Genome-wide real-time in vivo transcriptional dynamics during *Plasmodium falciparum* blood-stage development

Heather J. Painter\(^1,2\), Neo Christopher Chung\(^3,7\), Aswathy Sebastian\(^4\), Istvan Albert\(^1\), John D. Storey\(^3,5\) & Manuel Llinás\(^1,2,6\)

Genome-wide analysis of transcription in the malaria parasite *Plasmodium falciparum* has revealed robust variation in steady-state mRNA abundance throughout the 48-h intraerythrocytic developmental cycle (IDC), suggesting that this process is highly dynamic and tightly regulated. Here, we utilize rapid 4-thiouracil (4-TU) incorporation via pyrimidine salvage to specifically label, capture, and quantify newly-synthesized RNA transcripts at every hour throughout the IDC. This high-resolution global analysis of the transcriptome captures the timing and rate of transcription for each newly synthesized mRNA in vivo, revealing active transcription throughout all IDC stages. Using a statistical model to predict the mRNA dynamics contributing to the total mRNA abundance at each timepoint, we find varying degrees of transcription and stabilization for each mRNA corresponding to developmental transitions. Finally, our results provide new insight into co-regulation of mRNAs throughout the IDC through regulatory DNA sequence motifs, thereby expanding our understanding of *P. falciparum* mRNA dynamics.
Hundreds of millions of annual malaria cases are caused by intracellular protozoan *Plasmodium* parasites resulting in almost half a million deaths and an enormous economic impact on almost half the world’s population annually. Efforts to eradicate this pathogen have been at the forefront of infectious disease research for decades. Although it has been over 15 years since the completion of the whole genome sequence of *Plasmodium falciparum*, transcriptional regulation throughout the parasite’s complex life cycle remains poorly understood.

The malaria parasite life cycle commences upon the transfer of *Plasmodium* sporozoites from the saliva of an infected *Anopheles* spp. female mosquito vector to a human host. These sporozoites travel to and colonize a small number of hepatocytes within the liver where they replicate, forming thousands of merozoites. Liver-stage merozoites are released into the circulatory system and initiate the blood stage of infection, which is responsible for the clinical symptoms of malaria. Development during the 48 h intraerythrocytic development cycle (IDC) is comprised of parasite maturation and cell division, resulting in the formation of up to 32 daughter cells, which are released and invade new erythrocytes. A small proportion of asexual parasites commit to an alternative cellular fate of sexual differentiation during the IDC, eventually resulting in the development of male and female gametocytes. Maturation of *P. falciparum* sexual stage parasites takes 10–12 days, and it is these sexual forms that are taken up by a mosquito vector for eventual onward transmission to another human host, thus completing the *Plasmodium* full life cycle. Differentiation of the malaria parasite in these various cell types and tissues is regulated by coordinated programs of gene expression.

Gene expression during the IDC is highly periodic and follows a transcriptional cascade with the majority of genes expressed in a “just-in-time” fashion. Although canonical transcriptional machinery is present, the mechanisms that underlie the regulation and specificity of active transcription remain largely uncharacterized. To date, only a single family of specific transcription factors, the 27-member Apicoplastan AP2 (ApiAP2) protein family, has emerged as transcriptional regulators with functions across all developmental stages. However, the challenge of regulating roughly 5500 *Plasmodium* genes with a small repertoire of transcription regulators suggests a significant role for epigenetic control and post-transcriptional regulation in *Plasmodium* spp.

The control of gene expression at the RNA level is evidenced by several studies reporting a significant delay between transcription and translation, ribosomal influence on the timing of mRNA translation, and a lack of coordination between active transcription and mRNA abundance during parasite development. Bioinformatic analyses have suggested that between 4 and 10% of *Plasmodium* genome encodes RNA-binding proteins (RBPs) while the majority of predicted RBPs remain uncharacterized.

Several post-transcriptional regulatory factors have been well-studied such as DOZI (DDX-6 class DEAD box RNA helicase), CITH (Sm-like factor homolog of CAR-1 and Trailer Hitch), and pumilio family proteins (PUF1 and PUF2) that have been demonstrated to play critical roles in the translational repression of genes essential for *Plasmodium* transmission. Additionally, the acetylation lowers binding affinity (ALBA) domain-containing proteins (PAlba1-4) Bruno/CELF family members (CELF1 and CELF2) and PfC7 have been demonstrated to interact with RNA; however, the exact role has yet to be defined for a majority of these RBPs in parasite development.

In all living systems, cellular mRNA levels are determined by the interplay of tightly regulated processes of RNA production, processing, stabilization, and degradation. Thus, changes in transcript abundance measured by DNA microarray, RNA-sequencing, or RT-qPCR studies reflect the relative rates of both transcription and mRNA stabilization. Since the advent of the complete *P. falciparum* genome sequence, numerous studies have utilized whole-genome methods to measure transcript abundance from population of parasite-infected red blood cells as well as, more recently, single cells. In addition, other studies have measured transcriptional activity by comprehensively capturing nascent mRNA transcription, stabilization, and total abundance of mRNA in vivo throughout the IDC using 4-thiouracil (4-TU) biosynthetic RNA labeling.

Bioinformatic analysis has revolutionized the capture and measurement of in vivo cellular mRNA dynamics during development and in response to various stressors across a diverse array of organisms including *Plasmodium* and human and yeast cells, short periods of metabolic labeling into nascent RNA has been used to capture RNA synthesis, decay, and splicing rates. However, thiol-modified RNA capture and kinetic profiling of RNA dynamics are dependent upon pyrimidine salvage, and the evolutionary lineage of *Plasmodium* has lost this biochemical capacity. We have previously demonstrated that genetic engineering of the pyrimidine salvage pathway into *Plasmodium* parasites enables the efficient uptake of 4-TU and biosynthetic capture of thiol-labeled RNAs to examine RNA dynamics during parasite development.

Herein, we report the application of biosynthetic RNA labeling and capture to measure genome-wide nascent mRNA transcription and stability in vivo throughout the IDC of the malaria parasite. This method relies on the genetically facilitated scavenging of 4-TU and genome-wide DNA microarray measurements of labeled and unlabeled RNAs. Using this method, we have captured mRNA dynamics during asexual development at hourly resolution and apply a quantitative normalization model to define gene-specific mRNA metabolism. These data provide dynamic information and stability information for all mRNAs expressed during the IDC and also recapture known features of the “just-in-time” cascade of mRNA transcription. Additionally, modeling of the mRNA dynamics reveals regulatory mechanisms that cooperate during parasite development, which we used to predict functional roles for putative transcriptional regulators. Our results provide an updated in-depth view of mRNA transcriptional dynamics and gene regulation in *P. falciparum* throughout the IDC.

**Results**

**Genetic modification of *P. falciparum* for pyrimidine salvage.** In vivo mRNA labeling is dependent upon the salvage of 4-TU; however, *P. falciparum* is not capable of pyrimidine precursor uptake and incorporation into RNA. We previously reported the use of episomal expression of the bifunctional fusion gene encoding cytosine deaminase and uracil phosphoribosyltransferase (FCU) from yeast to enable parasite salvage of 4-TU. To avoid cell-to-cell heterogeneity in plasmid copy number expressing FCU, this study uses a single copy cassette that is stably integrated into the attB
locus of 3D7<sup>attB</sup><sup>52</sup> to constitutively express a GFP-tagged FCU protein (Supplementary Fig. 1A)<sup>25</sup>. Stable integration of 3D7attB::fcu-gfp (3D7<sup>attB</sup>::fcu-gfp<sup>10</sup>) was verified by PCR (Supplementary Fig. 1A). Expression of FCU-GFP was confirmed by Western blot and localized to the cytoplasm by immunofluorescent imaging analysis (Supplementary Fig. 1B and C). Salvage and incorporation of a 4-TU into total cellular RNA by the 3D7attB::fcu-gfp transgenic parasites was confirmed by Northern blot (Supplementary Fig. 1D).

### Nascent in vivo mRNA labeling and capture throughout the IDC

The goal of this study was to produce high-resolution RNA dynamics data for each transcript expressed during the 48 h <i>P. falciparum</i> IDC transcriptome. To capture nascent RNA transcription and stabilization, we exposed highly synchronized 3D7attB::fcu-gfp to 10 min pulses of 40 µM 4-TU followed by total RNA extraction of 4-TU labeled RNA is separated from the total RNA resulting in two fractions, transcribed (labeled) and stabilized RNAs were quantified by Northern blot (Supplementary Fig. 1D).

#### Statistical model of <i>P. falciparum</i> asexual mRNA dynamics.

Because standard two-color DNA microarray analysis normalizes log<sub>2</sub>(green/red) values within individual timecourses, the log<sub>2</sub> values do not distinguish the relative contributions of the various pools of RNA (labeled, unlabeled) to the total RNA signal (Fig. 1b; raw log<sub>2</sub> ratios) when timecourses are directly compared.

**Fig. 1** Biosynthetic mRNA labeling and capture throughout the <i>P. falciparum</i> IDC reveals complex developmental mRNA dynamics. **a** Biosynthetic mRNA capture timecourse with 4-TU (40 µM) labeling (10 min pulse) of highly synchronous transgenic 3D7<sup>attB</sup>::fcu-gfp parasites every hour throughout the 48 h IDC followed by total RNA extraction. 4-TU labeled RNA is separated from the total RNA resulting in two fractions, transcribed (labeled) and stabilized (unlabeled). These RNAs and the total cellular RNA pool were quantified by DNA microarray analysis (Supplementary Fig. 2). **b** Data obtained by cDNA microarray was normalized and modeled to determine the contribution of nascent transcription (red circles) and stabilization (blue circles) to the total abundance profile (black circles) for each gene. The modeled expression values of nascent transcription (5207 genes, red) and stabilization (5236 genes, blue) values for each gene were ordered based on the peak timing of total abundance (5428 genes) during the 48 h IDC as determined by Lomb-Scargle analysis and displayed in a heatmap as the log<sub>2</sub> expression value. The summation of the expression values for both transcription and stabilization is displayed as a heatmap representing the estimated total mRNA abundance (5207 genes, black), recapitulating the canonical phaseogram of measured total mRNA abundance (left, RGB heatmap).
to each other. To circumvent this issue, we developed a statistical model to quantitatively determine the contribution of nascent transcription (labeled) and mRNA stabilization (unlabeled) to the total abundance profile of each gene by comparing the signal intensities for each probe on the DNA microarrays across all three timecourses (Methods). After the three timecourses were separately normalized by Rnits [33] (Supplementary Fig. 3), the mRNA dynamics were estimated by negative least squares averaging across the gene to which they mapped using Tukey’s biweight [54, 55] and plotted across the 48 h IDC (Fig. 1c, calculated values), resulting in mRNA dynamic profiles for most P. falciparum genes (Supplementary Data 1). Gene plots can be found on individual gene pages on http://www.PlasmoDB.org [36] and https://doi.org/10.6084/m9.figshare.6200792.v1 where both the raw log 2 (green/red) ratios and normalized expression values are also available for download.

Based on the calculated normalized nascent mRNA expression values, we found that active transcription from the nuclear genome occurs throughout the IDC (Fig. 1c and Supplementary Data 1). We also found that transcription of nascent mRNAs from both the mitochondrial and apicoplast organelle genomes is tightly coordinated and begins 8–10 hpi (Supplementary Fig. 4A and 4B). Transcription of these non-nuclear genes peaks at 31 hpi (Supplementary Fig. 4B), coinciding with the start of organelar genome replication and division [37, 38]. In contrast, nuclear-encoded genes destined for the apicoplast are transcribed prior to apicoplast encoded genes (Supplementary Fig. 4C), likely due to their role in transcription and translation of the apicoplast genome [39]. Interestingly, we also observe that the capture of nascent transcription from the apicoplast genome reveals a wide range of expression levels regardless of their genome position (Supplementary Fig. 4D).

Based on the statistical model of the data, we found that the abundance profiles of individual genes are dictated by a combination of nascent transcription and mRNA stability as the parasite develops (Fig. 1b). When considering the global transcriptome, the calculated contribution (from the statistical model) of nascent transcription and mRNA stabilization to total abundance profiles varies as the parasite progresses through the IDC (Fig. 1c and Supplementary Fig. 5). To confirm the accuracy of the model, we summed the expression values of both nascent transcription and stabilization to simulate the total RNA values (Fig. 1c). When the estimated total mRNA abundances are plotted in the same order as the observed values, the canonical phaseogram of the P. falciparum transcriptome observed by traditional methods is recapitulated (Fig. 1c [52, 9, 10, 36]), with a median correlation of 0.94 (Supplementary Fig. 5A). Additionally, the gene-specific patterns of mRNA stabilization or transcription correlate with the overall transcript abundance profiles (Supplementary Fig. 5B and 5C; transcription versus total abundance median corr = 0.86, stabilization median corr = 0.85), confirming that our method and model capture the mRNA dynamics that dictate the abundance of a given transcript throughout the IDC. In contrast, gene-specific transcription and stabilization dynamic profiles are not highly correlated with each other (median r = 0.69) (Fig. 2a), likely due to the differences in peak timing of nascent transcription versus stabilization. By calculating the difference in the maximum transcription and stabilization to the total abundance, we estimate that peak transcription precedes the peak abundance (~80 min), while peak stabilization follows approximately 41 min after (Fig. 2b). This supports the current notion that the majority of genes are transcribed in a “just-in-time” manner with nascent transcription peaking prior to mRNA abundance, while post-transcriptional regulation likely influences the delay from peak time of mRNA abundance.

Variation in temporal mRNA dynamics throughout the IDC. To determine the stage-specific effects of mRNA dynamics on transcript abundance, we binned genes into six groups based on their peak timing of abundance during the IDC as captured by the total RNA timecourse (Fig. 2c). These six groups reflect the major morphological stages of development during the IDC including early ring (0–10 hpi), mid-ring (11–15 hpi), late ring/early trophozoite (16–21 hpi), mid-trophozoite (22–26 hpi), late trophozoite (27–32 hpi), and schizont (33–48 hpi). We then predicted the predominant (r ≥ 0.50) mechanism acting on each gene based on the correlation of either transcription (Fig. 2c, red) or stabilization (blue) to the total abundance profile (Fig. 2c, pie charts). Overall, the influence of transcriptional activity or mRNA stability on total transcript abundance across the genome is relatively equal (transcription = 2767 genes; stabilization = 2660 genes) (Fig. 2c). However, there is a higher proportion of contribution from nascent transcription to mRNA abundance profiles at a majority of developmental stages (early ring = 55%, late ring/early trophozoite = 68%, mid-trophozoites = 61%, late trophozoites = 52%), while stabilization dominates at the mid-ring (64%) and schizont (62%) stage transitions (Fig. 2c and Supplementary Fig. 5D). These results reveal that there are stage-specific regulatory mechanisms that influence the patterns of mRNA abundance throughout both transcription and mRNA stability during the IDC.

Previous studies have suggested that both transcriptional activity and mRNA half-life increase as the parasite progresses through the IDC [39, 40]. These results were based on methodologies that relied on measuring ex vivo transcription from isolated nuclei [26, 39] or inhibition of transcription with Actinomycin D [40] followed by RNA-seq or DNA microarray to quantify mRNA. Because our dataset was generated without physical or chemical perturbations to the parasite, we analyzed our results in the context of these earlier studies [39, 40]. To do this, the rate of transcription for each gene was calculated based on the slope of the expression value of the labeled mRNA just prior to the peak of nascent transcript levels, while the rate of decay was determined by the slope of the unlabeled mRNA expression value following the peak of stabilization (Supplementary Data 2). The transcription and decay rates of genes with peak expression during the early ring, mid-ring, late ring/early trophozoite, mid-trophozoite, late trophozoite, and schizont stages were grouped together. The rate of transcription and decay are represented in Fig. 2d, e for each phase of development. These data show that genes expressed during the early and mid-ring stage are transcribed at a slower rate (median = 1.77 and 2.6 gene transcripts/min, respectively), compared to the rest of the life-cycle (median = 4.1–6.3 gene transcripts/min) (Fig. 2d, Supplementary Data 2). Conversely, the mean rate of decay varies significantly as the parasite progresses through the IDC (Fig. 2e; Supplementary Data 2), suggesting that the regulation of mRNA turnover is stage-specific. Surprisingly, the median rate of decay is slowest during the mid- to late-ring stages and becomes significantly more rapid during the mid-
trophozoite stage of parasite development (Fig. 2c). However, there are an increased number of genes that are stabilized during the later stages of parasite development (Fig. 2e positive outliers 36–48 hpi = 21 genes, Supplementary Data 2). Although the increase in stability of a small number of genes during the late IDC stage confirms earlier findings by Shook et al.40, it is important to note that the vast majority of genes with peak abundance during the late trophozoite and schizont stage do not display a significantly lower rate of decay (Fig. 2e).

**IDC stage transitions coincide with “bursts” in mRNA dynamics.** Although it is conventionally accepted that transcription during the *Plasmodium* IDC occurs in a “just-in-time” fashion10, this is largely based on transcript abundance measurements and annotated gene function. Using the measured mRNA dynamics, we examined the timing of active transcription and stabilization of the parasite as it develops throughout the IDC. By binning the genes based on their peak time of nascent transcription or peak time of stabilization every 30 min throughout the IDC, our data reveal peaks in the numbers of genes being transcribed and/or stabilized at four major time periods during parasite development (Fig. 3 and Supplementary Data 3). Interestingly, these “bursts” of mRNA regulatory activity coincide with major morphological transitions during the IDC (Fig. 3). Next, we determined the total number of genes within each “burst” by placing the boundaries at the nearest minimum trough adjacent to each peak and then performed GO-term enrichment analysis on all genes. The first peaks of both transcription and stability occur during the mid-ring stage of development (12 hpi), and is characterized by peak transcription of genes whose products are enriched for RNA processing, translation, and antigenic variation (Fig. 3 and Supplementary Data 3). During this stage, genes encoding for variant surface antigens and cell adhesion also reach their peak stability (Fig. 3 and Supplementary Data 3). As expected, peaks of gene regulation that occur in the early-, mid-, and late-trophozoite are enriched for genes whose functions are necessary and unique for each phase of trophozoite development (Supplementary Data 3). For example, the mid-trophozoite is known to be the most metabolically active stage of development, and during this phase, there is an enrichment in transcription and stabilization of genes, which encode for amino acid, tRNA, ncRNA, pyruvate, glycolytic, and carbohydrate metabolic processes (Fig. 3 and Supplementary Data 3). In contrast, the late-trophozoite phase is characterized by a peak in the regulation of genes involved in DNA metabolism and replication in preparation for mitotic production of daughter cells.

Beyond the identification of genes enriched during developmental transitions, we also identified a number of genes whose peak timing of transcription and stabilization occurred at either 1 or 48 hpi. As expected, at 48 hpi the number of stabilized genes (378 genes) outnumbers those that are actively transcribed (140 genes) (Fig. 3 and Supplementary Data 3). This corroborates earlier studies demonstrating the increased presence of mRNA-interacting proteins during late schizogony27 and the increased half-life of genes at this stage40. As expected, these genes are enriched in processes necessary for parasite invasion and the establishment of a niche within the erythrocyte that will allow for the uptake and utilization of nutrients from the extracellular space (Supplementary Data 3), biological process GO-term enrichment, vesicle-mediated transport (GO:0006888,
Due to the importance of parasite-derived erythrocyte signaling and microtubule formation. Conversely, genes with major morphological stage transitions of the IDC. Separate histograms displaying the number of genes which peak in transcription (red) and stabilization (blue) at either 1 or 48 hpi highlighting the reliance on a specific regulatory mechanism at each stage of development. The transcribed (red) or stabilized (blue) genes within each group were analyzed for GO-term enrichment and significance values of each gene at the peak hour of transcription for each population of transgenic parasites, we determined the expression of each gene family to determine the peak time of nascent transcription during the IDC.

**Fig. 3** Capture of *P. falciparum* mRNA dynamics reveal “bursts” of regulation during the IDC. Histogram plot displaying the number of genes with peak transcription (left) or stabilization (right) timing binned into 30 min intervals throughout the IDC representing the hours in which there are greater than five genes enriched (5–45 hpi). The histogram is divided into five distinct groups based on the distribution of the modes and adjacent minima, corresponding with major morphological stage transitions of the IDC. Separate histograms displaying the number of genes which peak in transcription (red) and stabilization (blue) at either 1 or 48 hpi highlighting the reliance on a specific regulatory mechanism at each stage of development. The transcribed (red) or stabilized (blue) genes within each group were analyzed for GO-term enrichment and significant terms (Bonferroni corrected *p*-value ≤ 0.01) are displayed in pie charts and listed in Supplementary Data 3

Nascent transcription captures clonally variant expression. One advantage of biosynthetic mRNA capture is the ability to detect and compare the level of transcriptional activity for individual genes. Due to the importance of parasite-derived erythrocyte surface antigens in modulating immunity to malaria infections, transcriptional regulation of these antigens is one of the most highly-studied areas of *Plasmodium* biology. It has long been established that members of surface antigen gene families (vars, rifins, stevors, pfmc-2tms, and surfinfs) are transcribed during the late ring to early trophozoite stages of development and exhibit variation in their expression among parasite clones. To identify the timing of transcription, we calculated the average normalized expression values (log2 mean centered) for each gene family to determine the peak time of nascent transcription during the IDC. Transcription of the *var* and *rifin* gene families peaked at 11 hpi, while *stevor* reached their peak transcription at 12 hpi. Interestingly, the peak expression of *surfin* and *pfmc-2tm* genes occur at 10 and 16 hpi, respectively (Supplementary Fig. 6A), which may be explained by the demonstration that the transcriptional regulation of these two families occurs through different factors. To determine which member(s) of these surface antigen gene families was being transcribed in this clonal population of transgenic parasites, we determined the expression values of each gene at the peak hour of transcription for each family of surface antigens (Fig. 4b). Our data demonstrate transcription of a single dominant *var*, as well as multiple *rifin* and *stevor* genes (Fig. 4b), supporting the mutually exclusive
Fig. 4 Clonal expression of *P. falciparum* variant surface antigens. a Transcription of surface expressed gene families: *vars* (78/78), *rifins* (155/185), and *stevors* (35/36) is displayed as the mean (log₂ expression value) and ±s.e.m. throughout the 48 h IDC. The hpi where peak transcription of each gene family is displayed as a dropline in each graph. b At the peak time of transcription, the expression value for each *var* (sub-groups and pseudogenes are labeled), *rifin*, and *stevor* genes

Fig. 5 mRNA dynamics of invasion-related genes. a The calculated total abundance (black), transcription (red), and stabilization (blue) expression values of selected genes demonstrated to be involved in parasite erythrocyte invasion are represented and clustered based on Euclidian distance metric. Highlighted in red text are selected genes representative of the various mRNA dynamics, which regulate expression. b The profiles of the expression of values of transcription (red circles) and stabilization (blue circles) for selected genes representing the dominant mode of regulation through nascent transcription (*msp7* and *msp2*), a balance of both transcription and stabilization (*sera3* and *sera4*), or stabilization alone (*rhoph3* and *rhoph2*) are plotted as throughout the IDC. The insets in *rhoph2* and *rhoph3* plots are the expression values of transcription from 0 to 48 hpi
expression of only the var gene family of surface antigens. Although pfmc-2tm and surfin genes are also considered surface antigens, they do not demonstrate clonal variation (Supplementary Fig. 5B), as reported previously\(^6\). Taken together, the biosynthetic labeling and capture of nascent transcription detects transcriptional variation within clonally variant gene families.

**mRNA regulation of invasions genes is predictive of function.** A closer examination of the mRNAs stabilized during the schizont stage of development revealed that a majority encode proteins that function in the invasion of merozoites into erythrocytes. Because of this observation, we determined if these mRNAs are co-regulated by comparing their similarity in total abundance, transcription, and stabilization (Fig. 5a). As previously demonstrated\(^1\), we found that the total mRNA abundance profiles of invasion genes peak during schizogony (Fig. 5a). However, the levels of active transcription and stabilization vary drastically (Fig. 5a, b), revealing that there are at least two distinct regulatory mechanisms that dictate these patterns of total abundance. For example, a majority of mRNAs that encode merozoite surface proteins (msps), including msp7 and msp2, are predominantly regulated at the level of transcription and appear to be minimally influenced by stabilization (Fig. 5a, b). This may suggest that msp5s are rapidly transcribed and translated during schizogony for presentation on the surface of the merozoite prior to rupture from the red blood cell\(^6\). Conversely, although rhoptry genes rhoph2 and rhoph3 are also transcribed during the schizont stage, their peak abundance is contributed by mRNA stabilization (Fig. 5b), suggesting that these mRNAs may be translationally repressed or are slow to turn over until their gene products are necessary. Previous reports have demonstrated a “step-wise” role for proteins involved in invasion, which is required after the merozoite has formed a tight interaction with the erythrocyte surface\(^6\). More recently, invasion proteins (such as RhopH2 and RhopH3) have been shown to assist in the establishment of nutrient transporters across the parasitophorous vacuolar membrane following invasion\(^3,6,5-6\). Although these genes are transcribed during the schizont stage and their protein products are necessary for invasion, stabilization guarantees that the mRNAs will be available for translation immediately upon invasion to promote proper nutrient transport for the developing parasite as recently demonstrated\(^6,66\). Interestingly, the mRNA dynamics of both sera3 and sera4 reveal a balance of transcription and stabilization (Fig. 6b), potentially due to their dual role in two stages of the IDC, rupture as well as reinvasion. The SERA proteins have been demonstrated to be essential for schizont rupture and remain associated with the merozoite surface until attachment and invasion of a new red blood cell\(^68\).

**mRNA dynamics are predictive of cis-regulatory sequence motifs.** High resolution capture of mRNA dynamics provides a more refined set of data for the identification of DNA regulatory motifs associated with co-transcription. Using the finding informative regulatory elements (FIRE) algorithm\(^69\), we used co-regulated mRNAs to predict DNA sequence motifs that are enriched in the 5’ untranslated regions of these genes (up to −1,000 bp from ATG start site) (Fig. 1c and Supplementary Data 1). We found enrichment of 21 DNA motifs associated with

---

**Fig. 5** Prediction of cis-regulatory motifs from *P. falciparum* mRNA dynamics. a) FIRE analysis prediction of enriched DNA and RNA motifs within the whole genome dynamics data. The profiles of transcription and stabilization are clustered into peak timing during the IDC (columns) predicted enriched motifs arranged into putative functional modules (rows). The yellow color map indicates overrepresentation of a motif in a given cluster; significant overrepresentation (P < 0.05 after Bonferroni correction) is highlighted with red frames. Similarly, the blue color map and blue frames indicate underrepresentation. For each motif, the ApiAP2 trans-factor demonstrated to interact with each predicted motif\(^70\) is listed. b) The nascent transcription profile of genes identified by FIRE (left panel) or ChIP-seq (right panel\(^71\)) with the enriched PfAP2-I motif GTGCA are plotted (gray, mean log2 expression value and ± s.e.m.) in comparison to the transcription profile of pfap2-i (red circle). Venn Diagram represents the overlap of genes identified by FIRE (blue) or ChIP-seq (green). c) The nascent transcription profile of genes identified by FIRE (left panel) or ChIP-seq (right panel\(^72\)) with the enriched PfAP2-G2 motif ACAAC are plotted (gray, mean log2 expression value and ± s.e.m.) in comparison to the transcription profile of pfap2-g2 (red). Venn Diagram represents the overlap of genes identified by FIRE (blue) or ChIP-seq (green).
co-transcription during the IDC (Fig. 6a). Of these 21 motifs, seven have been previously identified as ApiAP2 cis-interacting motifs. To verify the accuracy of our motif identification, we calculated the average transcription of genes enriched for the known DNA recognition motif (GTGCA), which is bound by the third AP2 domain of PfAP2-1, compared to the expression of Pfap2-i (Fig. 6b). As expected for a transcriptional activator, genes containing the GTGCA regulatory motif are actively transcribed following the peak expression of Pfap2-i (Fig. 6b). Additionally, our bioinformatic analysis identified 120 genes previously demonstrated to be directly regulated by PfAP2-I via ChIP-seq analysis (Fig. 6c). Conversely, transcription of genes enriched for the cis-regulatory motif ACAAC decreases following the transcription of ap2-g2 (Fig. 6c), which has been suggested to act as a transcriptional repressor during the IDC in Plasmodium berghei. We compared the target genes enriched with the ACAAC DNA-motif with the AP2-G2 ChIP-seq gene targets and found that capture of nascent transcription, combined with FIRE analysis, identified greater than 50% of the reported targets of AP2-G2 (623/1222) (Fig. 6b). Taken together, using motif identification analysis, we can predict the potential regulatory role of ApiAP2 proteins based upon the transcriptional patterns of genes containing their cognate DNA-interaction motif compared to the mRNA dynamics of individual ApiAP2s (Supplementary Fig. 7A). Finally, we used the nascent transcription data to predict the possible regulatory roles for ApiAP2 proteins for which no identified cis-motif has been reported. To do this, we compared the timing and intensity of all genes enriched for the orphan AGACA motif (Fig. 6a) to the transcriptional pattern of two functionally uncharacterized ApiAP2 protein-coding genes: pf3D7_1139300 and pf3D7_0404100. The peak timing for both of these apiap2 genes occurs just prior to a decrease in transcription of genes enriched with the AGACA motif (Supplementary Fig. 7B) (as seen for AP2-G2 (Fig. 6c)).

Discussion

Cellular RNA levels are subject to extensive regulation involving alterations in the rates of RNA synthesis (transcription), processing (e.g., splicing, polyadenylation, transport), stabilization, and decay. Data on gene expression levels in P. falciparum obtained by microarray or RNA-seq reflect steady-state quantities of intracellular RNA, which are a product of the rate of both genome-wide transcription and stabilization. However, a major constraint of these analyses is the inability of these methods to differentiate between transcription and stabilization, which contribute to global abundance profiles throughout the IDC. As we have demonstrated, this constraint can be overcome by 4-TU biosynthetic labeling of newly transcribed mRNA providing direct isolation and analysis of nascent transcripts with minimal perturbations and toxic effects to the cell.

Here, we genetically modified P. falciparum to constitutively express the pyrimidine salvage pathway by single-copy genome integration, enabling the rapid incorporation of 4-TU into nascent mRNA transcripts. Using these parasites, we labeled and captured nascent transcription and mRNA stability and determined their contribution to the total abundance profiles of each gene throughout intraerythrocytic development. Earlier transcriptome studies captured timing and mRNA abundance of each gene by DNA microarray analysis, revealing a periodicity to gene expression during the 48 h IDC. Here, we demonstrate that the periodicity of each gene is defined by contributions from both transcription and stabilization. However, mRNA regulation is also time-dependent and gene-specific. Given the complexities of genome-wide mRNA dynamics, we examined the contribution of mRNA regulation to gene expression as the parasite progresses through the IDC revealing that transcription and stabilization do not contribute equally to the patterns of abundance observed for a given gene. Our results identify active transcription throughout the IDC (Fig. 1c), in contrast to previous reports, which suggest that there is minimal transcriptional activity during the ring and schizont stages of development. We also find that there are distinct periods of transcriptional and post-transcriptional regulatory bursts, which coincide with breaks in protein production at the late-ring and trophozoite stages of development, supporting the suggestion that replenishment of proteins at these stages is regulated at the level of gene expression. Our data also support previous studies, which suggest that the half-life of the transcripts increases during the intraerythrocytic developmental cycle. However, this slow rate of decay applies to less than 25 genes and the majority of genes undergo a fairly consistent rate of decay throughout the IDC. Therefore, the degradation machinery is consistently active throughout the IDC and it is more likely that post-transcriptional regulation leads to stabilization of this small subset of genes.

Beyond the examination of genome-wide P. falciparum mRNA dynamics throughout the IDC, these global transcriptional dynamics provide gene-specific profiles of nascent transcription and stabilization. From this, we distinguished regulatory dynamics of genes with similar patterns of abundance and determine mutually exclusive expression of single genes within clonally variant families. Co-transcription and co-stabilization of genes provide a useful resource for the detection of cis-acting regulatory motifs and, the prediction of motifs with associated trans-factors. These orphan motifs can now be used to identify new DNA-binding proteins, which remain uncharacterized. Our data provides a comprehensive resource that will enable further investigation into the gene regulatory systems and mRNA metabolism of the malaria parasite.

Future studies will need to clarify the mechanisms regulating genome-wide rates of decay by evaluating the RNA-interacting proteins that regulate transcript stability throughout the IDC. By combining biosynthetic mRNA labeling and capture with cross-linking and mass spectrometry, a complete RNA–protein interactome of developmentally regulated transcripts can now be achieved. Additional future applications of this method will also focus on measuring immediate transcriptional responses to environmental or chemotherapeutic treatments. This data set not only provides the malaria research community with measurements of the RNA dynamics for all individual genes throughout the IDC, but will also serve to enhance the bioinformatics analysis of transcription models, co-expression networks, and will aid in defining the timing of biological processes.

Methods

Transgene construction. The open reading frame of the yeast FCU gene was PCR amplified from P. falciparum vector pCC1, and cloned as a translational fusion into the unique AvrII and BsiWI sites of the pLN-ENR-GFP vector, resulting in pLN-5 cam-FCU-GFP plasmid and placing transcriptional control under the P. falciparum calmodulin promoter (pf3D7_1434200). This plasmid was transformed into, replicated in, and isolated from DH5α Escherichia coli for transfection into P. falciparum 3D7attBα.

Strains and culture maintenance. P. falciparum strain 3D7attBα was obtained from BEI Resources (cat # MRA-845) as a mycoplasma-free frozen stock and has been described in previous studies. Parasite cultures were maintained under standard conditions at 5% hematocrit of O. volvulus blood, 95% relative humidity, and 5% carbon dioxide (Biological Specialty Corporation) in RPMI1640 containing hypoxanthine, NaHCO3, HEPES, and glutamine and 5% AlbuMAX II (Life Technologies). The 3D7attBα line has been genetically altered to contain an acceptor attR site (from Mycobacterium smegmatis) into the non-essential ceg5 locus. This recombinant locus is maintained by continuous selection with 2.5 mM WR99210 (Jacobus Pharmaceuticals, Princeton, NJ).

Generation of genetically modified parasite lines. Transfection of P. falciparum strain 3D7 (BEI Resources, Cat # MRA-102) was performed as previously described. Briefly, 5–7% ring-stage parasite cultures were washed three times with
times the pellet volume of cytotax. The parasitized RBC pellet was resuspended to 50% hematocrit in cytotax. In preparation for transfection, 100 µg of both P.l. S-gen and the plasmid, expression vector containing the 3xHA-tag that mediates attP and attB recombination, were precipitated and resuspended in 100 µL cytotax. The plasmid and 250 µL of the 50% parasitized RBC suspension were combined and transferred to a 0.2 cm electroporation cuvette on ice. Electroporation was carried out using a BioRad GenePulser set at 0.31 kV, 960 µF. The electroporated cell suspensions were immediately transferred to a 2-mL Eppendorf tube on ice for 20 min with rotation. The RNA-bead slurry was placed on a magnetic stand for 1 min, and liquid carefully removed and saved for RNA precipitation. This sample contained RNA that was not thioltated or biotinylated. Then the RNA-bound beads underwent four rounds of stringent washes with buffer containing 0.5 M NaCl, 50 mM Tris–HCl (pH 7.5), 500 µM EDTA in DEPC-treated water. 4-TU-labeled biotinylated RNA was eluted from the beads with 5% 2-mercaptoethanol incubated for 10 min at room temperature, placed back on the magnetic stand, liquid removed and saved for RNA precipitation. The RNA in this fraction contained mRNA transcribed in the presence of 4-TU from FCU-GFP expressing cells.

Reverse transcription and Agilent DNA microarray analysis. Total RNA, unbound, and bound fractions of mRNA are precipitated by traditional methods (0.1 vol NaCl, 0.1 vol linear acrylamide and 3x volume 100% EtOH) followed by suspension in 20 µL DEPC-treated water. The concentration of each sample is determined by NanoDrop ND-1000. Starting with 1.0–2.5 µg of RNA, single-strand DNA is synthesized using M-MLV reverse transcriptase supplied with random hexamers (Promega). Incorporation of 4-TU and biotinylation was determined by NanoDrop and confirmed using Dynabeads MyOne™ Streptavidin C1 magnetic beads (Life Technologies) at a concentration of 2 µg/µL of RNA. Beads were washed as per manufacturers protocol to remove any RNases and blocked with 16 µg of yeast tRNA (Life Technologies). 4-TU-labeled biotinylated RNA was eluted from the beads using Dynabeads MyOne™ Streptavidin C1 magnetic beads and fitc-derivatized RNA was collected from each bead suspension for application to Agilent DNA microarrays. The RNA preparation was incubated on ice with 350 µg tRNA (Life Technologies) and 300 µL 1 M Tris–HCl (pH 7.5) for 1 min, washed as per manufacturers protocol to remove any RNases and blocked with 10% normal goat serum and 3% BSA. To detect FCU-GFP localization, the membrane was incubated with a 1:1000 dilution of mouse anti-GFP antibody (Thermo Fisher Scientific) followed by secondary antibody AlexaFlour 488 rat anti-mouse (Molecular Probes) diluted to 1:500. Cells were resuspended in Slowfade Antifade reagent with DAPI to stain parasite nuclei (Thermo Fisher Scientific) and mounted on slides using Fluoromount-GTM (Southern Biotech). Fluorescent microscopic images were obtained with an Olympus BX61 system and deconvoluted using SlideBook 5.0 software (Intelligent Imaging Innovations).

Assessment of FCU-GFP transgene expression and localization. Western blot analysis of FCU-GFP protein expression was carried out on mixed stage transgenic P. falciparum (10% parasitemia, 5% hematocrit), isolated by saponin (0.01%) lysis. Protein extracts from both 3D7attB and 3D7attBUFCU-GFP (clone A2) transgenic parasites were run on an SDS-10% polyacrylamide gel and transferred to a nitrocellulose membrane. The membrane was blocked with 5% milk for 30 min. To detect protein, the membrane was exposed to sheep anti-cytochrome c oxidase primary antibody (Thermo Fisher Scientific, cat # PA1-85363) diluted 1:500 in 3% bovine serum albumin (BSA) and incubated overnight at 4 °C followed by a 1 h incubation with donkey anti-sheep HRP secondary antibody (Thermo Fisher Scientific, cat# A16401) diluted 1:1000 in 3% BSA. The membrane was incubated in 1 min in Pierce® ECL-reagent (Thermo Fisher Scientific) and protein detected by exposure to film.

Verification of in vivo FCU-GFP protein expression by IFA. FCU-GFP expression in transgenic parasite lines was verified by immunofluorescence imaging analysis. Mix stage transgenic P. falciparum-infected erythrocytes were fixed with 4% formaldehyde/0.0075% glutaraldehyde (Polysciences, Inc) using the method of Tonkin et al.[9]. Fixed cells were permeabilized with 0.1% Triton-X 100 (Sigma-Aldrich) and blocked with 10% normal goat serum and 3% BSA. To detect FCU-GFP localization, the suspension was incubated with a 1:1000 dilution of a mouse anti-GFP antibody (Thermo Fisher Scientific) followed by secondary antibody AlexaFlour 488 rat anti-mouse (Molecular Probes) diluted to 1:500. Cells were resuspended in Slowfade Antifade reagent with DAPI to stain parasite nuclei (Thermo Fisher Scientific) and mounted on slides using Fluoromount-GTM (Southern Biotech). Fluorescent microscopic images were obtained with an Olympus BX61 system and deconvoluted using SlideBook 5.0 software (Intelligent Imaging Innovations).

Biosynthetic modification and isolation of labeled mRNA. Briefly, transgenic P. falciparum was incubated with 4-TU for 7 days and total RNA was extracted as described above. All RNA was subjected to the biosynthetic modification and magnetic separation as previously published with a few adjustments[16][17]. