbit anti-proROP4 (G.E Ward), mouse anti-IMC1 (M.J. Gubbels) and rabbit anti-GRA6 (C. Mercier). To quantify the percentage of vacuoles presenting apically positioned rhoptries and proROP4-positive vacuoles in control parasites or induced AP μ 1-KO and DD μ 1 parasites, a total of 150 vacuoles were monitored for each condition in 3 independent assays. Data are presented as mean \pm Standard Error Mean (SEM).

Structured illumination microscopy (SIM)

SIM was used to obtain high-resolution images using an ElyraPS1 microscope system (Zeiss) with a 100x oil-immersion lens (alpha Plan Apochromat 100x, NA 1.46, oil immersion) and a resolution of 120 nm along the x-y axis and 500 nm along the z-axis (PSF measured on 100 nm beads; Sampling voxel size: 0,050 μ m*0,050 μ m*0,150 μ m). Three lasers (405, 488, and 561 nm) were used for excitation. SIM images were acquired with an EMCCD camera (Andor Technology Ltd, UK) and processed with ZEN software, where exposure times varied between 100 and 150 ms. Three-dimensional images were generated using a z-step of 150 nm (total thickness \sim 5 μ m). The acquisition was done sequentially using Zeiss Filter Sets 43HE, 38HE and BP 420–480. 15 frames were acquired to reconstruct one image (5 rotations x 3 phases, with a SIM Grating period of 51 μ m for the blue channel, 42 μ m for the green channel, 34 μ m for the red channel). 100 nm beads were imaged to measure the chromatic mis-alignment of our system (fit procedure by the Zen software); this parameter enabled correcting the alignment on each acquired multi-channel stack. Image reconstructions and co-localization quantification were determined with IMARIS software (Bitplane).

Intracellular growth assay

AP μ 1-KO parasites were allowed to invade HFF monolayers for 3 h and treated with 50nM Rapamycin for 6 h. After 3 washes with warm medium, parasites were allowed to grow for additional 16 h before fixation with 4% PFA. For the DD μ 1 strain, parasites were inoculated onto HFF monolayers for 3 h and treated with or without shield-1 (1 μ M) for 16 h, before fixation with 4% PFA. In both cases, intracellular parasites were counted after staining with anti-GAP45 antibodies. The numbers of parasites per vacuole were counted for more than 200 vacuoles for each condition performed in duplicate. Data are presented as mean values of three independent assays \pm SEM.

Invasion assay

Intracellular DD μ 1 transfected parasites induced with or without 1 μ M Shield for 16 h or intracellular AP μ 1-KO parasites induced with 50 nM rapamycin for 6 hours and allowed to grow for an additional 16 hours were mechanically released from host HFF cells. Two million parasites were then allowed to adhere to host cell monolayers by centrifugation for 3 min at 1200rpm then shifted to 37°C for 45 min. Non adherent parasites were washed away with PBS followed by fixation with 4% PFA for 10min. The red-green invasion staining procedure was followed as described earlier [49]. Briefly, adherent external parasites were labeled with mouse anti-SAG1 antibodies, followed by secondary anti-mouse antibodies coupled to Alexa594. After cell permeabilisation with Triton 0.1% for 10 min, invaded intracellular parasites were detected using rabbit anti-GAP45 antibodies followed by secondary anti-rabbit antibodies coupled to Alexa488. For AP μ 1-KO parasites, YFP-positive parasites were counted for their

invasion capacity compared to non-YFP neighbouring parasites present on the same coverslips. At least, 300 parasites (for DD μ 1 and control parasites) and 150 parasites (AP μ 1-KO) were counted for each condition. Data are presented as mean values of three independent assays \pm SEM.

Egress assay

Host cells (HFFs) were seeded in 8-well chambers (Nunc® Lab-Tek® II chambered cover-glass). 5×10^4 freshly egressed parasites per well were seeded onto HFF monolayers and allowed to invade for 2 hours. Parasites were then treated with 50nM Rapamycin for six hours. Subsequently medium was changed and parasites were allowed to grow further for 24 hours. The chamber was then placed on an inverted microscope (Axio-observer, Zeiss) equipped with an incubation chamber set at 37°C, and supplied with 5% CO₂. Egress was induced with 2 μ M calcium ionophore A23187 (Sigma-Aldrich). The movies were captured using a 40X Plan apochromat NA 1.4 objective. Image acquisition was performed using AxioVision Software (Zeiss) for up to 10 minutes on each well. A total of 50 YFP-positive and 150 YFP-negative vacuoles were monitored. Data are presented as mean values of three independent assays \pm SEM.

Immunoprecipitation

For immunoprecipitation assays, a minimum of 0.6 billion parasites of AP μ 1-HA, EpsL-cMyc / pLIC-AP μ 1-HA, and DD μ 1 strains were lysed on ice for 30 min in modified RIPA buffer (50mM TrisHCl pH8.0, 2mM EDTA, 75mM NaCl, 0.65% NP40, 0.005%SDS, 0.5mM PMSF) and centrifuged at 14 000 rpm for 15 min to remove cell debris. Protein concentration was determined using the BCA protein assay kit (Pierce™). 500 μ g of total lysate were immunoprecipitated by binding to 50 μ l of anti-cMyc agarose beads (Pierce™) or anti-HA agarose beads (Pierce™) overnight. After five washes of 10 min each with modified RIPA buffer, bound proteins were eluted by boiling the samples in laemmli buffer. Samples were then subjected to SDS PAGE and western blotting or gel-extracted for tryptic digestion and mass spectrometry analysis.

