5-methylcytosine modification by *Plasmodium* NSUN2 stabilizes mRNA and mediates the development of gametocytes

Meng Liu^{a,1}, Gangqiang Guo^{a,b,1}, Pengge Qian^{c,1}, Jianbing Mu^{d,1}, Binbin Lu^b, Xiaoqin He^e, Xiaomin Shang^b, Guang Yang^a, Shijun Shen^a, Wenju Liu^a, Liping Wang^a, Liang Gu^a, Quankai Mu^f, Xinyu Yu^e, Yuemeng Zhao^{b,g}, Richard Culleton^{h,i}, Jun Cao^e, Lubin Jiang^g, Thomas E. Wellems^{d,2}, Jing Yuan^{c,2}, Cizhong Jiang^{a,2}, Qingfeng Zhang^{b,2}

^a Key Laboratory of Spine and Spinal Cord Injury Repair and Regeneration of Ministry of Education, Orthopaedic Department of Tongji Hospital, Shanghai Key Laboratory of Signaling and Disease Research, School of Life Sciences and Technology, Tongji University, Shanghai 200065, China

^b Unit of Molecular Parasitology, Research Center for Translational Medicine, East Hospital, Tongji University School of Medicine, Shanghai 200120, China

^c State Key Laboratory of Cellular Stress Biology, Innovation Center for Cell Signaling Network, School of Life Sciences, Xiamen University, Xiamen, Fujian, China

^d Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rockville, MD 20892-8132, USA

^e National Health Commission Key Laboratory of Parasitic Disease Control and Prevention, Jiangsu Provincial Key Laboratory on Parasite and Vector Control Technology, Jiangsu Institute of Parasitic Diseases, Wuxi, 214064, China

^f Department of Immunogenetics, Institute of Tropical Medicine (NEKKEN), Nagasaki University, 1-12-4 Sakamoto, Nagasaki 852-8523, Japan

^g Unit of Human Parasite Molecular and Cell Biology, Key Laboratory of Molecular Virology and Immunology, Institut Pasteur of Shanghai, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai 200031, China

26 ^h Division of Molecular Parasitology, Proteo-Science Centre, Ehime University,
27 Matsuyama, Ehime 790-8577, Japan

28 ⁱ Department of Protozoology, Institute of Tropical Medicine, Nagasaki University,
29 Nagasaki, Japan

30

31

32 ¹These authors contributed equally to this work.

33 ²To whom correspondence may be addressed. Email: qfzhang@tongji.edu.cn,
34 twellems@niaid.nih.gov, czjiang@tongji.edu.cn, or yuanjing@xmu.edu.cn

Abstract

5-methylcytosine (m⁵C) is an important epitranscriptomic modification involved in mRNA stability and translation efficiency in various biological processes. However, it remains unclear if m⁵C modification contributes to the dynamic regulation of the transcriptome during the developmental cycles of *Plasmodium* parasites. Here, we characterize the landscape of m⁵C mRNA modifications at single nucleotide resolution in the asexual replication stages and gametocyte sexual stages of rodent (*P. yoelii*) and human (*P. falciparum*) malaria parasites. While different representations of m⁵C-modified mRNAs are associated with the different stages, the abundance of the m⁵C marker is strikingly enhanced in the transcriptomes of gametocytes. Our results show that m⁵C modifications confer stability to the *Plasmodium* transcripts and that a *Plasmodium* ortholog of NSUN2 is a major mRNA m⁵C methyltransferase in malaria parasites. Upon knock-out of *P. yoelii* *nsun2* (*pynsun2*), marked reductions of m⁵C modification were observed in a panel of gametocytogenesis-associated transcripts. These reductions correlated with impaired gametocyte production in rodent and human malaria parasites. Restoration of the *nsun2* gene in the knock-out parasites rescued the gametocyte production phenotype as well as m⁵C modification of the gametocytogenesis-associated transcripts. Together with the mRNA m⁵C profiles for two species of *Plasmodium*, our findings demonstrate a major role for NSUN2-mediated m⁵C modifications in mRNA transcript stability and sexual differentiation in malaria parasites.

55

56 **Significance**

57 Modifications of RNA including methylations of cytosine (m⁵C) and adenosine (m⁶A) have
 58 important roles in RNA metabolism, cellular responses to stress, and biological processes
 59 of differentiation and development. Here, we report on the profiles of m⁵C mRNA
 60 modifications in malaria parasites that infect rodents (*Plasmodium yoelii*) and humans
 61 (*Plasmodium falciparum*). These parasites have genes that encode homologs of human and
 62 plant NSUN2 methyltransferases (m⁵C “writers”). We show that one of these homologs,
 63 termed PyNSUN2, stabilizes mRNA transcripts in *P. yoelii* and mediates m⁵C-associated
 64 development of the parasite sexual stages (gametocytes). Further research on m⁵C and
 65 other epitranscriptomic modifications may yield new insights into molecular pathways of
 66 gametocyte development and mosquito infectivity that can be exploited to interrupt malaria
 67 transmission.

68

69

70 **Keywords:** RNA methyltransferase; gene knock-out; gametocytogenesis;
 71 epitranscriptomic modifications; m⁵C transcript profiles; RNA bisulfite sequencing

72

73

Introduction

Malaria is a mosquito-borne infectious disease caused by unicellular apicomplexan protozoa of the genus *Plasmodium*. Consequences of this disease in 2019 included an estimated 229 million infections and 409,000 deaths globally (1). Among the species infecting humans, *Plasmodium falciparum* is the most virulent and accounts for most of these deaths. *Plasmodium* parasites have a complex life cycle that alternates between mosquito and vertebrate hosts, in which the parasites complete numerous rounds of asexual proliferation in the haploid state and pass through a brief diploid period following obligatory mating of male and female gametes in the mosquito midgut (2). During the proliferation of intraerythrocytic parasites in the vertebrate bloodstream, a small fraction of the asexual population commits to gametocytogenesis, producing sexual stage gametocytes that are taken up by mosquitos in which they emerge from red blood cells as gametes (3, 4). These cellular developments are associated with stage-specific and highly dynamic transcriptomes under the hierarchical control of transcription factors and epigenetic regulators (3, 5, 6).

Epigenetic regulation at the transcriptional level plays a critical role in genome expression events and their global outcome. For instance, heterochromatin protein 1 (HP1)-dependent heterochromatin modification provides a transcriptionally repressive microenvironment for the silencing of antigenically variant genes and *ap2-g*, a master regulator of sexual commitment, thus determining parasite adaptation and development in

the human host and transmission into mosquitoes (5, 7, 8). Regulatory pathways such as RNase-mediated nascent decay, m6A modification-mediated mRNA stability, and RNA helicase-associated translational repression confer other important mechanisms for fine-tuning post-transcriptional regulation in malaria parasites (9-11). These findings demonstrate that the highly dynamic transcriptome of malaria parasites is controlled by a complex but well-organized multi-layered regulatory network.

While DNA methylation has been widely studied as an epigenetic phenomenon, characterizations of mRNA methylation in the modulation of transcriptome stability and expression are more recent (12-14). Reversible mRNA modifications are now known to be involved in the post-transcriptional regulation of gene expression in eukaryotes (11). For example, N6-methyladenosine (m⁶A) is the most prevalent RNA modification and has been investigated extensively in many eukaryotic organisms, particularly in its association with various cellular processes such as cell differentiation, embryonic development, and stress responses. In these processes m6A epitranscriptomic modifications are involved in regulation of mRNA metabolism, translation, decay, and miRNA biogenesis (11, 15, 16). In malaria parasites, the epitranscriptome has likewise been recognized as an important modulator of post-transcriptional gene expression. Baumgarten *et al.* (12) described highly regulated features of m⁶A mRNA methylation in *P. falciparum*, found that these features are associated with stage-specific fine-tuning of the transcriptional cascade, and suggested

that widely distributed m⁶A modifications may shape transcriptome expression through effects on mRNA stability during blood-stage development.

In addition to m⁶A, RNA modifications such as N¹-methyladenosine (m¹A), m⁵C, 5-hydroxymethylcytosine (hm⁵C), pseudouridine (Ψ), inosine (I), uridine (U) and methylation of ribose at the 2' position (2'-O-Me) are also known to regulate transcription (15). Modification of mRNA by m⁵C is important for diverse biological processes including the progression of bladder cancer, HIV-1 infection, and developmental progressions in zebrafish, *Arabidopsis thaliana*, and rice (13, 17-19), suggesting that cytosine methylation is a powerful regulator of cellular processes at the epitranscriptomic level (16). In other studies, pathways involving m⁵C-modified transcripts have been found to include DNA repair and homologous recombination (20). Specific methyltransferases (m⁵C writers) in various species include NSUN2 in humans, TRM48 in *Arabidopsis thaliana*, and OsNSUN2 in rice. NSUN2 can drive human urothelial carcinoma pathogenesis by targeting the m⁵C methylation site in the HDGF 3' untranslated region (18). TRM4B, by promoting mRNA stability through m⁵C modifications, influences the transcriptional levels of genes involved in the root development of *Arabidopsis thaliana* (13). Mutation of OsNSUN2 affects photosynthesis efficiency in rice (19).

Among binding proteins (m⁵C “readers”) that recognize m⁵C-modified transcripts are the Aly/REF export factor (ALYREF) in humans and the Y-box binding protein 1 (Ybx1) in zebrafish. ALYREF is an adaptor that regulates the transport of m⁵C-modified transcripts

from the nucleus to the cytoplasm (14), whereas Ybx1 promotes the stability of its target mRNAs in an m⁵C-dependent manner (21). A potential m⁵C demethylase, ALKBH (an AlkB homolog), was reported more recently from *Arabidopsis* and is attracting study for its possible role as an m⁵C “eraser” (22, 23).

