bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# *Plasmodium* condensin core subunits (SMC2/SMC4) mediate atypical mitosis and are essential for parasite proliferation and transmission

Rajan Pandey[1,8], Steven Abel[2,8], Matthew Boucher[1,8], Richard J Wall[3], Mohammad Zeeshan[1], Edward Rea[1], Aline Freville[1], Xueqing Maggie Lu[2], Declan Brady[1], Emilie Daniel[1], Rebecca R. Stanway[4], Sally Wheatley[1], Gayani Batugedara[2], Thomas Hollin[2], Andrew R. Bottrill[5], Dinesh Gupta[6], Anthony A. Holder[7], Karine G. Le Roch[2*] and Rita Tewari[1*]

[1]*School of Life Sciences, Queens Medical Centre, University of Nottingham, Nottingham, NG7 2UH, UK,* [2]*Departement of Molecular, Cell and Systems Biology, University of California Riverside, 900,University Ave, Riverside, CA, 92521, USA,* [3]*Wellcome Trust Centre for Anti-Infectives Research, School of Life Sciences, University of Dundee, Dundee, DD1 5EH, UK,* [4]*Institute of Cell Biology, University of Bern, Bern 3012, Switzerland,* [5]*School of Life Sciences, Gibbet Hill Campus, University of Warwick, Coventry, CV4 7AL, UK,* [6]*Translational Bioinformatics Group, International Center for genetic engineering and Biotechnology, New Delhi, 110067, India,* [7]*Malaria Parasitology Laboratory, The Francis Crick Institute, London, NW1 1AT, UK*

[8] These authors contributed equally to this work

[*]For correspondence

Karine G. Le Roch: karine.leroch@ucr.edu

Rita Tewari: rita.tewari@nottingham.ac.uk

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Summary

Condensin is a multi-subunit protein complex that regulates chromosome organization, segregation and condensation during cell division in eukaryotes. In *Plasmodium* spp., the causative agent of malaria, cell division is atypical and the role of condensin is unclear. Here we examine the role of SMC2 and SMC4, the core subunits of condensin during endomitosis in schizogony and endoreduplication in male gametogenesis. SMC2 and SMC4 localize at discrete foci during schizogony, and with a diffuse nuclear distribution during male gametogenesis. ChIP-seq analyses suggest a centromeric location of SMC2/SMC4 only during schizogony. Co-immunoprecipitation data reveal the presence of both condensin complex I and II during male gametogenesis, but only the SMC2/SMC4 heterodimer during schizogony. Finally, knockdown of *smc2* and *smc4* gene expression revealed their essential roles in parasite proliferation and transmission. This study shows that condensin core subunits (SMC2/SMC4) have differential complex and distinct functions at different stages of the parasite life cycle.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Introduction

Cellular proliferation in eukaryotes requires chromosome replication and segregation, followed by cell division, to ensure that daughter cells have identical copies of the genome. During classical open mitosis in many eukaryotes, chromosome condensation, centrosome migration and formation of the mitotic spindle are followed by dissolution of the nuclear envelope(Guttinger et al., 2009). In contrast, in some unicellular organisms such as the budding yeast, *Saccharomyces cerevisiae*, mitosis is closed: the nuclear membrane remains intact and chromosomes are separated by spindles assembled within the nucleus(Sazer et al., 2014). The mechanisms and the various regulatory molecules involved in cell division have been well studied in many eukaryotes. The cell division regulators include cyclins, cyclin-dependent kinases (CDKs), components of the anaphase promoting complex (APC), and other protein kinases and phosphatases(Chang et al., 2014; Fisher et al., 2012; Harashima et al., 2013).

An essential component of chromosome dynamics is a family of 'Structural Maintenance of Chromosomes' proteins, originally described in budding yeast as 'stability of minichromosomes' (SMC) proteins, which are implicated in chromosome segregation and condensation(Hirano, 2016; Uhlmann, 2016). Most eukaryotes have at least six genes encoding SMC proteins (each 110-170 kDa, with a central hinge region and N and C-terminal globular domains with Walker A and Walker B motifs forming the ATPase head domain). The six SMCs can be classified as subunits of condensin (SMC2 and SMC4, required for chromosomal condensation), cohesin (SMC1 and SMC3, required for chromosomal segregation) and the SMC5-SMC6 complex (involved in DNA repair and homologous recombination)(Hirano, 2016; Uhlmann, 2016).

Higher eukaryotic organisms have two different condensin complexes: condensin I and condensin II, whereas many single celled organisms such as yeast have only one condensin complex. SMC2 and SMC4 form the core structure for both condensin I and condensin II in higher eukaryotes(Hirano, 2016), and interact with three additional non-SMC components, including one kleisin(Schleiffer et al., 2003) and two heat protein subunits(Neuwald and Hirano, 2000). Kleisin Iγ (CAP-H), Heat IA (CAP-D2) and Heat IB (CAP-G) form the condensin I complex, whereas Kleisin IIβ (CAP-H2), Heat IIA (CAP-D3) and Heat IIB (CAP-G2) form the condensin II complex(Hirano, 2016; Uhlmann, 2016) (Figure 1A). Electron microscopy and protein-protein interaction studies have revealed the characteristic architecture and geometry of condensin complexes(Anderson et al., 2002; Onn et al., 2007). Condensin plays a vital role in cell division processes such as chromosomal condensation, correct folding and organization of chromosomes prior to anaphase and proper chromosome segregation and

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

separation(Hirano, 2016; Kschonsak et al., 2017; Ono et al., 2013; Rawlings et al., 2011; Uhlmann, 2016). Indeed, both SMC and non-SMC components are necessary for its full function, for example, chromosomal condensation is not observed in the absence of kleisin, showing its critical role for complex formation and condensation(Cuylen et al., 2011; Rawlings et al., 2011).

*Plasmodium,* the apicomplexan parasite that causes malaria, undergoes two types of atypical mitotic division during its life cycle: one in the asexual stages (schizogony in the liver and blood stages within the vertebrate host, and sporogony in the mosquito gut), and the other in male gametogenesis during the sexual stage(Arnot et al., 2011; Sinden, 1991b). Division during schizogony/sporogony resembles closed endomitosis with repeated asynchronous nuclear division without cytokinesis, resulting in a multinucleated syncytium. An intact nuclear envelope is maintained, wherein the microtubule organizing center (MTOC), known as the centriolar plaque or spindle pole body (SPB), is embedded, and rounds of mitosis and nuclear division proceed without chromosome condensation(Arnot et al., 2011; Francia and Striepen, 2014; Gerald et al., 2011; Sinden, 1991a, b; Sinden et al., 1976). In male gametogenesis, exposure of the male gametocyte to the mosquito midgut environment leads to activation, which results in endoreduplication with three rounds of rapid chromosome replication within 8-10 min and chromosomal condensation, followed by nuclear and cell division to produce eight motile male gametes (exflagellation)(Guttery et al., 2012b; Sinden, 1991b; Sinden et al., 1976; Sinden et al., 2010). During exflagellation, each condensed haploid nucleus and associated MTOC, together with a basal body, axoneme and flagellum form the microgamete that egresses from the main cellular body)(Guttery et al., 2012b; Sinden, 1991b; Sinden et al., 1976; Sinden et al., 2010).

The atypical cell division and proliferation of malaria parasites is controlled by, amongst others, unique and divergent Apicomplexa-specific CDKs, aurora-like kinases (ARKs), and mitotic protein phosphatase 1 (PP1), as well as only four APC components(Guttery et al., 2014; Roques et al., 2015; Wall et al., 2018; Ward et al., 2004; Wilkes and Doerig, 2008). However, there are no known classical group 1 cyclins, polo-like kinases (that are major regulators in mitotic entry) or classical mitotic protein phosphatases (CDC14 and CDC25) encoded in the genome(Guttery et al., 2014; Tewari et al., 2010; Ward et al., 2004; Wilkes and Doerig, 2008).

In *Plasmodium,* the role of condensin during cell division as well as general chromosome dynamics is unknown. Here, we investigated the location and function of the core subunit of condensin (SMC2 and SMC4) during two different mitotic division stages in the *Plasmodium* life cycle; during schizogony in the host blood, and during male gametogenesis in the mosquito vector. This was performed using the rodent malaria model *Plasmodium berghei.* For this

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

analysis we used a combination of cell biology, proteomics, transcriptomics and reverse genetics approaches. Spatiotemporal localization using live cell imaging indicates that both SMC2 and SMC4 have a dynamic profile, with either discrete foci during schizogony or a more diffused nuclear localization during male gametogenesis. Genome-wide distribution studies using ChIP-seq experiments, suggested that both components (SMC2/SMC4) are located at or near the centromeres during schizogony, but this strong interaction is not observed during gametogenesis. Interestingly, we identified a differential composition of the condensin complex between the distinct mitotic stages, suggesting divergent mechanisms at the molecular level. Our data demonstrate that the condensin core subunits (SMC2/SMC4) have distinct functions at different stages of the parasite life cycle. Functional analyses using a conditional gene knockdown approach indicate that condensins are required for parasite proliferation and transmission.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Results

### Bioinformatic analysis identifies both core subunit components (SMC2/SMC4) of condensin encoded in the *Plasmodium* genome

In order to identify condensin in *Plasmodium* we screened for condensin core subunit genes in the *P. berghei* genome using PlasmoDB version 42(Bahl et al., 2002). Domain analysis revealed a conserved domain architecture for both core SMC components of condensin, SMC2 and SMC4 (Figure 1A and 1B). A comparative sequence analysis revealed low sequence similarity and identity (~29-34%, except for SMC4 with *Arabidopsis thaliana* (65%), Figure 1C), though there was similarity in size and overall domain structure when compared with the other studied organisms. Interestingly, we found two N-terminus ATPase domain sparsed by 44 amino acid sequences for SMC4 in *P. berghei*. Similar pattern has also been observed in other studied *Plasmodium* spp. Subsequently, we generated a 3D model of the *P. berghei* SMC2 and SMC4 ATPase head domain and partial coiled region using homology-based 3D structure modeling (Figure 1D). Root mean square deviation (RMSD) analysis of the 10ns molecular dynamics (MD) simulation trajectory of the proteins showed a stable conformation comparable to pre-simulation energy minimized structures. Radius of gyration analysis also confirmed stable conformation for the predicted SMC2 and SMC4 domain structures during the 10ns MD simulation (Figure S1). In this model of the SMC subunits, the N- and C- terminal ABC ATPase head and coiled-coil arms connecting the Hinge domain (Figure 1D) are present, as in other organisms. It is most likely that the heads of *Plasmodium* SMC2 and SMC4 undergo ATP-dependent engagement and disengagement, and may perform similar chromosomal functions as in other eukaryotes(Hirano, 2016).

### Condensin core subunits are expressed at every proliferative stage of the parasite life cycle and have a centromeric location during schizogony

To locate the condensin SMC subunits during two proliferative stages (schizogony and male gametogenesis) of the *Plasmodium* life cycle, transgenic parasite lines were created to express GFP-tagged SMC2 and SMC4, using single crossover homologous recombination (Figure S2A). Integration PCR and western blot experiments were used to confirm the successful generation of transgenic lines (Figure S2B and S2C). We found that SMC2 and SMC4 are expressed during both schizogony and male gametogenesis. In dividing schizonts within host red blood cells, we observed discrete foci adjacent to the nuclear DNA for both SMC2 and SMC4 in the parasite cell (Figure 2A and 2B), whereas during male gametogenesis the localization was more diffuse, throughout the nucleus (Figure 2C and 2D). To further examine whether these foci are centromeric or centrosomal, we used co-localization

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

immunofluorescence assays using anti-centrin and anti-α-tubulin together with anti-GFP antibodies. In schizonts, immunofluorescence assays with anti-centrin antibodies revealed that the SMC4 location is within close proximity to centrin suggesting non-centrosomal localization (Figure 2E). However, partial co-localization was observed with anti-α-tubulin antibodies in schizonts (Figure 2E) but not during male gametogenesis (Figure 2F). In the absence of *Plasmodium* centromere-specific antibodies, our preliminary results with live imaging of kinetochore/centromeric proteins NDC80GFP and SMC2mCherry expressed in the same cell, shows a centromeric localization of SMC2 protein with NDC80 (Results not shown, unpublished). In addition, we observed either a nuclear or discrete focal location adjacent to nuclear DNA for SMC4GFP throughout the parasite life cycle including female gametocytes, oocyst development and the liver stages (Figure S3), suggesting that condensin core subunits are likely involved at all proliferative stages during the parasite life cycle.

To identify the SMC2 and SMC4 DNA binding sites in a genome-wide manner, we performed ChIP-seq experiments for schizont (after 8 hours in culture) and gametocyte (6 min post activation) stages using SMC2GFP and SMC4GFP-tagged parasites. A wild type strain (WTGFP) was used as a negative control. Binding of the SMC2 and SMC4 subunits was restricted to a region close to the previous computationally annotated centromeres (centromere locations for *P. berghei* chromosomes were predicted based on using *P. falciparum* as reference and conservation of genomic sequences among *Plasmodium* spp.) of all 14 chromosomes at the schizont stage (Figure 3A)(Iwanaga et al., 2012). This restriction was not observed during gametocytogenesis with a random distribution of condensin core subunits, suggesting a distinct function for the core subunits in these two mitotic stages. Identical patterns were obtained between biological replicates for each condition analyzed, confirming the reproducibility of the ChIP-seq experiments.

The chromosome binding sites of SMC2 and SMC4 in the schizont were slightly away from the locations previously annotated as centromeres(Iwanaga et al., 2012). However, at the centers of the ChIP-seq peaks for all 14 chromosomes, we identified distinct regions of very low GC-content that are not present in the previously annotated centromeric regions (shown for chromosomes 11 and 14 in Figure 3B). AT-rich troughs have been associated with centromeres in yeasts(Lynch et al., 2010) The peaks are also centered within extended intergenic regions suggesting that the SMC binding sites are the first experimentally validated centromeres for all 14 *P. berghei* chromosomes (Table S1), and these are different to the previously computationally annotated centromeres, which are slightly different(Iwanaga et al., 2012).

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## The full condensin complex is present during male gametogenesis, but ancillary proteins are absent during schizogony

To examine the co-localization of SMC2 and SMC4 proteins, we generated transgenic parasite lines expressing either SMC2mCherry or SMC4GFP and crossed them genetically. The progeny, expressing both SMC2mCherry and SMC4GFP, showed co-localization of the two proteins during schizogony and gametogenesis (Figure 4A) consistent with SMC2 and SMC4 heterodimer complex formation at both stages.

Next, we directly investigated the interaction between SMC2 and SMC4, and the presence of other interacting partner proteins, such as other condensin components (Figure 1A). We immunoprecipitated SMC2GFP and SMC4GFP from lysates of cells undergoing mitosis during either schizogony (after 8 hours of incubation in schizont culture medium *in vitro,* when most parasites are undergoing nuclear division) or during gametogenesis (at 6 min after activation when cells are in the last phase of cell division before cytokinesis and the chromosomes are beginning to condense). Immunoprecipitated proteins were then digested with trypsin and the resultant peptides analyzed by liquid chromatography and mass spectrometry (LC/MS). At both stages of the life cycle that were investigated, we recovered peptides from both SMC subunits confirming SMC2-SMC4 heterodimer formation during schizogony and male gametogenesis (Figure 4B). In schizonts, only SMC2 and SMC4 derived peptides were recovered, whereas from gametocyte lysates we detected kleisins and other components of canonical Condensin I and II complexes together with the SMC subunits (Figure 4B and Table S2). In some cases, condensin II HEAT subunit (CAP-G2 and CAP-D3) were observed, however since kleisin was never recovered in 5 experimental replicates of schizogony, we assume that condesin II complex formation is absent at this stage (Table S2). Analysis of the *Plasmodium* genome identified genes coding for both condensin I and condensin II complex subunits, in agreement with the immunoprecipitation in which we detected Kleisin Iγ (CAP-H), Heat IA (CAP-D2) and Heat IB (CAP-G) from the condensin I complex, and Kleisin IIβ (CAP-H2), Heat IIA (CAP-D3) and Heat IIB (CAP-G2) from the condensin II complex. The domain architecture of these subunits from *P. berghei* is shown in Figure 4C. The presence of both condensin I and II in male gametocytes is consistent with the potentially atypical chromosomal condensation has been previously observed in male gametocytes just before exflagellation(Sinden, 1991b; Sinden et al., 1976; Sinden and Hartley, 1985). In other organisms, kleisin is essential for formation of the full condensin complex, and in the absence of kleisin, chromosomal condensation is ablated(Cuylen et al., 2011; Rawlings et al., 2011). The presence of kleisins and other components of both condensin complexes in gametocytes

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

is consistent with the potential role these complexes may play in chromosome condensation during exflagellation.

In view of the importance of kleisin to the structure and function of the condensin complexes, we examined the evolutionary relationships of kleisin among some Apicomplexa and other organisms including *S. cerevisiae*, *A. thaliana* and *Homo sapiens*. The phylogenetic analysis indicated that kleisin is clustered into two groups, which correspond to components of condensin I and condensin II, respectively (Figure 1A and 4D). The presence of both condensin I and II component genes only in *Plasmodium* and *Cryptosporidium* shows that requirement for both condensin complexes is not a universal feature of Apicomplexa (Figure 4E). This suggests that these two genera have unique aspects of chromosome segregation and cytokinesis.

## Knockdown of condensin (SMC2 and SMC4) expression affects parasite proliferation and impairs parasite transmission

To examine further the functions of SMC2 and SMC4, we first attempted to delete the two genes. In both cases we were unable to produce gene knockout (KO) mutants (Figure S4A). Similar results have been reported previously from large scale genetic screens in *P. berghei*(Bushell et al., 2017; Schwach et al., 2015). Together these data indicate that the condensin subunits SMC2 and SMC4 are likely essential for asexual blood stage development (schizogony). To investigate the function of SMC2 and SMC4 during cell division in male gametogenesis, we used a promoter trap double homologous recombination (PTD) approach to down-regulate gene expression at this stage by placing each of the two genes under the control of the AMA1 promotor. AMA1 is known to be highly expressed in asexual blood stages but not during sexual differentiation. This strategy resulted in the successful generation of two transgenic parasite lines in which SMC2 and SMC4 subunit expression decreased significantly in sexual stages: $P_{ama1}$*smc2* (SMC2PTD) and $P_{ama1}$*smc4* (SMC4PTD), respectively (Figure S4B and S4C).

Since SMC2PTD and SMC4PTD had similar phenotypes, we performed a global transcriptome analysis only on SMC4PTD to identify other affected genes and regulators involved in cell division and proliferation. This analysis of SMC4PTD gametocytes 30 min after activation (when chromosome condensation and exflagellation is complete) confirmed the nearly complete ablation of *smc4* gene expression (Figure 5A). For pairs of the two biological samples (wild-type [WTGFP] and SMC4PTD), the Spearman correlation coefficients of 0.97 and 0.99 respectively, demonstrated the reproducibility of this experiment. (Figure 5B). In addition to SMC4, expression of a further 104 genes was also significantly downregulated,

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

while expression of only 5 genes was significantly upregulated (Figure 5C and Table S3). Gene Ontology (GO) enrichment analysis of the downregulated genes identified several associated with microtubule and cytoskeletal function (Figure 5D). The reduced expression levels of ten of these genes was also examined by qRT-PCR (Figure 5E). By this method, there was a statistically significant difference in the level of expression of nine of these genes when comparing samples from WTGFP and SMC4PTD (Figure 5E). Of particular interest are AP2O2 (an AP2 domain transcription factor) and HMG (putative high mobility group protein B3), which act as transcription regulators in ookinetes; a putative SET domain protein which is known to be involved in methyl group transfer from S-adenosyl-L-methionine (AdoMet) to a lysine residue in histones and most likely associated with transcriptional repression; and finally a RCC; a protein predicted to be involved in chromosome condensation and chromosomal dynamics(Bahl et al., 2002). Other genes that were significantly downregulated include FRM2, involved in cytoskeleton organization; CCRNOT2 and NOT that form the CCR4-NOT complex, a key regulator of eukaryotic gene expression; and SEC7, involved in regulation of ARF protein signal transduction (SEC7). Some of the other significantly downregulated genes, include AP2 transcription factor AP2-Sp; molecular motor kinesin-4, a putative regulator of chromosome condensation (PBANKA_0820800); and SMC1, a member of the SMC family, all known to be involved in either gene expression or chromatid segregation.

While we were unable to detect any particular impaired phenotype in the SMC2PTD and SMC4PTD lines at the asexual blood stage (schizogony), and the parasite formed a similar number of schizonts and nuclei as compared to the WTGFP line (Figure 6A), we observed a ~50% reduction in the number of exflagellation centers during male gametogenesis (Figure 6B). Fertilization and zygote formation leading to ookinete conversion was reduced to 10-15% (Figure 6C). The ookinete motility assay showed normal movement of SMC4PTD ookinetes as compared to WT (Video S1 and S2). In the mosquito gut on days 9, -14 and -21 post infection, we detected significantly fewer oocysts in the SMC2PTD and SMC4PTD lines (Figure 6D and 6E). Furthermore, the oocysts were considerably smaller compared to those of WTGFP (Figure 6D and Figure 6F) with unequal distribution and clusters of DNA in some of the oocysts at 14 and 21 days post infection. No sporogony or endomitosis was observed within oocysts (Figure 6G). We were also unable to detect any sporozoites in the mosquito salivary glands (Figure 6H) and hence no parasite transmission from infected mosquitoes to mice was observed for either SMC2PTD or SMC4PTD parasite lines, in bite-back experiments (Table S4) indicating that condensins are required for parasite transmission.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Discussion

Condensins are multi-subunit complexes that are involved in chromosomal condensation, organization and segregation, and have been widely studied in many eukaryotes(Hirano, 2016; Uhlmann, 2016). Their role in many unicellular protozoans such as *Plasmodium* is unknown. Here we describe the structure, localization and functional role of the condensin core subunits (SMC2/SMC4) in the mouse malaria-causing parasite, *P. berghei* using molecular dynamics, live cell imaging, protein pulldown and conditional gene knockdown approaches. *Plasmodium* shows atypical features of closed mitotic division during schizogony (with no observed chromosomal condensation and extensive nuclear division prior to cytokinesis) and male gametogenesis (with atypical chromosomal condensation and endoreduplication observed prior to exflagellation). Our previous studies have shown that the parasite has an unusual repertoire of proteins involved in the regulation of the cell cycle and cell proliferation: it has no identifiable centrosome, no obvious complement of cell-cycle cyclins, a small subset of APC components, a set of divergent and *Plasmodium*-specific CDKs, and a complete absence of polo-like kinases and CDC24 and CDC14 phosphatases, as compared to most organisms that have been studied(Arnot et al., 2011; Francia et al., 2015; Guttery et al., 2012a; Guttery et al., 2014; Roques et al., 2015; Tewari et al., 2010).

Here, using bioinformatics screening, we identified both condensin I and II complex subunit components encoded in the *Plasmodium* genome. The two core subunits of condensin, SMC2 and SMC4 have low sequence similarity to the proteins in model organisms but a similar protein structure was predicted by molecular modelling.

Protein localization studies in different stages of the parasite life cycle using live cell imaging of SMC2GFP and SMC4GFP show distinct patterns during the mitotic divisions of schizogony and male gametogenesis. Whereas discrete foci were detected during endomitosis in schizogony, a more diffused nuclear localization was observed during male gametogenesis. The discrete foci during schizogony were not coincident with centrin, marking the SPB but very close to it. The IFA and ChIP-seq analyses suggest that SMC2 and SMC4 form a complex that binds at or near the centromere of all 14 chromosomes, and that these protein-DNA complexes are located in the nucleus near the SPB in schizonts. Genome-wide studies of condensin distribution in mammalian or yeast cells have shown that the complex is non-randomly distributed across the chromosomes, and often found at the boundaries of topologically associating domains (TAD) within chromosome territories, which supports the proposed role in transcriptional regulation and global chromosomal organization(Kim et al., 2016; Yuen et al., 2017). The *Plasmodium* genome lacks classical TADs(Ay et al., 2014;

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Bunnik et al., 2018; Bunnik et al., 2019), and our results suggest that condensins are restricted to binding centromeric regions in the highly proliferative asexual stage, where they may have a constrained role in sister chromatid cohesion and segregation(Iwasaki and Noma, 2016).

In activated gametocytes, where the condensin complexes do not bind specifically to the centromeres of chromosomes, the proteins are distributed throughout the nucleus. This is in concordance with centromere location in many other organisms(Bachellier-Bassi et al., 2008; Fujiwara et al., 2013). The diffuse pattern observed during male gametogenesis was distinct to that in schizogony, and since both SMC2 and SMC4 are core subunits of condensin complexes, it was surprising that they had a unique distribution in the two different mitotic stages in *Plasmodium*. Interestingly, the protein interaction analysis showed that different complexes formed during the two mitotic divisions. We detected only the SMC2-SMC4 heterodimer in schizogony, and we observed both condensin I and condensin II complexes during male gametogenesis. These results substantiate previous electron microscopy results in which atypical and limited chromosomal condensation was observed during male gametogenesis but not in schizogony(Sinden, 1991b; Sinden et al., 1976; Sinden and Hartley, 1985). Our bioinformatics analysis shows that two parasites, *Plasmodium* and *C. parvum,* both with unusual modes of cell division*(Francia and Striepen, 2014)*, are the only Apicomplexa where components of both condensin I and II complexes are encoded in the genomes, similar to what is observed in higher eukaryotes. In other parasites, *Trypanosoma brucei* encodes only condensin I complex(Hammarton, 2007) with closed mitosis and no chromosome condensation, whereas in *Giardia intestinalis*, both condensin complexes (condensin I and condensin II) are present and chromosome condensation occurs(Tumova et al., 2015). However, Giardia lacks one of the non-SMC HEAT subunits (HEAT IB and HEAT IIB). Another protist, *Tetrahymena thermophila*, exhibits noncanonical cell division between somatic and germline nucleus and has expanded paralogs of condensin I, with different kleisin components between germline (Cph1 and Cph2) and somatic cells (Cph3, Cph4 and Cph5)(Howard-Till and Loidl, 2018; Howard-Till et al., 2019).

For condensin I and II complexes, a differential localization pattern has been observed in various organisms. In the red alga *Cyanidioschyzon merolae*(Fujiwara et al., 2013), condensin II has a centromeric location during metaphase, whereas condensin I distributes more broadly along the chromosome arms. In higher eukaryotes including *Drosophila melanogaster*(Oliveira et al., 2007), *Caenorhabditis elegans*(Collette et al., 2011) and HeLa cells(Hirota et al., 2004; Ono et al., 2004) condensin I is present in the cytoplasm and has a nuclear localization after nuclear envelope breakdown in open mitosis, whereas the nuclear localization of condensin II is observed in interphase, it is stabilized on chromatin during prophase, and the complex

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

remains associated with chromosomes throughout mitosis, at least in HeLa cells. Budding yeast and fission yeast which undergo closed mitosis like *Plasmodium* have only a single condensin complex, but there is a differential localization pattern in each species. In budding yeast, the condensin I complex is located in the nucleus throughout the cell cycle, a pattern observed for condensin II in higher eukaryotes, despite the greater protein sequence similarities of the yeast complex to higher eukaryote condensin I(Thadani et al., 2012). Also, within the nucleus, the condensin localization at the kinetochore is cell cycle dependent(Bachellier-Bassi et al., 2008). In fission yeast, the single condensin complex is predominantly cytoplasmic during interphase and nuclear during mitosis, with the location dependent on CDK phosphorylation at Thr19 of SMC4/Cut3(Sutani et al., 1999).

The present study shows that the SMC2-SMC4 complex plays an essential role during schizogony, as we and previous genome-wide functional screens were unable to disrupt the genes(Bushell et al., 2017). Our conditional knockdown approach using the promoter trap suggests that reduction in SMC2 or SMC4 affects male gametogenesis and zygote differentiation and causes total impairment of endomitotic cell division in the oocyst, thereby blocking parasite transmission.

The partial defect observed in male gamete formation (exflagellation) in the PTD parasite lines may be due to the necessity of condensin complex formation for proper chromosomal condensation during exflagellation. RNA-seq analysis of the SMC4PTD line confirmed the reduced expression of the *smc4* gene and identified dysregulated transcripts that are likely either critical for gene expression, or chromosomal segregation and condensation, microtubule assembly and male gametocyte activation. This further demonstrates that SMC2 and SMC4 complexes are essential for proper chromosome condensation and separation during exflagellation. Among the significantly dysregulated genes, deletion of AP2-O2, has been shown to strongly impair ookinete and oocyst development, leading to an absence of sporozoite formation and a complete blockage of transmission(Modrzynska et al., 2017). The SET protein, which is a post translational modification protein, is essential for parasite survival(Schwach et al., 2015). RCC is predicted to be a regulator of chromosome condensation, is essential for parasite survival, and acts as anchor for both parasite kinase (CDPK7) and phosphatase (PP1)(Lenne et al., 2018). The phenotype observed in SMC4PTD parasites may therefore reflect contributions from all these differentially regulated genes.

The reduction in mature ookinete formation in the SMC4PTD line, which is the stage when meiosis takes place, suggests an important role of condensin during meiosis as well. During this stage, chromosomal condensation has been observed(Sinden and Hartley, 1985), and this may be similar to the situation in *Arabidopsis*, where condensin is important in

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

chromosomal condensation during meiosis(Smith et al., 2014). A severe defect in number and size of oocyst formation in the mosquito gut was also observed, and at this stage multiple rounds of endomitotic division give rise to thousands of sporozoites. The process requires ten or more rounds of DNA replication, segregation, and mitotic division to create a syncytial cell (sporoblast) with thousands of nuclei over a period of several days(Francia and Striepen, 2014; Gerald et al., 2011). The proper organization of nuclei into individual sporoblasts is organized by putative MTOCs(Roques et al., 2019; Sinden and Strong, 1978). As condensin has been shown to play an important role in organizing MTOCs(Kim et al., 2014) it may be that in the absence of condensin the endomitotic division is impaired and no sporozoites are formed. Deletion of a number of genes that affect oocyst maturation results in a phenotype similar to that of the SMC2PTD and SMC4PTD parasite lines. For example *Pb*MISFIT, causes reduced oocyst size(Bushell et al., 2009), and *Pb*CYC3 has smaller oocysts but sporogony, sporozoite formation and transmission are normal(Roques et al., 2015), and disruption of the gene for a G-actin sequestering protein affects oocyst maturation and sporozoite formation(Hliscs et al., 2010). A *Plasmodium* mutant lacking a protein phosphatase (*Pb*PPM5) shows a similar oocyst and sporozoite development pattern but in a global transcriptome analysis no significant differential expression of SMC4PTD was observed(Guttery et al., 2014).

In summary, the present study shows that the condensin core subunits SMC2 and SMC4 play crucial roles in the atypical mitoses of the *Plasmodium* life cycle. Their removal or depletion causes impaired parasite development and blocks transmission. Further analyses of non-SMC components of condensin I and II will give us better insight into the function of condensin during *Plasmodium* cell proliferation.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

# Methods

## Ethics statement

All animal work done at the University of Nottingham has passed an ethical review process and has been approved by the United Kingdom Home Office. The work was carried out under UK Home Office Project Licenses (40/3344 and 30/3248).

## Bioinformatics analysis

Condensin complex protein sequences were retrieved from PlasmoDB(Bahl et al., 2002), EuPathDB(Aurrecoechea et al., 2010) and from NCBI databases for model organisms (File S1). An NCBI conserved domain database (CDD) search was used to identify conserved domains. PHYRE2(Kelley et al., 2015) was used to generate 3D structure models. GROMACS 4.6.3(Van Der Spoel et al., 2005) with CHARMM27(Sapay and Tieleman, 2011) force field was used to perform molecular dynamics simulation in an aqueous environment. The energy minimization was performed using steepest descent minimization till maximum force reached below 1000KJ/mol/nm. Temperature and pressure equilibrium were done for 1ns, respectively before performing the 10 ns production simulation. ClustalW was used to generate multiple sequence alignments of the retrieved sequences(Larkin et al., 2007). ClustalW alignment parameters included gap opening penalty (GOP) of 10 and gap extension penalty (GOE) of 0.1 for pairwise sequence alignments, and GOP of 10 and GOE of 0.2 for multiple sequence alignments, gap separation distance cut-off value of 4 and the Gonnet algorithm in protein weight matrix. Other parameters like residue-specific penalty and hydrophobic penalties were “on” whereas end gap separation and use of negative matrix were set to “off”. The phylogenetic tree was inferred using the neighbor-joining method, computing the evolutionary distance using the Jones Taylor Thornton (JTT) model for amino acid substitution with the Molecular Evolutionary Genetics Analysis software (MEGA 6.0) (Tamura et al., 2013). Gaps and missing data were treated using a partial deletion method with 95% site-coverage cut-off. We performed 1000 bootstrap replicates to infer the final phylogenetic tree.

## Generation of transgenic parasites

GFP-tagged vectors were designed using the p277 plasmid vector and transfected as described previously(Guttery et al., 2014). Targeted gene deletion vectors were designed using the pBS-DHFR plasmid(Tewari et al., 2010). Conditional gene knockdown constructs (SMC2PTD and SMC4PTD) were designed using $P_{ama1}$ (*pSS368*)(Sebastian et al., 2012). *P. berghei* ANKA line 2.34 (for GFP-tagging) or ANKA line 507cl1 (for gene deletion

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

and promoter swap) parasites were transfected by electroporation as described previously(Wall et al., 2018). Genotypic analysis was performed using diagnostic PCR reaction and Western blot. All of the oligonucleotides used to confirm genetically modified tag and mutant parasite lines can be found in Table S5.

## ChIP-seq and Global transcriptomic analysis

For the ChIP-seq analysis, libraries were prepared from crosslinked cells (using 1% formaldehyde) and for the global transcriptome analysis, libraries were prepared from lyophilized total RNA, both using the KAPA Library Preparation Kit (KAPA Biosystems). Libraries were amplified for a total of 12 PCR cycles (15 s at 98°C, 30 s at 55°C, 30 s at 62°C) using the KAPA HiFi HotStart Ready Mix (KAPA Biosystems). Libraries were sequenced using a NextSeq500 DNA sequencer (Illumina), producing paired-end 75-bp reads. FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), was used to analyze raw read quality. Any adapter sequences were removed using Trimmomatic (http://www.usadellab.org/cms/?page=trimmomatic). Bases with Phred quality scores below 25 were trimmed using Sickle (https://github.com/najoshi/sickle). The resulting reads were mapped against the *P. berghei* ANKA genome (v36) using Bowtie2 (version 2.3.4.1) or HISAT2 (version 2-2.1.0), using default parameters. Reads with a mapping quality score of 10 or higher were retained using Samtools (http://samtools.sourceforge.net/), and PCR duplicates were removed by PicardTools MarkDuplicates (Broad Institute). For the transcript analysis, raw read counts were determined for each gene in the *P. berghei* genome using BedTools (https://bedtools.readthedocs.io/en/latest/#) to intersect the aligned reads with the genome annotation and for the ChIP-seq to obtain the read coverage per nucleotide. For the transcriptomic analysis, read counts were normalized by dividing by the total number of mapped reads for the library. Genome browser tracks were generated and viewed using the Integrative Genomic Viewer (IGV) (Broad Institute). Proposed centromeric locations were obtained from Iwanaga and colleagues(Iwanaga et al., 2012). GC content was calculated using a sliding window of 30 bp across the peak region as described previously[34]. SMC2 gametocyte sample is shown at half height due to higher level of background compared to other samples. Differential expression analysis was done in two ways: (1) the use of R package DESeq2 to call up- and down-regulated genes, and (2) manual analysis, in which raw read counts were normalized by library size, and genes above a threshold level of difference in normalized read counts between conditions were called as up- or down-regulated. Gene ontology enrichment was done using PlasmoDB (http://plasmodb.org/plasmo/) with repetitive terms removed by REVIGO (http://revigo.irb.hr/).

## Immunoprecipitation and Mass Spectrometry Analysis

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Schizonts, following 8 hours in *in vitro* culture, and male gametocytes 6 min post activation were used to prepare cell lysates. Purified parasite pellets were crosslinked using formaldehyde (10 min incubation with 1% formaldehyde, followed by 5 min incubation in 0.125M glycine solution and 3 washes with phosphate buffered saline (PBS) pH7.5). Immunoprecipitation was performed using crosslinked protein and a GFP-Trap®_A Kit (Chromotek) following the manufacturer's instructions. Proteins bound to the GFP-Trap®_A beads were digested using trypsin and the peptides were analysed by LC-MS/MS.

## Phenotypic Analysis

Phenotypic analyses of the mutant parasite lines were performed at different points of parasite life cycle as described previously(Guttery et al., 2014). Briefly, infected blood was used to analyse asexual blood stages and gametocytes. Schizont culture was used to analyse different stages of asexual development. *In vitro* cultures were prepared to analyse activated gametocyte, exflagellation, zygote formation and ookinete development. For *in vitro* exflagellation studies, gametocyte-infected blood was obtained from the tails of infected mice using a heparinised pipette tip. Gametocyte activation was performed by mixing 100 µl of ookinete culture medium (RPMI 1640 containing 25 mM HEPES, 20% fetal bovine serum, 10 mM sodium bicarbonate, 50 µM xanthurenic acid at pH 7.6) to the gametocyte infected blood. Microgametogenesis was monitored at two time points to study mitotic division (6 and 15 min post activation [mpa]). The material was fixed and processed for immunofluorescence assay (IFA) using a range of different antibody markers. For mosquito transmission and bite back experiments, triplicate sets of 40-50 *Anopheles stephensi* mosquitoes were used. The mosquito guts were analysed on different days post infection (dpi); 9 dpi, 14 dpi and 21 dpi to check oocyst development and sporozoite formation. Parasites were visualised on a Zeiss AxioImager M2 microscope fitted with an AxioCam ICc1 digital camera (Carl Zeiss, Inc).

## Quantitative RT-PCR

RNA was isolated from different parasite life stages, which include asexual stages, purified schizonts, activated and non-activated gametocytes, ookinetes and sporozoites, using an RNA purification kit (Stratagene). cDNA was prepared using an RNA-to-cDNA kit (Applied Biosystems). Primers for qRT-PCR were designed using Primer3 (Primer-blast, NCBI). Gene expression was quantified from 80 ng of total cDNA. qRT-PCR reactions used SYBR green fast master mix (Applied Biosystems) and were analysed using an Applied Biosystems 7500 fast machine. Experiments used *hsp70* and *arginine-tRNA synthetase* as reference genes. The primers used for qRT-PCR can be found in Table S5.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Statistical analysis

Statistical analysis was performed using Graph Pad Prism 7 software. An unpaired t-test was conducted to examine significant differences between wild-type and mutant strains.

## Data Availability

Sequence reads have been deposited in the NCBI Sequence Read Archive with accession number PRJNA542367.

## Acknowledgements

We thank Prof. Frank Uhlmann, The Francis Crick Institute, for stimulating discussions and advice on condensins and Julie Rodgers for insectary assistance. This project was funded by MRC project grants and MRC Investigators grants awarded to RT (G0900109, G0900278, MR/K011782/1) and BBSRC (BB/N017609/1). RP was supported by MRC grant MR/K011782/1. AAH was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001097), the UK Medical Research Council (FC001097), and the Wellcome Trust (FC001097). DG was supported by the Department of Biotechnology (DBT), Government of India, grant BT/BI/25/066/2012, KGLR was supported by the National Institutes of Allergy and Infectious Diseases and the National Institutes of Health (grants R01 AI06775, R01 AI136511 and R21 AI142506-01) and the University of California, Riverside (NIFA-Hatch-225935).

## Author contribution

RT, AAH and KGLR conceived and designed all experiments. RT, RP, SA, MB, RJW, MZ, ER, AF, DB, ED and SW performed the GFP tagging and conditional knockdown experiments; RRS performed liver stage imaging; RP, MZ, ED and RT performed protein pulldown experiments; ARB performed the mass spectroscopy; RP and DG performed the phylogenetic analysis and molecular dynamics; SA, RP XML, GB, TH, KGLR and RT performed the RNA-seq and ChIP-Seq experiments; RP, SA, AAH, KGLR and RT analyzed the data; RP, SA, AAH, KGLR and RT wrote the manuscript and all others contributed to it.

## Declaration of Interests

The authors declare no competing interests.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## References

Anderson, D.E., Losada, A., Erickson, H.P., and Hirano, T. (2002). Condensin and cohesin display different arm conformations with characteristic hinge angles. J Cell Biol *156*, 419-424.

Arnot, D.E., Ronander, E., and Bengtsson, D.C. (2011). The progression of the intra-erythrocytic cell cycle of Plasmodium falciparum and the role of the centriolar plaques in asynchronous mitotic division during schizogony. Int J Parasitol *41*, 71-80.

Aurrecoechea, C., Brestelli, J., Brunk, B.P., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O.S., Heiges, M.*, et al.* (2010). EuPathDB: a portal to eukaryotic pathogen databases. Nucleic Acids Res *38*, D415-419.

Ay, F., Bunnik, E.M., Varoquaux, N., Bol, S.M., Prudhomme, J., Vert, J.P., Noble, W.S., and Le Roch, K.G. (2014). Three-dimensional modeling of the P. falciparum genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. Genome Res *24*, 974-988.

Bachellier-Bassi, S., Gadal, O., Bourout, G., and Nehrbass, U. (2008). Cell cycle-dependent kinetochore localization of condensin complex in Saccharomyces cerevisiae. J Struct Biol *162*, 248-259.

Bahl, A., Brunk, B., Coppel, R.L., Crabtree, J., Diskin, S.J., Fraunholz, M.J., Grant, G.R., Gupta, D., Huestis, R.L., Kissinger, J.C.*, et al.* (2002). PlasmoDB: the Plasmodium genome resource. An integrated database providing tools for accessing, analyzing and mapping expression and sequence data (both finished and unfinished). Nucleic Acids Res *30*, 87-90.

Bunnik, E.M., Cook, K.B., Varoquaux, N., Batugedara, G., Prudhomme, J., Cort, A., Shi, L., Andolina, C., Ross, L.S., Brady, D.*, et al.* (2018). Changes in genome organization of parasite-specific gene families during the Plasmodium transmission stages. Nat Commun *9*, 1910.

Bunnik, E.M., Venkat, A., Shao, J., McGovern, K.E., Batugedara, G., Worth, D., Prudhomme, J., Lapp, S.A., Andolina, C., Ross, L.S.*, et al.* (2019). Comparative 3D genome organization in apicomplexan parasites. Proc Natl Acad Sci U S A *116*, 3183-3192.

Bushell, E., Gomes, A.R., Sanderson, T., Anar, B., Girling, G., Herd, C., Metcalf, T., Modrzynska, K., Schwach, F., Martin, R.E.*, et al.* (2017). Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes. Cell *170*, 260-272 e268.

Bushell, E.S., Ecker, A., Schlegelmilch, T., Goulding, D., Dougan, G., Sinden, R.E., Christophides, G.K., Kafatos, F.C., and Vlachou, D. (2009). Paternal effect of the nuclear formin-like protein MISFIT on Plasmodium development in the mosquito vector. PLoS Pathog *5*, e1000539.

Chang, L.F., Zhang, Z., Yang, J., McLaughlin, S.H., and Barford, D. (2014). Molecular architecture and mechanism of the anaphase-promoting complex. Nature *513*, 388-393.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Collette, K.S., Petty, E.L., Golenberg, N., Bembenek, J.N., and Csankovszki, G. (2011). Different roles for Aurora B in condensin targeting during mitosis and meiosis. J Cell Sci *124*, 3684-3694.

Cuylen, S., Metz, J., and Haering, C.H. (2011). Condensin structures chromosomal DNA through topological links. Nat Struct Mol Biol *18*, 894-901.

Fisher, D., Krasinska, L., Coudreuse, D., and Novak, B. (2012). Phosphorylation network dynamics in the control of cell cycle transitions. J Cell Sci *125*, 4703-4711.

Francia, M.E., Dubremetz, J.F., and Morrissette, N.S. (2015). Basal body structure and composition in the apicomplexans Toxoplasma and Plasmodium. Cilia *5*, 3.

Francia, M.E., and Striepen, B. (2014). Cell division in apicomplexan parasites. Nat Rev Microbiol *12*, 125-136.

Fujiwara, T., Tanaka, K., Kuroiwa, T., and Hirano, T. (2013). Spatiotemporal dynamics of condensins I and II: evolutionary insights from the primitive red alga Cyanidioschyzon merolae. Molecular biology of the cell *24*, 2515-2527.

Gerald, N., Mahajan, B., and Kumar, S. (2011). Mitosis in the human malaria parasite Plasmodium falciparum. Eukaryot Cell *10*, 474-482.

Guttery, D.S., Ferguson, D.J., Poulin, B., Xu, Z., Straschil, U., Klop, O., Solyakov, L., Sandrini, S.M., Brady, D., Nieduszynski, C.A.*, et al.* (2012a). A putative homologue of CDC20/CDH1 in the malaria parasite is essential for male gamete development. PLoS Pathog *8*, e1002554.

Guttery, D.S., Holder, A.A., and Tewari, R. (2012b). Sexual development in Plasmodium: lessons from functional analyses. PLoS Pathog *8*, e1002404.

Guttery, D.S., Poulin, B., Ramaprasad, A., Wall, R.J., Ferguson, D.J., Brady, D., Patzewitz, E.M., Whipple, S., Straschil, U., Wright, M.H.*, et al.* (2014). Genome-wide functional analysis of Plasmodium protein phosphatases reveals key regulators of parasite development and differentiation. Cell Host Microbe *16*, 128-140.

Guttinger, S., Laurell, E., and Kutay, U. (2009). Orchestrating nuclear envelope disassembly and reassembly during mitosis. Nat Rev Mol Cell Biol *10*, 178-191.

Hammarton, T.C. (2007). Cell cycle regulation in Trypanosoma brucei. Molecular and biochemical parasitology *153*, 1-8.

Harashima, H., Dissmeyer, N., and Schnittger, A. (2013). Cell cycle control across the eukaryotic kingdom. Trends Cell Biol *23*, 345-356.

Hirano, T. (2016). Condensin-Based Chromosome Organization from Bacteria to Vertebrates. Cell *164*, 847-857.

Hirota, T., Gerlich, D., Koch, B., Ellenberg, J., and Peters, J.M. (2004). Distinct functions of condensin I and II in mitotic chromosome assembly. J Cell Sci *117*, 6435-6445.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Hliscs, M., Sattler, J.M., Tempel, W., Artz, J.D., Dong, A., Hui, R., Matuschewski, K., and Schuler, H. (2010). Structure and function of a G-actin sequestering protein with a vital role in malaria oocyst development inside the mosquito vector. J Biol Chem *285*, 11572-11583.

Howard-Till, R., and Loidl, J. (2018). Condensins promote chromosome individualization and segregation during mitosis, meiosis, and amitosis in Tetrahymena thermophila. Molecular biology of the cell *29*, 466-478.

Howard-Till, R., Tian, M., and Loidl, J. (2019). A specialized condensin complex participates in somatic nuclear maturation in Tetrahymena thermophila. Molecular biology of the cell, mbcE18080487.

Iwanaga, S., Kato, T., Kaneko, I., and Yuda, M. (2012). Centromere plasmid: a new genetic tool for the study of Plasmodium falciparum. PLoS One *7*, e33326.

Iwasaki, O., and Noma, K.I. (2016). Condensin-mediated chromosome organization in fission yeast. Curr Genet *62*, 739-743.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc *10*, 845-858.

Kim, J.H., Shim, J., Ji, M.J., Jung, Y., Bong, S.M., Jang, Y.J., Yoon, E.K., Lee, S.J., Kim, K.G., Kim, Y.H.*, et al.* (2014). The condensin component NCAPG2 regulates microtubule-kinetochore attachment through recruitment of Polo-like kinase 1 to kinetochores. Nat Commun *5*, 4588.

Kim, K.D., Tanizawa, H., Iwasaki, O., and Noma, K. (2016). Transcription factors mediate condensin recruitment and global chromosomal organization in fission yeast. Nat Genet *48*, 1242-1252.

Kschonsak, M., Merkel, F., Bisht, S., Metz, J., Rybin, V., Hassler, M., and Haering, C.H. (2017). Structural Basis for a Safety-Belt Mechanism That Anchors Condensin to Chromosomes. Cell *171*, 588-600 e524.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Lenne, A., De Witte, C., Tellier, G., Hollin, T., Aliouat, E.M., Martoriati, A., Cailliau, K., Saliou, J.M., Khalife, J., and Pierrot, C. (2018). Characterization of a Protein Phosphatase Type-1 and a Kinase Anchoring Protein in Plasmodium falciparum. Front Microbiol *9*, 2617.

Lynch, D.B., Logue, M.E., Butler, G., and Wolfe, K.H. (2010). Chromosomal G + C content evolution in yeasts: systematic interspecies differences, and GC-poor troughs at centromeres. Genome Biol Evol *2*, 572-583.

Modrzynska, K., Pfander, C., Chappell, L., Yu, L., Suarez, C., Dundas, K., Gomes, A.R., Goulding, D., Rayner, J.C., Choudhary, J.*, et al.* (2017). A Knockout Screen of ApiAP2 Genes

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Reveals Networks of Interacting Transcriptional Regulators Controlling the Plasmodium Life Cycle. Cell Host Microbe *21*, 11-22.

Neuwald, A.F., and Hirano, T. (2000). HEAT repeats associated with condensins, cohesins, and other complexes involved in chromosome-related functions. Genome Res *10*, 1445-1452.

Oliveira, R.A., Heidmann, S., and Sunkel, C.E. (2007). Condensin I binds chromatin early in prophase and displays a highly dynamic association with Drosophila mitotic chromosomes. Chromosoma *116*, 259-274.

Onn, I., Aono, N., Hirano, M., and Hirano, T. (2007). Reconstitution and subunit geometry of human condensin complexes. The EMBO journal *26*, 1024-1034.

Ono, T., Fang, Y., Spector, D.L., and Hirano, T. (2004). Spatial and temporal regulation of Condensins I and II in mitotic chromosome assembly in human cells. Molecular biology of the cell *15*, 3296-3308.

Ono, T., Yamashita, D., and Hirano, T. (2013). Condensin II initiates sister chromatid resolution during S phase. J Cell Biol *200*, 429-441.

Rawlings, J.S., Gatzka, M., Thomas, P.G., and Ihle, J.N. (2011). Chromatin condensation via the condensin II complex is required for peripheral T-cell quiescence. The EMBO journal *30*, 263-276.

Roques, M., Stanway, R.R., Rea, E.I., Markus, R., Brady, D., Holder, A.A., Guttery, D.S., and Tewari, R. (2019). Plasmodium centrin PbCEN-4 localizes to the putative MTOC and is dispensable for malaria parasite proliferation. Biol Open *8*.

Roques, M., Wall, R.J., Douglass, A.P., Ramaprasad, A., Ferguson, D.J., Kaindama, M.L., Brusini, L., Joshi, N., Rchiad, Z., Brady, D.*, et al.* (2015). Plasmodium P-Type Cyclin CYC3 Modulates Endomitotic Growth during Oocyst Development in Mosquitoes. PLoS Pathog *11*, e1005273.

Sapay, N., and Tieleman, D.P. (2011). Combination of the CHARMM27 force field with united-atom lipid force fields. J Comput Chem *32*, 1400-1410.

Sazer, S., Lynch, M., and Needleman, D. (2014). Deciphering the evolutionary history of open and closed mitosis. Curr Biol *24*, R1099-1103.

Schleiffer, A., Kaitna, S., Maurer-Stroh, S., Glotzer, M., Nasmyth, K., and Eisenhaber, F. (2003). Kleisins: a superfamily of bacterial and eukaryotic SMC protein partners. Mol Cell *11*, 571-575.

Schwach, F., Bushell, E., Gomes, A.R., Anar, B., Girling, G., Herd, C., Rayner, J.C., and Billker, O. (2015). PlasmoGEM, a database supporting a community resource for large-scale experimental genetics in malaria parasites. Nucleic Acids Res *43*, D1176-1182.

Sebastian, S., Brochet, M., Collins, M.O., Schwach, F., Jones, M.L., Goulding, D., Rayner, J.C., Choudhary, J.S., and Billker, O. (2012). A Plasmodium calcium-dependent protein kinase

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

controls zygote development and transmission by translationally activating repressed mRNAs. Cell Host Microbe *12*, 9-19.

Sinden, R.E. (1991a). Asexual blood stages of malaria modulate gametocyte infectivity to the mosquito vector--possible implications for control strategies. Parasitology *103 Pt 2*, 191-196.

Sinden, R.E. (1991b). Mitosis and meiosis in malarial parasites. Acta Leiden *60*, 19-27.

Sinden, R.E., Canning, E.U., and Spain, B. (1976). Gametogenesis and fertilization in Plasmodium yoelii nigeriensis: a transmission electron microscope study. Proc R Soc Lond B Biol Sci *193*, 55-76.

Sinden, R.E., and Hartley, R.H. (1985). Identification of the meiotic division of malarial parasites. J Protozool *32*, 742-744.

Sinden, R.E., and Strong, K. (1978). An ultrastructural study of the sporogonic development of Plasmodium falciparum in Anopheles gambiae. Trans R Soc Trop Med Hyg *72*, 477-491.

Sinden, R.E., Talman, A., Marques, S.R., Wass, M.N., and Sternberg, M.J. (2010). The flagellum in malarial parasites. Curr Opin Microbiol *13*, 491-500.

Smith, S.J., Osman, K., and Franklin, F.C. (2014). The condensin complexes play distinct roles to ensure normal chromosome morphogenesis during meiotic division in Arabidopsis. Plant J *80*, 255-268.

Sutani, T., Yuasa, T., Tomonaga, T., Dohmae, N., Takio, K., and Yanagida, M. (1999). Fission yeast condensin complex: essential roles of non-SMC subunits for condensation and Cdc2 phosphorylation of Cut3/SMC4. Genes Dev *13*, 2271-2283.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol *30*, 2725-2729.

Tewari, R., Straschil, U., Bateman, A., Bohme, U., Cherevach, I., Gong, P., Pain, A., and Billker, O. (2010). The systematic functional analysis of Plasmodium protein kinases identifies essential regulators of mosquito transmission. Cell Host Microbe *8*, 377-387.

Thadani, R., Uhlmann, F., and Heeger, S. (2012). Condensin, chromatin crossbarring and chromosome condensation. Curr Biol *22*, R1012-1021.

Tumova, P., Uzlikova, M., Wanner, G., and Nohynkova, E. (2015). Structural organization of very small chromosomes: study on a single-celled evolutionary distant eukaryote Giardia intestinalis. Chromosoma *124*, 81-94.

Uhlmann, F. (2016). SMC complexes: from DNA to chromosomes. Nat Rev Mol Cell Biol *17*, 399-412.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E., and Berendsen, H.J. (2005). GROMACS: fast, flexible, and free. J Comput Chem *26*, 1701-1718.

Wall, R.J., Ferguson, D.J.P., Freville, A., Franke-Fayard, B., Brady, D., Zeeshan, M., Bottrill, A.R., Wheatley, S., Fry, A.M., Janse, C.J.*, et al.* (2018). Plasmodium APC3 mediates

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

chromosome condensation and cytokinesis during atypical mitosis in male gametogenesis. Sci Rep *8*, 5610.

Ward, P., Equinet, L., Packer, J., and Doerig, C. (2004). Protein kinases of the human malaria parasite Plasmodium falciparum: the kinome of a divergent eukaryote. BMC Genomics *5*, 79.

Wilkes, J.M., and Doerig, C. (2008). The protein-phosphatome of the human malaria parasite Plasmodium falciparum. BMC Genomics *9*, 412.

Yuen, K.C., Slaughter, B.D., and Gerton, J.L. (2017). Condensin II is anchored by TFIIIC and H3K4me3 in the mammalian genome and supports the expression of active dense gene clusters. Sci Adv *3*, e1700191.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

## Legends:

**Figure 1:** Architecture of condensin (SMC2/SMC4) in *Plasmodium berghei.* (A) Composition of two conventional condensin complexes (condensin I and condensin II) which are comprised of heterodimeric core subunits, SMC2 and SMC4, along with non-SMC regulatory subunits, Kleisin and a pair of HEAT subunits specific for either condensin I or II (modified from Hirano T., 2016[8] and Uhlmann F., 2012[53]). (B) Domain architecture of *P. berghei* SMC2 and SMC4. (C) Sequence coverage and amino acid identity of *P. berghei* SMC2 and SMC4 with *H. sapiens*, *S. cerevisiae* and *A. thaliana* proteins. (D) Homology-based predicted three-dimensional structures of *P. berghei* SMC2 and SMC4 showing coiled backbone extension without hinge domain and with ATPase head formation, required for condensin complex. See also Figure S1.

**Figure 2:** Temporal dynamics of condensin (SMC2 and SMC4) in two distinct *Plasmodium* proliferative stages (schizogony and male gametogenesis) undergoing atypical mitotic division. (A-D) Live cell imaging of SMC2GFP and SMC4GFP expressed during schizogony (100X magnification) and male gametogenesis (63X magnification). The white arrows in Panels C and D indicate discrete localization in male gametes. DIC: Differential interference contrast, Merge: Hoechst and GFP. Scale bar = 2µm. (E-F) Immunofluorescence fixed-cell imaging of SMC4GFP and colocalization with antibodies specific for centrin and α-tubulin in mitotic cells (Early schizonts 100X magnification and male gametocytes 63X magnification). The white arrow in Panel F indicates exflagellating male gamete. Scale bar = 2µm. See also Figure S2, S3

**Figure 3:** ChIP-seq analysis of SMC2GFP and SMC4GFP profiles. (A) Genome-wide ChIP-seq signal tracks for SMC2GFP and SMC4GFP for all 14 chromosomes in schizont and gametocyte stages. The localization of previously annotated centromeres is indicated by blue circles. SMC2 and SMC4 proteins bind near the putative centromere in each of the 14 chromosomes (distance from centromere is shown in ±kb). (B) Zoom-in on regions associated with the ChIP-seq peak identified, low GC-content at the centers of peaks shown for chromosome 11 and chromosome 14, which suggest association of the proteins with these newly defined centromeres in the schizont stage. Signals are plotted on a normalized read per million (RPM) basis. See also Table S1.

**Figure 4:** Differential condensin complex formation during schizogony and male gametogenesis, and phylogenetic analysis of kleisin. (A) Co-localization of SMC4GFP (green) and SMC2mCherry (red). Merge: Hoechst, GFP and mCherry. Scale bar = 2µm. (B) Venn diagram displays the unique and shared proteins in the condensin complex of schizonts and gametocytes. Analysis of SMC2GFP and SMC4GFP protein complexes by tryptic

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

digestion and liquid chromatography-mass spectrometry following GFP-specific immunoprecipitation from a lysate of schizonts maintained in culture for 8 hour and gametocytes activated for 6 min. The list of all identified proteins is provided as Table S2. (C) Different domain architecture for subunits of Condensin I and Condensin II complexes. Schematic figure displays domain composition and protein length in the respective complex subunits. (D) Maximum likelihood phylogeny based on the alignment of kleisin subunits from apicomplexan species (*Plasmodium* spp, *Toxoplasma gondii*, *Cryptosporidium parvum, Babesia bovis* and *Eimeria tenella)* and other selected organisms. Topological support from bootstrapping is shown at the nodes. The protein sequences for selected organisms have been provided in File S1. (E) Distribution of condensin components across Apicomplexa and other organisms. Presence (filled circle) or absence (empty circle) of condensin complex genes in each genome. * represent 4 *Plasmodium* spp., namely, *P. falciparum, P. vivax, P. berghei* and *P. yoelii*; # denotes *H. sapiens, A. thaliana* and *D. melanogaster*. See also Table S2 and File S1.

**Figure 5:** Global transcriptomic analysis for SMC4PTD in activated gametocytes by RNA-seq. (A) Confirmation of successful depletion of SMC4 transcript in the SMC4PTD line. (B) Log-normalized scatterplots demonstrating high correlation between replicates for genome-wide expression. Spearman correlation calculated using read counts normalized by number of mapped reads per million. (C) MA plot summarizing RNA-seq results. M: Log ratio and A: Mean average. Every gene is placed according to its log fold expression change in the SMC4PTD line as compared to the WT line (y-axis), and average expression level across replicates of both lines (x-axis). Red color indicates statistical significance of differential expression at the false positive threshold of 0.05. 105 genes are downregulated in the SMC4PTD line, and 5 genes are upregulated. (D) GO enrichment analysis of genes with $\log_{10}$ fold expression change of -0.5 or lower in the SMC4PTD line. (E) qRT-PCR analysis of selected genes identified as down-regulated in (C), comparing transcript levels in WT and SMC4PTD samples. Error bar = ±SEM, n= 3. Unpaired t-test was performed for statistical analysis: *p<0.05 **p<0.001, ***p<0.0001 and ****p<0.00001. See also Table S3, S5 and Figure S4.

**Figure 6:** Phenotypic analysis of conditional gene expression knockdown in SMC2PTD and SMC4PTD transgenic lines at various proliferative stages during life cycle. (A) Average number of nuclei per schizont (mitotic division within red cell). N = 5 (minimum 100 cells). (B) Number of exflagellation centers (mitotic division during male gametogenesis) per field at 15 minutes post-activation. N = 3 independent experiments (10 fields per experiment) (C) Percentage ookinete conversion from zygotes. Minimum of 3 independent experiments

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(minimum 100 cells) (D) Live cell imaging of WTGFP, SMC2PTD and SMC4PTD oocysts (endomitosis in parasite within mosquito gut) at 9, 14 and 21 days post infection, using 10X and 63X magnification to illustrate differences in size and frequency. Scale bar = 5µm (63X) and 20µm (10X). (E) Number of oocysts at 9, 14 and 21 days post-infection (dpi). N = 3 independent experiments with a minimum of 5 mosquito guts. (F) Oocyst diameter at 9, 14 and 21 days post-infection (dpi). N = 3 independent experiments. (G) Number of sporozoites at 14 and 21 dpi in mosquito gut. N = 3 independent experiments with a minimum of 5 mosquito guts. (H) Number of sporozoites at 21 dpi in mosquito salivary gland. Minimum of 3 independent experiments. Error bar = ±SEM. Unpaired t-test was performed for statistical analysis: *p<0.05 **p<0.01 and ***p<0.001. See also Table S4 and Figure S4.

**Figure S1:** Molecular dynamics simulation showing stable predicted 3D structure for SMC2 and SMC4. Related to Figure 1. Root mean square deviation (RMSD) calculation using protein backbone structure, during the 10ns production simulation for SMC2 (A) and SMC4 (B), respectively. Post 8ns and 2ns MD simulation, RMSD fluctuations becomes comparable and constant for SMC2 and SMC4 respectively. (C and D) Radius of gyration fluctuations within 2Å suggest correct and stable protein fold for SMC2 and SMC4, respectively. Additionally, during a 10ns molecular dynamics run the protein structure did not break, confirming stable predicted 3D structure.

**Figure S2:** Generation and genotype analysis of SMC2GFP and SMC4GFP parasite lines. Related to Figure 2. (A) Schematic representation of the endogenous *smc(2/4)*, the GFP-tagging construct and the recombined *smc(2/4)* locus following single homologous recombination. Arrows 1, 2 and 3 indicate the position of PCR primers used to confirm successful integration of the construct. (B) Diagnostic PCR of SMC2GFP, SMC4GFP and WT parasites using primers IntT138 (SMC2, Arrow 1), IntT143 (SMC4, Arrow 1) and ol492 (Arrow 3). IntT138 and T1382 (SMC2, Arrow 2), IntT143 and T1432 (SMC4, Arrow 2) primers were used as control. Integration of the SMC tagging construct gives a band of 995 bp and 1006 bp for SMC2GFP and SMC4GFP parasite lines. (C) Western blot of SMC2GFP (172 kDa), SMC4GFP (198kDa) and WTGFP (29 kDa) protein to illustrate SMC2GFP and SMC4GFP in schizont stage extracts.

**Figure S3:** Localisation of SMC4GFP throughout the *Plasmodium* life cycle as detected by live cell imaging. Related to Figure 2. (A) Asexual blood stages and sexual stages at different time points. Merge: Hoechst (blue, DNA), GFP (green) and P28 (red, cell surface marker during female gamete activation, zygote and ookinete stages). (B) Sporogony in the mosquito oocyst; 9 days post infection (dpi), 14 dpi, 21 dpi and mature single sporozoite at 21

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

dpi. Merge: Hoechst and GFP. (C) Liver stage schizont at 60 hours post infection. Merge: DAPI (blue) and GFP (green). Scale bar = 2µM.

**Figure S4.** Generation of knockout and conditional knockdown of SMC2 and SMC4 genes using a *dhfr* drug-selectable marker or *ama1* promoter trap double homologous recombination (PTD), and genotype analysis. Related to Figure 6. (A) Schematic representation of the endogenous *smc(2/4)* locus, the targeting gene deletion construct and the recombined *smc(2/4)* locus following double homologous recombination. (B) Schematic representation of the promoter trap strategy (SMC2PTD and SMC4PTD), placing *smc(2/4)* under the control of the blood stage *ama1* promoter by double homologous recombination. Arrows 1 and 2 indicate the primers position used to confirm 5' integration and arrows 3 and 4 indicate the primers used for 3' integration. Primers 1 and 4 were also used for Knock-In PCR. (C) Integration PCR of the promotor trap construct into the *smc(2/4)* locus. Primer 1 (5'-IntPTD18 [SMC2], 5'-IntPTD008 [SMC4]) with primer 2 (5'-IntPTD) were used to determine successful 5' integration of the selectable marker resulting in a band of 1518 and 1655 bp for SMC2PTD and SMC4PTD, respectively. Primer 3 (3'-intPTama1) and primer 4 (3'-IntPTD18 (SMC2) and 3'-IntPTD008 (SMC4)) were used to determine the successful 3' integration of *ama1* promoter resulting in a band of 1004 bp and 877 bp for SMC2PTD and SMC4PTD, respectively. Primer 1 (5'-IntPTD18 and 5'-IntPTD008) and primer 4 (3'-IntPTD18 and 3'-IntPTD008) were used to show complete knock-in of the construct with a band at 4698 bp (SMC2PTD) and 4653 bp (SMC4PTD), and the absence of a band at 1538 bp (for SMC2, endogenous) and 1493 bp (for SMC4, endogenous) resulting in complete knock-in of the construct.

**Table S1**: Experimentally validated coordinates of the *P. berghei* centromeres. Related to Figure 3.

**Table S2**: List of main protein hits in the SMC2GFP and SMC4GFP co-immunoprecipitation experiments. Related to Figure 4.

**Table S3**: List of differentially expressed genes between SMC4PTD and WT activated gametocytes. Related to Figure 5.

**Table S4:** Mosquito bite back analysis of WTGFP, SMC2PTD and SMC4PTD parasites. Related to Figure 6.

**Table S5**: Primers used in this study (5' – 3'). Related to Figure 2, Figure 5 and Figure 6.

**File S1**: Protein sequences used for phylogenetic analysis. Related to Figure 4.

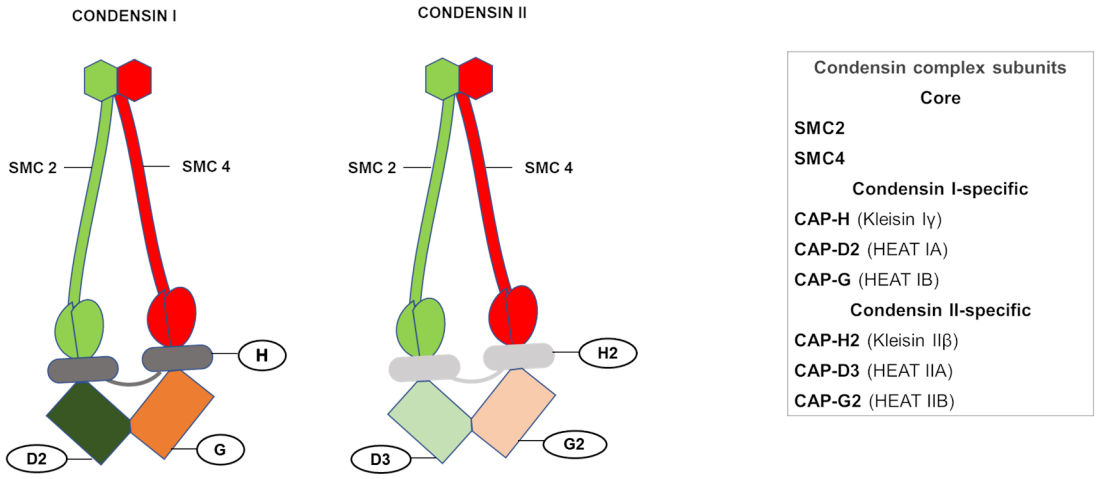
**Video S1:** Ookinete motility assay for SMC4PTD parasites. Related to Figure 6.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

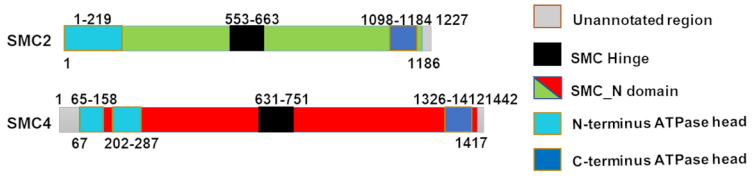
**Video S2:** Ookinete motility assay for WT parasites. Related to Figure 6.

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**A**



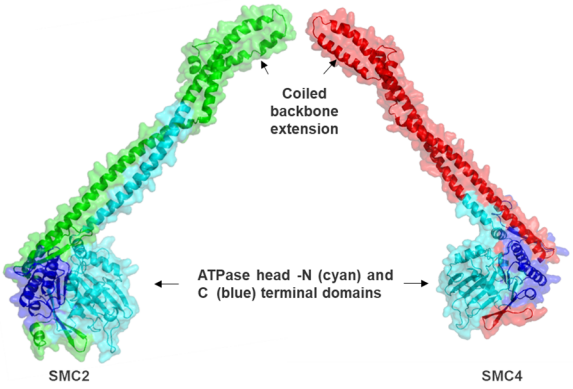

**B**



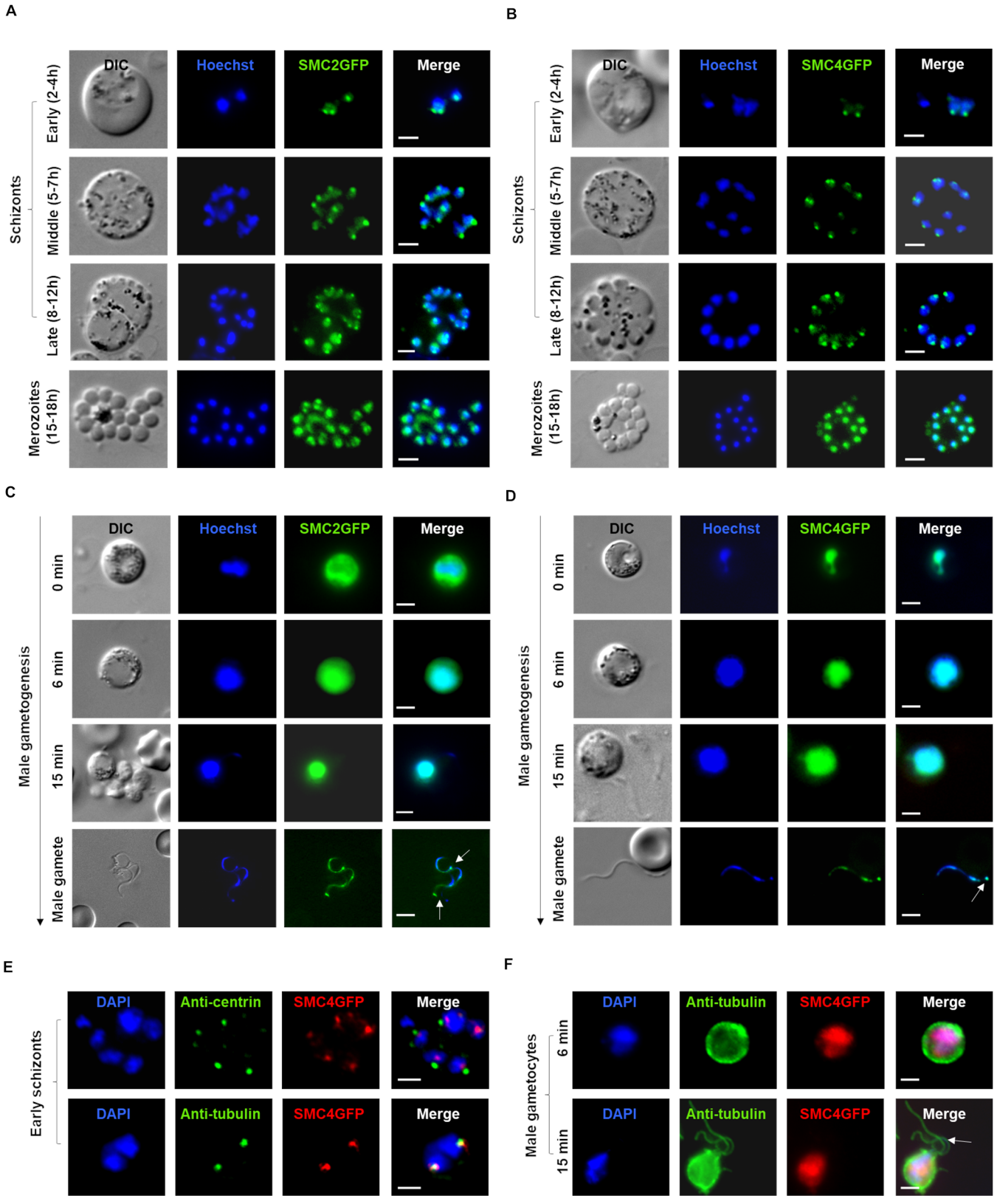

**C**

| *P. berghei* SMC | Protein Length | Sequence coverage / amino acid Identity (%) | | |
|---|---|---|---|---|
| | | *H. sapiens* | *S. cerevisiae* | *A. thaliana* |
| SMC2 (PBANKA_1416900) | 1227 | 98/34 | 97/34 | 97/34 |
| SMC4 (PBANKA_1108700) | 1442 | 87/29 | 94/31 | 85/65 |