GST pull-down

The C-terminal ear appendage domain of the β (BAE) and γ (GAE) subunits were GST tagged by cloning into a pGEX6p3 vector (Pharmacia). Expression of GST-BAE and GST-GAE in BL21 competent cells was achieved by induction with 1mM IPTG at 37°C for 4 h. Bacteria lysates expressing GST-BAE, GST-GAE, and GST (control) were bound to 100 μ l of Protino Glutathione agarose 4B beads (Machery Nagel) in GST-lysis/binding buffer (Tris HCl (pH 7.6) 50mM, EDTA 1mM, EGTA 1mM, 2-mercaptoethanol 10mM, NaCl 150mM, TritonX-100 0.5%, and PMSF 0.5mM) overnight at 4°C. The beads were washed 5 times with wash buffer A (Tris HCl (pH 7.6) 50mM, 2-mercaptoethanol 10mM, NaCl 500mM, Triton 0.5% and PMSF 0.5mM) and 3 times with wash buffer B (Tris HCl (pH 7.6) 20mM, NaCl 150mM, NP40 0.65%, SDS 0.005%, PMSF 0.5mM) sequentially. Beads containing 150 μ g of the recombinant proteins and the control GST protein were incubated with a lysate from 0.4 billion EpsL-cMyc / μ 1-HA intracellular parasites overnight at 4°C. Parasites were lysed using modified RIPA (TrisHCl (pH8.0) 50mM, EDTA 2mM, NaCl 75mM, NP40 0.65%, SDS 0.005%, PMSF 0.5mM). After 3 washes with the lysis buffer, the proteins bound to the beads were eluted with 1x Laemmli blue buffer by boiling. The samples were subject to western blot and mass spectrometric analyses.

Mass spectrometry proteomic analysis

After denaturation at 100°C in 5% SDS, 5% βmercaptoethanol, 1 mM EDTA, 10% glycerol, 10 mM Tris buffer pH 8 for 3 min, protein samples were fractionated on a 10% acrylamide SDS-PAGE gel. The electrophoretic migration was stopped as soon as the protein sample entered 1 cm into the separating gel. The gel was briefly stained with Coomassie Blue, and five bands, containing the whole sample, was cut. In gel digestion of gel slices was performed as previously described [65]. An UltiMate 3000 RSLCnano System (Thermo Fisher Scientific) was used for separation of the protein digests. Peptides were automatically fractionated onto a commercial C18 reversed phase column (75 μm×150 mm, 2 μm particle, PepMap100 RSLC column, Thermo Fisher Scientific, temperature 35°C). Trapping was performed during 4 min at 5 μl/min, with solvent A (98% H₂O, 2% ACN and 0.1% FA). Elution was performed using two solvents A (0,1% FA in water) and B (0,1% FA in ACN) at a flow rate of 300 nl/min. Gradient separation was 3 min at 5% B, 37 min from 5% B to 30% B, 5 min to 80% B, and maintained for 5 min. The column was equilibrated for 10 min with 5% buffer B prior to the next sample analysis. The eluted peptides from the C18 column were analyzed by Q-Exactive instruments (Thermo Fisher Scientific). The electrospray voltage was set at 1.9 kV, and the capillary temperature was set at 275°C. Full MS scans were acquired in the Orbitrap mass analyzer over m/z 300–1200 range with resolution 35,000 (m/z 200). The target value was 5.00E+05. Ten most intense peaks with charge state between 2 and 4 were fragmented in the HCD collision cell with normalized collision energy of 27%, and tandem mass spectrum was acquired in the Orbitrap mass analyzer with resolution 17,500 at m/z 200. The target value was 1.00E+05. The ion selection threshold was 5.0E+04 counts, and the maximum allowed ion accumulation times were 250 ms for full MS scans and 100 ms for tandem mass spectrum. Dynamic exclusion was set to 30 s.

Proteomic data analysis

Raw data collected during nanoLC-MS/MS analyses were processed and converted into *.mgf peak list format with Proteome Discoverer 1.4 (Thermo Fisher Scientific). MS/MS data was interpreted using search engine Mascot (version 2.4.0, Matrix Science, London, UK) installed on a local server. Searches were performed with a tolerance on mass measurement of 0.2 Da for precursor and 0.2 Da for fragment ions, against a composite targetdecoy database (50620 total entries) built with 3 strains of *Toxoplasma gondii* ToxoDB.org database (strains ME49, GT1 and VEG, release 12.0, September 2014, 25264 entries) fused with the sequences of recombinant trypsin and a list of classical contaminants (46 entries). Cysteine carbamidomethylation, methionine oxidation, protein N-terminal acetylation and cysteine propionamidation were searched as variable modifications. Up to one trypsin missed cleavage was allowed.