In *P. falciparum*, the DNA methyltransferase homologue TRDMT1 has been shown to modify position 38 of aspartic acid tRNA, where the presence of m⁵C may modulate translation of proteins with effects on infection and pathogenesis (24). More generally, a number of RNA (cytosine-5)-methyltransferase (RCMT) orthologs have been predicted from genome analysis of *Plasmodium* species including *P. falciparum* and the rodent malaria parasite *Plasmodium yoelii* (25-27). Findings from these studies suggest that m⁵C mRNA methylations occur, but the patterns and roles of m⁵C modifications in the biological processes of parasite infection, virulence, and transmission remain largely unknown. Among *Plasmodium* species, *P. falciparum* and *P. yoelii* have comparatively low G+C contents (both approximately 20%) (25). Considering these low G+C contents, it is interesting to speculate whether epitranscriptomic effects from m⁵C modifications may differ in these parasites compared to those in other *Plasmodium* species that have higher G+C contents.

The present study was designed to characterize the overall transcriptome profile and stage-specific dynamics of m⁵C mRNA modifications in the asexual replicating forms and gametocytes of *P. yoelii* and *P. falciparum*. Our results demonstrate comparatively high

levels of m⁵C mRNA methylation in the gametocytes of both species. Enhanced stability of m⁵C transcripts is likely to play an important role in the processes of sexual stage development and transmission. *P. yoelii* and *P. falciparum* each carry an ortholog of NSUN2 (PyNSUN2; PfNSUN2) that functions as a methyltransferase with a major role in these m⁵C modifications of the transcriptome. We show that *P. yoelii* gametocytogenesis is disrupted by knock-out of the *Pynsun2* gene and can be reestablished by restoration of *Pynsun2* expression.

Results

Global profiles of m⁵C modification in the *P. yoelii* and *P. falciparum* transcriptomes

To test for detectable m⁵C modifications, dot blot assays were performed with mRNA preparations from *P. yoelii* and *P. falciparum* mRNAs. Using an anti-m⁵C antibody previously shown to detect mRNA m⁵C modifications in *Oryza sativa* (19), our experiments demonstrated the presence of m⁵C modifications in schizont-stage parasites of both *Plasmodium* species (Fig. 1A).

For quantitative assessments of modified nucleotides in the *P. yoelii* transcriptome, purified mRNA preparations from *P. yoelii* were completely digested to mononucleotides, which were then dephosphorylated and subjected to analysis by triple quadrupole liquid chromatography–mass spectrometry (QQQ LC–MS). Results identified well-known mRNA modifications including m⁶A, the most abundant mRNA modification in eukaryotes

including *P. falciparum* (12, 28). Other modifications including m⁵C, hm⁵C, m⁵U, and m⁷G were detected at lower levels (Fig. 1B).

To obtain a transcriptome-wide maps of m⁵C modification at single-base resolution, we performed RNA-BisSeq on mRNA preparations from asexual schizont (S) and sexual gametocyte (G) stages of *P. yoelii*, and from synchronized schizonts and three stages of induced gametocytes (5–6 day stage III, 8–9 day stage IV, and 12–13 day stage V) of *P. falciparum*. For each stage, two biological replicates were used for high-confidence site calling (SI Appendix, Table S1). The m⁵C sites identified between the independent replicates displayed variation and reproducibility comparable to those of high quality studies from other organisms such as zebrafish, human, and mouse (21, 29). About one third of m⁵C sites were shared between replicates (SI Appendix, Fig. S1A). The methylation levels of these shared (common) m⁵C sites were significantly higher than those of the m⁵C sites unique to one replicate or the other (SI Appendix, Fig. S1A, box plots). Median fractional methylation levels at the shared sites were 0.77 for *P. yoelii* schizonts and 0.15 for *P. falciparum* schizonts. In comparison, the median m⁵C methylation levels were 0.39 for *P. yoelii* day 3 gametocytes, and 0.2, 0.48, and 0.39 for the stage III, IV, and V *P. falciparum* gametocytes, respectively (SI Appendix, Fig. S1B).

Examination of complete transcripts harboring m⁵C modifications showed that individual transcripts more often carried these modifications at two or more sites than at just one site alone, and that a greater fraction of *P. yoelii* transcripts carried multiple m⁵C

modifications than did *P. falciparum* transcripts (SI Appendix, Fig. S1C and SI Appendix, Table S1). Frequency analyses of the nucleotides neighboring the modifications found that most m⁵C sites were in CHH regions (where H = A, C, U) (SI Appendix, Fig. S1D) and often in the context of AU-rich segments (Fig. 1C). The large majority of modifications were identified within the transcript coding sequences (CDSs; SI Appendix, Fig. S1E), where most of these m⁵C sites were located upstream of stop codons (Fig. 1D) in a different pattern than in human HeLa cells (14) and zebrafish embryos (21). Fig. 1E presents representative examples of *P. yoelii* and *P. falciparum* transcripts that carry m⁵C modifications in the 3' regions of their CDSs.

Profiles of m⁵C modifications in the asexual and sexual stages of *P. yoelii* and *P. falciparum*

Results of RNA-BisSeq analysis identified 7409 m⁵C modifications in 527 mRNA transcripts from the *P. yoelii* schizont set (Py17X_S; SI Appendix, Table S2) and 9168 m⁵C modification in 796 transcripts from the *P. yoelii* gametocyte set (Py17X_G; SI Appendix, Table S2). In contrast, results from the *P. falciparum* datasets showed lower levels of 335 m⁵C modifications in 220 asexual stage transcripts, and 781, 419, and 891 m⁵C modifications in 422 GIII, 224 GIV, and 438 GV transcripts, respectively (SI Appendix, Table S2).

To further investigate potential roles of mRNA m⁵C modifications in the *Plasmodium* life cycle, we performed Gene Ontology (GO) analysis on the sets of methylated transcripts from *P. yoelii* and *P. falciparum*. For *P. yoelii* gametocytes, results from this analysis identified enriched numbers of transcripts relating to metabolic/cellular processes and sexual development. Correspondingly, for *P. falciparum* gametocytes, transcripts relating to cellular and metabolic processes were enriched in GIII stages and transcripts relating to sexual development were enriched in GV stages (Table 1). In the asexual stage parasites including schizonts, m⁵C modifications were enriched with transcripts having assigned GO terms of nucleoside transport, nucleic acid metabolic processes, and chromatin organization (Table 1). These features were also evident in heatmaps of the m⁵C-modified transcripts, which identified similarities as well as distinct differences between the asexual stages and gametocytes both within and between the *P. yoelii* and *P. falciparum* populations (SI Appendix, Fig. S2).

m⁵C-mediated mRNA stability correlates with gametocyte development

Having found that m⁵C methylated transcripts were significantly enriched for GO identifications of sexual development in *P. yoelii* and *P. falciparum* gametocytes, we focused next on genes previously reported to be involved in gametocyte development (30). Our search of the *P. yoelii* data identified 848 genes with transcript abundances of more than 3-fold ($\log_2 \geq 1.585$) in gametocytes relative to asexual parasites (SI Appendix, Table

S3). Similarly, 661 gametocyte development-related genes were identified with ≥ 3 -fold higher transcript abundance in *P. falciparum* gametocytes (stage GIII) vs. asexual stages (SI Appendix, Table S3). On comparative analysis of these gametocyte transcripts, significantly increased m⁵C levels were observed relative to the schizont transcripts of *P. yoelii* and *P. falciparum*, particularly in the comparison using m⁵C-modified stage V gametocyte transcripts (Fig. 2A and B).

Our comparisons confirmed the presence of m⁵C modifications in transcripts with well-established roles in sexual commitment (3, 5, 6, 31). For example, m⁵C methylation of AP2-O (PY17X_0907300) transcripts was not detected in schizonts but was present at a level of 0.34 along with higher expression of these transcripts in gametocytes (SI Appendix, Fig. S3A). In another example, the DEAD/DEAH helicase (PY17X_0313200), an AP2-G-induced transcript involved in male gametocyte development in *P. berghei* (5), was found to be m⁵C methylated in both asexual and sexual stages of *P. yoelii*. Similarly, data from *P. falciparum* showed stage-specific m⁵C methylation of three transcripts that possibly regulate sexual commitment and gametocytogenesis (2, 3, 23), including genes whose disruption is known to lead to greatly reduced gametocyte numbers (GIG, AP2-G2, Pfs16) (2, 3) (SI Appendix, Fig. S3A). However, no detectable m⁵C modification was found for AP2-G transcripts in either the *P. yoelii* or *P. falciparum* malaria parasites.

In view of the above results and m⁵C stabilization of transcripts known from other systems (18, 21), we asked if m⁵C might be associated with mRNA stability in *Plasmodium*

parasites. To address this question, we first assessed the abundance of mRNAs whose m⁵C levels differed in sexual and asexual parasites, relative to the abundance of mRNAs that showed no change of m⁵C levels between these stages. Results from these comparisons suggested a positive correlation between m⁵C level and mRNA abundance in malaria gametocytes relative to schizonts (SI Appendix, Fig. S3B).

We next studied the half-lives of mRNAs in *P. yoelii* gametocytes that were treated with actinomycin D as a transcription inhibitor (18). Results from RNA-Seq analysis indicated that mRNAs with m⁵C modification had significantly longer half-lives than those without m⁵C modification (Fig. 2C and SI Appendix, Table S4). Protein quantification with Tandem Mass Tags (TMT) (SI Appendix, Table S5) showed that slightly greater protein abundance in *P. yoelii* was associated with m⁵C-modified than unmodified transcripts (Fig. 2D).

PyNSUN2 functions as a mRNA m⁵C methyltransferase in *P. yoelii*

To identify potential mRNA m⁵C methyltransferases in *P. yoelii*, we searched for orthologs of hNSUN2 and identified four candidate sequences: PY17X_0804600, PY17X_0938600, PY17X_0920500, PY17X_1447800 (named PyNSUN1–PyNSUN4, respectively) (Fig. 3A). Corresponding searches of the www.plasmoDB.org database identified PF3D7_0704200, PF3D7_1111000, PF3D7_1129400, and PF3D7_1230600 (PfNSUN1–PfNSUN4) as candidate RNA methyltransferases in *P. falciparum*. Sequence comparisons showed that individual members of this family are highly conserved between *P. yoelii* and

P. falciparum, and that the NSUN1 and NSUN2 sequences have close affinities to orthologs in humans, *Arabidopsis thaliana*, *Saccharomyces cerevisiae*, and *Oryza sativa* (Fig. 3A).

To study the role of these putative m⁵C methyltransferases in *Plasmodium*, we performed experiments to disrupt *Pynsun2* in the *P. yoelii* 17XNL line and *Pfnsun2* in the *P. falciparum* NF54 line by CRISPR-Cas9 gene editing (Fig. 3B). Functional disruption attempts of the *sun1*, *sun3*, and *sun4* homologs were performed similarly. After a minimum of three independent transfection attempts with 2 to 3 different single guide RNA (sgRNA) sequences for each individual homolog (SI Appendix, Table S6), we obtained knock-out lines for three of the eight genes (*Pynsun1*, *Pynsun2* and *Pfnsun2*) (Fig. 3C and SI Appendix, Fig. S4A).

Comparative RNA-BisSeq analysis on the schizont stages of 17XNL wild-type (WT) and *Pynsun2* knock-out parasites demonstrated a marked decrease in the density of m⁵C modifications in mRNA transcripts (Fig. 3D). By RNA-BisSeq, 7409 m⁵C sites were detected in 527 mRNAs of the WT parasites whereas this number was reduced to 1845 m⁵C sites in 265 mRNAs of the knock-outs (SI Appendix, Table S2).

To confirm the evidence from knock-out parasites that *Pynsun2* functions as a major m⁵C methyltransferase, we re-introduced *Pynsun2* to its original locus using CRISPR-Cas9 allelic modification (Fig. 3E and SI Appendix, Fig. S4B). Comparative analysis of the restored line (*Pynsun2*-KO-RC) confirmed the return of substantial m⁵C levels, reflecting restored methylation activity (Fig. 3D).

Gametocytogenesis is impaired by disruption of *Plasmodium nsun2* and can be restored by genetic complementation

Considering the differentially greater m⁵C levels in gametocyte vs. schizont mRNA transcripts and the successful knock-outs of *Pynsun1*, *Pfnsun2*, and *Pfnsun2* in blood stage parasites, we asked if expression of these genes might be associated with phenotypes of gametocyte production or parasite stages development in the mosquito. Comparative experiments with the knock-out and WT lines demonstrated that disruptions of *Pynsun2* in *P. yoelii* and *Pfnsun2* in *P. falciparum* resulted in dramatically reduced gametocyte production (Fig. 4A). By contrast, gametocyte production in *Pynsun1* knock-out line was not significantly reduced (Fig. 4B). The disruption of *Pynsun2* affected the production of male and female gametocytes similarly (Fig. 4A) while it showed no detectable effect on the propagation of asexual parasites (SI Appendix, Fig. S5A).

Consistent with the effect of gene disruption on gametocyte production, *in vitro* ookinete conversion rates to mature forms were significantly reduced in the knock-outs relative to those of the WT parasites (Fig. 4C). Also, the numbers of midgut oocysts at day 7 and salivary gland sporozoites at day 14 were markedly decreased in *Pynsun2* knock-out parasites (Fig. 4D), suggesting that *Pynsun2* knock-out effects may impair the development of parasite stages in the mosquito. As with m⁵C modification of the mRNA transcripts,

gametocyte production in the *Pynsun2*-KO-RC line was successfully rescued after restoration of *Pynsun2* expression (Fig. 4A).

To further investigate the association of decreased numbers of gametocytes with diminished mRNA m⁵C modification, we identified transcripts whose m⁵C levels were significantly reduced in *Pynsun2*-KO relative to WT schizonts (SI Appendix, Table S7).

Many of these genes showed significant *PyNsun2*-KO associated expression changes, including 145 instances of upregulation and 473 instances of downregulation (Fig. 4E), both of which were reversed after restoration of *Pynsun2* expression in the knock-out line (Fig. 4F). Among the downregulated genes, we identified transcripts that have been reported to be AP2-G-mediated and involved in gametocyte development (5) (SI Appendix, Fig. S5B). These included transcripts of Py17x_0313200 (DEAD/DEATH), disruption of which can greatly reduce the number of gametocyte parasites in *P. berghei* (5). Interestingly, no significant change of AP2-G expression itself was evident with the *Pynsun2* knock-out (SI Appendix, Fig. S5C).

Discussion

The results of this study illuminate a landscape of m⁵C mRNA transcriptome methylations in the schizonts and gametocytes of two *Plasmodium* species, *P. yoelii* and *P. falciparum*, and show that these epitranscriptomic modifications are linked to development of gametocytes. Further, the searches for possible RNA methyltransferases and findings from

experiments to knock out and repair the candidate genes indicate that a homolog of the NSUN2 family plays a major role in these m⁵C mRNA modifications. Although little or no effect on the growth of asexual blood stage parasites was detected in these experiments, markedly decreased gametocyte production was found after disruption of the *Pynsun2* homolog in *P. yoelii* (*Pynsun2*-KO) and disruption of the *Pfnsun2* homolog in *P. falciparum* (*Pfnsun2*-KO). Moreover, in evaluations of *Pynsun2*-KO infectivity, the conversion rate of ookinetes from immature to mature forms was also decreased, suggesting that m⁵C mRNA modifications may be important both to gametocyte production and to parasite development following zygote formation in the mosquito.

Dynamic m⁵C modifications are vital to gene regulatory networks in a diverse variety of biological processes including embryogenesis in zebrafish (21) and stages of development in *Arabidopsis* (13). Global transcriptome fluctuations of 30% in m⁵C methylation occur with altered photosynthesis efficiency and heat tolerance in rice (19); and m⁵C fluctuations have likewise been associated with the maternal-to-zygotic transition and the development of bladder carcinoma (18, 21). The m⁵C modification patterns observed in the present study, including increased levels of methylation in gametocytes and their loss following knock out of *Plasmodium* NSUN2, are also associated with the essential processes of male and female gametocytogenesis and progression to mature ookinetes following fertilization.

m⁵C modifications can stabilize mRNA transcripts and increase half-life, and they may also influence export of the transcripts for translation in the cytoplasm (14, 21). In the present work, several transcripts known to be involved in sexual commitment and gametocyte development were found to be m⁵C-modified, suggesting a role for this form of post-transcriptional methylation in the timing and control of their expression. Interestingly, no m⁵C modification was found for AP2-G, a master regulator of sexual commitment and development (32-34), nor was there any evidence for a change of AP2-G expression in NSUN2-disrupted parasites. These observations suggest that pathways downstream of AP2-G are epigenetically modulated at the post-transcriptional level. Knock-outs of NSUN2 and loss of m⁵C methylation may adversely affect the stability and translation of mRNA transcripts in these pathways, resulting in dysregulation and impairment of effective gametocytogenesis.

As in other organisms (13, 14, 21), the m⁵C sites of transcripts targeted by NSUN2 are found predominantly in the protein coding regions of *Plasmodium* mRNAs. However, the profiles of m⁵C distributions within these regions differ among organisms. In the two *Plasmodium* species studied here, a peak of m⁵C levels in CDSs occurs just upstream of the stop codon. In contrast, the m⁵C levels in cells from human and mouse cells or from zebrafish embryos are more uniformly distributed across the CDSs or are enriched near the translation initiation sites. *Arabidopsis* shows two peaks of m⁵C enrichment that flank the stop codon. Differences in the consensus site logos for m⁵C modification are also present

between *Plasmodium* spp. and other eukaryotes (18, 19, 35). In *P. yoelii* and *P. falciparum*, the target C nucleotide is flanked on each side by a highly AU-rich sequence, which may in part reflect the correspondingly high AT content of the genomes and transcriptomes of these two species.

Our homology searches for sequences related to NSUN2 identified other potential methylases, termed *Plasmodium* NSUN1, NSUN3, and NSUN4 in this work. Multiple attempts with CRISPR/CAS9 strategies to knock out these different homologs in *P. yoelii* and *P. falciparum* yielded only one parasite line with a disruption of *Pynsun1*. However, unlike the *Pynsun2* and *Pfnsun2* knock-outs, the *Pynsun1*-KO line showed little or no reduction of gametocyte production relative to that of WT *P. yoelii*. Further investigations are needed to establish phenotypes associated with *Plasmodium* NSUN1 as well as NSUN3 and NSUN4, which, in light of the multiple unsuccessful knock out attempts, are likely to be essential for parasite survival. We note that the levels of m⁵C methylation in schizonts of the *Pynsun2*-KO and *Pfnsun2*-KO lines were decreased by only 30–50% relative to WT schizonts, as measured by both RNA-BisSeq and LC-MS/MS. Writer(s) in addition to NSUN2 may thus be involved in m⁵C transcript modifications of asexual as well as sexual stage parasites. These possibilities, in addition to the potential activities of m⁵C methyltransferases on other RNAs such as tRNA and ncRNAs (36) also remain to be explored.

We asked if m⁵C modifications of schizonts and gametocytes might be subject to an ‘eraser’ (demethylase) activity, but currently no such m⁵C demethylase has been identified in any eukaryote in which the landscape and biological roles of m⁵C mRNA modification have been well studied. While this does not rule out the possibility that erasers may yet be found, it is possible that m⁵C levels may simply rise and fall due to different rates of m⁵C writer activity and subsequent degradation of m⁵C-modified transcripts. This possibility is supported by the positive correlation between m⁵C status and mRNA stability (half-life) in gametocytes, as well as the greater m⁵C/C levels detected by LC-MS/MS in gametocytes relative to shorter-lived schizonts.