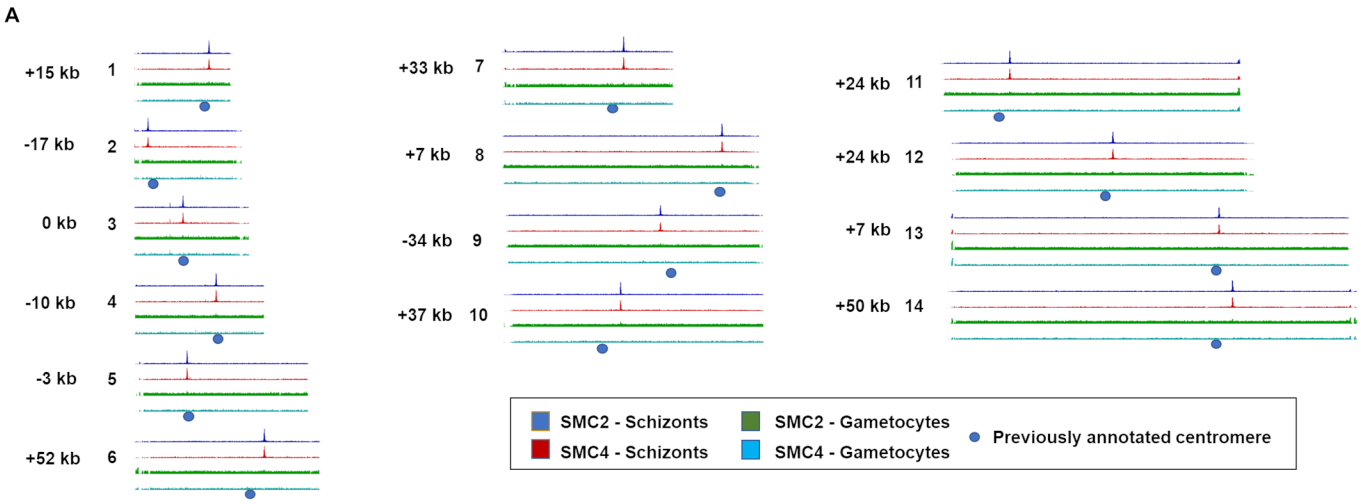
**D**

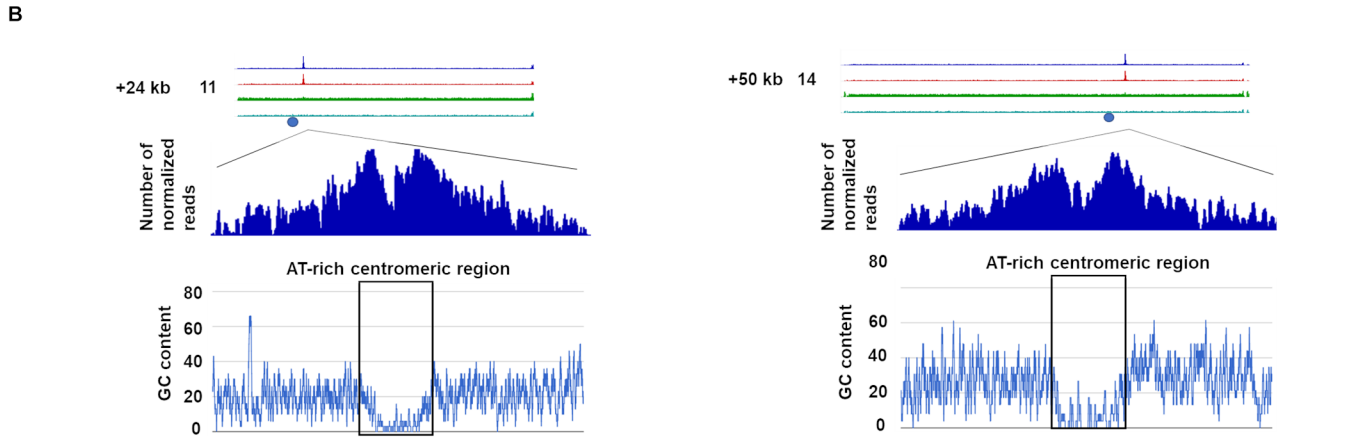


bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

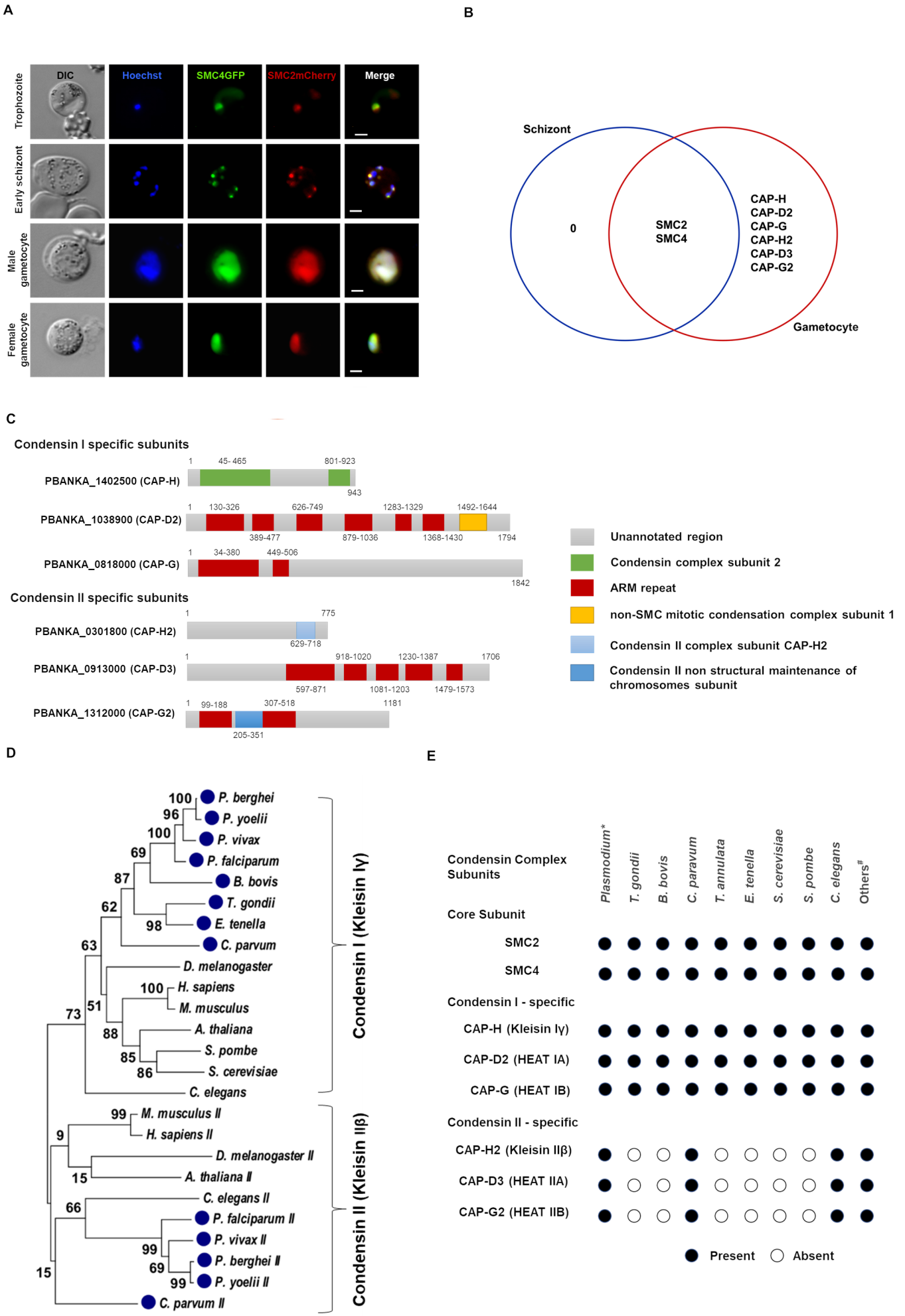


bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

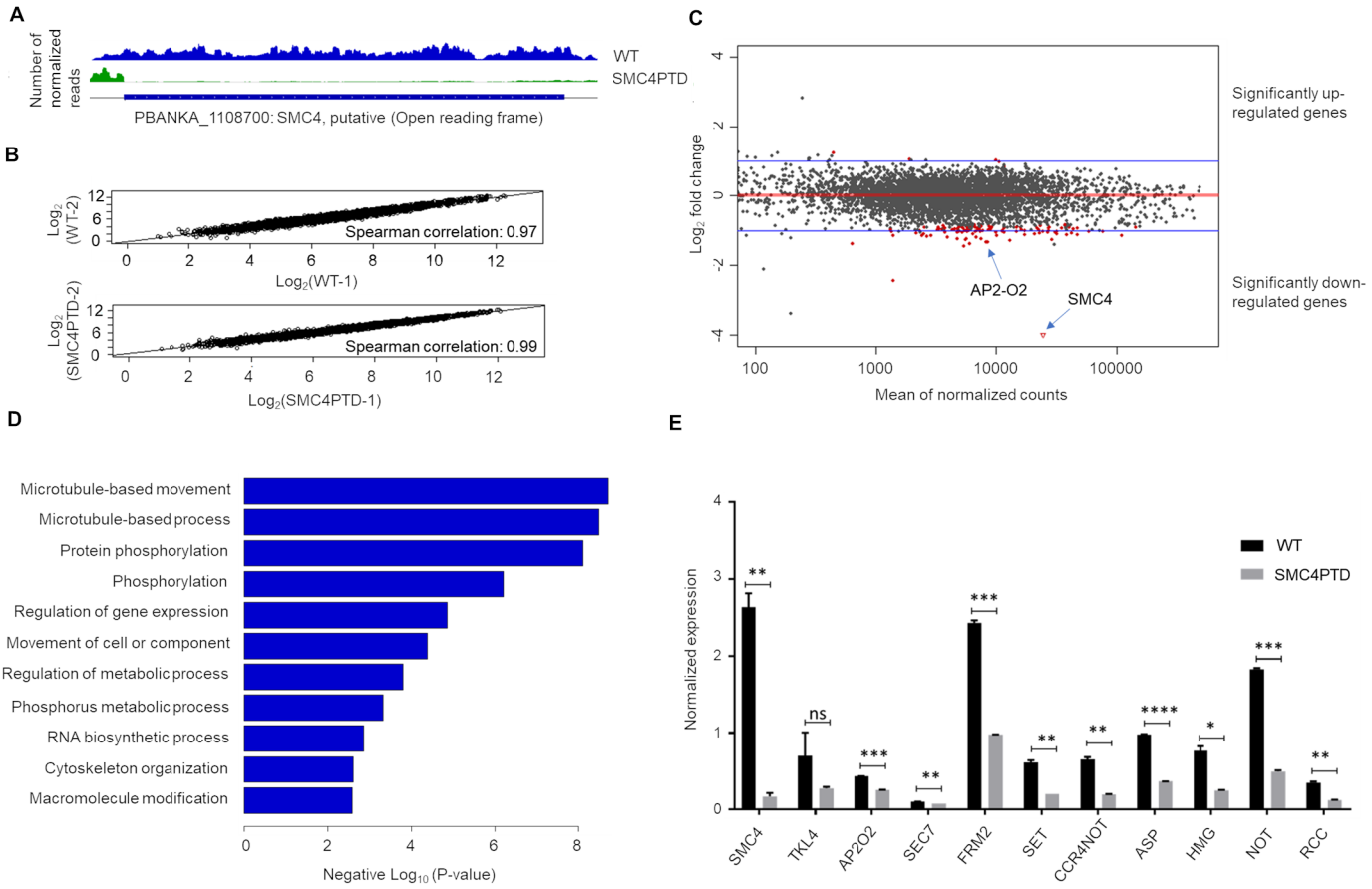


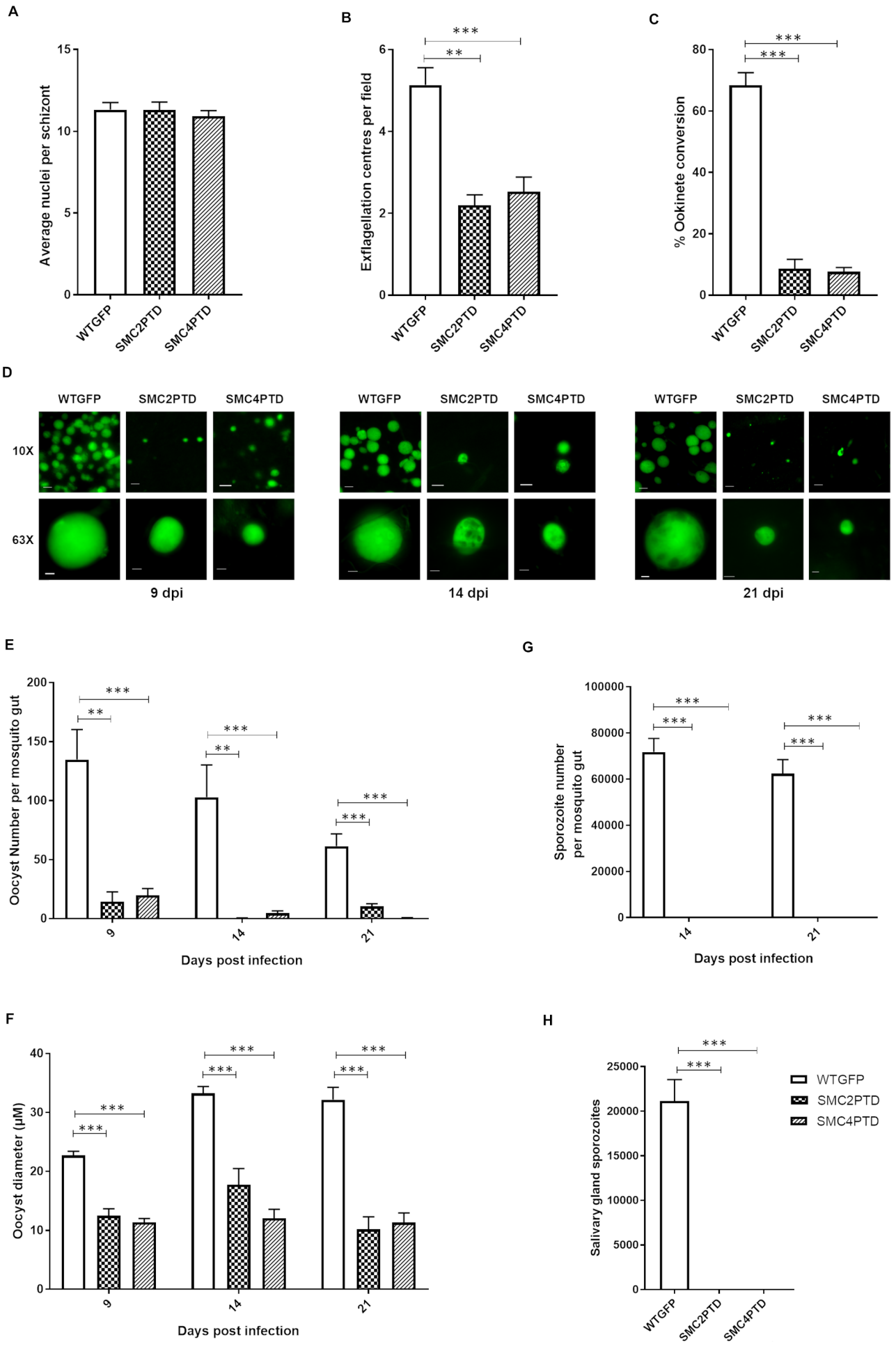
bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.



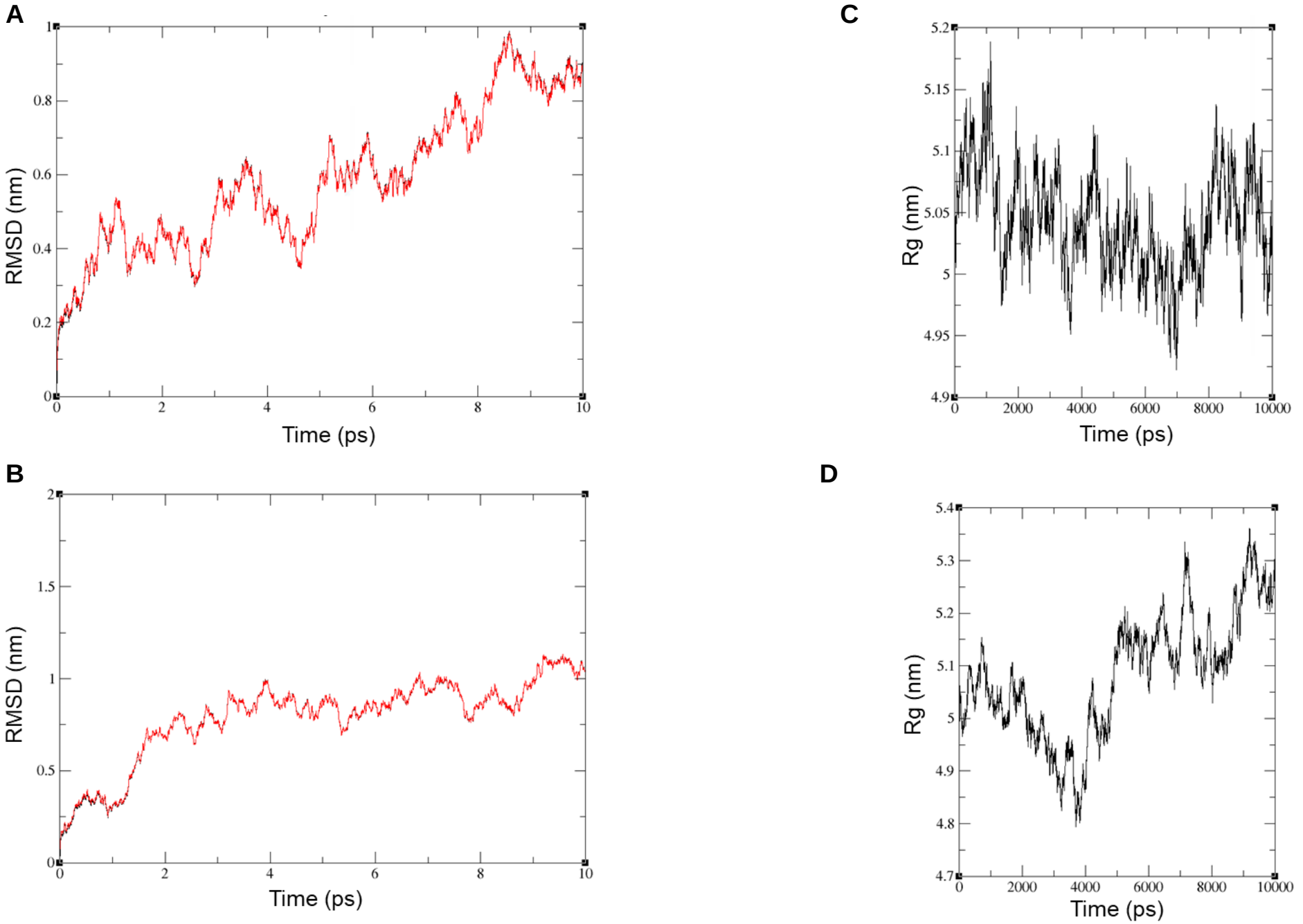



**Figure S1**

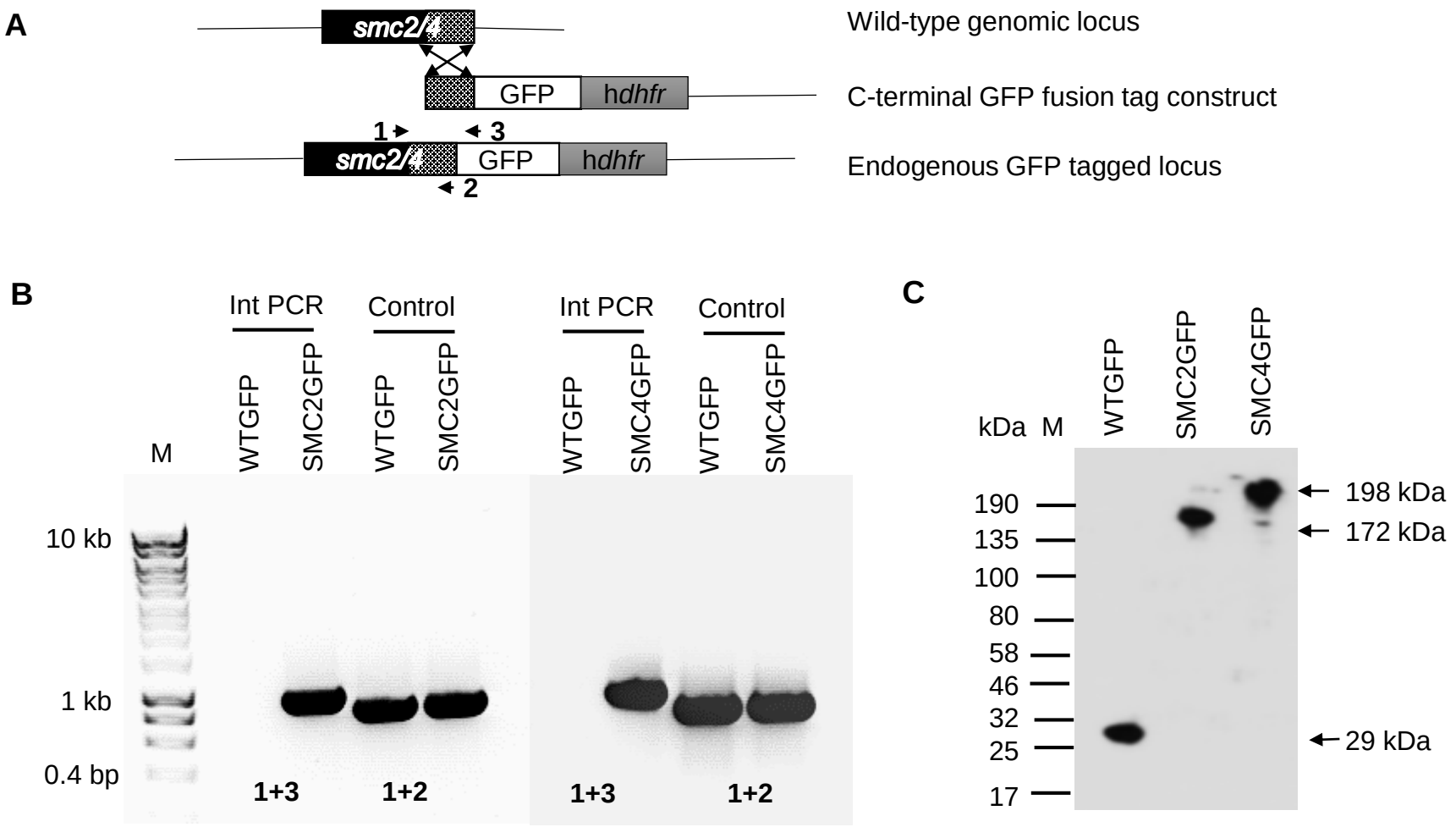

A
smc2/4
GFP
hdhfr
Wild-type genomic locus
C-terminal GFP fusion tag construct
1
3
2
Endogenous GFP tagged locus
B
Int PCR
Control
Int PCR
Control
M
WTGFP
SMC2GFP
WTGFP
SMC2GFP
WTGFP
SMC4GFP
WTGFP
SMC4GFP
10 kb
1 kb
0.4 bp
1+3
1+2
1+3
1+2
C
kDa
M
WTGFP
SMC2GFP
SMC4GFP
190
135
100
80
58
46
32
25
17
198 kDa
172 kDa
29 kDa

**Figure S2**

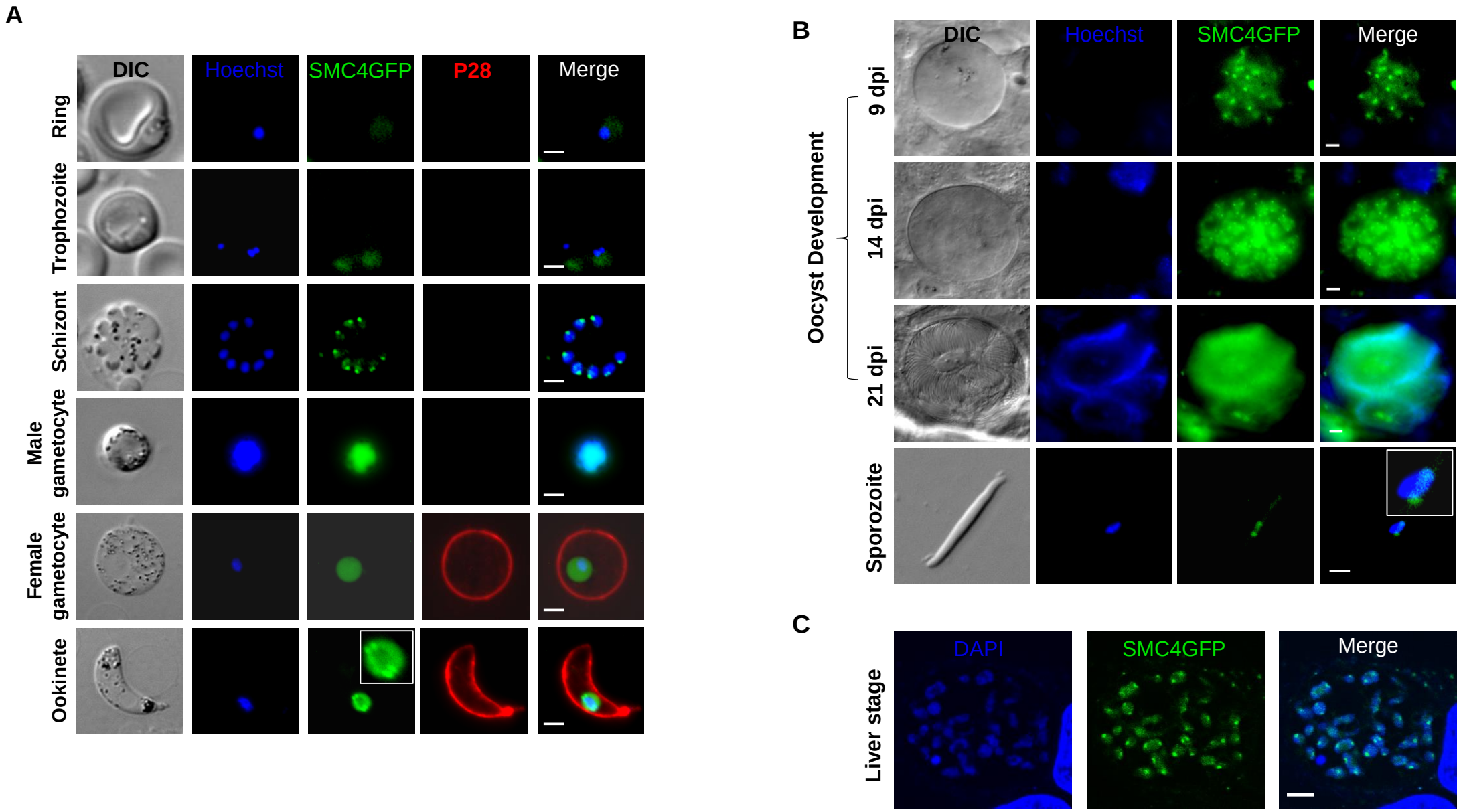

A
DIC
Hoechst
SMC4GFP
P28
Merge
Ring
Trophozoite
Schizont
Male gametocyte
Female gametocyte
Ookinete
B
DIC
Hoechst
SMC4GFP
Merge
Oocyst Development
9 dpi
14 dpi
21 dpi
Sporozoite
C
DAPI
SMC4GFP
Merge
Liver stage

**Figure S3**

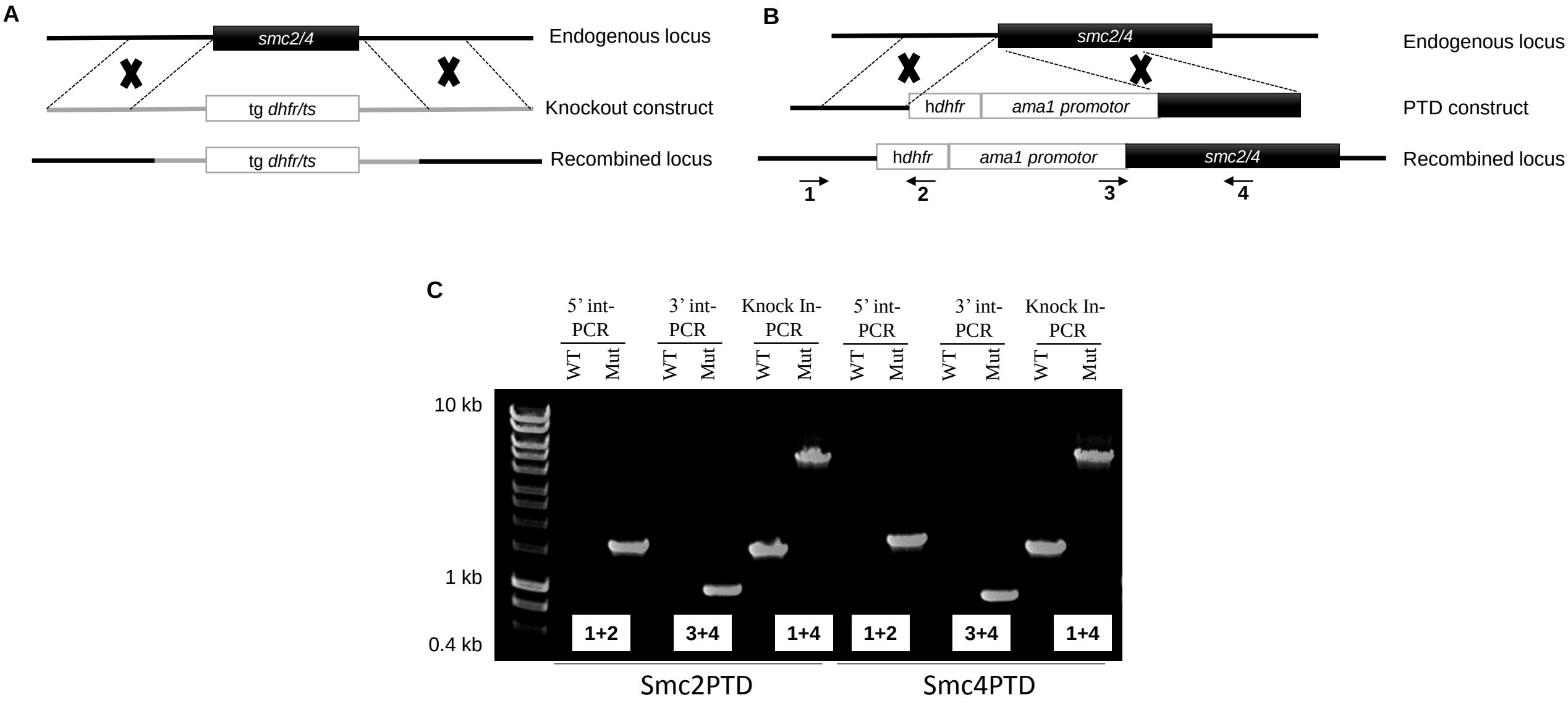

A
smc2/4
Endogenous locus
tg dhfr/ts
Knockout construct
tg dhfr/ts
Recombined locus
B
smc2/4
Endogenous locus
hdhfr
ama1 promotor
PTD construct
hdhfr
ama1 promotor
smc2/4
Recombined locus
1
2
3
4
C
5' int-PCR
3' int-PCR
Knock In-PCR
5' int-PCR
3' int-PCR
Knock In-PCR
WT
Mut
WT
Mut
WT
Mut
WT
Mut
WT
Mut
WT
Mut
10 kb
1 kb
0.4 kb
1+2
3+4
1+4
1+2
3+4
1+4
Smc2PTD
Smc4PTD

**Figure S4**

| Experiment | | Bite back (dpi) |
|---|---|---|
| 1 | WTGFP | 4 |
| | SMC2PTD cl 5 | - |
| | SMC4PTD cl 3 | - |
| 2 | WTGFP | 4 |
| | SMC2PTD cl 5 | - |
| | SMC4PTD cl 3 | - |
| 3 | WTGFP | 4 |
| | SMC2PTD cl 5 | - |
| | SMC4PTD cl 3 | - |

**Table S4**

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

**File S1**

>sp|G5EGE9|DPY26_CAEEL Condensin complex subunit dpy-26 OS=Caenorhabditis elegans OX=6239 GN=dpy-26 PE=1 SV=1

MDVPSSSNVTGRRKRQVLDDDEDDGFRSTPLRKVRGTKKIRPADVVPETIMTKIGAHIDD

IVNKKKVGELNCFEYKSPLEIHTIEDMIKAKASIQEMAVVLEGAQCIIGYRVDRLHHDVR

QIDSALSSGTVMRDSNGEEIHLTLESRKAKKKMAVVDGMNGMLDFLNNMDDALTTTELDA

DNDKNWKEDEENIAGEPRIDFKANSKDVDAFLQRDIFPEKLIYALSIKRATDLRADLLSD

VSNYISADDTAHDLKDANIDWLRANPTFQKATKGSVCNSSNSFHSLNYYGIHSPDGRTLM

LHNRIADKNADDRFFTSDVSVSLVKNTRALLTNSLDKKPRILDNYLMLEVKDRPVIGRYK

IMSKDVKKSTLPLAESSREKDLANLTFAEMNHRPSNLDMTVAGASDMSMLPGNQGLPLAQ

GENDETIALDRLTPPLQSSVSQKASDEYVLPPLEANDLDEHLIGKLPIEPNEMDQTLANM

FDKKLEVFNTSDTLESKVWKNGIRAEEWGEDDEAIMKNDTKHPRQAGIEGWIKATDAWTN

YDVVKMNVNREARSQLDENAIDEQESYRNMVPEIGKNLFLVKSDDYMNNYPGDRPADFTV

NDEVSDVMKMWSGEDSTAEDDVPLEQIQQEIREQVQQQDVDMEPIEEMDYDAGGAAFDV
D

FDDRLAAPVEVEEMEGDNNRNDGRVADILFNEQMDETEVEERNEQDVQRELEDIALAADE

VAELMTSAPPPQLVGPSAEMREEIQNIGKNDNAHWVPPVVGDQERQAAVTAQRKRREKKA

KSRKATVEDFVHYFRDIPDDEIEREITAAKCSKIADEKSTFLSEQQLYLPTLGIENKPHV

AFEMGLLGNSGMFFKKSYGKIRLERVKNQKAEQDLFIDEARGNKDSDCLNWLLSFSGFRC

MENPEPITGSDSDENLRTAVEQPFDDDFANDYYDEDRYDPNYEQQLAAAQMGPDMQRKL
A

LTASHINQMFPNIHSKRYGGEYGDSDDEFDDSFDRQSIQAKNLDAAKHKKCLAEILKTDS

LSMPSIQYVLEQLTSNQTLRMNNTTIRAADDRNETGRPATPTMEADKTLTSVFDYRSPNK

SNHDVNETMKALTEMPDYQAADERPNNQPTTSTYGTANTENRKVHINGCHTLLSLALSMP

SRMGETVRPSSIVSFLLHIANENNLQIVQDRSKRSWMSDFIVLNSSESLPRGLKMGRIED

QDEFWKRTQDPDAIEGTASDANNVFSNLMRRPKAVPVRKGRGAGGQPTTSDLGAIVEEEE

MEE

>sp|Q15003|CND2_HUMAN Condensin complex subunit 2 OS=Homo sapiens OX=9606 GN=NCAPH PE=1 SV=3

MGPPGPALPATMNNSSSETRGHPHSASSPSERVFPMPLPRKAPLNIPGTPVLEDFPQNDD

EKERLQRRRSRVFDLQFSTDSPRLLASPSSRSIDISATIPKFTNTQITEHYSTCIKLSTE

NKITTKNAFGLHLIDFMSEILKQKDTEPTNFKVAAGTLDASTKIYAVRVDAVHADVYRVL

GGLGKDAPSLEEVEGHVADGSATEMGTTKKAVKPKKKHLHRTIEQNINNLNVSEADRKCE

IDPMFQKTAASFDECSTAGVFLSTLHCQDYRSELLFPSDVQTLSTGEPLELPELGCVEMT

DLKAPLQQCAEDRQICPSLAGFQFTQWDSETHNESVSALVDKFKKNDQVFDINAEVDESD

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

CGDFPDGSLGDDFDANDEPDHTAVGDHEEFRSWKEPCQVQSCQEEMISLGDGDIRTMCP
L

LSMKPGEYSYFSPRTMSMWAGPDHWRFRPRRKQDAPSQSENKKKSTKKDFEIDFEDDIDF

DVYFRKTKAATILTKSTLENQNWRATTLPTDFNYNVDTLVQLHLKPGTRLLKMAQGHRVE

TEHYEEIEDYDYNNPNDTSNFCPGLQAADSDDEDLDDLFVGPVGNSDLSPYPCHPPKTAQ

QNGDTPEAQGLDITTYGESNLVAEPQKVNKIEIHYAKTAKKMDMKKLKQSMWSLLTALSG

KEADAEANHREAGKEAALAEVADEKMLSGLTKDLQRSLPPVMAQNLSIPLAFACLLHLAN

EKNLKLEGTEDLSDVLVRQGD

>sp|Q9Y7R3|CND2_SCHPO Condensin complex subunit 2 OS=Schizosaccharomyces pombe (strain 972 / ATCC 24843) OX=284812 GN=cnd2 PE=1 SV=1

MKRASLGGHAPVSLPSLNDDALEKKRAKENSRKQRELRRSSALHSITPRRESLNNSSPFN

SSHQVPVLSNFEEWIKLATDNKINSTNTWNFALIDYFHDMSLLRDGEDINFQKASCTLDG

CVKIYTSRIDSVATETGKLLSGLANDSKVLQQTEEGEDAENDDEDLQKKKERKRAQRSVK

TLVKDFESIRAKKFELECSFDPLFKKMCADFDEDGAKGLLMNHLCVDQHGRIVFDSSDTV

IKDLENKDVEAESQEAVVAAPIESHDTEMTNVHDNISRETLNGIYKCYFTDIDQLTICPS

LQGFEFDSKGNLDVSLLKSLSDEVNMITTTSLVDNTMEKTDADAASLSSDSDGEEGHIVH

ALEEMAYDEENPYVDVVPKAMDESENPDFGVDTEVNMADGSTMNENYSIISTAAANGVYE

YFDKSMKKNWAGPEHWRIQALRKNINNASTVFNSSNTAESSDNVSRSLSSTERKKRRELD

NAIDFLQEVDVEALFTPATSSLKLPKSHWKRHNRCLLPDDYQYDSKRLLQLFLKPKMSVL

PNADGEGQLQLNKALDDENDLDGIQPHGFDSDGSDNVDEGIPPYGFGDSDSPKQTPLLTP

PSSSGFGDNLLLTARLAKPDMLNYAKRAKKVDVRVLKEKLWKCLDLENTIKENSINSHIE

GSEMESEETNMPVKSFFSTVNQLEETYEKKELKDISTSFAFICVLHLANEHNLELTSNED

FSDVFIRPGPNLTTLEALENDV

>sp|P38170|CND2_YEAST Condensin complex subunit 2 OS=Saccharomyces cerevisiae (strain ATCC 204508 / S288c) OX=559292 GN=BRN1 PE=1 SV=3

MTTQLRYENNDDDERVEYNLFTNRSTMMANFEEWIKMATDNKINSRNSWNFALIDYFYDL

DVLKDGENNINFQKASATLDGCIKIYSSRVDSVTTETGKLLSGLAQRKTNGASNGDDSNG

GNGEGLGGDSDEANIEIDPLTGMPISNDPDVNNTRRRVYNRVLETTLVEFETIKMKELDQ

ELIIDPLFKKALVDFDEGGAKSLLLNTLNIDNTARVIFDASIKDTQNVGQGKLQRKEEEL

IERDSLVDDENEPSQSLISTRNDSTVNDSVISAPSMEDEILSLGMDFIKFDQIAVCEISG

SIEQLRNVVEDINQAKDFIENVNNRFDNFLTEEELQAAVPDNAEDDSDGFDMGMQQELCY

PDENHDNTSHDEQDDDNVNSTTGSIFEKDLMAYFDENLNRNWRGREHWKVRNFKKANLV
N

KESDLLEETRTTIGDTTDKNTTDDKSMDTKKKHKQKKVLEIDFFKTDDSFEDKVFASKGR

TKIDMPIKNRKNDTHYLLPDDFHFSTDRITRLFIKPGQKMSLFSHRKHTRGDVSSGLFEK

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

STVSANHSNNDIPTIADEHFWADNYERKEQEEKEKEQSKEVGDVVGGALDNPFEDDMDGV

DFNQAFEGTDDNEEASVKLDLQDDEDHKFPIRENKVTYSRVSKKVDVRRLKKNVWRSINN

LIQEHDSRKNREQSSNDSETHTEDESTKELKFSDIIQGISKMYSDDTLKDISTSFCFICL

LHLANEHGLQITHTENYNDLIVNYEDLATTQAAS

>sp|Q8C156|CND2_MOUSE Condensin complex subunit 2 OS=Mus musculus OX=10090 GN=Ncaph PE=1 SV=1