Correlative light electron microscopy (CLEM)

Host cells were cultured on alphanumeric gridded-glass bottom dishes (P35G-1.5-14-CGRD, MatTek Corporation, Ashland, MA, USA) until 50% confluence was reached. Parasites were allowed to invade for 2 h, washed twice with warm medium, then induced for 6 h with Rapamycin 50n M, washed thrice with PBS and allowed to grow for additional 16hrs. Cells were then fixed with 4% PFA / 0.5% glutaraldehyde in PBS over-night. YFP-positive *TgAPμ1*-KO parasites were imaged using a Zeiss LSM880 confocal microscopy and localized on the alphanumeric grid using transmitted light. After observation, cells were fixed with 2% glutaraldehyde in 0.1 M sodium cacodylate buffer over-night. After washing with water, cells were sequentially stained with 1% osmium tetroxide reduced with 1.5% potassium hexacyanoferrate

(III) for 1 hour, 1% thiocarbohydrazide for 30 minutes, 1% osmium tetroxide, 1% uranyl acetate overnight at 4°C, and finally lead aspartate for 3 h. All stains were made in water, in the dark and at room temperature unless otherwise indicated. All stains were also washed with water. After staining, cells were dehydrated in graded ethanol solutions, infiltrated with epoxy resin and cured at 60°C for 48 h. After separation of the resin from the glass, cells of interest were relocated with the imprinted-alphanumeric grid at the surface of the resin. Small blocks of resin containing the cells of interest were prepared for sectioning parallel to the resin surface. Serial sections of 80 nm thickness were set down on carbon/formvar-coated slot grids. Sections were observed with a Hitachi H7500 TEM (Elexience, France), and images were acquired with a 1 Mpixel digital camera from AMT (Elexience, France).

Transmission electron microscopy

After infection of a confluent HFF monolayer, cells containing replicating shield-1 induced DD μ 1 or control parasites were detached with a scraper, spun down and fixed with 1% glutaraldehyde in 0.1 M sodium cacodylate pH 6.8 overnight at 4°C. Cells were post-fixed with 1% osmium tetroxide and 1.5% potassium ferricyanide for 1 hour, then with 1% uranyl acetate for 45 minutes, both in distilled water at room temperature in the dark. After washing, cells were dehydrated in graded ethanol solutions then finally infiltrated with epoxy resin and cured for 48 hours at 60°C. Sections of 70–80 nm thickness on formvar-coated grids were observed with a Hitachi H7500 TEM (Elexience, France), and images were acquired with a 1 Mpixel digital camera from AMT (Elexience, France).

Statistics

Means and SEM and SD were calculated in Excel. *P*-values were calculated in Excel using the Student's *t*-test assuming equal variance, unpaired samples and using two-tailed distribution.

Supporting information

S1 Fig. A- Confocal microscopy images showing the co-localization of μ 1-HA (green) and cMyc-tagged Rab5A (red) at the duplicated Golgi during the G1/S phase of the cell cycle in μ 1-HA KI parasite. Bar: 2 μ m. **B-** μ 1-HA KI parasites (upper panel, green) or RH parasites transiently transfected with cMyc-tagged Rab5A (lower panel, green) were treated (or not) with Brefeldin A for 1 h before fixation and processing for IFA. SORTLR was used as a marker for the TGN compartment (red). The images show the dispersion of μ 1-HA and Rab5A in vesicles and aggregates also positive for SORTLR. Bars: 2 μ m.

(TIF)

S2 Fig. (Supplementary data for Table 1). Detailed list of proteins identified by mass spectrometry after immunoprecipitation of the μ 1-HA subunit using anti-HA antibodies in KI parasites expressing μ 1-HA under the endogenous promoter. The parental strain RH Δ KU80 has been used as a control for non-specific binding on the anti-HA antibody coated beads.

(XLSX)

S3 Fig. A- WB image showing the expression of the HA-tagged σ 1 subunit of TgAP1 at the expected size (20 kDa) in KI parasites and its absence in the parental RH Δ KU80 strain using anti-HA antibody. **B-** SIM images showing the localization of σ 1-HA (green) at the TGN together with SORTLR (red). The white frame indicates the region zoomed in and shown as insets in the lower panel of images. Bar: 2 μ m. **C-** Sequence alignment of the unique *T. gondii* ENTH-domain containing protein (TG GT1_214180) with Human epsin1 (*Hs*, UniProtKB: Q9Y6I3) and epsinR (CLINT1/Epsin4, UniProtKB: Q14677), with *Arabidopsis thaliana*

epsinR2 (*At*, UniProtKB: Q67YI9) and *Plasmodium falciparum* (*Pf*) unique ENTH-domain containing protein (PF3D7_1245800). A scheme illustrating the positions of the identified conserved domains in *TgEpsL*, such as the ENTH domain, the clathrin binding site (DLL/LXD) and the NPF motif (the latter being predicted to mediate the association of epsin proteins with clathrin adaptor complexes), is shown below.

(TIF)

S4 Fig. (Supplementary data for Table 2). Detailed list of proteins identified by mass spectrometry following the IP of EpsL-cMyc protein in double KI parasites expressing EpsL-cmyc and μ 1-HA proteins. The single KI parasites expressing μ 1-HA was used as a control for non-specific binding to the anti-cMyc antibody-coated beads.

(XLSX)

S5 Fig. The schemes illustrate the conserved identified domains of the AP1- γ (TGGT1_240870) and AP1- β (TGGT1_313670) subunits of the AP1 complex, including the N-terminal Adaptin domain, the clathrin binding sites present in the hinge domain and the C-terminal appendage ear (AE) domains. Sequence alignments of the GAE and BAE domains of *T. gondii* AP1 (*Tg*) with the corresponding sequences found in the AP1 complex of *Plasmodium falciparum* (*Pf*), *P. berghei* (*Pb*), *Arabidopsis thaliana* (*At*) and humans (*Hs*) are also shown. The accession numbers of analyzed proteins are indicated in the [S1 Table](#).

(TIF)

S6 Fig. A, B- Confocal images showing the localization of AMA1 (A), MIC1 (A) and SORTLR (B) proteins (in red) in control (YFP-negative) and AP μ 1-KO parasites (YFP-positive). SORTLR was not found mis-localized in AP μ 1-KO. **C-** Confocal images showing the localization of GRA3 (green) and GRA6 (red) in YFP-positive AP μ 1-KO parasites (yellow). No defect in dense granule biogenesis was observed upon AP μ 1 depletion. All bars: 2 μ m.