Control of mRNA turnover and translational repression has been shown to be essential for *Plasmodium* sexual development and is mediated by a member of the DDX6-family of DEAD box RNA helicases (37). Translational repression also occurs in mature sporozoites and is associated with their quiescence in mosquito salivary glands until blood feeding and salivation (38-40).

Here we have demonstrated that m⁵C modifications contribute to the stabilization of mRNA transcripts in *Plasmodium*, play an essential role in sexual stage development, and may be important for the maturation of ookinetes. We also speculate that m⁵C “reader” molecules may recognize these modifications and facilitate the transport of methylated transcripts to the cytoplasm for eventual translation. Further investigation of these m⁵C-mediated processes and those of other epitranscriptomic modifications such as m⁶A (12)

could yield new insights into gametocytogenesis, mosquito infectivity, and molecular pathways that might be exploited to interrupt malaria transmission.

Methods

Animal use and ethics statement

All animal experiments were performed in accordance with approved protocols (XMULAC20140004) by the Committee for Care and Use of Laboratory Animals of Xiamen University. ICR mice (female, 5 to 6 weeks old) were purchased and housed in the Animal Care Center of Xiamen University and kept at room temperature under a 12 h light/dark cycle at a constant relative humidity of 45% and were used for parasite propagation, drug selection, parasite cloning, and mosquito feedings.

Parasite culture and mouse infections

P. falciparum strain 3D7 (NF54) parasites were cultured and synchronized as previously reported (12). Human erythrocytes were obtained under approved clinical protocol at the Shanghai Blood Center. Microscopy of Giemsa-stained thin blood films was used to monitor parasite development.

P. yoelii 17XNL strain parasites were obtained from the Malaria Research and

Reference	Reagent	Resource	Center	(MR4)
-----------	---------	----------	--------	-------

(<https://www.beiresources.org/About/MR4.aspx>). The parasites were injected intravenously or intraperitoneally into mice for infection, and parasitemia was monitored by Giemsa-stained thin blood films.

Dot blot detection of m⁵C Levels

Dot blot detection was performed as described previously (41). Biotin-labeled RNA oligonucleotides synthesized with or without m⁵C were used as positive and negative controls. mRNA was isolated from preparations of total RNA by the Dynabeads® mRNA Purification Kit (Thermo Fisher Scientific). mRNA was denatured at 95 °C in a heat block for three min and chilled on ice immediately to prevent the re-formation of secondary structures of mRNA. A sample containing 0.5µg mRNA was applied directly onto the Hybond-N+ membrane (Millipore) and secured to the membrane by two treatment rounds in a HL-2000 Hybrilinker™ Hybridization Oven/UV Crosslinker (Analytik Jena US) operating in Autocrosslink mode (setting of 1,200 microjoules [x100]; 25-50 sec). After UV crosslinking, the membrane was incubated in 10 ml of 5% of non-fat milk in PBST (PBS with 0.05% Tween-20) for 1 h at room temperature with gentle shaking. The primary mouse anti-m⁵C antibody (1:1000; Abcam, ab10805) was then added in 10 ml of antibody dilution buffer (non-fat milk in PBST) and incubated overnight at 4 °C. After three times washing (5 min each) in 10 ml of PBST, membrane was finally incubated with HRP-conjugated anti-mouse IgG (1:5000; GE Healthcare) overnight at 4 °C. Signals from m⁵C

were developed with an enhanced chemiluminescence (ECL) western blotting kit (GE Healthcare).

mRNA modification determinations by LC-MS/MS

Total RNA was extracted from synchronized parasites samples using Trizol (ThermoFisher Scientific) followed by treatment with DNase I (ThermoFisher Scientific). mRNA was isolated from total RNA samples using NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB). Purified mRNA was quantified using the Qubit RNA HS Assay kit (ThermoFisher Scientific). mRNA was hydrolysed to single nucleotides and the pretreated nucleoside solution was deproteinized using a Satorius 10,000-Da MWCO spin filter. Analysis of nucleoside mixtures was performed on an Agilent 6460 QQQ mass spectrometer with an Agilent 1260 HPLC system. Multi reaction monitoring (MRM) mode was performed due to its high selectivity and sensitivity attained through working with parent-to-product ion transitions. LC-MS data was acquired using Agilent Qualitative Analysis software. MRM peaks of each modified nucleoside were extracted and normalized to the amount of mRNA purified. Detailed analysis of mRNA modifications was performed as previously described (12, 35).

Transcriptome sequencing

Total RNA was isolated from parasite preparations by the Direct-zol RNA Kit (Zymo Research). The poly(A) RNA fraction was selected using KAPA mRNA Capture Beads (KAPA), and fragmented to about 300–400 nucleotides (nt) in length. A KAPA Stranded mRNA-Seq Kit (KAPA) was used for RNA-seq library preparation for strand-specific RNA-seq. The libraries were sequenced on an Illumina NovaSeq 6000 platform to generate 150 bp pair-end reads.

RNA-BisSeq data collection from transcripts and analysis

Parasites were isolated by treatment of 200 – 300 µl of loosely packed red blood cells from blood or culture at 3% parasitemia with 0.15% saponin in 1 × PBS on ice for 10 min. The parasite pellets were washed twice with precooling 1 × PBS and then resuspended in 1 mL TRIzol (Invitrogen). Total RNA was extracted using the Direct-zol RNA Kit (Zymo Research) according to the manufacturer's instructions. mRNA enrichment from the total RNA was performed using the Dynabeads™ mRNA Purification Kit (Invitrogen) according to the manufacturer's instructions.

RNA bisulfite treatment and purification of converted RNA was performed using the Methylamp™ RNA Bisulfite Conversion Kit (Epigentek Group Inc.) according to the manufacturer's instructions. One µg of mRNAs along with 5 ng of dihydrofolate reductase (*dhfr*) RNA (1:2000) were used as an input RNA and unmethylated control, respectively.

The resulting RNA was subsequently used for fragmentation and library construction using the KAPA Stranded mRNA-Seq Kit (KAPA) according to the instructions provided by the manufacturer. Sequencing was performed on an Illumina NovaSeq 6000 instrument with paired end 150-bp read length.

Raw sequencing data were filtered to remove low-quality reads and adapter contaminations with cutadapt (42) as detailed in Methods. Reads with average quality score ≥ 20 and length ≥ 50 bp were kept. Clean reads were mapped to the *P. falciparum* 3D7 genome (Pf 3D7 v32, obtained from PlasmoDB) or to the *P. yoelii* 17X genome (Py17x v45, obtained from PlasmoDB) using meRanGh, and only unambiguously aligned reads were used to call m⁵C sites using the meRanCall tool from the meRanTK toolkit (43) (FDR < 0.01). Analysis of the *dhfr* spike-in revealed C to T conversions were around 98% (SI Appendix, Table S1). At least 20 reads were obtained for the candidate cytosine positions, and were recorded if methylated cytosine depth was greater than 5 and m⁵C level was greater than 0.1. mRNA-BisSeq data collection and analysis were performed in two replicate experiments. Detection of an m⁵C site in both replicates was required for its inclusion in further analysis. The m⁵C sites were mapped to the regions of introns, exons, 5'-UTRs, and 3'-UTRs as previously described (14). The meRanCall tool from meRanTK was used to localize m⁵C sites preferences with an input parameter of -sc 10. Logo plots were generated with R package ggseqlogo (44). PyNSUN2 targeted genes were defined as

transcripts with m⁵C modifications in WT schizont parasites that are lost in *Pynsun2*-KO parasites.

Plasmid construction and transfection

To disrupt the *PfNSUN2* opening reading frame (ORF) CDS in WT 3D7 and NF54 strains, we used CRSIPR/Cas9 to generate knock-out vectors. Briefly, the pL6 plasmid (45) was modified to contain a complementary sequence of sgRNA between the Xho I and Avr II sites. For each knock-out assay, at least two sgRNAs were individually incorporated into plasmids to target the CDS of their target gene (SI Appendix, Table S6). Sequences (~1kb) homologous to the target genes were designed to include several base pair deletions and were inserted into the pL6 plasmid between Asc I and Afl II sites by the In-Fusion PCR Cloning System (primer pairs are listed in SI Appendix, Table S6). Recombinant plasmids with were simultaneously transfected along with pUF-Cas9 vectors into ring-stage parasites of 3D7-G7 and NF54-F5 by electroporation as described previously (46). Drug-resistant parasites were obtained after three or four weeks of continuous culture under 2.5 nM WR99210 pressure (12). Selected knock-out or knock-in parasites were confirmed by PCR detection and cloned by limiting dilution.

P. yoelii genetic modifications were performed using the CRISPR/Cas9 plasmid pYCM as previously described (47). To obtain gene knock-outs, 5' and 3'-genomic fragments (400 to 700 bp) of target genes were amplified using the corresponding primers (SI Appendix,

Table S6). The PCR products were restriction-digested and cloned into matched sites of the pYCM vector. SgRNA oligonucleotides were annealed and inserted into the pYCM vector. At least two sgRNAs were designed to disrupt the CDS of a target gene for each deletion modification using the online program EuPaGDT (48). For PyNSUN2 tagging, a 400 to 800 bp fragment from C-terminal of the CDS and 400 to 800 bp segment from the 3'-UTR were amplified and fused with a DNA sequence 6HA in frame at C-terminal of the gene. There were at least two sgRNAs designed to target sites close to the C-terminal of the gene CDS region for incorporation of the insert. Five micrograms of circular recombinant plasmid DNA were electroporated into infected RBCs using a Lonza Nucleofector, and the transfected parasites were intravenously injected into a naïve ICR mouse. Infected mice were treated with pyrimethamine (6 µg/ml) in drinking water; drug-resistant parasites were obtained following 5 to 7 days of drug selection.