MRIPRSETMNSSFLKARGQQDVLSSPLERVPPASRPGKAPLGTPKTPVLEDFPQNDDEKE

RMQRRRSRVFDLQFSTDSIHLASPNRNIDVSTTISKFTNTQITEHYSTCIKLSSENKITT

KNAFGLHLIDFMSEILKQKDAEPTNFKVAAGTLDASTKIYAVRVDAVHADVYRVLGGLGK

DTPPQGEESHSGDGSTLETERTKKPAKPKKKQSCKTIEQNLSNINVSEADGKCAVDPMFQ

KTAASFDECSTTGVFLSTLHCQDYRSELLFPSDMQTLSSGEPLELPDLGFVDMTDLEASL

QQCVEDRPLCPSLAGFQFTKWDSETHNESVSALVDKFKKNDQVFDINAEAEDDEEDVPDG

PLVEDFVDNDEPDLSAAGDHEEFRSWKELCQVQSNQEEVISLEDRDIQVMCSFLSMKPGE

YSYFSPRTMKMWAGPDHWRFRPRPKQDATSCTEHKKKSAKKDFEINFDDDIDFDAYFQKT

KAATILTKSTLENQNWKATTLPTDFHYETDNLIQLHLKPGKRSLKMDQDQKAKTEHYEEI

EDYDYNNPNDTSNYCPGLQAADSDYEEADDLFADPVGTLDLESDPKTTQENGHISPENQG

VDITTYQELNLVAEPQKVNKIEIHYAKTAKKMDMKKLKQSMWSLLTKFSRKEADTEANHT

ESGQEGAPEEVADEKKLSGLTKDLQTRLPPLMAQNLSIPLAFACLLHLANEKNLKLEGTE

DLSDVLVMQGD

>tr|P91663|P91663_DROME Condensin complex subunit 2 OS=Drosophila melanogaster OX=7227 GN=barr PE=2 SV=1

MTLPRLETPLRRSAVGSYQEGVSRMLTPFNDDEAERREARRRTLLQQHHRSSTLESIEDN

ETIKNCLELYNGNKVSKDNAWNLMLIDSLANLLDHHHKRMSNFKMAGSSLEASSKVYGLR

VDSIYLDAMRISAGLSARTLTDKQINAAEDDDGPQGEQATGEGQDSAQQAAKEAAPKPKR

QKKPISTVTKNRETLNSRLDTAPLQDPVFGKLNSTCGASINASNRLMHNILPSFDSELRL

RTTYNFWNSEESTEEVQDHTTLNAEMEQWPATSLMSTNLMRKLLPHAERSNLRPLHTGYI

ITSAPNPKSANEKAAEVVQDEDHDEGLDNADDVCVNKISMAFDINAECEPMPDLDGPPPL

VLEVDSNELEELTAEEQMVINNCRRLRKQTEFIEDLRPVDGNSKLEYSYRPMDQISQFWA

GPSHWKFKRTRPRSTFSQTNGQVDTQPIRTQRAKKSAHLNANRRAKALDYGNVTENFFQQ

LDTTIRQRKANFQKKWDPRKLILPTKFELDPDLFFKYESAPSIKLSKRAGEPDSDEGGDL

GIDMDADMHHDDDNDQELFNNEHFTDAVPANVSVIAAIAAEQAAEASMMNVSAGEIGLTQ

MNATCNNTVFEIGTEFEGAPSQVAKVIVPFAKRAKVIDMKNLKKSCNSLIQKQLLNAVPE

ETIPSHPKKKGEHYSKGFASFQQVYQKLPDLLTTKMSDSLSPSVAFYAVLHLANDLKLRL

IPQEDLEDFQIRQVLD

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

>sp|Q564K3|CND2_ARATH Condensin complex subunit 2 OS=Arabidopsis thaliana OX=3702 GN=CAPH PE=1 SV=1

MDESLTPNPKQKPASTTTRIQAPTSPFFLGSNDDRLEREQARAARAAASRRRSVIFARGS

QPETESDPCFDKQQILELFQNCIKLASENKINQKNTWELNLIDHLCEIIKVEDENNTETN

FQKASCTLEAGVKIYSMRVDSVHSEAYKVLGGITRAGHDDGGDHEDAAGAVENATNQKKQ

PEKKISPLSTLEPSFDALNVKKFDVAFAVDPLYHQTSAQFDEGGAKGLLLNNLGVYGGCQ

VLFDSQEIPGKLVSSANKHDKSETIDLSFVKECVEQMVLNMRKKDEIVPSLRAIINQFDE

ENQRPSDTFSCGQQTTESFDISHGNDASYADDDEGYENFGTSFDYEGQSGDVDENFGPN
E

AEPIYSNFHEEVEPASLQDMDSDDRLENVDDYLFLSLGISSKQNSWAGPDHWKYRKTKGP

DVQPASEIKSSPPAKKTRKKKQAEPELDFAKALEEEMPDIFAPPKNPKTLLLPASRTPCQ

TKLPEDCHYQPENLIKLFLLPNVMCLGRRRRKNSGETSRQQPDDYEHGESWGNDNVYDD
D

DGPFDDNENDQSDAEDTNTLISQPRQVNKIDVQYDKASKQVDVQVLKETLWECLQESHQP

PIQDEEHQQEPPESRSFKVLLASFPDDCQAAERTQDISPHLCFICLLHLANEHNLSLIGS

QNLDDLTIHLA

>sp|Q8BSP2|CNDH2_MOUSE Condensin-2 complex subunit H2 OS=Mus musculus OX=10090 GN=Ncaph2 PE=1 SV=1

MEDVEVRFAHLLQPIRDLTKNWEVDVAAQLGEYLEELDQICISFDEGKTTMNFIEAALLI

QGSACVYSKKVEYLYSLVYQALDFISGKRRAKQLSLVQEDGSKKTVNSETPCETENEFLS

LDDFPDSRANVDLKNDQASSELLIIPLLPMALVAPDEVEKNSSPLYSCQGDILASRKDFR

MNTCMPNPRGCFMLDPVGMCPVEPVVPVEPYPMSRSQKDPEDAEEQPMEVSRNGSPVP
VP

DISQEPDGPALSGGEEDAEDGAEPLEVALEPAEPRTSQQSAILPRRYMLRERQGAPEPAS

RLQETPDPWQSLDPFDSLESKVFQKGKPYSVPPGVEEAPGQKRKRKGATKLQDFHKWYL
D

AYAEHPDGRRARRKGPTFADMEVLYWKHVKEQLETLQKLRRRKINERWLPGAKQDLWPT
E

EDRLEESLEDLGVADDFLEPEEYVEEPAGVMPEEAADLDAEAMPESLRYEELVRRNVELF

IATSQKFIQETELSQRIRDWEDTIQPLLQEQEQHVPFDIHIYGDQLASRFPQLNEWCPFS

ELVAGQPAFEVCRSMLASLQLANDYTVEITQQPGLEAAVDTMSLRLLTHQRAHTRFQTYA

APSMAQP

>sp|Q6IBW4|CNDH2_HUMAN Condensin-2 complex subunit H2 OS=Homo sapiens OX=9606 GN=NCAPH2 PE=1 SV=1

MEDVEARFAHLLQPIRDLTKNWEVDVAAQLGEYLEELDQICISFDEGKTTMNFIEAALLI

QGSACVYSKKVEYLYSLVYQALDFISGKRRAKQLSSVQEDRANGVASSGVPQEAENEFLS

LDDFPDSRTNVDLKNDQTPSEVLIIPLLPMALVAPDEMEKNNNPLYSRQGEVLASRKDFR

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

MNTCVPHPRGAFMLEPEGMSPMEPAGVSPMPGTQKDTGRTEEQPMEVSVCRSPVPALG
FS

QEPGPSPEGPMPLGGGEDEDAEEAVELPEASAPKAALEPKESRSPQQSAALPRRYMLRE
R

EGAPEPASCVKETPDPWQSLDPFDSLESKPFKKGRPYSVPPCVEEALGQKRKRKGAAKLQ

DFHQWYLAAYADHADSRRLRRKGPSFADMEVLYWTHVKEQLETLRKLQRREVAEQWLRP
A

EEDHLEDSLEDLGAADDFLEPEEYMEPEGADPREAADLDAVPMSLSYEELVRRNVELFIA

TSQKFVQETELSQRIRDWEDTVQPLLQEQEQHVPFDIHTYGDQLVSRFPQLNEWCPFAEL

VAGQPAFEVCRSMLASLQLANDYTVEITQQPGLEMAVDTMSLRLLTHQRAHKRFQTYAAP

SMAQP

>sp|Q9LUR0|CNDH2_ARATH Condensin-2 complex subunit H2 OS=Arabidopsis thaliana OX=3702 GN=CAPH2 PE=2 SV=1

MTSHGGGEVRGERIHTVQPERDLVANWEVDLSEKLEEYLLKICSGEITGNEEDGQIPVNF

AEAALLLQGSVQVYSKKVEYLYNLVLRTLEFLSKQRDQEQSKGTSNENEASSSRQVDEEE

NDLFWNVDDIPVDTKNRLDSSVGGDTCPSQFVKPPANLVVLEGDCLDTSGDGGELESYLL

ATTHLYRDFILLDPCDAVAVNEFLGDNYGGKGRNSAHRGSSVRKSFHSSVGRSGGSARKS

SVGKNQGTNVHLSPICGNGPNDQNCDQGSQPPVFEDNDHGFDMDNEYGGAMDFSDTDA
DE

DDPWKPLNPYEPGKLKVKPFKKVKILKKIGWSITKDHMTSMFPLARPNGPISSELIEIWK

MHGCASKDEQASQDIPYYEKLREMLVNGGNQPCGANGNYNDNDKDNHDEANNGDFHDF
GE

HDGDDAEHPFMDEDVLNMNDGGAAEFHNYDGFENGESNCQESLEDLCRSHLDALLANIAK

SEKQTDLAARVSTWKQKIEQNLEEQELHPPFDIQEYGDRIINKLTVEESGNVETFTDLMK

DQEKHEVARAFSALLQLVNNGDVDLEKPGNSTNEPMCYTAVKPFSVRLLKVHNRKNEKRG

IHLPQKRAKSPITKGKSHESPPPKKRNTCSVSSQTRKVSLKISKINGVGVRCTPNSKKRR

KGRSDDVTEVTEVASIEKSLGKL

>sp|P34341|KLE2_CAEEL Condensin-2 complex subunit kle-2 OS=Caenorhabditis elegans OX=6239 GN=kle-2 PE=1 SV=2

MTRNAPPGQESTDLAWLVTPAKDLVENFSIDVLKALAGYLEVIRQESEDTDNQVDAATTY

RLFDFQRACRIIQGSCAVYGRKVDHVYELTISVVDLVENKGQDDGNTGSRRGAGRRKNFN

LGSTNYDLADIDSLKQEALANFEKTVKEEKKSIDAVRMVENAEVIESQYERKSCLVAKPT

QFMFKLNYGQLNRTDEQILNAKSRPDVIGKVKDFEIKKSKVKHDQQILYSHDCYRGNLDQ

FTLPGARWMPDNKELAANFGVADLEVELDLEQEHEKISAYGPFKDPLSGREVVPPPRWFI

EQEAVRQNQEIQSRATSRAITIAAKTLRDSQGFGSQPTRLSQPFVERHRQSNHLNDFLSF

VEGRVNKNRPSTHLTTGLVDMFVDNFGSVMQNDEPNTSRRPDENYAPMDFDDDFGGGG
DD

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

DDDDYIRNLSRRDEKRAPAPWDELDKNHIIWYTGDENLPVVSKPVKKITKFQPKPAEMLA

RKQRREEKINKSRRDEFMETHDYLQDYYYWRSAARINPIKDWKIESLRTAILAEKKRRIK

EKTAKIREARIQNMQRKRTARVIPVEQFEPVTEDIPTSNRRTLGAEYDDVVDEDLAAEVE

LSMFGGGFDDDEEDVRPRGERPPMAPNNLEFDALQTDFDIPPAEYVPLRFEDIDDAELNS

VINLPGNLLIDKALPLLKKFAENRTDREQMAYEMAKAYEDVDVAVSTLQEHVDKWHSRME

PILEEGETRKEYDVHAVGRAVIGQYDIEGGTKRLLDLVMDRPWYEISRYFLSCLFMCNVG

NVMVSEDMELPLEERINSMKITLLKRDMHCEMFKEAGALDA

>tr|Q8INL2|Q8INL2_DROME Chromosome associated protein H2, isoform E OS=Drosophila melanogaster OX=7227 GN=Cap-H2 PE=4 SV=2

MERILPEEAEAAYLSEAREQILEIAKNRPGTQVAKCIRAYDEQQDLASLVVLLEELSKNT

DYSTDLRISLGYYIEELLRHCLGRNDVSRRSAIVAAGSALQYCGKIYGDRVEYLCQVVEH

QIEALLTSELQKETPSGSAAPEKNERRPEETRKRHPKKLTNKEVDPYLLTLEPKRFKTMS

DDKRFNAAGFVKCTRNRTIEYLYQDHTPPNLWKHAPIVDPHNPYDQDEKKQYKMFTYHVE

HRYNTLLPDIPFERLNLIKEYVHTNQVNTTEILNEHMTTKEYLDEYIALENQMLAARYGA

IVTRRRRLVDSARFMEDNLDEGLAKKMCMDKNLPMDTNETVLIDQSLVDENSRSLTTAES

TMGISQAENSTLKSSEVEATLSVSHIENSLSDSQEENPPLSSTLAINESSVLDSTRVDPI

KDLTLDELLLIDSGISMEELSDIQMHTAGQSFDDEGVVLSDLEDQRQLSPMLQMVSPSME

AKTLNIIEMDADLIMNVPREVSYPILLNVMGLPIKRLRRKCIFKLPPEFDLFRQARLPIK

REGQQKSPTTPRTVLQIGREAPQNEREPGSPCSLEFDEDLNFLGFRQRRPTFDSGFDIEE

PRVSSCVSTIGEVKTELEDEIEANTNNLTDASASDALNENVNDTKESGLGDSLAQELNAS

TEPNAENNTANSSQLEVTATESSGTLSGLETTVDLGMESALVSSVLDLSTEPSAMDSNII

NVTQPDNSDLEISTVQDQDTLHDEADANDLSAETPNPIADSAVDDCSLIRDWHRRLAPAL

EAAHERQNFNIKDLGTEILDICKAGNRTATLADVMADKDPSIMCRYMLASLVLTNHGNVS

LDFENRDKSKPIDMSQFRMHLKSMKRMEINPEDDVGNINAAQSKSTPRSKQLNDSAHKPR

TSTTASAPTKRKSAENSLSEVFAKTVRLIQPIPKMWPTPSDADSGISSMGSSLASTARLK

>PF3D7_1304000 | Plasmodium falciparum 3D7 | condensin complex subunit 2, putative | protein | length=1024

MKKLGVNNAGENKGNSIQNKDGTNKKTIEVNKNMRRLTFLNNESNDSIEGNKNKYDKNKV

KEINDVFKNCMVALSHNKICTRNAFDIHIIEHLEDLINLNDEEIPEELNDEMIENGEFNL

SFTRASKAIEGATKVYGYRVEAIYDQTYNFLTNMNLAKQFELDNDMMDDNKNTIDPLNKR

MRKRKLTYLQESSTLAKSSDITVDSLSLSNISVDTFFLKLNSTYDHSCGQKYLLPNLNLN

NDLSIQFDGDIDVCEYKKRRKTLDDIGKDGVKDKEDGIHIIMKNKNDDNCDTYDKIICNN

NDINKVDEKKNNDVDICENRNKKNESCNIYVDTFNVNEHKEFFDIYKKKLFLNADILKEI

LYGGPDDFNSLHICPELDYFKEELKKHKLKRTDSKDTDDVDKEDGDGNLYNDDDIGNTDK

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

NNIINFYDENLGCNMSNNMLMESIYNSDALNTDRKNIFPENLGMSCGNLNEEKKNNNMNS

SNNNLNNCNYDNNDLYLGDYRMDDLNIENVMQESLGFDNLNYSCDKNMNFSNNNGMLLQQ

SINMMDGMNFPELIKSENKELDLGNTQINNEHYLLKNNYDIHKRHSIISEVPDDDTLWNR

SVMTFENRINAIDINNELNYHYYMPNKLMMMGNFRNLIDINKNVTNNINNNNKSTILQNI

IMQKKMKYAFDVTSIDFENLYTENNNIELSTYDLWKKEKKKYVSNALFCIDQTSYIFDTR

ENCVNCVNTVTDKVMKFSKFCVYPDFLHACNKNRNKHNNNNINKNHMNVILNEVNNDNIL

NDIDPINQDDFNNNIDNFDMENMDDMNDGHLHEGLNEAIDKYYNMDFDNIWANNDNKNND

NINNNNNLNNSTYNNMNSHNNNNNNNNNNLSLFKKHSNFFNTALFQFNHSNTFGNVPFEN

VSKFVDVAKIKKILCDIVKPNEEQKSNTNDSLCEYQTEKDNQIVTYTAEKTTTFEEIVKE

TTAKLNESEASSTSIHMLFVCLLYTCNDQELLLEKIPNQNNFYVRYGLPVECHVNHDDIP

MLKN

>PBANKA_1402500 | Plasmodium berghei ANKA | condensin complex subunit 2, putative | protein | length=943

MKKLGVSNKTNTNFTQIKGPDPIFKNPIEFNKNLRRLSFLNNKDEDLNKSEKNKVKEIND

VFKNCMAALSHNKICTRNAFDIRIIDHLEDLVNLNDEEINEELNDELLETGDFNLSFTRA

SKAIEGATKVYGYRVEAIYDQTYNFISNMNIAKKSETNDDVIDEKKHANEITNKKVKKRK

LEFFQESSTLAKPSDITIESVSVSNISVDTFFLKLNITYDHSSGISYLLPNLTLNNDLSI

QFDGDIDTCEYKKKMKFEEKLGKQNIERRNEKNCGSNDELNDMLQKNESPVKVGEYVTFN

DNDKIMAREYKSKLYTNSDILREMFFGNEIEEFNNLNICPELDYFKEEIKNLKLKRSDSK

ILDDEDIDNYNDDDTGLDKFSKKKNELNLDGNMSMDNLLGDMDGNNNDDGNNEHMNDNAN

YGDNNKILESSFNNNLNFDDCNIDDLNIENVMQESMVFDNMNLNDSLNNNLNLSQNILSL

HHNSNILGSSIPLPELMKSENKDFSLINSFTGNFNFSSQNMLFKNQDKGSPSKPKNMLSI

IPDDDTLWNRQAISFENRLNAIDVNSKFNYHYYNPSKLMINGNFANLMSMAKVAFKNKQG

PLNALTNKKLKTSFDITYINFENLYIEVNDVELSAYDLWNKEKKKYISNSLFAIDQTSYI

FETKDNCINCVNTVIDRIMKFARSPFIESQNFNTDIKTNVILNEINNDYIGNDLDQINNF

ERKQSENMFMDNMDDGQDYQMHEGLNDSIDKFYNMDFEDIWQNENKNDITKFGSKNDNTS

IFQLHQSMGHTNSLGSVIAPDNLPKFVDVSKIKKILFNIVKPDENEENVENGKSEENGKS

EENGKSEENEENKKNSDSSKQIVPYEGEKTTTFKSIINKTKTKLTESEVNGTSIHMLFVC

LLYTCNDQELLLEKIENEEDFYVHYGLPVEFHIKQDDTRMLKN

>PY01534 | Plasmodium yoelii yoelii 17XNL | CCAAT-box DNA binding protein subunit B | protein | length=886