(TIF)

S7 Fig. A- Confocal images showing the endogenous localization of MIC2, M2AP, MIC3 and MIC8 (all in green) in micronemes of DD μ 1 parasites over-expressing the cMyc μ 1 protein (red) after shield-1 induction for 24 hours (+S). Bars: 2 μ m. **B-** WB image showing no differences in the expression and proteolytic maturation of MIC2, MIC8, MIC3 and M2AP proteins in shield-1 induced control parasites (left) and DD μ 1 parasites (right) for 24 h. Actin (ACT1) was used as loading control. **C-** Confocal images showing the co-localization of SORTLR (white) with proROP4 (green) and cMyc μ 1 (red) in control (RH) and DD μ 1 parasites incubated with Shield-1 for 24 h. A zoom of the Golgi region indicated by a white frame on the merged image is shown as inset. Bars: 2 μ m. **D-** Histogram indicating the percentage of co-localization between the proROP4 signal and the SORTLR signal after image acquisition by airyscan microscopy. Data are indicated as average \pm SD, $n \geq 15$ vacuoles (Student's t-test).

E. Confocal images displaying the localization of the endogenous σ 1-HA protein (green) in cMyc μ 1-overexpressing parasites (red) induced with shield-1 for 16 hours. The endogenous σ 1-HA subunit was found localized at the Golgi together with the cMyc μ 1 protein. Bars: 2 μ m.

(TIF)

S8 Fig. (Supplementary data for Table 3). Detailed list of proteins identified following immunoprecipitation of the over-expressed cMyc μ 1 in DD μ 1-induced parasites using anti-cMyc antibodies. The parental strain RH was used as a control for non-specific binding on the anti-cMyc antibody coated beads.

(XLSX)

S9 Fig. Confocal images showing daughter cell bud formation revealed by a tubulin staining (green, upper panel, arrows) and centromere duplication labeled with the protein chromo1 (green, lower panel, arrows) in shield-1 induced DD μ 1 parasites (cMyc μ 1-overexpressed protein in red). A zoom of the region indicated by a white frame in the merge image is shown in the inset on the right. Note the accumulation of the cMyc μ 1 protein at the basal pole of connected parasites. Bars: 2 μ m.

(TIF)

S10 Fig. A- upper panel: strategy used to perform CLEM microscopy. HFF cells were allowed to grow (50% confluent) on alphanumerical coverslips. Confocal microscopy images were taken (40X objective) to spot the YFP positive parasites corresponding to AP μ 1-KO parasites or control non-YFP vacuoles. A mosaic of 8*8 microscopy fields centred on the parasites of interest was then acquired to determine the vacuole position on the grid as illustrated for the region 3T, corresponding to the AP μ 1-KO vacuole shown in Fig 9. The YFP-positive vacuoles at the positions 6H and 3K (arrows) correspond to vacuoles analyzed by CLEM and shown in B (image c: 6H and images d, e, f: 3K). **B-** Typical mature rhoptries and apical / lateral micronemes were detected in non-YFP control parasites (images a and b, arrows). By contrast, in YFP-positive AP μ 1-KO, no mature rhoptries were visualized at the apical pole, which instead displayed numerous big vesicles and only apical micronemes were detected (images c and d, arrows). Drastically affected AP μ 1-KO vacuoles showed abnormal division, with parasites that have broken out or not fully segregated (image e). Image f corresponds to a zoomed region of image e (white frame) showing incomplete pellicle formation and separation between two neighbouring cells (arrows). Rh: rhoptries, MIC: micronemes, DG: dense granules. Bars: 500nm.

(TIF)

S11 Fig. Transmission electron microscopy images showing the formation of enlarged lucent vesicles upon over-expression of AP μ 1. In control parasites (A-D), lucent vesicles were not often visualized or detected as small vesicles resembling endosomes (C and D, arrows). C represents a zoom of the region indicated by a white frame in B. In opposite, in induced DD μ 1 parasites (E-L), large lucent vesicles accumulated mainly at the post-nuclear anterior region of the parasite (arrows). We often detected membranous or vesicle-like material within their lumen (E-F and I-J, arrow heads). In some cases, this material seems to be generated by internal budding of the limiting membrane (J, arrow heads). F, H and J represent a zoom of the region indicated by a white frame in E, G, and I, respectively. Bars: 100 nm or 500 nm as indicated.

(TIF)

S1 Table. Accession numbers of the genes used for the sequence alignment analysis showed in S5 Fig of the Appendage Ear domain of the Gamma subunit (GAE) and the Beta subunit (BAE).

(PDF)

S1 Movie. 3D-SIM image reconstruction showing the localization of the Rab5A compartment (green) compared to AP μ 1 (red) localized in the TGN area. The KI parasite line expressing Rab5-YFP under the endogenous promoter was used for the study.

(AVI)

S2 Movie. 3D-SIM image reconstruction showing the localization of the Rab5A compartment (green) compared to the TGN marker SORTLR (red). The KI parasite line expressing Rab5-YFP under the endogenous promoter was used for the study.

(AVI)

S3 Movie. Time-lapse movie (one frame every 10 min for a duration of 5 hours) showing MORN1-cherry localization during the budding process in rapamycin induced AP μ 1-KO parasites (YFP-positive). Note that the basal complex is normally assembled and contractile rings developed towards the basal pole of budding daughter cells. Movies were captured using an inverted microscope (Axio-observer, Zeiss) equipped with a 40X Plan apochromat NA 1.4 objective. Image acquisition was performed using AxioVision Software (Zeiss).

