Transcriptome-wide dynamics of extensive m^6^A mRNA methylation during *Plasmodium falciparum* blood-stage development

Sebastian Baumgarten\(^{1,2,3,}\), Jessica M. Bryant\(^{1,2,3}\), Ameya Sinha\(^{4,5}\), Thibaud Reyser\(^{1,2,3,7}\), Peter R. Preiser\(^{4,5}\), Peter C. Dedon\(^{5,6}\) and Artur Scherf\(^{1,2,3}\)

Malaria pathogenesis results from the asexual replication of *Plasmodium falciparum* within human red blood cells, which relies on a precisely timed cascade of gene expression over a 48-h life cycle. Although substantial post-transcriptional regulation of this hardwired program has been observed, it remains unclear how these processes are mediated on a transcriptome-wide level. To this end, we identified mRNA modifications in the *P. falciparum* transcriptome and performed a comprehensive characterization of N^6^-methyladenosine (m^6^A) over the course of blood-stage development. Using mass spectrometry and m^6^A RNA sequencing, we demonstrate that m^6^A is highly developmentally regulated, exceeding m^4^A levels known in any other eukaryote. We characterize a distinct m^6^A writer complex and show that knockdown of the putative m^6^A methyltransferase, PfMT-A70, by CRISPR interference leads to increased levels of transcripts that normally contain m^6^A. In accordance, we find an inverse correlation between m^6^A methylation and mRNA stability or translational efficiency. We further identify two putative m^6^A-binding YTH proteins that are likely to be involved in the regulation of these processes across the parasite's life cycle. Our data demonstrate unique features of an extensive m^6^A mRNA methylation programme in malaria parasites and reveal its crucial role in dynamically fine-tuning the transcriptional cascade of a unicellular eukaryote.

\(^{1}\)Biology of Host-Parasite Interactions Unit, Department of Parasites and Insect Vectors, Institut Pasteur, Paris, France. \(^{2}\)CNRS, ERL 9195, Paris, France. \(^{3}\)INSERM, Unit U1201, Paris, France. \(^{4}\)School of Biological Sciences, Nanyang Technological University Singapore, Singapore, Singapore. \(^{5}\)Antimicrobial Resistance Interdisciplinary Research Group, Singapore-MIT Alliance for Research and Technology, Singapore, Singapore. \(^{6}\)Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. \(^{7}\)Present address: Laboratoire de Chimie de Coordination, Toulouse, France.

\(^{*}\)e-mail: sebastian.baumgarten@pasteur.fr
or directly recruit specific m’A-binding proteins\(^1\). In mammalian cells, m’A-binding proteins include proteins of the YTH (YTH521-B homology) family and elf3 (eukaryotic initiation factor 3). m’A-mediated recruitment of elf3 initiates translation independent of the canonical cap-binding elf4E\(^2\), whereas YTHDF (YTH domain family) proteins can decrease mRNA stability\(^3\) and enhance translation\(^4\). The cellular and organismal processes affected by m’A are widespread and include cell differentiation and cancer cell progression in mammalian cells\(^5\), sex determination in Drosophila\(^6\), and meiosis in yeast\(^7\).

The *P. falciparum* genome has the highest AT content (~80%) of any organism sequenced so far\(^8\). More intriguing yet is the even stronger adenosine bias in the mRNA transcriptome, with adenosines constituting ~45% of all mRNA nucleotides, compared to ~32% in yeast or ~27% in humans. The reason for this high AT bias remains puzzling, but given this unique nucleotide composition and high level of post-transcriptional regulation during the IDC, we hypothesized that mRNA nucleotide modifications, especially on adenosines, could provide an additional layer of gene regulation in *P. falciparum*.

Here, we use mass spectrometry to identify *P. falciparum* mRNA modifications and show that m’A mRNA methylation is a crucial layer of post-transcriptional regulation during the IDC. We identify unique features of the parasite’s methylation program, including a distinct m’A writer complex, preferential m’A site localization, m’A-mediated translational repression and a diverse set of putative m’A-binding proteins. Using CRISPR interference, we characterize the methyltransferase PfMT-A70 as an integral component in m’A methylation and show that m’A dynamically fine-tunes a hardwired transcriptional program across *P. falciparum* blood-stage development.

**Results**

**Identification of a dynamic m’A methylation program during *P. falciparum* blood-stage development.** We first identified m’A mRNA modifications in an unbiased manner with chromatography-coupled triple quadrupole mass spectrometry (LC-MS/MS) from synchronized parasites harvested over the course of the IDC (Supplementary Fig. 1a,b). Ten modifications previously identified in eukaryotic mRNAs were detected in our analysis over the course of the IDC (Fig. 1b,c and Supplementary Fig. 1c,d). These modifications include N’-methylguanosine (m’G), pseudouridine (Ψ), 5-methylcytidine (m’C), and N’-methyladenosine (m’A) (Supplementary Table 1 and Fig. 1d). m’A was previously reported to be the most abundant internal mRNA modification in mammalian cells\(^9\). While global m’A/A levels in *P. falciparum* are comparable to those in human mRNA (Fig. 1e and Supplementary Table 2), overall m’A levels in the parasite’s mRNA (45% adenosine) are higher throughout almost the entire IDC (Fig. 1e). Moreover, m’A mRNA methylation in *P. falciparum* is highly developmentally regulated, with a continuous increase from 0.3% at early stages of the IDC up to 0.7% m’A/A at 30 h.p.i. (Fig. 1e and Supplementary Fig. 1e).

The *P. falciparum* genome encodes a highly conserved putative orthologue (PF3D7_0729500) of the m’A mRNA methyltransferase\(^10\) characterized in other eukaryotes (METTL3 in mammalian cells, Ime4 in *Saccharomyces cerevisiae*, and MTA in *Arabidopsis thaliana*) (Fig. 2a). The encoded protein contains the functional domain of its putative orthologues (MT-A70) and has been designated PfMT-A70. To investigate a potentially conserved role in *P. falciparum*, we generated a cell line expressing an HA-tagged

---

**Fig. 1** Global dynamics of mRNA modifications during the *P. falciparum* IDC. a, Diagram illustrating the asexual replicative cycle of *P. falciparum* inside human RBCs including RBC invasion, host cell remodelling, replication via schizogony and RBC egress. b, Heatmap of normalized abundances (modification/canonical nucleotide) of ten mRNA modifications measured by LC-MS/MS at 6-h intervals over the course of the IDC (see also Supplementary Fig. 1d). Average of four biological replicates for each modification and timepoint, except 30 h.p.i. (n = 2). Blue indicates a high z-score while yellow indicates a low z-score (see Methods). c, LC-MS/MS extracted ion chromatograms of modified nucleotides analysed in parasite mRNA collected at 36 h.p.i. rC, cytosine; rU, uracil; rG, guanosine; rA, adenosine; m’A, N’-methyladenosine; m’G, N’-methyl-2’-guanosine; Ψ, pseudouridine; Am, 2’-O-methyladenosine; m’C, 5-methylcytosine; Gm, 2’-O-methylguanosine; mC, (3 or 4)-methylcytosine; m’U, 5-methyluridine; m’Am, 2’-O-methyladenosine; Cm, 2’-O-methylcytosine. d, Analysis of m’A by high-resolution mass spectrometry. Product ion spectrum of m’A in *P. falciparum* mRNA (grey) using heavy (C13) labelled E. coli tRNA (blue) as a standard. Fragmentation of m/z 282.12 (C12-red) and m/z 293.16 (C13-blue) leads to the breakage of the glycosidic bond resulting in the loss of the ribose sugar (132 Da) to form a daughter ion having a m/z 150.08 (C12-red) and m/z 156.10 (C13-blue). e, Global dynamics of calibrated m’A/A (blue) and m’A/AUGC (grey) levels across the asexual replicative cycle in 6-h intervals. As a reference, human A549 mRNA was analysed in parallel\(^11\). n = 4 for all time points except 30 h.p.i. (n = 2); n = 2 for human A549 samples. Points represent individual biological replicates with bars showing the mean.
PfMT-A70 protein, which localized to the nucleus and the cytoplasm (Fig. 2b).