MKKLGVSNKANANFTQIKGPDPIFKNPIEFNKNLRRLSFLNNKDEDLNKSEKNKVKEIND

VFKNCMAALSHNKICTRNAFDIRIIDHLEDLVNLNDEEINEELNDELLETGDFNLSFTRA

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

SKAIEGATKVYGYRVEAIYDQTYNFISNMNIAKKSDTNDDVIDEKKHVNEITNKKMKKRK

LEFFQESSTLAKPSDITIESVSVSNISVDTFFLKLNITYDHSAGISYLLPNLTLNNDLSI

QFDGDIDTCEYKKKMKFEEKLGKQNIERRNEKNLGSNDELNDMLQKNESPVKVGEYVTFN

DNDKIMAREYKSKLYTNSDILREMFCGNEIEEFNNLNICPELDYFKEEIKNLKLKRSDSK

TLDDEDIDNYNEDDDTGLDKFSKKKNELNLDENMSMDNLLGDMDGNNNDDCNNENMNDN
G

NYGDNNKILESSFNNNLNFDDCNIDDLNIENVMQESMAFDNMNLNDSLNNNLNLSQNMLS

LHHNSNILGGNIPLPELMKSENKDFSLINSFTGNFNFSSQNMLFKNQDKGSPSKPKHMLS

IIPDDDTLWNRQAISFESRLNAIDVNSKFNYHYYNPSKLMINGNFSNLMSMAKAAFKNKQ

GPLNALANKKLKTSFDITDINFENLYIEVNDVELSVYDLWNKEKKKYISNSLFAIDQTSY

IFETKDNCINCVNTVIDRIMKFARSPFIESQNFNTDIKTNVILNEINNDYIGNDLDQINN

FERRQTENMFMDNMDDGQDYQMHEGLNDSIDKFYNMNFEDIWQNENKNDITKFGSKNDN
T

SIFQLHQSIGHANSLGSVVAPDNLPKFVDVSKIKKILFNIVKPDENEENAENENEENKKN

SDSSKQIVPYEGEKTTTFKNIINKDANFVNRAHELYICILYIYFLT

>PVX_122040 | Plasmodium vivax Sal-1 | hypothetical protein, conserved | protein |
length=928

MKKLGVNTKAGANLTRIKENENLFKKPIEFNKNLRRLSFLNNKNNDANQKDHQDKCDKTK

VKEINDVFKNCMVALSHNKICTRNAFDIRIIDHLEDLVNLNDEEINEELNDEMLETGDFN

LSFTRASKAIEGATKVYGYRVEAIYDQTYNFLSNMNIAKQSEVNEELVEEKKNANEISNR

KIKKRKLEFLQESSTLAKSSDITMDSVTVSNISVDTFFLKLNSTYDHSSSNSYLLPNLIL

NNDLSIQFDGDIDACEYKKRKKMEEATEEGVDEDTAATQTSGKFPPMIRSSSSYDNDCSD

AMVKCYKKKLFLHSDVLRDILFVAGNEDFNSLNICPELDYFKAEISKHKMKRSESKELED

GEQGNDEDDDDDDEEDDRRYAYNLNEEEHMSMHKGDDYGGERAKNNLASSVGGNLTMD
NL

LDDCDFHNASVGDGRMLNSSVNNNSMHFNDYKIEDLNIENVMQESLAFDNMNLNDSVGNN

LNYSQSILSFQQGSNTMTGLPLPELMKSENKDFSLMNSLGGNFPLSTQNSLFFKNQGGSP

TRKGLISTIPDEDTLWNRNVMTFENRLNAIDVNNKFNYHYYVPSKLMAHGNFRNLMDIAK

GTHKNKHMVLQSVVQKKVKASFDVSLIDFENLYKEINDVELSTYDLWKKEKKKYVSNALF

SIDQTSYIFETKDNCINCVNTVTDRIMKFSKAPMIAPGDYCGDLKLNVVLNEVNSDFVTN

DMGQIHYGEGRNMESTFPENMDEHQDCHMQEGLNDAIDKFYNMDFDDIWQNENEHNLSK
Q

GSKNENASALQLRQRVSHGTTLGNANMENVAKFVDVSKIKKILCDIVKPAGKEATGSEGP

AQRNDQIVPYVEEKTTTFKDIIIATKSKMNPEEVNGTSIHMLFVCLLYTCNDQELLLEKI

PNEDNFYVRYGLPVEYHVKPDDMLMLEN

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

>TGME49_288930 | Toxoplasma gondii ME49 | hypothetical protein | protein | length=1185

MVARATDGGDTGAPAGSETHPVVRLGFRPDGKRRLSVVSGPASLSTGDGAPSGESRRGV
G

GTTGASLAGHAPGALARQASSLNSASSSFSSFASSSAFPASLPFSSSANRPGFARPTARA

ASGAGRALAGRLLVRAFFDDMKTVSMQRINQKNAFQVDLIDRLALVVHQQLVKTDNAVLT

DGTDGQAPALLDGASSPDLSSTSELSRTGRGEEGEEGITFTHVSSAVEGATRVYGYRVEA

VYDQTYHVLNLMSSSRQGQGGEDAGDGDAAAHPTRRGRHHQQLALFKKGGASTLAPASE
I

TESQIEKDSCVDPYFLKISGMFDQAGAKGLLLANLEVDTSLRMKLDGECRAFPTGVLRNG

REKGNAAREETEANHAPEETAEVAEPSVDCAFLRDLLLAGETPASVLALDICEKPIGHFR

ELLQSLRRGRDVKAGDALSEAEEETNELDVDDDMLDVDMQGTTLGDSGFQQATQEDEELL

KIGESYMDDEAASGGLREVLEAQEAESREVTQDDELFGDGVDFFCDGHVGDADDDRDGS
W

DRPRGEIRERDAEGERILYEVAADESSASSCTHPKSYSFDRLVPAFSQMLGVPAGGHGAQ

LAAKRAFGLVPTAAGSQGGGDPEGRETGRRRGSGEETSSSGCVDFFPLSAKGSAAAVGR
P

SAPSARLTVKERQLRKQEQFREMLDPFRVDMSNLDTKSGGLPLGSKFQCYVPKPQENVTV

ASDFTSSPFLHSPHLLVSLAMVPSKEIRLLRHTGPSVGDGWAPDGAGNSGGRAGRDLGRG

DSQAPDGPWEATFIDAGDDGGLEDDSFWSAAAHGPLLKESALEAADAKAQTLLEGGGGLD

RDWSDTAWAGGENDQASSYIAHLFERGDAASRGGAAGLSLGDIVLLAEPQKVGVSELRLS

IPSRYVDVAVVKKALKTALGVVPDFSADQDAENSALRRKRKRDMLTAPGEEEPQEEDDVD

SCLDSVDGETKEPCRDLALWNPRGRRDGIDFETLCAETTRKLPSTEKANLSPQMLFVCLL

YTCNEETLHLEQSADFSSFSAFVHAPTDWHMRDDAEAVRVLDTVRHYTPLALPAVPEKKK

KETLSLEAPSAESAADDGGAAGDADDAQSPAKRRGALGDEERRQKRRREGRGGKHEEG
NR

KGHRDEETGSKKRRGKKRRKEDSDESSDGSDDEEAAASEDRDGSC

>cgd3_3960 | Cryptosporidium parvum Iowa II | Condensin complex subunit 2 | protein |
length=901

MARLSLHPRIFESEYSAETDNYLEENVSNLNKNGWSGRKIQKSVLPNKRRSISIGQSRND

DSETTRLSRTSINLKGQIDQNKSFFSKSSSDQLPNEELDKLCNQCLNLLRQNKISSKNAF

DILLIDHLNDIVNVQDSQNEEEKSNIIKKENKENVSKINNNSKEKNIEIDKSKMNKGEFH

NEMISSETTSQDTNFQKFQRAAVTLEASARIYGYRVDSTFDNAYRILSNIKSGQILARND

SDCEEDEKDQQEDDSSGNRLEESSKKKKYKRNLIFSGENNTIVSNPESITLKEFEKTKEM

SEKPFFVNFLNDKVKIDNIRFGLDYNNIGSISSMLMNNLELKDSLYNRDKSASLTFNISG

RNNLINSNNFSCSTICKDELFKVNKKTLDLLLPNPNENFGKDISNIVCPNLSRIFDSIQE

ITENYGDIKQVKDSTSIEEYQLNEQVIRRFFEFTDPSFENDSLESENNFLKKHSQSMDDI

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

EYVNENQDLEFLNNFEEEEKVELNMGLQDFQMHVERVIKDLNKTESSEHILEDLDQNKDI

EIEKKSGEFKFSKIGDVLDFLAKKSTKSLEFFNERKKGEKSNSVFGQKLRSKSENTNFEI

PSPLLTNNIFNGIEWMNNLPKIKQTQINSSRSKTKNLSSNKLFVGKNSTTSIFNYNLNHL

FCLANLTGVKINLRSIHKAENSEETAGQNYITNDILIKTNNSQHINKNILNNSMSEQTDL

MVNHCEIEDLMQVESYLQPLSQEEFSIDNGKDILIQSKLLLPFDEDKISDTDILLEKSSK

SLHVRDEYNSSLKFSKSSNHVDIDLIKNVLKNSIQILQSDDSNSLTLFNIINQSRKLLQQ

DGLNSVSTGIFFICILHICNENNYSLNLSDQNANSDSPIELNSENYVHIMLNLNSSVDLQ

K

>ETH_00039510 | Eimeria tenella strain Houghton | Ccaat-box DNA binding protein subunit B, related, related | protein | length=122

GFDEEEPLEAAAAAAAERVSFCHLSSAIEGATKVYGYRVEAVYDQTYHLLNGLSSLKNSS

NGVEDDEEQQQQQQQQGSKRKQRQQQQQLALFFKGGSSTLADPQDILNEQQDTGVLIDP
F

FV

>BBOV_I002180 | Babesia bovis T2Bo | hypothetical protein | protein | length=726

MASRKRSRDVDSSEISEAPQPVSKKVTGPSGNTQDLLTLFTDCMSALSTNKICSRNAFEVGI
IDHMTDLVHLDDGSVDDD

VVELLPEDTSSGASRRLNFTRASKVVESASKIYGYRIEAIYDQTFNVLMSMNSANQADGTSG
STSATKPRGRHRVKIDLT

TSSRTLAPESEVTLTEIPMDNVILDPYFLKISSMFDHSGAMGLLLINLQVTDDLSLDLDGDSLV
FPPPRARSEHDDSNVI

LSRDAVKKCFFGNTEPKRMEILPEAAYFRQELERLREMRRRKECGDVIASDSDGEINPEPE
SKHIDFFKVRQPIAPIEDS

LEPMDAMGHIDDVPEDTVHDMPISGTPPDTGSDTGSVPNSTGIALRQRLAEIDLIGGSQFSY
YTSVPVKTHVPSSKNKGS

DTEPDDTNLKTQRPVKSRQNITNDLESYLRNIDLEEKIASVELSTVFTTPKAVGKKAPSTFAA
SDFTGAIYRFNDTFLTR

LGLLGNRCLRFVSDHEWRNGPNTDSSVPILQVHYCADIDAPYSTMRISENFSWRLDEQVGL
LDEDDMWNASNDFEDPELP

LTQDGPVVANIESMYESEPGEPLLALSQHADPAAISQWPREGAIVPAPVAAYVDIFKIKKTLC
GVIVPPPAFHEALEDKQ

QIDKEHGSGNTQSEYSCGFQDAVTDTVSRLMDSDVAALTSHILFVCLLHVCNEQDLLLKQS
RPLEDFVICAGAPKEQHLG

DVQGSL

>PBANKA_030180| Plasmodium berghei ANKA | condensin-2 complex subunit H2, putative | protein | length=775

MSTQEDEVRLLIQNLQKCNNTNECINFDLASTIQEFLNSLDKNSFEDIDKTIRENEKDKD

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

LMNSFTSAAIFLENCVKILGLKIEHLHNLAHNTLYNIYKENKNSNSNKKQLLIIDEEEYL
YINEIKNLKNTITENDIIEEDLLVKTIPLPTFLFTDHIRVKNKTENHKNIIKYDEKDNFL
INKKKLLNEPDIENIPYIDTLGNESINSIETKIMDLNSTNSYKNMDNLNLIDNKQTIEIS
SVNSLNFDKLFLENDGIILLDINDYNIFINDEYDFTLQNKNSTILFEKYEFFSRNSIYLS
NNLTEYIHEQNTIQHTYKINNIYDITSLRLCTDTLLFKTDFYSYDIALDIIKNKNYLINK
FERQKKKLYILDETIHKDHKYNTIYKQNADYCDYCGSIITSIIEPNEPNNRCIYCNNNTQ
IDRNEKNGLYKKLPGYYNLSCYNITETEDFLTYMQPNKIIDIIMKNEINDINLDSNSTTN
KFEENIIDPIFHQNNDFTDKGDNYRLSNDQKLFRQIKIPSLYIQKLGLNIDYYYLEPLIY
NLIKNLKKEKNVDRFFSINFYDHNENYDIEILKDDYYHEIKDEQNKTIQETLNMDTFINI
KSIDNHVKNFPTSILKKTDSNTSLVFSFEDKIQDRVNAWSNFLEEKLVILKRQPQYNVEY
YKKKILKYIINNGDNIYFPDLIKNDEKYQIYRNFLTTLMLINTNKLNISEIDQQKHSNNI
TNYQINIKNINVNEYMNISSAFDNTKFTINDKKRKITEKSDNIDNSFHLEKKNHI

>PFB0185w | Plasmodium falciparum 3D7 | condensin-2 complex subunit H2, putative | protein | length=797

MSTTDELNLLIQNLQKCNNTNECINFDLSSTIQGFLNCLDRNVLENIDKGLGENEYEKEV
VDNFTSAAIFVENCVKIFSQKIEHLHNLAHNTLYNIYKENKHNSSSKKNQLIMSDEEEYL
YINEIKNMKNTQHDNDIIEDDILIKTIPFPTFLFSDNIKKTKDINEDKRKTNFNNNEEDK
EKDNKNKDNDIDAINEFEITDNNSVNTLNFEKIFIENDGILLLDINDYNVFIDDPYNFSI
QNKNSTILFEKYDFFSRRSTYLSSNTLSKYVVENKNMDHIYKLYNHITDIINKNICFDIF
LFKQDFFDYDFSLGILKNKKSILNKFKQQQKKLHPLEENTHMDTHHINNNHHLQKYDLNR
PLPNYYMLHCYNIKNYQDFFRYMQPNYILEIMKRHIIKEIYNTNQQERAIQKEAYEIYNE
QTKKKNDHKENNNIDVPKYKDNTKCYDSPFYNYYISNNIIQFDHLIDDDMIYFDEYFYKS
LILYNTNINDLHKNTNNNQTNDETNIINNMKDEKQKNLIIYSNINNFSNDQKLFNQIKIP
ELYIQKLGLNFSYYHLEPLIYNFIKTLKKKNDFEKFFSVNLFDDKPIYEFDILRDDEYDE
QKNEDNKNHIEENINFENITDKNILNDEMDNIPIAIFENDHLDNTFIMNDDQELQDRVSK
WNAFLEEKLEILKRQPKYDLDLYKKNIINYTINNGENILFTKLIKNKDKFEISRNFLTTL
MLINADILNIKKINKHKKSNNISNYEIHIKKENLQQYLSISKQVQNKSFLIKEKKRKKNK
QHLTNGMKDTSKKKQKI

>PVX_003630| Plasmodium vivax Sal-1 | hypothetical protein | protein | length=784

MSTPDELTALIQNLQKCNNTSESINFDLASTIQEFLNCLDRNVLEELEGGAHEGEKERES
DREREKEGEREREQDVANSFTSAAIFVENCVKIFGLKIEHLHNLAHNTLYNIYRENKHSN

bioRxiv preprint doi: https://doi.org/10.1101/674028; this version posted June 17, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

AGKKHMVMSDEEEYLFINEVKNLKSCPGESEPCMEEDLLVKTIPLPTFLFSENVKKKEGA
AEVDSELGIDLSADLDTPRRSDNGSGTDDRRGHSNLSKESEEAAREANGETDSIASSPAW
RKGENDRLTEEAEEVLEQNSLKPLNFDKMYLENDGILLLDINDYNVFINDEYDMSILNQN
SSMLFEKYDFSCSRHSTYLSPANLTKYIEREKTVGDIYRTYHFNDILSDDLCSDVFLFKS
DFSHYDLALGIIKSKKYLLSRFKEQKKYLYVLDENAHMDKDRVGTSTIEKNDIKRKLPHY
YMLNCQNIKRAQDFFTYMQPSQIIDIITRNARRRYLSCLPSGETSQGGQSGNAAGADSPV
GGAASHTSPPATPAPPPPGRSTDQKLFEQIKIPDIYVQKLGLNFRYYHLEPLIYNLIKEL
KKKKNVEKYFSLNLFDEQNNYDFDILQDEEYADEQNEGNAAAEGEGNLTDENFLDVRSLN
GDMVGGMMGGMMGGMMDDIPLDVFEKKDSSDAFFASFDEDIHDRVNKWNAFLEEKLQLL
R
SHPKYDVDQYKKNIIHHTLNSGAKTPLCNLIKDREPYQVCRNFLTTLMLINTNMLQISEV
NQHSQSNDVSNYQINVKKENVQEYLGSSKRFKNASFAIKDKKRKTASKGGTRNAKPAKKK
PHKD

>PY01684 | Plasmodium yoelii yoelii 17XNL | hypothetical protein | protein | length=773
MSTQEDEVRLLIQNLQKCNNTNECINFDLASTIQEFLNSLDKNSFEDIDKTIRENEKDKD
LMNSFTSAAIFLENCVKILGLKIEHLHNLAXNTXYNIYKENKNNNSNKKQLLIIDEEYLY
INEIKNLKNTITENDIIEEDLLIKTIPLPTFLFTDHIKVKNKIDNHKNIIKNDEKDNSLL
NKKKNLNKSDIENIPYIDTLENESITSIQKNVMDLNSEYSYQNMDNMDNMNSMENKQTLE
ISSVNSLNFDKLFLENDGIILLDINDYNIFINDEYDFTLQNKNSTILFEKYEFFSRNSIY
LSNNLSEYIYEQNTIQHTYKINNIYDITSLRLCTDTLLFKTDFYSYDLALDVIKNKNYLI
NKFERQKNKLYILDETTHKNHKYNTIYKQNANYCDYCGSIIEPNESNNHCIYCNNNNEKN
GFYKKLPGYYNLSCYNITETQDFLTYMQPNKIIDIIMQNEINDTNLDANSTTNKLDQNII
DPIFHENNDSSDKSDDQKLPNDPKLTNDQKLFKQIKIPPLYIQKLGLNIDYYYLEPLLYN
LIKSLKKEKNVDRFFSINFYDHNENYDTEILKDDDYQEIKDEQNKTIQETLTMGTFINIK
SIDNHVKNLPTSILKKTDSNTSLAFSFEDKIQDRVNKWRNFLEKKLDILKRQPPYNVEYY
KKKILKYMISNGDNIYFPDLVNDNEKYKIYRNFLTTLMLINTNKLDITEIEQNNSNNITN
YKINVKNMNVNEYINFPNSFDNIKFTINDKKRKITQNFDNKNDSLYLQKKHHI