(AVI)

S4 Movie. Egress was triggered in rapamycin induced AP μ 1-KO parasites (YFP-positive) after incubation with the calcium ionophore A23187. Time-lapse movies were recorded for 10 minutes. While control parasites immediately initiated vacuole egress, some AP μ 1-KO parasites appeared tethered and were unable to escape the vacuolar space though vacuole lysis has occurred.

(AVI)

Acknowledgments

We thank A. Bongiovanni for his help in image analysis, P.J. Sloves and M. Cossa for their help in performing experiments and Eik Hoffmann for critical reading of the manuscript. We are grateful to M. Meissner for providing the DD overexpression and CreLOX plasmids and the DiCre parasite strain. We thank J.F. Dubremetz, V. Carruthers, D. Soldati-Favre, D.S. Roos, M. Gissot, M.J. Gubbels and G.E. Ward for providing us with antibodies and plasmids.

Author Contributions

Conceptualization: SM.

Funding acquisition: SM GL.

Investigation: KV AP LH NB EW JMS FS.

Methodology: KV EW NB.

Project administration: SM.

Resources: SM MAH FL.

Supervision: GL.

Validation: SM.

Visualization: KV SM.

Writing – original draft: KV SM.

Writing – review & editing: GL FL MAH.

References

1. Saadatnia G, Golkar M. A review on human toxoplasmosis. *Scand J Infect Dis.* 2012 Nov; 44(11):805–14. <https://doi.org/10.3109/00365548.2012.693197> PMID: 22831461
2. Besteiro S, Dubremetz J-F, Lebrun M. The moving junction of apicomplexan parasites: a key structure for invasion. *Cell Microbiol.* 2011 Jun; 13(6):797–805. <https://doi.org/10.1111/j.1462-5822.2011.01597.x> PMID: 21535344
3. Lamarque MH, Roques M, Kong-Hap M, Tonkin ML, Rugarabamu G, Marq J-B, et al. Plasticity and redundancy among AMA-RON pairs ensure host cell entry of Toxoplasma parasites. *Nat Commun.* 2014 Jun 17; 5:4098. <https://doi.org/10.1038/ncomms5098> PMID: 24934579