We next performed immunoprecipitation of PfMT-A70-HA followed by LC-MS and identified 16 specific proteins (Fig. 2c,d and Supplementary Table 3), one of which (PF3D7_1230800) is highly conserved with eukaryotic orthologues of WTAP (Fig. 2d and Supplementary Fig. 2a). The P. falciparum orthologue contains a WTAP/MUM2 domain and was designated PfWTAP. We identified a second putative mRNA methyltransferase containing an MT-A70 domain (PF3D7_1235500, designated PfMT-A70.2) that shows high sequence conservation with orthologues of the bilaterian METTL14 (Fig. 2d and Supplementary Fig. 2a). All three P. falciparum orthologues reach peak transcription levels between 12 and 24 h.p.i., shortly before maximal m’A/A ratios are reached (Figs. 1e and 2b). Among the other co-immunoprecipitated proteins are known RNA/DNA binding proteins such as Alba3, an ATP-dependent RNA helicase, and a FoP-domain-containing protein. Three proteins identified with high confidence (PF3D7_0707200, PF3D7_1036900 and PF3D7_1366300) have unknown functions and show no homology to other eukaryotic proteins (Supplementary Table 3).

CRISPR interference of PfMT-A70 decreases global m’A mRNA methylation. We targeted PfMT-A70 to gain insight into the functional role of m’A mRNA methylation in P. falciparum. As attempted CRISPR/Cas9-targeted deletion was unsuccessful, we utilized CRISPR interference29 (CRISPRi; Supplementary Fig. 2b), which has recently been adapted for P. falciparum29. For CRISPRi of PfMT-A70, we targeted its promoter downstream of the transcription start site and ~100 bp upstream of the translation start site with a gRNA complementary to the non-template strand (gPfMT-A70’, Fig. 3a). A cell line expressing a non-specific gRNA served as a negative control (gControl, Fig. 3a). gPfMT-A70-targeted dCas9 chromatin immunoprecipitation followed by sequencing (ChiP-seq)30 showed a robust and highly specific enrichment at the target site, which overlaps with a nucleosome-depleted region (FAIRE)32 enriched for H3K9ac and H2A.Z32 (that is, hallmarks of promoter regions in P. falciparum32; Fig. 3a,c).

To determine the effects of PfMT-A70 CRISPRi, we harvested mRNA from synchronized gPfMT-A70 and gControl parasites at 12, 24 and 36 h.p.i. RT-qPCR revealed a significant downregulation of PfMT-A70 transcripts throughout the IDC, with knockdown levels ranging from ~15% at 12 h.p.i. to 40% at 24 and 36 h.p.i. (Fig. 3d, top). While PfMT-A70 CRISPRi did not have a discernible effect on transcription rates or lower transcript degradation rates. Because transcription rates or lower transcript degradation rates. Because

m’A mRNA methylation might provide a means of modulating the transcriptionome independently of initial gene transcription. Throughout the IDC, there is a modest increase in the total number of transcribed genes that have at least one m’A peak (Fig. 4f). However, the m’A enrichment (that is, m’A/IP/m’A-input), or fraction of transcripts from one gene containing a specific m’A peak, changes extensively throughout the IDC (Fig. 4g). Most m’A peaks reach maximum m’A enrichment at only one time point, suggesting that transcript-specific methylation is indeed an actively regulated mechanism.

m’A inversely correlates with mRNA stability and translational efficiency. To determine the effect of m’A on individual transcripts, we first compared m’A enrichment of individual m’A peaks between gControl and gPfMT-A70 parasites and found a significant reduction in the gPfMT-A70 cell line (Fig. 5a,b, top) at all three time points, mirroring the decrease in global m’A/A levels in gPfMT-A70 parasites as measured by LC-MS/MS (Fig. 3d, bottom). These decreases are consistent, with two replicates each of gControl and gPfMT-A70 showing a high degree of correlation at each time point (Fig. 5b, bottom, and Supplementary Tables 4–6).

We performed strand-specific mRNA sequencing in gControl and gPfMT-A70 parasites at 12, 24 and 36 h.p.i. and observed minimal changes in global transcript abundance following PfMT-A70 knockdown, with only a small percentage of genes showing significant differential expression (Supplementary Tables 9 and 10). However, at 12 and 24 h.p.i., we found a significant increase in overall abundance of transcripts that show the most pronounced decrease in m’A enrichment (≥two-fold) following PfMT-A70 depletion (Fig. 5c).

Theoretically, this observed increase in transcript abundance following PfMT-A70 knockdown could be due either to higher transcription rates or lower transcript degradation rates. Because
m^A methylation was previously shown to affect mRNA degradation\(^2\), we compared mRNA stability data (measured as mRNA half-life) previously obtained from wild-type parasites\(^1\) between m^A-methylated and non-methylated transcripts found in our study. Strikingly, m^A-methylated transcripts have significantly lower mRNA stabilities than non-methylated transcripts (Fig. 6a, left). These differences are significant at 12 (\(P = 5.3 \times 10^{-5}\)) and 24 h.p.i. (\(P = 9.1 \times 10^{-5}\)), but not at the 36 h.p.i. time point (\(P = 0.59\)). We found even lower mRNA stabilities for the 'm^A-sensitive' transcripts (at least two-fold decrease in m^A enrichment and increased transcript abundance (log, fold-change >0)) at 12 and 24 h.p.i. (Fig. 6a, right). These data suggest that the increase of transcript abundances following PfMT-A70 knockdown might indeed reflect decreased rates of m^A-mediated mRNA degradation at the 12 and 24 h.p.i. time points. A further comparison of m^A methylation with additional mRNA stability data\(^6\) revealed a similar pattern, showing lower mRNA stabilities for m^A-methylated than non-methylated transcripts (Supplementary Fig. 4a).

We next compared translational efficiencies (TE) between m^A-methylated and non-methylated transcripts throughout the IDC. TEs were calculated previously for \(P. falciparum\) as the ratio of ribosome-bound RNA versus steady state mRNA\(^6\). At each developmental stage, m^A methylation correlates with significantly lower translational efficiencies for a transcript (Fig. 6b, left). This correlation is most pronounced at the 12 (\(P = 4.7 \times 10^{-4}\)) and 24 h.p.i. (\(P = 5 \times 10^{-3}\)) time points, but is still significant at 36 h.p.i. (\(P = 1.5 \times 10^{-3}\)). Similarly to the pattern seen for mRNA stability, we observed an additional significant decrease in translational efficiencies between m^A-methylated and the subset of 'm^A-sensitive' transcripts at 12 and 24 h.p.i. (Fig. 6b, right). Notably, the 'm^A-sensitive' transcripts show no preference for m^A peak location across the coding sequence (Supplementary Fig. 4b).

A GO enrichment analysis of the 'm^A-sensitive' transcripts shows that these are enriched for biological processes that are highly relevant for their respective time points of development. At 12 h.p.i., these processes include regulation of gene expression (\(Q = 3 \times 10^{-3}\)), DNA-templated transcription and regulation of RNA biosynthesis (\(Q = 3 \times 10^{-3}\)). Among these gene transcripts are four members of the AP2 transcription factor family, two histone deacetylases, one histone-lysine methyltransferase, and one bromodomain-containing protein (Supplementary Table 11). At 24 h.p.i., binding to chromatin (\(Q = 6.1 \times 10^{-3}\)) is among the most significantly enriched molecular functions, including SNF2 helicases and two SET- and one bromodomain-containing protein (Supplementary Table 11).

**Fig. 2 | Characterization of the \(P. falciparum\) m^A writer complex.** a. Phylogenetic tree and protein alignments of PfMT-A70 with orthologues in other eukaryotes. The location of the MT-A70 domain is shown below the alignment (yellow bar). The grey-scale gradient indicates the percentage similarity across all proteins at an aligned position. b. Western blot analysis showing the enrichment of HA-tagged PfMT-A70 in the cytoplasmic (C) and nuclear (N) cell fractions at 12 h.p.i. PfAldolase and histone H3 serve as controls for the cytoplasmic and nuclear fraction, respectively. Numbers on the left indicate molecular weight in kilodaltons. i, input; s, supernatant; ip, co-ip fraction. Data are representative of three independent biological experiments.

c. Western blot analysis of PfMT-A70-HA co-IP with anti-HA antibodies. PfAldolase and histone H3 demonstrate efficacy of the anti-HA co-IP. Numbers on the left indicate molecular weight in kilodaltons. i, input; s, supernatant; ip, co-ip fraction. Data are representative of three independent biological experiments. d. Schematic of the human m^A writer complex (left) and putative, non-conserved co-factors of \(P. falciparum\) (right) are shown in grey. e. Gene transcription (in reads per kilobase of exon per one million mapped reads (RPKM))\(^{27}\) of putative orthologues of the \(P. falciparum\) m^A writer complex at 6-h intervals over the course of the IDC.
high conservation of key residues known to bind mRNA and the methylated adenosine in other eukaryotes24 (Supplementary Fig. 4c). PfYTH1 also features the C-terminal arrangement of the YTH domain and an N-terminal low-complexity disordered region that is characteristic of metazoan YTHDF proteins (Supplementary Fig. 4d)24. In contrast, PfYTH2 does not show overall similarity to other known mA reader proteins. Both proteins are highly expressed early during the IDC (Fig. 6d), and ectopic expression of PfYTH1 revealed localization of the protein to both the nucleus and cytoplasm (Fig. 6e). In comparison to the dynamic expression during the IDC, PfYTH2 shows even higher expression in the transmission stages of the parasite, namely gametocytes and salivary gland sporozoites, the latter of which also highly expresses PIMT-A70 (Fig. 6d).

**Discussion**

In this study, we characterize a dynamic and extensive mA mRNA methylation program that fine-tunes the transcriptional cascade during blood-stage development of *P. falciparum*. We generated a high-resolution map of mA methylation sites throughout the IDC and identify PIMT-A70 as an integral member of the putative mA methyltransferase complex.

Although gene knockout was unsuccessful, CRISPRi allowed us to reduce PIMT-A70 transcript levels to observe consistently decreased mA levels via LC-MS/MS and mA-seq, as well as increased abundance of transcripts with decreased mA enrichment. Higher knockdown levels might be possible with multiple gRNAs, and reveal an effect on parasite growth or more drastic changes in global transcript abundances. However, even using a dCas9-repressor fusion protein did not achieve higher knockdown efficiencies of an essential gene in *P. falciparum*25. Thus, it is likely that long transfection and cloning periods favour transgenic clones with lower knockdown levels and possibly lower phenotypic defects. Notably, knockdown or even knockout of the mA-methyltransferase in *D. melanogaster*26 or mouse cells27, respectively, resulted in only moderate changes in gene expression.

The *P. falciparum* mA methylation program is set apart from those in other eukaryotes by several key aspects. First, while MT-A70 methyltransferases can be found throughout eukaryotes28 and the core members of the *P. falciparum* mA writer complex are conserved, we did not identify other co-factors that are similarly conserved in metazoans and plants (for example, HAKAI and VIRMA29–31) in our PIMT-A70 co-immunoprecipitation. Thus, the putative P. falciparum co-factors are likely to be lineage-specific, and several of them do not share any sequence homology to other eukaryotic proteins. In line with this divergence of the extended mA writer complex, we find that the mA motif differs from those found in other eukaryotes27,28, showing a higher prevalence of adenosine at positions –1 and +2 that mirrors the higher adenosine content of the parasite's mRNA transcriptome. Most identified co-factors show nuclear localization or are found exclusively in the nuclear proteome25 (for example, PIMT-A70.2), suggesting that mA methylation in *P. falciparum* occurs prior to nuclear mRNA export. Unlike in other eukaryotes, there is no preference for mA methylation sites close to the stop codon and the 3' UTR in *P. falciparum*. It was proposed that this 3' enrichment largely depends on the protein VIRMA, which in mammalian cells specifically recruits the mA writer complex to the 3' UTR32; however, we did not identify an orthologue in the *P. falciparum* genome.

With the high adenosine content in the parasite's protein-coding transcriptome, overall mA levels markedly exceed those measured in human mRNA, suggesting a key role for this modification during asexual replication. It is tempting to speculate that mA methylation might represent one evolutionary driving force of high adenosine content observed at gene loci. In addition, the functional diversity of transcripts containing mA suggests that this modification is involved in various important cellular processes during the IDC. While mA may directly affect the fate of an mRNA transcript and its encoded protein, the modulation of transcripts encoding major regulatory proteins such as AP2 transcription factors or chromatin remodelers might further amplify its modulatory potential.

In mammalian cells, mA levels are mostly static even during key developmental stages such as stem cell differentiation33,34, with deregulation of METTL3 and increases in global mA of only 25% being sufficient to inhibit differentiation, promote cell growth, and maintain tumorigenicity35. Although the mA-demethylases Alkbh5 and FTO may modulate methylation levels in some cases36, mA is generally deposited early and maintained throughout the life of an mRNA transcript37. In contrast, the most unique feature of the *P. falciparum* mA methylation program is its dynamic nature. Parallel LC-MS/MS measurements and transcript-specific mA-seq demonstrated similar mA dynamics, with global mA/A levels doubling and the number of identified mA peaks gradually increasing towards the end of the IDC. Since we did not find a potential mA demethylase in the *P. falciparum* genome, mA methylation dynamics during the IDC could result from different rates of active mA methylation by the writer complex and the subsequent rate of mA-methylated transcript degradation. Indeed, the observed gradual increase of global mA/A levels measured by LC-MS/MS might be partially due to the mA-mediated mRNA degradation at earlier but not later stages of the IDC, as is evidenced by the inverse correlation of mA status and mRNA stability at 12 and 24 h.p.i., but not 36 h.p.i. It follows that transcripts with decreased mA enrichment are overabundant at only the earlier (and not the late) time point following PIMT-A70 depletion.

**Fig. 3 | Knockdown of the PIMT-A70 mA methyltransferase by CRISPRi.** a. Diagram of the CRISPR interference system for targeted transcriptional knockdown of PIMT-A70. A non-specific RNA (gControl) without a binding site in the *P. falciparum* genome is used as a negative control in all experiments (left). The specific gRNA (gPIMT-A70) targets the PIMT-A70 promoter on the non-template strand downstream of the putative transcription start site, blocking the elongating RNA polymerase II (RNP) and silencing the target gene (right). b, dCas9-ChIP sequencing (12 h.p.i.) shows enrichment for gPIMT-A70-targeted dCas9 (blue), but non-targeted dCas9 (gControl in grey) at the targeted upstream region of the PIMT-A70 gene (indicated at the top with a grey bar). The blue triangle indicates the location of the gRNA target site. Genomic features investigated in previous studies that define the putative promoter region are shown below: transcription start site (TSS)32,33, histone 2A variant (H2A.Z)34,35, acetylation of histone H3 at lysine 9 (H3K9ac)36, nucleosome depletion as determined by FAIRE-seq21 (FAIRE), and RNA-seq coverage (RNA) at the genomic locus. c, Circos plot of gPIMT-A70 dCas9 ChIP-seq (normalized to gControl) across all 14 nuclear chromosomes (exterior grey bars) shows no substantial enrichment of dCas9-binding other than at the PIMT-A70 promoter (blue arrowhead), demonstrating the high specificity of the method. Chromosome name and length (in megabases) are indicated at the exterior of the plot. The y axis indicates ChIP enrichment on a scale from 0 to 150 RPKM in 1,000 nt windows. dCas9 ChIP-seq data (b, c) are representative of two independent biological experiments. d, Comparison of PIMT-A70 transcription measured by RT-qPCR (top) and mA/A measured by LC-MS/MS (bottom) between gControl (grey) and gPIMT-A70 (blue) parasites at three different time points over the IDC (indicated at the top). PIMT-A70 transcript levels were normalized to those of the housekeeping gene serine RNA ligase (PF3D7_0717700). The number of biological replicates is indicated on the bottom of the graph. P values were calculated using two-sided independent-samples t-test (RT-qPCR) and a two-sided Mann-Whitney U-test (LC-MS/MS). The average of PIMT-A70 transcription and mA/A levels in gControl parasites were set to 1. Horizontal lines represent median and interquartile ranges.
These differences in m^6A-dependent mRNA degradation rates might be due to m^6A reader proteins mediating mRNA turnover. Both PfYTH.1 and PfYTH.2 transcription decreases over the course of the IDC (Fig. 6d), providing a potential link between the gradual increase in global m^6A/A levels and the decrease in m^6A-mediated transcript degradation over the course of parasite development. Similarly, we find the strongest association between m^6A and translational efficiency at 12 h.p.i., with a continuous decrease during the IDC (Fig. 6b, left). Thus, the timely action of m^6A writers and the combinatorial or exclusive expression of putative reader proteins tightly regulate m^6A dynamics, probably ensuring the maximal modulatory potential of m^6A at each step of the lifecycle.

Mammalian YTHDF proteins 1 to 3 have been linked to mRNA degradation and translational activation. However, based on their sequence similarity and the overlap of binding sites, it was suggested that these proteins might be partially redundant. In contrast,
Fig. 4 | Differential m6A methylation during the P. falciparum IDC. a, Fold enrichment of m6A levels as measured by LC-MS/MS in different fractions of the m6A-IP experiment, showing a three-fold depletion in the supernatant and 63-fold enrichment in the eluate. m6A levels were normalized to the total number of adenosines (that is, m6A/A) for each fraction, and fold enrichments were calculated as ratios over the m6A/A levels in the input mRNA sample (collected at 36 h.p.i.). Points represent individual biological replicates, with bars showing the mean. b, Distribution of m6A peak location along a normalized transcript coding sequence (CDS), 500 nt upstream of the start (5’ UTR) and 500 nt downstream of the stop codon (3’ UTR) for peaks identified at 12 (grey), 24 (yellow) and 36 (blue) h.p.i. The total number of identified m6A sites by m6A-seq is likely to be an underestimate. Besides those that were not called due to the conservative annotation approach, m6A sites in transcripts reaching peak expression at different time points or sites in low complexity regions are likely to have been missed. c, Exemplary RNA sequence coverage of m6A-IP (blue) and input (grey) gControl samples showing m6A peaks identified at 12 h.p.i. in PF3D7_0506800 (transcription factor 25; RPKM, 107), 24 h.p.i. in PF3D7_1250200 (CSC1-like protein; RPKM, 53) and at 36 h.p.i. in PF3D7_1017700 (conserved protein of unknown function; RPKM, 10). The location of the m6A peak is indicated by the yellow bar. Data are representative of two independent biological experiments. d, Comparison of the deduced sequence motif of m6A methylation sites in P. falciparum (top), S. cerevisiae (middle) and human (bottom). The relative height of each nucleotide indicates frequency and the total height at each position represents sequence conservation (‘bits’). The N6-methylated adenosine is located at position 0. e, Density plot showing the occurrence of the consensus motif within ±50 nt of all significant m6A-seq peak summits (blue) and size-matched random control sequences (grey) in 5 nt windows. f, Bar charts indicating the number of transcripts (RPKM ≥5) with no (grey) or at least one significant (blue) m6A peak at three different time points over the IDC. g, Heatmap of average row Z score normalized m6A enrichments in gControl parasites (average derived from n = 2 biologically independent experiments) displaying changes in transcript-specific methylation at three time points throughout the IDC (top). Each row represents an m6A peak (n = 840) in a transcript expressed with RPKM ≥ 5 at all three time points. Bottom: heatmap showing Pearson’s r correlation coefficients of m6A enrichments, demonstrating a high degree of reproducibility between the two replicates (R1 and R2) at each time point. Blue indicates a high Z score while yellow indicates a low Z score (see Methods).

P. falciparum encodes only two YTH proteins, and while their roles are unknown, the divergence of these two proteins beyond their YTH domain and the distinct expression patterns during the parasite life-cycle suggest that they mediate two distinct functions. m6A has not been associated with translational repression in any eukaryote yet, but this post-transcriptional regulation is of particular importance for P. falciparum during the IDC15 and especially during transmission between the human host and mosquito vector14,49. During the latent gametocyte and salivary gland sporozoite stages, mRNAs are kept in a translationally repressed state and protein synthesis is only initiated following transmission, readily ensuring rapid continuation of development12,49. The co-expression of PiMT-A70 and PiYTH.2 in
sporozoites suggests that m^A could also contribute to translational repression during this transmission stage (Fig. 6d).

The major factor controlling mRNA transcript abundance during the IDC is certainly the transcriptional activity at the genomic locus^18. However, here we show that the fate of individual transcripts from the same gene can be changed, possibly by decreasing mRNA stability and/or repressing translation through specific m^A methylation. Our findings that (1) transcripts with decreased m^A enrichment following PfMT-A70 knockdown were consistently more abundant and (2) there is an increased association between m^A sensitivity and mRNA stability and translational repression suggest that m^A serves as a balance to modulate the outcome of protein synthesis independent of initial transcriptional rates. Thus, temporally dynamic changes of transcript-specific m^A methylation rates during the IDC have, to a certain extent, the potential to fine-tune a possibly imprecise, hardwired transcriptional program.

Collectively, our study reveals an additional layer of dynamic and widespread post-transcriptional modulation of gene expression in...
**Correlation of m6A with mRNA stability and translational efficiency.**

**a.** Boxplots (left) depicting mRNA half-lives for transcripts without (grey; n = 1,734, 1,534 and 1,495 at 12, 24 and 36 h.p.i., respectively) or with (blue: n = 301, 377 and 383 at 12, 24 and 36 h.p.i., respectively) m6A peaks at 12, 24 and 36 h.p.i. Centre line, median; box limits, first and third quartile; whiskers, 1.5 interquartile range. P values were calculated with a two-sided Mann-Whitney U-test. Right: cumulative fraction plots of mRNA half-lives at 12 and 24 h.p.i. for transcripts without m6A peak (grey), with m6A peak (blue), and ‘m6A-sensitive’ (yellow) transcripts. 

**b.** Boxplots (left) depicting translation efficiencies for transcripts without (grey; n = 2,004, 1,975 and 1,986 at 12, 24 and 36 h.p.i., respectively) or with (blue: n = 317, 412 and 435 at 12, 24 and 36 h.p.i., respectively) m6A peaks at 12, 24 and 36 h.p.i. Centre line, median; box limits, first and third quartile; whiskers, 1.5 interquartile range. P values were calculated with a two-sided Mann-Whitney U-test against translation efficiencies of non-methylated transcripts. Right: cumulative fraction plots of translation efficiencies at 12 and 24 h.p.i. for transcripts without m6A peak (grey), with m6A peak (blue), and ‘m6A-sensitive’ (yellow) transcripts. P values were calculated with a two-sided Mann-Whitney U-test against translation efficiencies of non-methylated transcripts. 

**c.** Phylogenetic tree and protein alignment of the putative m6A reader protein PfYTH and orthologues in other eukaryotes. The location of the YTH domain is indicated below (yellow bar). The grey scale gradient indicates the percentage similarity across all proteins at an aligned position. YTHDC, YTH domain-containing protein; YTHDF, YTH domain-containing family protein. 

**d.** Transcription (RPKM) of the *P. falciparum* PfYTH proteins and PfMT-A70 over the course of the IDC6, stage V gametocytes (G), oocysts (O) and salivary gland sporozoites (S). Arrows indicate transmission from host to vector and vice versa. 

**e.** Western blot analysis showing the enrichment of HA-tagged PfYTH in the cytoplasmic (C) and nuclear (N) fractions at 12 h.p.i. PfAldolase and histone H3 serve as controls for the cytoplasmic and nuclear fractions, respectively. Numbers on the left indicate molecular weight in kilodaltons. Data are representative of two independent biological experiments.

*P. falciparum.* The conservation of core features of m6A mRNA methylation makes *P. falciparum* an excellent system to study the interplay between m6A methylation and gene transcription. Moreover, our results add m6A as a major player in the malaria parasite ‘epitranscriptome’ code and open new avenues for drug development in malaria parasites.
**Methods**

**Parasite culture.** Axenial blood-stage *P. falciparum* parasites (strain 3D7) were cultured as described previously. Briefly, parasites were cultured in human RBCs in RPMI-1640 culture medium (Thermo Fisher no. 31800-027) supplemented with 10% v/v ABhuman I (Thermo Fisher no. 11000028), hypoxantine (0.1 mM final concentration), CC-Pro no. Z-41-M) and 10 mg gentamicin (Sigma no. G1317-10mL) at 4% haematocrit and under 5% O2, 3% CO2 at 37°C. Parasite development was monitored by Giemsa staining. RBCs were obtained from the Etablissement Français du Sang with approval number HS 2015-24803. For synchronisation, late stage parasites were enriched with plasmaflag flotation followed by ring stage enrichment via sorbitol (5%) lysis 6 h later. For sampling of highly synchronous parasites during the IDC, the synchronous schizonts were harvested by centrifugation and wash once with D-PBS pH 7.4. The haemoglobin was removed by resuspending in a Guava easyCyte flow cytometer (Merck Millipore).

**Total RNA extraction.** For RNA extraction, infected RBCs (iRBCs) were collected by centrifugation and washed once with D-PBS pH 7.4 and resuspended in 700 µl QIAzol reagent (Qiagen no. 79603). Total RNA was extracted using the miRNeasy Mini Kit (Qiagen no. 217004), including an on-column DNase I digestion and exclusion of small RNA (<200 nt) according to the manufacturer’s protocol.

**Analysis of mRNA modifications by LC-MS/MS.** Total RNA from highly synchronous parasites was extracted from samples collected during the IDC at seven time points at 6 h intervals (that is, 6 h, 12 h, 18 h, 24 h, 30 h, 36 h and 42 h). To minimize human RNA and DNA contamination, parasites were cultured in leukocyte-depleted RBCs. Total RNA from human A549 cells was extracted as described above. Poly(A) RNA from *P. falciparum* or human A549 total RNA was enriched by two successive rounds of purification with the Dynabeads mRNA purification kit (Thermo Fisher no. 61006) (Supplementary Fig. 1b).

**Purified *P. falciparum* RNA was hydrolysed enzymatically as described with a slightly modified protocol using the following components in the buffer mix (10 mM Tris-HCl pH 7.9, 1 mM MgCl2, 5 µM benzamide (Merck no. 71206), 50 µM deoxersosine (Sigma no. D9533), 0.1 µg/µl pentostatin (Sigma no. SME0508), 100 µM butylated hydroxytoluene (Sigma no. W218405), 0.5 µg/µl tetrahydrodoruline (Calbiochem no. 584222), 5 µM bacterial alkaline phosphatase (Thermo Fisher no. 18010185), 0.5 U phosphodiesterase I (Sigma no. P23243) and [115N]-2′-deoxyadenosine ([115N]-dA) (Cambridge Isotope Laboratories no. NLM-3895-25) as the internal standard to account for variations in sample handling and mass spectrometer response.**

**Purified RNA was hydrolysed enzymatically as described with a slightly modified protocol using the following components in the buffer mix (10 mM Tris-HCl pH 7.9, 1 mM MgCl2, 5 µM benzamide (Merck no. 71206), 50 µM deoxersosine (Sigma no. D9533), 0.1 µg/µl pentostatin (Sigma no. SME0508), 100 µM butylated hydroxytoluene (Sigma no. W218405), 0.5 µg/µl tetrahydrodoruline (Calbiochem no. 584222), 5 µM bacterial alkaline phosphatase (Thermo Fisher no. 18010185), 0.5 U phosphodiesterase I (Sigma no. P23243) and [115N]-2′-deoxyadenosine ([115N]-dA) (Cambridge Isotope Laboratories no. NLM-3895-25) as the internal standard to account for variations in sample handling and mass spectrometer response.**

**Hi-Fusion HD cloning kit (Clontech no. 639849). The resulting plasmids were verified by Sanger sequencing. All PCR reactions were performed using the KAPA HiFi DNA Polymerase (Roche no. 07958846001) following the manufacturer’s protocol.**

**Protein fractionation and western blot analysis.** One ml iRBCs containing synchronous parasites expressing HA-tagged PMT-A70 or PYTH at ring stage (2–3% parasitaemia) were washed once with DPBS at 37°C, and RBCs were lysed with 0.5% saponin in DPBS. Parasites were then washed once again with DPBS. For separation of the cytoplasmic and nuclear protein fractions, the cell pellet was first resuspended in 1 ml cytoplasmatic lysis buffer (CLB: 25 mM Tris-HCl pH 7.5, 10 mM NaCl, 1% IGEPAL CA-630, 1 mM DTT, 1.5 mM MgCl2, 1X protease inhibitor (PI, Roche no. 118367001) and incubated on ice for 30 min. Cells were centrifuged, and the cytosolic fraction was collected and resuspended with centrifugation (13,500g, 10 min, 4°C). The pellet (containing the nucleus) was resuspended in 100 µl nuclear extraction buffer (25 mM Tris-HCl pH 7.5, 1 mM DTT, 1.5 mM MgCl2, 600 mM NaCl, 1% IGEPAL CA-630, PI) supplemented with 1 µl benzolase (Sigma Aldrich no. E1014-5KU) and sonicated for 10 cycles with 30 s (on/off) intervals (5 min total sonication time) in a Diagenode Bioruptor Pico (Bioruptor no. B0106090). This nuclear lysate was cleared with centrifugation (13,500g, 10 min, 4°C), and the nuclear fraction was diluted with 300 µl CLB. Protein samples were supplemented with NuPage Sample Buffer (Thermo Fisher no. NP0008) and NuPage Reducing Agent (Thermo Fisher no. NP0004) and denatured for 10 min at 70°C. The samples were separated on a NuPage 4–12% Bis-Tris gel using MOPS running buffer at 150 V for 1.5 h and transferred to a PVDF membrane overnight at 15 V at 4°C. The membrane was blocked for 1 h in 1% milk in 0.1% Tween20 in PBS (PBST). The HA-tagged proteins and histone H3 were detected with anti-HA (Abcam no. ab9110; 1:1,000 in 1% milk/PBST) and anti-H3 (Abcam no. ab1791; 1:1,000 in 1% milk/PBST) primary antibodies, followed by donkey anti-rabbit (Gen no. NA934-1ML) secondary antibodies conjugated to HRP (1:5,000). PFAldolase was detected using an HRP conjugated anti-PFAldolase (Abcam no. ab38905; 1:5,000 in 1% milk/PBST) antibody. The HRP signal was developed using the SuperSignal West Pico chemiluminescent substrate (Thermo Fisher no. 34580) and imaged with a ChemiDoc XRS+ (Bio-Rad).

**PMT-A70 co-immunoprecipitation.** *P. falciparum* PMT-A70-HA (n = 3 replicates) expressing parasites (cultured with 2.67 mM WR99210 (Jacobs Pharmaceuticals)) and wild-type *P. falciparum* 3D7 (as a negative control, cultured in standard growth medium) were sorbitol-synchronized. After 36 h, the parasites were washed twice after Percoll (Gen no. P94907) and resuspended in RPMI and co-crosslinked with 0.5 mM dithiobisuccinimidyl propionate (DSP) (Thermo Fisher no. 22585) in PBS for 60 min at 37°C. The reaction was quenched with PBS containing 25 mM Tris-HCl. These trophozoites were lysed with RIPA buffer (10 mM Tris HCl pH 7.5, 150 mM NaCl, 0.1% SDS, 1% Triton) containing protease and phosphatase inhibitor cocktail (Thermo Fisher no. 7840). The lysates were cleared by centrifugation at 16,000g for 10 min. Supernatants were incubated with 25 µl of anti-HA Dynabeads (Thermo Fisher no. 88368) overnight at 4°C. 20 µl of sample were taken prior to separation of the beads for Western blot analysis (Input, Fig. 2c and Supplementary Fig. 5). Beads were magnetically isolated and 20µl of sample was taken from the supernatant for Western blot analysis and immunofluorescence (Fig. 3). The beads were washed five times with 1 ml RIPA buffer following five washes with 1 ml PBS and one wash with 1 ml 100 mM ammonium bicarbonate (Sigma no. 09830). The beads were reduced with 10 mM dithiothreitol (Sigma no. D9779), alkylated with 55 mM iodoacetamide (Sigma no. 11149) and subjected to on-bead digestion using 50 µg of trypsin (Sigma Fisher no. 90039). The resulting peptides were desalted using C18 cartridges (Thermo Fisher no. 89852) and sent for MS analysis.

**Protein mass spectrometry.** Peptides were separated by reverse phase HPLC (Thermo Easy nLC1000) using a precolumn (made in house, 6 cm of 10 µm C18) and a self-pack 5 µm tip analytical column (12 cm of 5 µm C18, New Objective) on a gradient of 15–100% methanol gradient in water–acetonitrile (0.1% formic acid) at 150 µl/min. The eluate was transferred to the nanospray source and infused at 300 µl/min using a self-pack 5 µm tip analytical column (12 cm of 5 µm C18, New Objective) on a gradient of 15–100% methanol gradient in water–acetonitrile (0.1% formic acid) at 150 µl/min.
searched using Proteome Discoverer (Thermo Fisher) and Mascot (version 2.4.1). Mascot search parameters were: 10 ppm mass tolerance for precursor ions; 15 milli modifications: the Cas9 was mutated at the RuvC and HNH positions, as described above for 210 elution buffer (50 mM Tris-HCl pH 8, 10 mM EDTA pH 8, 1% SDS) and protein-DNA complexes were eluted by 30% Triton X-100 in an agitator (900 rpm 1 min, still 30°C). At this step, the input sample was recovered from storage, diluted with 170 μl elution buffer, and processed in parallel with the IP sample.

Protein-DNA complexes were reverse crosslinked by incubation at 65 °C for 6 h. Samples were diluted with 200 μl of TE buffer and incubated with RNaseA (0.2 mg/ml final concentration) (Thermo Fisher no. EN0531) at 37 °C for 2 h. Protease K (0.2 μg/ml final concentration) (NEB no. P8107S) was then added and samples were incubated at 55 °C for 2 h. DNA was purified by adding 400 μl phenol:chloroform:isoamyl alcohol, and phases were separated by centrifugation (10 min, 13,500 g, 4 °C) after vigorous vortexing. 16 μl of 5 M NaCl (200 mM final concentration) and 30 μg glycerol were added to the aqueous phase, and DNA was desalted by adding 800 μl 100% EtOH 4 °C and incubating for 30 min at −20 °C. DNA was pelleted by centrifugation (20,000g, 10 min, 4 °C) and washed with 500 μl 80% EtOH 4 °C. After centrifugation, the DNA pellet was air-dried and resuspended in 30 μl of 10 mM Tris-HCl pH 8.0.

m'A immunoprecipitation and sequencing. Two replicates were sampled for both cell lines expressing dCas9 and either a non-targeting control gRNA (gControl) or the PI-MT-A70 promoter-targeting gRNA (pMT-A70). Total RNA from highly synchronous parasites was collected at 12, 24 and 36 h.p.i. as described above, followed by poly(A) RNA enrichment using the Dynabeads mRNA purification kit (Thermo Fisher no. 61006).

For each IP, 500 ng purified mRNA was fragmented for 15 min at 70 °C to ~100 nt using the NEBNext fragmentation module (NEB no. E6150S) and purified using the Qiaqen MinElute kit (Qiaqen no. 74204) with slight modifications: to first purify small RNA (<200 nt, 100 μl of RNA sample was supplemented with 350 μl of lysis buffer (provided in the kit) followed by the addition of 700 μl of 100% ethanol. 20 ng fragmented mRNA were stored as input sample at ~80°C.

For each m'A-IP, 25 μl of protein A/G magnetic beads (Thermo Fisher no. 88802) were diluted in 250 μl reaction buffer (150 mM NaCl, 10 mM Tris-HCl pH 7.5, 0.1% Igepal CA-630 in nuclease-free H2O) and washed twice in 250 μl reaction buffer. The beads were resuspended in 100 μl reaction buffer and incubated with 2.5 μg anti-m'a antibody (Abcam no. ab51230) for 30 min at room temperature with constant rotation. The beads were subsequently washed three times with 0.5 ml reaction buffer. For each m'A-IP, 300–500 ng fragmented mRNA were diluted with nuclease-free H2O to a final volume of 395 μl supplemented with 5 μl RNAse inhibitor (Promega no. N2615), 100 μl 5X reaction buffer (750 mM NaCl, 50 mM Tris-HCl pH 7.5, 0.5% Igepal CA-630 in nuclease-free H2O) and added to the antibody-bead mixture. The sample was incubated at 4°C with constant rotation for 2 h. Following immunoprecipitation, the supernatant was removed and the beads were washed twice in 250 μl reaction buffer, twice in 250 μl high salt wash buffer (50 mM NaCl, 10 mM Tris-HCl pH 7.5, 1% Igepal CA-630 in nuclease-free H2O), and twice in 250 μl high salt wash buffer (500 mM NaCl, 10 mM Tris-HCl pH 7.5, 1% Igepal CA-630 in nuclease-free H2O) at room temperature. Following each wash, the beads were resuspended in 210 μl elution buffer (6.7 mM m'A, 5% Igepal CA-630 in nuclease-free H2O) and washed with 1,000 μl elution buffer (6.7 mM m'A, 5% Igepal CA-630 in nuclease-free H2O). RNA was purified using the Qiaqen MinElute kit (Qiaqen no. 74204) as described above.

RNA sequencing libraries were prepared using the Illumina TruSeq stranded RNA Library Prep Kit (Illumina no. RS-122-12101) according to the manufacturer's instructions with slight modifications. To account for the AT-richness of cDNA fragments, we used the KAPA HiFi polymerase (Roche no. 07958846001) at the library amplification step. The libraries were sequenced on an Illumina NextSeq 500 platform with a 1 x 150 bp single-end read layout.

For each sample, 25 μl of protein G magnetic beads (Thermo Fisher no. 10004D) were used and pre-washed three times in 500 μl ChIP dilution buffer. Beads were resuspended in 100 μl ChIP dilution buffer and incubated with 1 μg anti-m'a antibody (Abcam no. ab51230) and 2 μl of 10× IP buffer (500 mM NaCl, 1.2 mM EDTA pH 8, 1% Triton X-100, 0.1% SDS). Beads were washed twice in 1 ml of 10× IP buffer, incubated with 1 μl of antibody, and washed again with 1 ml of 10× IP buffer. Following immunoprecipitation, the beads were washed sequentially in 1 ml of low salt wash buffer (20 mM Tris-HCl pH 8, 150 mM NaCl, 2 mM EDTA pH 8, 1% Triton X-100, 0.1% SDS, PI), 1 ml high salt wash buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 2 mM EDTA pH 8, 1% Triton X-100, 0.1% SDS, PI), one ml LiCl wash buffer (10 mM Tris-HCl pH 8, 250 mM NaCl, 1 mM EDTA pH 8, 0.5% Igepal CA-630, 0.5% sodium deoxycholate, and 1% TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA pH 8)). For each wash, the beads were rotated for 5 min at 4 °C except for the last TE wash, which was performed at room temperature. The beads were subsequently resuspended in 210 μl elution buffer (50 mM Tris-HCl pH 8, 10 mM EDTA pH 8, 1% SDS) and protein-DNA complexes were eluted by 30% Triton X-100 incubation at 65 °C in an agitator (900 rpm 1 min, still 30°C). At this step, the input sample was recovered from storage, diluted with 170 μl elution buffer, and processed in parallel with the IP sample.

Phylogenetic analysis of the m'A writer and reader complex. Candidate proteins of the P. falciparum m'A writer complex were identified by searching
Articles

Known members of the m6A writer complex from other eukaryotes against the P. falciparum proteome using blastp. Protein alignments were generated using maft0 and visualized in Geneious3. For the phylogenetic analysis, gaps in the protein alignments were removed with trimal5 and the best phylogenetic model was calculated using protTest3. Phylogenetic trees were constructed using MEGA (v7) with 1,000 bootstrap replicates, and the bootstrap consensus trees were visualized in FigTree (v1.4.3, http://tree.bio.ed.ac.uk/software/figtree/).

Low-complexity disorder regions were annotated using D2P26.

Identification and quantification of mRNA modifications by LC-MS/MS. LC-MS data was extracted using the MassHunter Qualitative and Quantitative Analysis Software (version B06.00) and further analyzed as follows. Briefly, to account for any background signal that could be a contribution from the salts and enzymes in the digestion buffer, the signal intensity (that is area under the curve) for each ribonucleoside was first subtracted from a matrix sample (without any RNA). To calculate relative levels of each modified ribonucleoside and to adjust for different injection amounts of RNA in each sample, the matrix-corrected intensity of the modified ribonucleosides (for example m6A) was then divided by the intensity of the respective canonical ribonucleoside (for example rA). To adjust day-to-day fluctuation in retention time of the mass spectrometer, an internal standard ([15N]2′-deoxyadenosine) was spiked into each sample during RNA digestion. The absolute time point. Overlaps of m6A peaks and genic regions were identified using bedtools ‘multiinter’72, taking the strandedness of the underlying transcript into account. The central enrichment of the deduced sequence motif was calculated using homer2 ‘annotatePeaks.pl’.

m6A enrichment calculation. m6A enrichments were calculated following the approach originally reported66,68. Briefly, for each time point, the number of reads mapping to an m6A peak in the m6A-IP and m6A-input sample was calculated using bedtools ‘coverage -count’ and normalized to the total number of mapped reads in each sample. m6A enrichments were then calculated as m6A-IP/input for each peak and in each individual sample. To calculate decreases in global m6A enrichment in gPiMT-A70 (Fig. 5b, top) at each time point, m6A enrichments of every peak were averaged over the two replicates of each condition. Of note, m6A enrichments were highly reproducible (Fig. 5b, bottom) and the individual m6A enrichment values of each m6A peak and replicate are shown in Supplementary Tables 4–6.

For visualization of changes in m6A enrichments (Fig. 5a), we used cgat’s ‘bam2geneprofile’69. Briefly, the read coverage from m6A-IP and m6A-input libraries of each sample was calculated and normalized to the total size of mapped reads in each library in 10 nt windows and in a region ±2 kb surrounding the m6A peak summit. m6A enrichments in each 10 nt window were calculated as the ration of m6A-IP over m6A-input.

Differential gene expression analysis. RNA sequencing reads for three replicates of gPiMT-A70 and gControl cell lines were mapped to the P. falciparum genome and quality filtered as described above. Strand specific gene counts were calculated using htseq-count69.

Differential gene expression analysis was performed using DESeq270 using all replicates of gPiMT-A70 and gControl at each time point. Due to their monoallelic nature, genes encoding variant surface antigens (var, rifin and stevor) were excluded from the analysis. MA plots were generated using the ‘baseMean’ (normal mean normalized read count over all replicates and conditions) and ‘logFoldChange’ values (gPiMT-A70 over gControl) as determined by DESeq2 for each time point.

For the analysis of m6A enrichment, we first averaged the m6A enrichment of every m6A peak over the two replicates of each condition as described above. We then retained only those m6A-methylated transcripts that harbour m6A peaks with the most pronounced decreases in m6A enrichment (that is, ≥two-fold decrease). For the direct comparison of transcript log fold changes (Fig. 5c, right), the set of transcripts with reduced m6A enrichments was compared to non-methylated transcripts being expressed in the same range of mean normalized read counts (that is, baseMean value).

RPKM (reads per kilobase per exon per one million mapped reads) values were calculated in R using the command rpkml() from the package edgeR69.

mRNA stability and transcriotional efficiency analysis. For the comparative analysis of m6A methylation and mRNA stability or translation efficiency, we used m6A-methylated transcripts that were expressed in the top 50% of all transcripts at a given time point as calculated by RNA-seq of the m6A-input samples and with a peak density of ≥0.2 m6A peaks per kilobase of exon. Corresponding mRNA stability (measured as mRNA half-life) and transcriotional efficiency values of the same parasite developmental stages were extracted from refs. 12,13, respectively. Independent measurements of mRNA stability data were retrieved from ref. 12.

Statistical analysis. All statistical analysis were performed in R71. To test for a normal distribution of the data, the Shapiro–Wilks normality test was used. To test for significance between two groups, a two-sided independent-samples t-test or two-sided Mann–Whitney U-test were performed as indicated. Gene Ontology enrichments were calculated using the build-in tool at plasmoDB.org. Correlations between replicates were calculated in R using function cor() with default settings (calculation of Pearson correlation coefficient ρ).

All heatmaps were visualized using the function heatmap2() in the R gplots package with row z-score normalization. Z-scores were calculated over all samples (‘rows’) following the formula z = (x−μ)/s, with μ and s given value of a specific sample; u = average over all samples and s = standard deviation over all samples.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

All RNA sequencing data are accessible on the Gene Expression Omnibus database (https://www.ncbi.nlm.nih.gov/geo) under study accession number GSE123839. Raw sequence data are accessible on the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA473770. Proteomics and mRNA modification LC-MS/MS data are deposited at the Chorus database with accession number 1579.

Received: 10 March 2019; Accepted: 20 June 2019; Published online: 5 August 2019

Nature Microbiology | Vol 4 | December 2019 | 2246–2259 | www.nature.com/naturemicrobiology

2257
62. Capella-Gutierrez, S., Silla-Martinez, J. M. & Gabaldon, T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973 (2009).

63. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. ProTest 3: fast selection of best-fit models of protein evolution. Bioinformatics 27, 1164–1165 (2011).

64. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874 (2016).

65. Oates, M. E. et al. D5P: database of disordered protein predictions. Nucleic Acids Res. 41, D508–D516 (2013).

66. Aurrecoechea, C. et al. EuPathDB: the eukaryotic pathogen genomics database resource. Nucleic Acids Res. 45, D581–D591 (2017).

67. Li, H. Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009).

68. Li, H. et al. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).

69. Ramírez, E. et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44, W160–W165 (2016).

70. Krzywinski, M. et al. Circos: an information aesthetic for comparative genomics. Genome Res. 19, 1639–1645 (2009).

71. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008).

72. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010).

73. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589 (2010).

74. Marcis, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27, 764–770 (2011).

75. Crooks, G. E., Hon, G., Chandonia, J.-M. & Brenner, S. E. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190 (2004).

76. Meyer, K. D. et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149, 1635–1646 (2012).

77. Sims, D. et al. CGAT: computational genomics analysis toolkit. Bioinformatics 30, 1290–1291 (2014).

78. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31, 166–169 (2015).

79. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550 (2014).

80. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139–140 (2010).

81. R: A language and environment for statistical computing (R Development Core Team, 2012).

82. Broadhurst, K. M. et al. Strand-specific RNA sequencing in Plasmodium falciparum malaria identifies developmentally regulated long non-coding RNA and circular RNA. BMC Genom. 16, 454 (2015).

Acknowledgements

We thank A. Claes, C. Scheidig-Benatar and P. Chen for help with parasite culture. Protein mass-spectrometry was performed at the Biopolymers and Proteomics core of The Koch Institute Swanson Biotechnology Center. This work was supported by a European Research Council Advanced Grant (PlasmoSilencing 670301) and the French Parasitology consortium ParaFrap (ANR-11-LABX0024) to A.Scherf. Work in the labs of P.R.P. and P.C.D. was funded by the National Research Foundation Singapore under its Singapore-MIT Alliance for Research and Technology (SMART) Centre, Infectious Disease and Antimicrobial Resistance IRGs. S.B. and J.M.B. were supported by an EMBO fellowship (S.B.: ALTF 1444-2016; J.M.B.: ALTF 180-2015). A.Sc. acknowledges financial support from the Singapore-MIT Alliance (SMA) Graduate Fellowships.

Author contributions

P.R.P, P.C.D. and A.Sc. conceptualized the project. S.B., J.M.B. and A.Sc. conceived experiments. J.M.B. developed and performed CRISPR interference and dCas9 ChIP-seq experiments. S.B. performed m6A-seq and RT-qPCR experiments. A.S. performed and analysed LC-MS/MS and protein co-IP experiments. S.B., J.M.B. and T.R. generated constructs, transfectants and parasite material. S.B. performed bioinformatic analyses. P.R.P, P.C.D. and A.Sc. supervised and helped interpret analyses. All authors discussed and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41564-019-0521-7.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to S.B.

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

© The Author(s), under exclusive licence to Springer Nature Limited 2019
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted
- Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

- Libraries were sequenced on an Illumina NextSeq500 and collected in FastQ format. qPCR experiments were performed using BioRad CFX. Mass-spectrometry data were collected on a Agilent 6490 triple quadrupole mass spectrometer

Data analysis

- All software used is described in detail in the Methods section of the manuscript and include:
  - Mascot Server (version 2.4.1), Perkins et al. (1999)
  - Proteome Discoverer, Thermo Fisher Scientific
  - Agilent Masshunter Qualitative and Quantitative Analysis Software (version B06.00), Agilent Technologies
  - bwa (version 0.7.10), Li and Durbin, 2009
  - samtools (version 1.3), Li et al., 2009
  - bedtools (version 2.25.0), Quinlan, 2010
  - msc2 (version 2.1.1), Zhang et al., 2008
  - mafft (version 7.271), Katoh and Standley, 2013
  - FigTree (version 1.4.3), Rambault, 2012
  - MEGA (version 7.0.26), Kumar et al., 2016
  - Integrative Genomics Viewer (version 2.3.97), Thorvaldsdóttir et al., 2013
  - homer2 (version 4.10), Heinz et al., 2010
  - treeplot (version 2.3.5), Ramírez et al., 2016
  - cgat (version 0.3.3), Sims et al., 2014
  - bcftools (version 2.19), Illumina
  - D2P2, Dates et al., 2012
  - R (version 3.5.3), R Development Core Team, 2012
  - blastp (version 2.2.31+), Altschul et al., 1990
  - protTest3 (version 3.4.2), Darriba et al., 2011
  - Geneious (version 11.0.3), Kearse et al., 2012
  - trimal (version 1.4.rev15), Capella-Gutierrez et al., 2009
Data

Policy information about availability of data
All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw (fastq format) and processed (bedgraph format) data are deposited at the GEO database with accession number GSE123839 and at NCBI BioProject with accession number PRJNA473770. Proteomics and mRNA modification LC-MS/MS data are deposited at the Chorus database with accession number 1579

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

- Life sciences
- Behavioural & social sciences
- Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

- Sample size
- Data exclusions
- Replication
- Randomization
- Blinding

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
|     | Antibodies            |
|     | Eukaryotic cell lines |
|     | Palaeontology         |
|     | Animals and other organisms |
|     | Human research participants |
|     | Clinical data         |

Methods

| n/a | Involved in the study |
|-----|-----------------------|
|     | ChIP-seq              |
|     | Flow cytometry        |
|     | MRI-based neuroimaging |

Antibodies

Primary antibodies used for immunoprecipitations (concentrations mentioned in the Methods)
- anti-HA Dynabeads (Thermo Fisher # 88836)
- anti-m6A (Abcam # ab151230)
- anti-HA (Abcam #ab9110)
Primary antibodies used for Western Blots:
- anti-HA (Abcam #ab9110); Dilution: 1:1000
- anti-H3 (Abcam # ab1791); Dilution: 1:1000
- anti PfAldolase (Abcam # ab38905); Dilution: 1:5000

Secondary antibodies for Western blot:
- donkey anti-rabbit IgG-HRP (Sigma-Aldrich # GENA934-1ML)

Validation

All antibodies used in this study are commercially available.
- ab151230: Specificity for m6A-IP was tested by LC-MS/MS as reported in the manuscript (Fig. 4a, Supplementary Fig. 2f).
- Species reactivity (according to producer): Mouse, human, fly.
- Species reactivity (according to producer): independent.

Species reactivity (according to producer):
- ab9110: ELISA: The anti HA diluted 1:70,000 gave an O.D.=1.0 in a 15 minute reaction against peptide conjugated with a different carrier than used for anti peptide purification. HRP conjugated Goat anti rabbit IgG was used and TMB was the substrate.
- Validation application (according to producer): ChIP, IP, ELISA, IFA, WB among others.

Species reactivity for ab1791 (according to producer): Wide range of species, including rat, mouse, human, fly, yeast.

Species reactivity for ab38905 (according to producer): Species independent.

Species reactivity (according to producer): Mouse, human, fly.
- For ab1791 and ab38905, validation includes coincidence of Western Blot bands as described in other studies using these antibodies in P. falciparum (e.g. Chen et al. 2016, PMID 26902486 for ab1791 and ab38905) and absence of any bands of different size than the ones expected (Fig. 2b,c and Fig. 6e).
- Validate application (according to producer): Nucleotide Array, ChIP, IP, IFA, WB among others.

Eukaryotic cell lines

Cell line source(s)

All transfections were performed on bulk Plasmodium falciparum strain 3D7 (Walliker et al., 1987, PMID 3299700). All cell lines used for experiments (i.e. PfMT-A70-HA, PfYTH1-HA, gPfMT-A70, gControl) were generated in the laboratory as described in the Methods section of the manuscript.

Authentication

The cell lines used were authenticated by regular PCR amplification of the sgRNA from genomic DNA of gControl and gPfMT-A70 cell lines and subsequent Sanger sequencing. A ’phenotypic’ authentication was performed by measuring binding of dCas9 to the target locus using dCas9-ChIP-seq, measuring PfMT-A70 downregulation by RT-qPCR and m6A/A levels by LC-MS/MS (Fig. 3d)

Mycoplasma contamination

The cell lines were not tested for mycoplasma contamination.

Commonly misidentified lines

No commonly misidentified cell line was used in this study.

ChIP-seq

Data deposition

Confirm that both raw and final processed data have been deposited in a public database such as GEO.

Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links

May remain private before publication.

GEO: ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE123839
BioProject: ncbi.nlm.nih.gov/bioproject/?term=PRJNA473770
Chorus: chorusproject.org/pages/dashboard.html#projects/all/1579/experiments

Files in database submission

m6A-seq_12hpi_gControl_rep1.mRNAcorr.bedGraph.gz
m6A-seq_12hpi_gControl_rep2.mRNAcorr.bedGraph.gz
m6A-seq_12hpi_gPfMT-A70_rep1.mRNAcorr.bedGraph.gz
m6A-seq_12hpi_gPfMT-A70_rep2.mRNAcorr.bedGraph.gz
m6A-seq_24hpi_gControl_rep1.mRNAcorr.bedGraph.gz
m6A-seq_24hpi_gControl_rep2.mRNAcorr.bedGraph.gz
m6A-seq_24hpi_gPfMT-A70_rep1.mRNAcorr.bedGraph.gz
m6A-seq_24hpi_gPfMT-A70_rep2.mRNAcorr.bedGraph.gz
m6A-seq_36hpi_gControl_rep1.mRNAcorr.bedGraph.gz
m6A-seq_36hpi_gControl_rep2.mRNAcorr.bedGraph.gz
m6A-seq_36hpi_gPfMT-A70_rep1.mRNAcorr.bedGraph.gz
m6A-seq_36hpi_gPfMT-A70_rep2.mRNAcorr.bedGraph.gz
m6A-12hpi_peaks.bed.gz
m6A-24hpi_peaks.bed.gz
m6A-36hpi_peaks.bed.gz
Gene_counts_CrisprR1.txt.gz
gPfMT-A70-12hpi-m6A-IPrep1.fasta.gz
**Methodology**

**Replicates**
For each time point (i.e. 12, 24 and 36 hours post infection), m6A-seq was performed for two replicates of the gControl and gPfMT-A70 cell line.

**Sequencing depth**
Sequencing layout: 1x150bp

Sequencing Depth for each sample (ID: total number of reads/ uniquely mapped)
- gPfMTA70-12hpi-m6A-IP-rep1: 19430037/ 13963032
- gPfMTA70-12hpi-Input-rep1: 13441632/ 8719913
- gControl-12hpi-m6A-IP-rep1: 17321146/ 14293035
- gControl-12hpi-Input-rep1: 18402544/ 12551742
- gPfMTA70-24hpi-m6A-IP-rep1: 14832785/ 13002565
- gPfMTA70-24hpi-Input-rep1: 18337748/ 16470347
- gControl-24hpi-m6A-IP-rep1: 24141441/ 21828234
- gControl-24hpi-Input-rep1: 17482930/ 14648824
- gPfMTA70-36hpi-m6A-IP-rep1: 13516766/ 12385817
- gPfMTA70-36hpi-Input-rep1: 9840277/ 9560879
- gControl-36hpi-m6A-IP-rep1: 9211182/ 8488197
- gControl-36hpi-Input-rep1: 11425418/ 10446986
- gPfMTA70-12hpi-m6A-IP-rep2: 10348993/ 7464267
- gPfMTA70-12hpi-Input-rep2: 33755708/ 21878655
- gControl-12hpi-m6A-IP-rep2: 9300427/ 7683230
- gControl-12hpi-Input-rep2: 10141923/ 6954763
- gPfMTA70-24hpi-m6A-IP-rep2: 7546152/ 6599773
- gPfMTA70-24hpi-Input-rep2: 9258530/ 8307424
- gControl-24hpi-m6A-IP-rep2: 24255439/ 22058507
- gControl-24hpi-Input-rep2: 8434386/ 7072898
- gPfMTA70-36hpi-m6A-IP-rep2: 13336427/ 12283717
- gPfMTA70-36hpi-Input-rep2: 15567665/ 15439678
- gControl-36hpi-m6A-IP-rep2: 8044591/ 7538068
- gControl-36hpi-Input-rep2: 14031942/ 12887826

**Antibodies**
- anti-m6A (Abcam #ab151230)

**Peak calling parameters**
Peaks were called using macs2 with default settings but without prior peak modeling (option ‘--nomodel’) and the fragment size set to 150 bp (option ‘--extsize 150’). The genome size (-g) was adjusted to represent the total size of the protein-coding transcriptome.

**Data quality**
Only uniquely mapped reads with an alignment score ≥ 20 were used for peak calling. A single set of m6A peaks for each time point was generated by identifying overlaps of significant peaks (FDR ≤ 0.05) that were identified in ≥ two of the four samples collected at each time point. Relative m6A enrichments were correlated among replicates as shown in Fig. 4g and Fig. 5b. Number of peaks: 12 hpi: 603; 24hpi: 873; 36hpi: 996

**Software**
- Read mapping: bwa (version 0.7.10), Li and Durbin, 2009
- Alignment duplicate and quality filtering: samtools (version 1.3), Li et al., 2009
- Peak calling: macs2 (version 2.1.1), Zhang et al., 2008
Peak filtering: bedtools "multiinter" (version 2.25.0), Quinlan, 2010