***In vitro* transcription of mouse *dhfr* RNA**

Institute of Cancer Research (ICR) mouse heart tissue (0.05g) was obtained and then resuspended in 1 mL TRIzol (Invitrogen) for RNA extraction. The full-length *Mus musculus dhfr* CDS was amplified from cDNA and cloned between the XbaI and XhoI sites of the pBSKS vector (Stratagene, La Jolla, CA) containing a T7 promoter (forward primer: 5'-CACCGCGGTGGCGGCCGCTCTAGAATGGTTC GACCATTGAACTGC-3'; reverse primer: 5'-GTACCGGGCCCCCCTCGAGTTAGTCTTTCTTCTCGTAGACTTC-3'). The pBSKS-dhfr vector was linearized by FastDigest XhoI (Thermo Scientific™). Five

hundred nanograms of linearized pBSKS-dhfr vector was used as a DNA template for *in vitro* transcription with T7 RNA Polymerase (New England Biolabs) at 37 °C for 4 h in a 20 µl reaction mixture, according to the manufacturer's instructions. The DNA template was removed by DNase I (Invitrogen™) digestion for 15 min at room temperature.

Western blotting

Protein fractionation was performed with modifications according to Till *et al* (49) and Baumgarten *et al* (12). Synchronous middle-late stage parasites were released from 1 mL infected red blood cells by 0.15% saponin lysis and washed twice in 1 × pre-cooled PBS. The parasite pellet was first resuspended in 0.5 mL cytoplasmic lysis buffer (20mM Hepes pH7.9, 10mM KCl, 1.5mM MgCl₂, 1mM EDTA, 1mM EGTA, 0.65% NP-40, 1mM DTT, 1x Protease inhibitor cocktail) and incubated for 30 min on ice. Parasites were further homogenized for 100 strokes in a glass douncer, and the supernatant containing cytoplasmic proteins was isolated via centrifugation (10,000 rpm, 10 min, 4 °C). Protein samples were separated on SDS-PAGE and transferred onto PVDF membrane and visualized using ECL Western Blotting Detection Kit (GE Healthcare). The primary antibodies used in this study were mouse anti-HA (1:2000, Roche) and anti-GAPDH (1:1000, Servicebio, cat# GB12002), and secondary antibody was HRP-conjugated goat anti-mouse IgG (1:5000, Abcam, cat# ab6789) .

Gametocyte induction

For the *P. yoelii* 17XNL strain, gametocyte induction in mice was performed as previously described (47). Briefly, phenylhydrazine (80 µg/g mouse body weight) was administered to the ICR mice through intraperitoneal injection. The mice were infected three days later with 3.0×10^6 parasites through tail vein injection. Gametocytemia typically peaked at day 3 post injection. Giemsa staining of blood smears was used to count the number of male and female gametocytes. Gametocytemia was defined as the ratio of sex-specific gametocyte (male or female) over total infected erythrocytes. All experiments were repeated three times independently.

For *P. falciparum*, gametocyte induction was performed as described previously with minor modifications (50, 51). The *P. falciparum* population was synchronized and expanded to 8% parasitemia in culture (4% haematocrit). Medium changes were performed daily without dilution of the culture and without disturbing the red blood cell layer at the bottom of the culture dish. Giemsa's solution staining of thin blood smears was used to count the number of gametocytes from day 3. Stage III gametocytes were harvested on days 5–6, stage IV on days 8–9, and stage V on days 12–13. Gametocytemia determined on day 8-9 was used to calculate sexual commitment (%). Each experiment had three biological replicates.

Ookinete conversion rate determinations

Conversion rates to mature ookinetes were determined as recently described (52). Samples taken from culture 12 h after fertilization were stained by Giemsa solution (Sigma, cat# GS80). Each conversion rate was calculated as the number of mature ookinetes (stage V) over the number of total ookinetes (stages I–V).

RNA-Seq for mRNA half-life

ICR mouse blood with 4–6% gametocytemia was collected in anticoagulant tubes with heparin and immediately added to gametocyte culture medium (RPMI 1640 medium) in a blood/medium volume ratio of 1:10. Actinomycin D (Sigma-Aldrich) was added at 20 µg/ml (53), and samples were collected 0, 1, 3, and 5 h later. Total RNA was extracted by Direct-zol RNA Kit (Zymo Research) as per the manufacturer's instructions. Prior to construction of the library with KAPA Stranded mRNA-Seq Kit Illumina platform (KK8421), ERCC RNA Spike-In Control Mixes (Ambion) was added proportionally to the total RNA of each sample (0.1 µL per sample) (54). Two biological replicates were performed for gametocyte samples.

Mosquito maintenance and mosquito feeding

Sucrose solution (10%) was used to feed *Anopheles stephensi* mosquitoes (strain Hor) (55), which were reared at 28 °C, 80% relative humidity, in a 12 h light/dark cycle and blood

feeding assay was performed at 23°C. Each anaesthetized mouse with 4–6% gametocytemia was used to feed 80 female mosquitoes for 30 min. For oocyst detection, mosquito midguts were dissected seven days after blood-feeding and stained using 0.1% mercurochrome. For salivary gland sporozoite counting, salivary glands from 30 – 50 mosquitoes were dissected on day 14 post blood-feeding, and the number of sporozoites per mosquito was calculated.

RNA sequence analysis

Low-quality and adaptor sequences were trimmed using cutadapt (v1.18) (42) with parameters: -a AGATCGGAAGAGC -AAGATCGGAAGAGC –trim-n -m 50 -q 20,20. RNA sequencing reads were aligned using hisat2 (v2.1.0) (56) with strand specific mode (--rna-strandedness RF) to the *P. falciparum* 3D7 genome (Pf 3D7 v32, obtained from PlasmoDB) or to the *P. yoelii* 17X genome (Py17x v45, obtained from PlasmoDB). Mapped reads were subsequently assembled into transcripts guided by the PlasmoDB gff annotation files (Pf3D7 v32/ Py17x v45) using featureCounts (v1.6.1) (57) with parameters: -M -p -B -C -s 2. Read counts were obtained using featureCounts (v1.6.1). FPKM for genes were calculated according to the formula $FPKM = \frac{\text{readcount} \times 10^3}{\text{length} \times \text{totalcount} \times 10^{-6}}$. Genes with foldchange ≥ 3 were considered as differentially expressed genes for compared groups.

Gene homolog identification and phylogenetic analysis of the m⁵C writer

Candidate proteins of the malaria parasite m⁵C writer complex were identified by blast search against PlasmoDB database (<https://plasmodb.org/plasmo/>) using the amino acid sequence of human mRNA m⁵C methylase NSUN2 as a query. The amino acid sequences of human, rice, and *Arabidopsis* RNA (cytosine-5)-methyltransferases (RCMTs) domain family members were obtained from the NCBI database (<https://www.ncbi.nlm.nih.gov/>). Clustal X software (58) was used to construct the Neighbor-joining phylogenetic tree with 1,000 bootstrap replicates.

mRNA half-life analysis

Raw sequencing data were filtered to remove low-quality reads and adapter contaminations with cutadapt (v1.18) (42), as described above. Reads with average quality score ≥ 20 and length ≥ 50 bp were kept. Clean reads were mapped to the *P. yoelii* 17X genome (Py17x v45, obtained from PlasmoDB) with ERCC spike in sequence using hisat2 (v2.1.0) with strand specific mode (--rna-strandness RF). Uniquely mapped reads were used to calculate read count for each transcript using software featureCounts. To remove low abundance transcripts, transcripts with at least five read counts at 0 h were kept for downstream half-life analysis. The read count of each transcript in a time point was normalized with the sum of reads mapping to all ERCC spike-in transcripts. Half-lives of transcripts were calculated using R package MINPACK (59). The decay rates for transcripts were constant upon transcription inhibition. Transcripts are assumed to exponentially decay. Therefore, the

transcripts remaining at specific time point were calculated by the equation: $C = C_0 * e^{-kt}$.

Here, C represents transcript abundance as a function of time, C_0 is the transcript abundance at 0 h, and k is the first order rate constant, from which half-life values can be derived with the equation: $t_{1/2} = \ln(2)/k$

Proteomics analysis

TMT-based quantitative proteomics analysis was performed by Novogene Bioinformatics Technology Co Ltd. (Beijing, China), as described previously (60). Briefly, gametocytes of *P. yoelii* were collected and ground individually with liquid nitrogen and lysed in protein lysis buffer (100 mM NH_4HCO_3 , 8 M Urea and 0.2% SDS, pH = 8.0) followed by ultrasonication on ice for 5 min. Desalted peptides were labeled with TMT 6-plex reagents (TMT6plex™ Isobaric Label Reagent Set, Thermo Fisher) according to the manufacturer's instructions. Labeled peptides of all samples were mixed equally and desalted by peptide desalting spin columns (Thermo Fisher, 89,852). A C18 column (Waters BEH C18, 4.6 × 250 mm, 5 μm) on a Rigol L3000 HPLC was used to fractionate for TMT-labeled peptide mix. An EASY-nLC™ 1200 UHPLC system (Thermo Fisher) coupled with an Orbitrap Q Exactive™ HF-X mass spectrometer (Thermo Fisher) operated in the data-dependent acquisition (DDA) mode was used for shotgun proteomics analyses. The resulting spectra from each fraction were searched separately against *P. yoelii* 17X proteome (Py17x v45, obtained from PlasmoDB) by the search engines: Proteome Discoverer 2.2 (PD 2.2,

thermo). At least one unique peptide with false discovery rates (FDR) < 1.0 % on the peptide and protein level, respectively, was considered as a confident peptide spectrum match for further analysis. The intensities of the TMT reporter ions against those ones provided in the databases were compared for protein quantification.