4. Carruthers V, Boothroyd JC. Pulling together: an integrated model of Toxoplasma cell invasion. *Curr Opin Microbiol.* 2007 Feb; 10(1):83–9. <https://doi.org/10.1016/j.mib.2006.06.017> PMID: 16837236
5. Hunter CA, Sibley LD. Modulation of innate immunity by Toxoplasma gondii virulence effectors. *Nat Rev Microbiol.* 2012 Nov; 10(11):766–78. <https://doi.org/10.1038/nrmicro2858> PMID: 23070557
6. Mercier C, Dubremetz J-F, Rauscher B, Lecordier L, Sibley LD, Cesbron-Delauw M-F. Biogenesis of nanotubular network in Toxoplasma parasitophorous vacuole induced by parasite proteins. *Mol Biol Cell.* 2002 Jul; 13(7):2397–409. <https://doi.org/10.1091/mbc.E02-01-0021> PMID: 12134078
7. Mercier C, Adjogble KDZ, Däubener W, Delauw M-F-C. Dense granules: are they key organelles to help understand the parasitophorous vacuole of all apicomplexa parasites? *Int J Parasitol.* 2005 Jul; 35(8):829–49. <https://doi.org/10.1016/j.ijpara.2005.03.011> PMID: 15978597
8. Lopez J, Bittame A, Massera C, Vasseur V, Effantin G, Valat A, et al. Intravacuolar Membranes Regulate CD8 T Cell Recognition of Membrane-Bound Toxoplasma gondii Protective Antigen. *Cell Rep.* 2015 Dec 15; 13(10):2273–86. <https://doi.org/10.1016/j.celrep.2015.11.001> PMID: 26628378
9. Hakimi M-A, Bougdour A. Toxoplasma's ways of manipulating the host transcriptome via secreted effectors. *Curr Opin Microbiol.* 2015 Aug; 26:24–31. <https://doi.org/10.1016/j.mib.2015.04.003> PMID: 25912924
10. Pelletier L, Stern CA, Pypaert M, Sheff D, Ngô HM, Roper N, et al. Golgi biogenesis in Toxoplasma gondii. *Nature.* 2002 Aug 1; 418(6897):548–52. <https://doi.org/10.1038/nature00946> PMID: 12152082
11. Tomavo S, Slomianny C, Meissner M, Carruthers VB. Protein trafficking through the endosomal system prepares intracellular parasites for a home invasion. *PLoS Pathog.* 2013 Oct; 9(10):e1003629. <https://doi.org/10.1371/journal.ppat.1003629> PMID: 24204248
12. Carruthers VB, Tomley FM. Microneme proteins in apicomplexans. *Subcell Biochem.* 2008; 47:33–45. PMID: 18512339
13. Carruthers VB, Sherman GD, Sibley LD. The Toxoplasma adhesive protein MIC2 is proteolytically processed at multiple sites by two parasite-derived proteases. *J Biol Chem.* 2000 May 12; 275(19):14346–53. PMID: 10799515
14. El Hajj H, Papoin J, Cérède O, Garcia-Réguet N, Soëte M, Dubremetz J-F, et al. Molecular signals in the trafficking of Toxoplasma gondii protein MIC3 to the micronemes. *Eukaryot Cell.* 2008 Jun; 7(6):1019–28. <https://doi.org/10.1128/EC.00413-07> PMID: 18390648
15. Harper JM, Huynh M-H, Coppens I, Parussini F, Moreno S, Carruthers VB. A cleavable propeptide influences Toxoplasma infection by facilitating the trafficking and secretion of the TgMIC2-M2AP invasion complex. *Mol Biol Cell.* 2006 Oct; 17(10):4551–63. <https://doi.org/10.1091/mbc.E06-01-0064> PMID: 16914527
16. Meissner M, Reiss M, Viebig N, Carruthers VB, Torsel C, Tomavo S, et al. A family of transmembrane microneme proteins of Toxoplasma gondii contain EGF-like domains and function as escorts. *J Cell Sci.* 2002 Feb 1; 115(Pt 3):563–74. PMID: 11861763
17. Reiss M, Viebig N, Brecht S, Fourmaux MN, Soete M, Di Cristina M, et al. Identification and characterization of an escorter for two secretory adhesins in Toxoplasma gondii. *J Cell Biol.* 2001 Feb 5; 152(3):563–78. PMID: 11157983
18. Gaji RY, Flammer HP, Carruthers VB. Forward targeting of Toxoplasma gondii proproteins to the micronemes involves conserved aliphatic amino acids. *Traffic Cph Den.* 2011 Jul; 12(7):840–53.
19. Soldati D, Lassen A, Dubremetz JF, Boothroyd JC. Processing of Toxoplasma ROP1 protein in nascent rhoptries. *Mol Biochem Parasitol.* 1998 Oct 30; 96(1–2):37–48. PMID: 9851605
20. Morlon-Guyot J, Pastore S, Berry L, Lebrun M, Daher W. Toxoplasma gondii Vps11, a subunit of HOPS and CORVET tethering complexes, is essential for the biogenesis of secretory organelles. *Cell Microbiol.* 2015 Aug; 17(8):1157–78. <https://doi.org/10.1111/cmi.12426> PMID: 25640905
21. Kremer K, Kamin D, Rittweger E, Wilkes J, Flammer H, Mahler S, et al. An overexpression screen of Toxoplasma gondii Rab-GTPases reveals distinct transport routes to the micronemes. *PLoS Pathog.* 2013 Mar; 9(3):e1003213. <https://doi.org/10.1371/journal.ppat.1003213> PMID: 23505371
22. Sloves P-J, Delhaye S, Mouveaux T, Werkmeister E, Slomianny C, Hovasse A, et al. Toxoplasma sortilin-like receptor regulates protein transport and is essential for apical secretory organelle biogenesis and host infection. *Cell Host Microbe.* 2012 May 17; 11(5):515–27. <https://doi.org/10.1016/j.chom.2012.03.006> PMID: 22607804
23. Breinich MS, Ferguson DJP, Foth BJ, van Dooren GG, Lebrun M, Quon DV, et al. A dynamin is required for the biogenesis of secretory organelles in Toxoplasma gondii. *Curr Biol CB.* 2009 Feb 24; 19(4):277–86. <https://doi.org/10.1016/j.cub.2009.01.039> PMID: 19217293
24. Sakura T, Sindikubwabo F, Oesterlin LK, Bousquet H, Slomianny C, Hakimi M-A, et al. A Critical Role for Toxoplasma gondii Vacuolar Protein Sorting VPS9 in Secretory Organelle Biogenesis and Host Infection. *Sci Rep.* 2016 Dec 14; 6:38842. <https://doi.org/10.1038/srep38842> PMID: 27966671

25. Jackson AJ, Clucas C, Mamczur NJ, Ferguson DJ, Meissner M. Toxoplasma gondii Syntaxin 6 is required for vesicular transport between endosomal-like compartments and the Golgi complex. *Traffic Cph Den.* 2013 Nov; 14(11):1166–81.
26. Sangaré LO, Alayi TD, Westermann B, Hovasse A, Sindikubwabo F, Callebaut I, et al. Unconventional endosome-like compartment and retromer complex in Toxoplasma gondii govern parasite integrity and host infection. *Nat Commun.* 2016; 7:11191. <https://doi.org/10.1038/ncomms11191> PMID: 27064065
27. Robinson DG, Pimpl P. Clathrin and post-Golgi trafficking: a very complicated issue. *Trends Plant Sci.* 2014 Mar; 19(3):134–9. <https://doi.org/10.1016/j.tplants.2013.10.008> PMID: 24263003
28. Bonifacino JS, Rojas R. Retrograde transport from endosomes to the trans-Golgi network. *Nat Rev Mol Cell Biol.* 2006 Aug; 7(8):568–79. <https://doi.org/10.1038/nrm1985> PMID: 16936697
29. Johannes L, Wunder C. Retrograde transport: two (or more) roads diverged in an endosomal tree? *Traffic Cph Den.* 2011 Aug; 12(8):956–62.
30. Wang J-G, Li S, Zhao X-Y, Zhou L-Z, Huang G-Q, Feng C, et al. HAPLESS13, the Arabidopsis μ 1 adaptin, is essential for protein sorting at the trans-Golgi network/early endosome. *Plant Physiol.* 2013 Aug; 162(4):1897–910. <https://doi.org/10.1104/pp.113.221051> PMID: 23766365
31. Bonifacino JS. Adaptor proteins involved in polarized sorting. *J Cell Biol.* 2014 Jan 6; 204(1):7–17. <https://doi.org/10.1083/jcb.201310021> PMID: 24395635
32. Park M, Song K, Reichardt I, Kim H, Mayer U, Stierhof Y-D, et al. Arabidopsis μ -adaptin subunit AP1M of adaptor protein complex 1 mediates late secretory and vacuolar traffic and is required for growth. *Proc Natl Acad Sci U S A.* 2013 Jun 18; 110(25):10318–23. <https://doi.org/10.1073/pnas.1300460110> PMID: 23733933
33. Ma Y, Takeuchi M, Sugiura R, Sio SO, Kuno T. Deletion mutants of AP-1 adaptin subunits display distinct phenotypes in fission yeast. *Genes Cells Devoted Mol Cell Mech.* 2009 Aug; 14(8):1015–28.
34. Sosa RT, Weber MM, Wen Y, O'Halloran TJ. A single β adaptin contributes to AP1 and AP2 complexes and clathrin function in Dictyostelium. *Traffic Cph Den.* 2012 Feb; 13(2):305–16.
35. Teh O-K, Shimono Y, Shirakawa M, Fukao Y, Tamura K, Shimada T, et al. The AP-1 μ adaptin is required for KNOLLE localization at the cell plate to mediate cytokinesis in Arabidopsis. *Plant Cell Physiol.* 2013 Jun; 54(6):838–47. <https://doi.org/10.1093/pcp/pct048> PMID: 23543752
36. Bonnemaïson ML, Eipper BA, Mains RE. Role of adaptor proteins in secretory granule biogenesis and maturation. *Front Endocrinol.* 2013; 4:101.
37. Nevin WD, Dacks JB. Repeated secondary loss of adaptin complex genes in the Apicomplexa. *Parasitol Int.* 2009 Mar; 58(1):86–94. <https://doi.org/10.1016/j.parint.2008.12.002> PMID: 19146987
38. Edeling MA, Smith C, Owen D. Life of a clathrin coat: insights from clathrin and AP structures. *Nat Rev Mol Cell Biol.* 2006 Jan; 7(1):32–44. <https://doi.org/10.1038/nrm1786> PMID: 16493411
39. Di Cristina M, Spaccapelo R, Soldati D, Bistoni F, Crisanti A. Two conserved amino acid motifs mediate protein targeting to the micronemes of the apicomplexan parasite Toxoplasma gondii. *Mol Cell Biol.* 2000 Oct; 20(19):7332–41. PMID: 10982850
40. Hoppe HC, Ngô HM, Yang M, Joiner KA. Targeting to rhoptry organelles of Toxoplasma gondii involves evolutionarily conserved mechanisms. *Nat Cell Biol.* 2000 Jul; 2(7):449–56. <https://doi.org/10.1038/35017090> PMID: 10878811
41. Ngô HM, Yang M, Paprotka K, Pypaert M, Hoppe H, Joiner KA. AP-1 in Toxoplasma gondii mediates biogenesis of the rhoptry secretory organelle from a post-Golgi compartment. *J Biol Chem.* 2003 Feb 14; 278(7):5343–52. <https://doi.org/10.1074/jbc.M208291200> PMID: 12446678
42. Labesse G, Gelin M, Bessin Y, Lebrun M, Papoin J, Cerdan R, et al. ROP2 from Toxoplasma gondii: a virulence factor with a protein-kinase fold and no enzymatic activity. *Struct Lond Engl* 1993. 2009 Jan 14; 17(1):139–46.
43. Kaderi Kibria KM, Rawat K, Klinger CM, Datta G, Panchal M, Singh S, et al. A role for adaptor protein complex 1 in protein targeting to rhoptry organelles in Plasmodium falciparum. *Biochim Biophys Acta.* 2015 Mar; 1853(3):699–710. <https://doi.org/10.1016/j.bbamcr.2014.12.030> PMID: 25573429
44. Contento AL, Bassham DC. Structure and function of endosomes in plant cells. *J Cell Sci.* 2012 Aug 1; 125(Pt 15):3511–8. <https://doi.org/10.1242/jcs.093559> PMID: 22935651
45. Kibria KMK, Hossain MU, Oany AR, Ahmad SAI. Novel insights on ENTH domain-containing proteins in apicomplexan parasites. *Parasitol Res.* 2016 Jun; 115(6):2191–202. <https://doi.org/10.1007/s00436-016-4961-1> PMID: 26922178
46. Kalthoff C, Groos S, Kohl R, Mahrhold S, Ungewickell EJ. Clint: a novel clathrin-binding ENTH-domain protein at the Golgi. *Mol Biol Cell.* 2002 Nov; 13(11):4060–73. <https://doi.org/10.1091/mbc.E02-03-0171> PMID: 12429846

47. Mills IG, Praefcke GJK, Vallis Y, Peter BJ, Olesen LE, Gallop JL, et al. EpsinR: an AP1/clathrin interacting protein involved in vesicle trafficking. *J Cell Biol.* 2003 Jan 20; 160(2):213–22. <https://doi.org/10.1083/jcb.200208023> PMID: 12538641
48. Andenmatten N, Egarter S, Jackson AJ, Jullien N, Herman J-P, Meissner M. Conditional genome engineering in *Toxoplasma gondii* uncovers alternative invasion mechanisms. *Nat Methods.* 2013 Feb; 10(2):125–7. <https://doi.org/10.1038/nmeth.2301> PMID: 23263690
49. Herm-Götz A, Agop-Nersesian C, Münter S, Grimley JS, Wandless TJ, Frischknecht F, et al. Rapid control of protein level in the apicomplexan *Toxoplasma gondii*. *Nat Methods.* 2007 Dec; 4(12):1003–5. <https://doi.org/10.1038/nmeth1134> PMID: 17994029
50. Gissot M, Walker R, Delhay S, Huot L, Hot D, Tomavo S. *Toxoplasma gondii* chromodomain protein 1 binds to heterochromatin and colocalises with centromeres and telomeres at the nuclear periphery. *PLoS One.* 2012; 7(3):e32671. <https://doi.org/10.1371/journal.pone.0032671> PMID: 22427862
51. Agromayor M, Martin-Serrano J. Knowing when to cut and run: mechanisms that control cytokinetic abscission. *Trends Cell Biol.* 2013 Sep; 23(9):433–41. <https://doi.org/10.1016/j.tcb.2013.04.006> PMID: 23706391
52. Lorestani A, Sheiner L, Yang K, Robertson SD, Sahoo N, Brooks CF, Ferguson DJ, Striepen B, Gubels MJ. A *Toxoplasma* MORN1 null mutant undergoes repeated divisions but is defective in basal assembly, apicoplast division and cytokinesis. *PLoS One.* 2010; 5(8):e12302. PMID: 20808817
53. Agop-Nersesian C, Naissant B, Ben Rached F, Rauch M, Kretzschmar A, Thiberge S, et al. Rab11A-controlled assembly of the inner membrane complex is required for completion of apicomplexan cytokinesis. *PLoS Pathog.* 2009 Jan; 5(1):e1000270. <https://doi.org/10.1371/journal.ppat.1000270> PMID: 19165333
54. Parussini F, Coppens I, Shah PP, Diamond SL, Carruthers VB. Cathepsin L occupies a vacuolar compartment and is a protein maturase within the endo/exocytic system of *Toxoplasma gondii*. *Mol Microbiol.* 2010 Jun; 76(6):1340–57. <https://doi.org/10.1111/j.1365-2958.2010.07181.x> PMID: 20444089
55. Nakagawa T, Setou M, Seog D, Ogasawara K, Dohmae N, Takio K, et al. A novel motor, KIF13A, transports mannose-6-phosphate receptor to plasma membrane through direct interaction with AP-1 complex. *Cell.* 2000 Nov 10; 103(4):569–81. PMID: 11106728
56. Shiba Y, Takatsu H, Shin H-W, Nakayama K. Gamma-adaptin interacts directly with Rabaptin-5 through its ear domain. *J Biochem (Tokyo).* 2002 Mar; 131(3):327–36.
57. Mattera R, Arighi CN, Lodge R, Zerial M, Bonifacino JS. Divalent interaction of the GGAs with the Rabaptin-5-Rabex-5 complex. *EMBO J.* 2003 Jan 2; 22(1):78–88. <https://doi.org/10.1093/emboj/cdg015> PMID: 12505986
58. Kratzke M, Candiello E, Schmidt B, Jahn O, Schu P. AP-1/σ1B-Dependent SV Protein Recycling Is Regulated in Early Endosomes and Is Coupled to AP-2 Endocytosis. *Mol Neurobiol.* 2015 Aug; 52(1):142–61. <https://doi.org/10.1007/s12035-014-8852-0> PMID: 25128028
59. Pieperhoff MS, Schmitt M, Ferguson DJP, Meissner M. The role of clathrin in post-Golgi trafficking in *Toxoplasma gondii*. *PLoS One.* 2013; 8(10):e77620. <https://doi.org/10.1371/journal.pone.0077620> PMID: 24147036
60. Dou Z, McGovern OL, Di Cristina M, Carruthers VB. *Toxoplasma gondii* ingests and digests host cytosolic proteins. *mBio.* 2014; 5(4):e01188–01114. <https://doi.org/10.1128/mBio.01188-14> PMID: 25028423
61. Muñoz-Hernández S, Carmen MG del, Mondragón M, Mercier C, Cesbron MF, Mondragón-González SL, et al. Contribution of the residual body in the spatial organization of *Toxoplasma gondii* tachyzoites within the parasitophorous vacuole. *J Biomed Biotechnol.* 2011; 2011:473983. <https://doi.org/10.1155/2011/473983> PMID: 22190852
62. Ouologuem DT, Roos DS. Dynamics of the *Toxoplasma gondii* inner membrane complex. *J Cell Sci.* 2014 Aug 1; 127(Pt 15):3320–30. <https://doi.org/10.1242/jcs.147736> PMID: 24928899
63. Martins-Duarte ÉS, Carias M, Vommaro R, Suroliá N, de Souza W. Apicomplast fatty acid synthesis is essential for pellicle formation at the end of cytokinesis in *Toxoplasma gondii*. *J Cell Sci.* 2016 Sep 1; 129(17):3320–31. <https://doi.org/10.1242/jcs.185223> PMID: 27457282
64. Otegui MS, Herder R, Schulze J, Jung R, Staehelin LA. The proteolytic processing of seed storage proteins in *Arabidopsis* embryo cells starts in the multivesicular bodies. *Plant Cell.* 2006 Oct; 18(10):2567–81. <https://doi.org/10.1105/tpc.106.040931> PMID: 17012602
65. Miguet L, Béchade G, Fornecker L, Zink E, Felden C, Gervais C, et al. Proteomic analysis of malignant B-cell derived microparticles reveals CD148 as a potentially useful antigenic biomarker for mantle cell lymphoma diagnosis. *J Proteome Res.* 2009 Jul; 8(7):3346–54. <https://doi.org/10.1021/pr801102c> PMID: 19413345