Gene Ontology (GO) analysis

GO analysis of specific genes were performed using EnrichGO tool from Plasmodb (<https://plasmodb.org/>). Only terms in the biological process category were shown. GO terms with p-value ≤ 0.05 were considered as statistically significant terms. Genes enriched for terms sexual development (GO: GNF0004) were obtained from a previous study (61) and used for enrichment analysis..

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC) [81630063, 81971959, 31671353] to Q.Z., the Division of Intramural Research at the U.S. National Institute of Allergy and Infectious Diseases (J.M. and T.E.W.), and National Key R&D Program of China Grants [2018YFA0507300] to Q.Z., NSFC [31771419, 31721003] to C.J., NSFC [31772443, 31970387] to J.Y.

Author Contributions

QZ, CJ, JY, JM and TW conceived and designed experiments. GG, PQ, BL and XH performed the majority of experiments with the help from XS, QM, XY and YZ. ML, GY, SS, WL, LW, LG and JM conducted most bioinformatics analysis. QZ, CJ, JY, JM and TW wrote the manuscript with contribution from ML, GG, LJ, JC and RC. All authors discussed and approved the manuscript.

Conflict of Interest Statement

The authors declare no competing financial interests.

Codes and Data Availability

All data is available in the main text or the supplementary materials. The raw and processed high-throughput sequencing data of this study have been deposited into the Gene Expression Omnibus (GEO) and are publicly accessible under accession number GSE159127 (token: gjcrewgkdnijrsh).

Figure legends

Figure 1. Features of m⁵C-modified mRNA in malaria parasites. A, Dot blot assays show detection of m⁵C modifications in schizont-stage mRNAs from *P. falciparum* and *P. yoelii*. m⁵C-modified and unmodified RNA oligonucleotides served as positive and

negative controls, respectively. **B**, LC-MS/MS signals indicating modified nucleotides from *P. yoelii* schizont-stage mRNA. rC, cytosine; rU, uracil; rG, guanosine; rA, adenosine; m⁶A, N⁶-methyladenosine; m⁷G, N⁷-methyl-2'-guanosine; Am, 2'-O-methyladenosine; m⁵C, 5-methylcytosine; Gm, 2'-O-methylguanosine; m⁵U, 5-methyluridine; Cm, 2'-O-methylcytosine; hm⁵C, 5-Hydroxymethylcytosine. **C**, Frequency logo displays of the nucleotides proximal to mRNA m⁵C sites in the schizonts (S) and gametocytes (G) of *P. yoelii* (Py), and the schizonts (S) and stage III, IV, and V gametocytes (GIII, GIV, GV) of *P. falciparum* (Pf). **D**, Density distribution of the m⁵C sites in mRNA transcripts of *P. yoelii* and *P. falciparum*. The moving averages (10-bp window) of percentage mRNA cytosine content (light blue) are lower in the 5' and 3' UTR regions than in the CDS, as expected for these species. **E**, Integrative Genomics Viewer snapshots of the transcript sequence levels and m⁵C sites of expressed genes in *P. yoelii* (left) and *P. falciparum* (right). Vertical red bars indicate the m⁵C levels in specific parasite stages.

Figure 2. High levels of m⁵C methylation and its association with transcript longevity and protein expression in gametocytes. **A-B**, Heatmaps indicate the expression of transcripts whose levels are ≥ 3 -fold higher in gametocytes than in schizonts of *P. yoelii* and *P. falciparum*. Boxplots show corresponding m⁵C levels of those genes at different stages. (***: p-value < 0.001, Wilcoxon test). **C**, Boxplot and cumulative fraction plot indicate the longer mRNA half-lives of m⁵C methylated (red) relative to non-methylated

(blue) transcripts in *P. yoelii* gametocytes (***: p-value < 0.001, Wilcoxon test.). **D**, Boxplot and cumulative fraction plot compare the levels of proteins translated from m⁵C-methylated (red) and non-methylated (blue) transcripts in *P. yoelii* gametocytes (***: *P* < 0.001, Wilcoxon test).

Figure 3. Identification and verification of PyNSUN2 as a m⁵C methyltransferase. A,

Phylogenetic analysis of putative RNA (cytosine-5)-Methyltransferases (NSUNs) of *P. yoelii* and *P. falciparum*. Sequences were aligned using ClustalX 2.1. The neighbor joining (NJ) phylogeny was performed using MEGA 5.2.2 with 1000 replicates. **B**,

Strategies to generate *Pynsun2*-KO and *Pfnsun2*-KO parasites. For CRISPR/Cas9-mediated deletion of the *Pynsun2* CDS, Left Arm (LA) and Right Arm (RA) match

sequences in the 5' UTR and 3' UTR of *Pynsun2*, respectively. For disruption of *Pfnsun2*,

sgRNA targeting was used to delete a portion of the CDS (red box) and introduce a stop

codon. PCR products with indicated primer pairs confirmed the expected differences

between the WT and allelically-manipulated parasites. Sequence data confirming the

deletion and stop codon in the *Pfnsun2*-KO parasites is shown in **SI Appendix, Fig. S4A**.

C, Summary results from experiments to disrupt homologs of the NSUN family in *P. yoelii*

and *P. falciparum*. NT: not tested. **D**, Histogram shows numbers of mRNA m⁵C sites in

schizont stages of WT, *Pynsun2*-KO and *Pynsun2*-KO-RC *P. yoelii* clones. Box plot shows

the corresponding m⁵C levels in WT, *Pynsun2*-KO and *Pynsun2*-KO-RC lines (right panel).

(***: P -value < 0.001, Wilcoxon test.) **E**, Strategy for genetic complementation repair of the *Pynsun2*-KO with the gene CDS fused with a C-terminal 6HA. Upper panel shows the design for the CRISPR/Cas9 mediated gene knock-in. Left Arm (LA) and Right Arm (RA) match sequences in the 5' UTR and 3' UTR of *Pynsun2*, respectively. Red thunderbolt indicates the site for sgRNA targeting. Bottom panel shows the PCR products in the three complemented clones (RC) by the indicated primer sets.

Figure 4. Loss of gametocytogenesis in *Plasmodium* NSUN2 knock-out lines and its restoration by gene complementation. **A**, Gametocytogenesis is markedly decreased in *Pynsun2*-KO and *Pfnsun2*-KO parasites relative to high levels in the original WT lines. Repair of the knock-out in *Pynsun2*-KO parasites restores gametocyte production. Error bars represent median and 95% CIs. (***: P -value < 0.001; **: P -value < 0.01; N.S.: no significant difference, Wilcoxon test.) **B**, No significant effect on gametocyte production detected after knock-out of the *Pynsun1* gene. Error bars represent median and 95% CIs. (N.S.: no significant difference, Wilcoxon test.) **C**, Markedly decreased ookinete conversion rates in the *Pynsun2*-KO relative to WT parasites. Data were obtained from two independent replicates. Error bars represent median and 95% CIs. (***: P < 0.001, Wilcoxon test). **D**, Number of oocysts in mosquito midguts 7 days after blood feeding with WT or *Pynsun2*-KO gametocytes (left), and number of sporozoites in mosquito salivary glands 14 days after blood feeding in the same experiments (right). The number of

mosquitoes dissected in each group is indicated. Medians and 95% ICs are indicated. (***: $P < 0.0001$, Wilcoxon test). **E**, Identification of transcripts that are differentially expressed between *Pynsun2*-KO and WT parasites at the schizont stage. Transcripts that were found to have ≥ 3 -fold up- or down-regulation are marked in red or blue, respectively. **F**, Upregulation and downregulation of transcripts in the knock-out clone are reversed by restoration of *Pynsun2* expression in the *Pynsun2*-KO-RC clone. (***: $P < 0.001$, Wilcoxon test).

Figure S1. Detection and transcriptome-wide mapping of m⁵C mRNA modifications

in *Plasmodium*. **A**, Venn diagrams display the numbers of mRNA m⁵C sites that are shared or unique in two independent replicate determinations (rep1, rep2) from schizont and gametocyte stages of *P. yoelii* (Py.S, Py.G) and *P. falciparum* (Pf.S, Pf.GIII, Pf.GIV, Pf.GV). **B**, Box plots indicate m⁵C levels at the shared and unique sites in each replicate pair. (***: p-value < 0.001 , Wilcoxon test). Maximum, minimum, and median values are indicated in the box plots. **C**, Histograms and box plots indicate the distributions of modification levels at m⁵C sites in the schizont and gametocyte stages of *P. yoelii* and *P. falciparum*. **D**, Frequency histograms indicate the number of transcripts by the m⁵C sites they carry in schizonts and gametocytes of *P. yoelii* (Py.S, Py.G) and *P. falciparum* (Pf.S, Pf.GIII, Pf.GIV, Pf.GV). **E**, Proportions of m⁵C sites in each parasite stage are displayed by sequence context: CHH (orange), CHG (red) and CGH (green), where H represents A, C or U. **F**,

Transcriptome-wide distributions of mRNA m⁵C sites in the predicted CDS, introns, and flanking regions of genes expressed in schizonts and gametocytes of *P. yoelii* (Py.S, Py.G) and *P. falciparum* (Pf.S, Pf.GIII, Pf.GIV, Pf.GV). The presence of m⁵C in predicted intron regions suggests the methylation of incompletely or alternatively spliced transcripts.

Figure S2. Heatmaps representing m⁵C modifications in schizont and gametocyte stages of *P. yoelii* (Py.S, Py.G) and *P. falciparum* (Pf.S, Pf.GIII, Pf.GIV, Pf.GV). Vertical bars mark regions with sequences that match with GO terms for sexual development and nucleic acid metabolic processes.

Figure S3. Comparatively high m⁵C methylation levels correlate with expression of gametocytogenesis-associated genes. **A**, Heat map displays are aligned to indicate transcription levels alongside the relative m⁵C levels in *P. yoelii* and *P. falciparum* schizonts and gametocytes. Data on transcript sequences and m⁵C levels were each obtained from two biological replicates using RNA-Seq and RNA-BisSeq, respectively. **B**, Abundance of transcripts in gametocytes relative to schizonts correlates with their stage-specific m⁵C methylation levels. Levels of individual transcripts in gametocytes of *P. yoelii* (G) or *P. falciparum* (stage V, GV) relative to their levels in schizonts (S) are plotted by log₂ ratio. Results from transcripts with lesser or greater m⁵C levels in gametocytes than

schizonts are shown in blue ($S > G$) and red ($S < G$), respectively, while those with little or no difference of m^5C levels are shown in yellow ($S = G$).

Figure S4. Disruption of the *Pfnsun2* gene in *P. falciparum* and restoration of *Pynsun2*

expression in *P. yoelii* knock-out parasites. **A**, Predicted protein sequences and electropherograms from the *P. falciparum* WT and *Pfnsun2*-KO parasites. An early stop codon is introduced by the genetic modification. Red arrow indicates the deletion site. **B**, Immunoblot detection of 6×HA tagged PyNSUN2 protein in the asexual blood stages of *Pynsun2*-KO-RC and in a separate parasite strain (*Pynsun2*-6HA) containing a 6×HA tagged version of endogenous PyNSUN2. Control bands of GAPDH detection verify comparable loading of the lanes.

Figure S5. PyNSUN2 deficiency leads to a reduction in gametocyte production. **A**,

Results demonstrating comparable development of parasitemias in WT- and *Pynsun2*-KO-infected mice. Data points are means $\pm$ SEM from three independent biological experiments. **B**, Decreased expression of AP2-G induced genes associates with dramatically reduced m^5C levels in the *Pynsun2*-KO (KO) parasites. Error bars represent SEM for two biological replicates. **C**, Histogram of AP2-G (PY17X_1440000) transcript levels in WT and *Pynsun2*-KO parasites. Medians and 95% ICs are indicated. (N.S.: no significant difference, Wilcoxon test.)

824

825 **Table 1** Gene Ontology (GO) for m⁵C-modified transcripts in schizonts and gametocytes

826 of *P. yoelii* (E) and *P. falciparum* (F).

827

828 **SI Appendix, Table S1.** Summary of RNA-BisSeq mapping results and statistics of m⁵C

829 sites per transcript from schizonts and gametocytes of *P. yoelii* and *P. falciparum*

830 **SI Appendix, Table S2.** m⁵C levels of transcripts in schizont and gametocyte stages of *P.*

831 *falciparum* and *P. yoelii* (WT, *Pynsun2*-KO, and *Pynsun2*-KO-RC lines)

832 **SI Appendix, Table S3.** Transcripts with greater than 3-fold upregulation in gametocytes

833 relative to schizonts of the *P. falciparum* (NF54) and *P. yoelii* (17XNL) lines

834 **SI Appendix, Table S4.** Individual half-lives of transcripts in *P. yoelii* gametocyte stages

835 **SI Appendix, Table S5.** Protein abundances in schizonts and gametocytes of *P. yoelii* (two

836 replicates, TMT quantification)

837 **SI Appendix, Table S6,** Primer and gRNA sequences used for plasmid construction and

838 gene knock-out experiments

839 **SI Appendix, Table S7.** Genes exhibiting ≥ 3 -fold expression changes between the

840 *P. yoelii* WT, *Pynsun2*-KO, and *Pynsun2*-KO-RC parasites.

841

842 **References:**

- 843 1. World Health Organization (2020) World Malaria Report 2020. Geneva,
844 Switzerland: World Health Organization. Report No.: ISBN 978-92-4-
845 001579-1 (electronic version) ISBN 978-92-4-001580-7 (print version).
846 Available at: [https://www.who.int/teams/global-malaria-](https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2020)
847 [programme/reports/world-malaria-report-2020](https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2020). Accessed 18 Dec 2020.
- 848 2. McKenzie FE, Ferreira MU, Baird JK, Snounou G, & Bossert WH (2001)
849 Meiotic recombination, cross-reactivity, and persistence in *Plasmodium*
850 *falciparum*. *Evolution* 55(7):1299-1307.
- 851 3. Josling GA & Llinas M (2015) Sexual development in *Plasmodium* parasites:
852 knowing when it's time to commit. *Nat Rev Microbiol* 13(9):573-587.
- 853 4. Venugopal K, Hentzschel F, Valkiunas G, & Marti M (2020) *Plasmodium*
854 asexual growth and sexual development in the haematopoietic niche of the
855 host. *Nat Rev Microbiol* 18(3):177-189.
- 856 5. Kent RS, *et al.* (2018) Inducible developmental reprogramming redefines
857 commitment to sexual development in the malaria parasite *Plasmodium*
858 *berghei*. *Nat Microbiol* 3(11):1206-1213.
- 859 6. Josling GA, Williamson KC, & Llinas M (2018) Regulation of sexual
860 commitment and gametocytogenesis in malaria parasites. *Annu Rev*
861 *Microbiol* 72:501-519.
- 862 7. Kafsack BF, *et al.* (2014) A transcriptional switch underlies commitment to
863 sexual development in malaria parasites. *Nature* 507(7491):248-252.
- 864 8. Yuda M, Iwanaga S, Kaneko I, & Kato T (2015) Global transcriptional
865 repression: An initial and essential step for *Plasmodium* sexual
866 development. *Proc Natl Acad Sci U S A* 112(41):12824-12829.
- 867 9. Cortés A & Deitsch KW (2017) Malaria Epigenetics. *Cold Spring Harb*
868 *Perspect Med* 7(7).

- 869 10. Hammam E, *et al.* (2020) Discovery of a new predominant cytosine DNA
870 modification that is linked to gene expression in malaria parasites. *Nucleic*
871 *Acids Res* 48(1):184-199.
- 872 11. Zhao BS, Roundtree IA, & He C (2017) Post-transcriptional gene regulation
873 by mRNA modifications. *Nat Rev Mol Cell Biol* 18(1):31-42.
- 874 12. Baumgarten S, *et al.* (2019) Transcriptome-wide dynamics of extensive
875 m(6)A mRNA methylation during Plasmodium falciparum blood-stage
876 development. *Nat Microbiol* 4(12):2246-2259.
- 877 13. Cui X, *et al.* (2017) 5-methylcytosine RNA methylation in Arabidopsis
878 Thaliana. *Mol Plant* 10(11):1387-1399.
- 879 14. Yang X, *et al.* (2017) 5-methylcytosine promotes mRNA export - NSUN2 as
880 the methyltransferase and ALYREF as an m(5)C reader. *Cell Res*
881 27(5):606-625.
- 882 15. Shi H, Chai P, Jia R, & Fan X (2020) Novel insight into the regulatory roles
883 of diverse RNA modifications: re-defining the bridge between transcription
884 and translation. *Mol Cancer* 19(1):78.
- 885 16. Trixl L & Lusser A (2019) The dynamic RNA modification 5-methylcytosine
886 and its emerging role as an epitranscriptomic mark. *Wiley Interdiscip Rev*
887 *RNA* 10(1):e1510.
- 888 17. Courtney DG, *et al.* (2019) Epitranscriptomic addition of m(5)C to HIV-1
889 transcripts regulates viral gene expression. *Cell Host Microbe* 26(2):217-
890 227 e216.
- 891 18. Chen X, *et al.* (2019) 5-methylcytosine promotes pathogenesis of bladder
892 cancer through stabilizing mRNAs. *Nat Cell Biol* 21(8):978-990.
- 893 19. Tang Y, *et al.* (2020) OsNSUN2-mediated 5-methylcytosine mRNA
894 modification enhances rice adaptation to high temperature. *Dev Cell*
895 53(3):272-286 e277.

- 896 20. Chen H, *et al.* (2020) m(5)C modification of mRNA serves a DNA damage
897 code to promote homologous recombination. *Nat Commun* 11(1):2834.
- 898 21. Yang Y, *et al.* (2019) RNA 5-methylcytosine facilitates the maternal-to-
899 zygotic transition by preventing maternal mRNA decay. *Mol Cell* 75(6):1188-
900 1202 e1111.
- 901 22. Esteve-Puig R, Bueno-Costa A, & Esteller M (2020) Writers, readers and
902 erasers of RNA modifications in cancer. *Cancer Lett* 474:127-137.
- 903 23. Huong TT, Ngoc LNT, & Kang H (2020) Functional characterization of a
904 putative RNA demethylase ALKBH6 in Arabidopsis growth and abiotic
905 stress responses. *Int J Mol Sci* 21(18).
- 906 24. Govindaraju G, Jabeena CA, Sethumadhavan DV, Rajaram N, & Rajavelu
907 A (2017) DNA methyltransferase homologue TRDMT1 in Plasmodium
908 falciparum specifically methylates endogenous aspartic acid tRNA. *Biochim*
909 *Biophys Acta Gene Regul Mech* 1860(10):1047-1057.
- 910 25. Carlton JM, *et al.* (2002) Genome sequence and comparative analysis of
911 the model rodent malaria parasite Plasmodium yoelii yoelii. *Nature*
912 419(6906):512-519.
- 913 26. Gardner MJ, *et al.* (2002) Genome sequence of the human malaria parasite
914 Plasmodium falciparum. *Nature* 419(6906):498-511.
- 915 27. Pavlopoulou A & Kossida S (2009) Phylogenetic analysis of the eukaryotic
916 RNA (cytosine-5)-methyltransferases. *Genomics* 93(4):350-357.
- 917 28. Meyer KD & Jaffrey SR (2017) Rethinking m(6)A readers, writers, and
918 erasers. *Annu Rev Cell Dev Biol* 33:319-342.
- 919 29. Huang T, Chen W, Liu J, Gu N, & Zhang R (2019) Genome-wide
920 identification of mRNA 5-methylcytosine in mammals. *Nat Struct Mol Biol*
921 26(5):380-388.

- 922 30. Hall N, *et al.* (2005) A comprehensive survey of the Plasmodium life cycle
923 by genomic, transcriptomic, and proteomic analyses. *Science*
924 307(5706):82-86.
- 925 31. Filarsky M, *et al.* (2018) GDV1 induces sexual commitment of malaria
926 parasites by antagonizing HP1-dependent gene silencing. *Science*
927 359(6381):1259-1263.
- 928 32. Josling GA, *et al.* (2020) Dissecting the role of PfAP2-G in malaria
929 gametocytogenesis. *Nat Commun* 11(1):1503.
- 930 33. Ralph SA & Cortes A (2019) Plasmodium sexual differentiation: how to
931 make a female. *Mol Microbiol* 112(6):1627-1631.
- 932 34. Llorca-Batlle O, *et al.* (2020) Conditional expression of PfAP2-G for
933 controlled massive sexual conversion in Plasmodium falciparum. *Sci Adv*
934 6(24):eaaz5057.
- 935 35. Guo G, *et al.* (2020) Disease activity-associated alteration of mRNA m(5)C
936 methylation in CD4(+) T cells of systemic lupus erythematosus. *Front Cell*
937 *Dev Biol* 8:430.
- 938 36. Sibbritt T, Patel HR, & Preiss T (2013) Mapping and significance of the
939 mRNA methylome. *Wiley Interdiscip Rev RNA* 4(4):397-422.
- 940 37. Mair GR, *et al.* (2006) Regulation of sexual development of Plasmodium by
941 translational repression. *Science* 313(5787):667-669.
- 942 38. Zhang M, *et al.* (2010) The Plasmodium eukaryotic initiation factor-2alpha
943 kinase IK2 controls the latency of sporozoites in the mosquito salivary
944 glands. *J Exp Med* 207(7):1465-1474.
- 945 39. Lindner SE, *et al.* (2019) Transcriptomics and proteomics reveal two waves
946 of translational repression during the maturation of malaria parasite
947 sporozoites. *Nat Commun* 10(1):4964.

- 948 40. Bogale HN, *et al.* (2021) Transcriptional heterogeneity and tightly regulated
949 changes in gene expression during *Plasmodium berghei* sporozoite
950 development. *Proc Natl Acad Sci U S A* 118(10):e2023438118.
- 951 41. Shen LS, Liang Z, & Yu H (2017) Dot blot analysis of N6-methyladenosine
952 RNA modification levels. *Bio-Protocol* 7(1):e2095.
- 953 42. Martin M (2011) Cutadapt removes adapter sequences from high-
954 throughput sequencing reads. *EMBnet J* 17(1):10-12.
- 955 43. Rieder D, Amort T, Kugler E, Lusser A, & Trajanoski Z (2016) meRanTK:
956 methylated RNA analysis ToolKit. *Bioinformatics* 32(5):782-785.
- 957 44. Wagih O (2017) ggseqlogo: a versatile R package for drawing sequence
958 logos. *Bioinformatics* 33(22):3645-3647.
- 959 45. Fan YT, *et al.* (2020) Rrp6 Regulates Heterochromatic Gene Silencing via
960 ncRNA RUF6 Decay in Malaria Parasites. *Mbio* 11(3).
- 961 46. Jiang L, *et al.* (2013) PfSETvs methylation of histone H3K36 represses
962 virulence genes in *Plasmodium falciparum*. *Nature* 499(7457):223-227.
- 963 47. Jiang Y, *et al.* (2020) An intracellular membrane protein GEP1 regulates
964 xanthurenic acid induced gametogenesis of malaria parasites. *Nat*
965 *Commun* 11(1):1764.
- 966 48. Wright DA, *et al.* (2006) Standardized reagents and protocols for
967 engineering zinc finger nucleases by modular assembly. *Nat Protoc*
968 1(3):1637-1652.
- 969 49. Voss TS, Mini T, Jenoe P, & Beck HP (2002) *Plasmodium falciparum*
970 possesses a cell cycle-regulated short type replication protein A large
971 subunit encoded by an unusual transcript. *J Biol Chem* 277(20):17493-
972 17501.

- 973 50. Delves MJ, *et al.* (2016) Routine in vitro culture of *P. falciparum* gametocytes
974 to evaluate novel transmission-blocking interventions. *Nat Protoc*
975 11(9):1668-1680.
- 976 51. Fan Y, *et al.* (2020) Rrp6 regulates heterochromatic gene silencing via
977 ncRNA RUF6 decay in malaria parasites. *mBio* 11(3):e01110-01120.
- 978 52. Wang X, Qian P, Cui H, Yao L, & Yuan J (2020) A protein palmitoylation
979 cascade regulates microtubule cytoskeleton integrity in *Plasmodium*.
980 *EMBO J* 39(13):e104168.
- 981 53. Shock JL, Fischer KF, & DeRisi JL (2007) Whole-genome analysis of mRNA
982 decay in *Plasmodium falciparum* reveals a global lengthening of mRNA half-
983 life during the intra-erythrocytic development cycle. *Genome Biol* 8(7):R134.
- 984 54. Wang X, *et al.* (2014) N6-methyladenosine-dependent regulation of
985 messenger RNA stability. *Nature* 505(7481):117-120.
- 986 55. Wang M & Wang J (2020) Glucose transporter GLUT1 influences
987 *Plasmodium berghei* infection in *Anopheles stephensi*. *Parasit Vectors*
988 13(1):285.
- 989 56. Kim D, Langmead B, & Salzberg SL (2015) HISAT: a fast spliced aligner
990 with low memory requirements. *Nat Methods* 12(4):357-360.
- 991 57. Liao Y, Smyth GK, & Shi W (2014) featureCounts: an efficient general
992 purpose program for assigning sequence reads to genomic features.
993 *Bioinformatics* 30(7):923-930.
- 994 58. Tamura K, *et al.* (2011) MEGA5: molecular evolutionary genetics analysis
995 using maximum likelihood, evolutionary distance, and maximum parsimony
996 methods. *Mol Biol Evol* 28(10):2731-2739.
- 997 59. Moré JJ (1978) The Levenberg-Marquardt algorithm: implementation and
998 theory. *Proceedings of the Biennial Conference Held at Dundee, June 28–*
999 *July 1, 1977*, Numerical Analysis. Lecture Notes in Mathematics, ed Watson
1000 (Springer, Berlin, Heidelberg), Vol 630, pp 105-116.

- 1001 60. Jia S, *et al.* (2020) TMT-based proteomic analysis of the fish-borne spoiler
1002 *Pseudomonas psychrophila* subjected to chitosan oligosaccharides in fish
1003 juice system. *Food Microbiol* 90:103494.
- 1004 61. Young JA, *et al.* (2005) The *Plasmodium falciparum* sexual development
1005 transcriptome: a microarray analysis using ontology-based pattern
1006 identification. *Mol Biochem Parasitol* 143(1):67-79.
- 1007

Table1 Gene Ontology for m⁵C methylated transcripts in schizonts and gametocytes of P.falciparum (Pf) and P.yoelii (Py)

ID	Biological processes	Result.count	P-value	stage
GO:0007049	cell cycle	5	9.12E-03	Py.S
GO:0015698	inorganic anion transport	2	1.18E-02	Py.S
GO:0090304	nucleic acid metabolic process	49	1.99E-02	Py.S
GO:0006352	DNA-templated transcription, initiation	4	4.45E-02	Py.S
GO:0006720	isoprenoid metabolic process	3	4.73E-02	Py.S
GO:0016310	phosphorylation	28	9.44E-04	Py.G
GO:0006796	phosphate-containing compound metabolic process	46	1.31E-03	Py.G
GO:0009166	nucleotide catabolic process	7	2.20E-03	Py.G
GO:0046031	ADP metabolic process	6	2.27E-03	Py.G
GO:0046939	nucleotide phosphorylation	6	7.41E-03	Py.G
GO:0006468	protein phosphorylation	21	1.24E-02	Py.G
GO:GNF0004	Sexual Development	78	1.48E-08	Py.G
GO:0006325	chromatin organization	5	6.59E-03	Pf.S
GO:0043543	protein acylation	4	7.73E-03	Pf.S
GO:0000280	nuclear division	3	2.06E-03	Pf.S
GO:0015858	nucleoside transport	2	4.82E-03	Pf.S
GO:0031365	N-terminal protein amino acid modification	2	9.38E-03	Pf.S
GO:0015858	nucleoside transport	3	4.94E-04	Pf.GIII
GO:0009987	cellular process	159	1.18E-03	Pf.GIII
GO:0031365	N-terminal protein amino acid modification	3	1.86E-03	Pf.GIII
GO:0006325	chromatin organization	8	1.78E-03	Pf.GIII
GO:1901642	nucleoside transmembrane transport	2	6.26E-03	Pf.GIII
GO:0008152	metabolic process	140	4.68E-03	Pf.GIII
GO:0042273	ribosomal large subunit biogenesis	5	8.46E-03	Pf.GIII
GO:0007155	cell adhesion	9	1.65E-03	Pf.GIV
GO:0006811	ion transport	9	9.05E-03	Pf.GIV
GO:0048870	cell motility	6	4.01E-03	Pf.GV
GO:0051674	localization of cell	6	4.01E-03	Pf.GV
GO:0000737	DNA catabolic process, endonucleolytic	2	5.92E-03	Pf.GV
GO:0006415	translational termination	3	7.60E-03	Pf.GV
GO:0043624	cellular protein complex disassembly	3	7.60E-03	Pf.GV
GO:GNF0004	Sexual Development	43	2.76E-05	Pf.GV

A

B

C

D

E

F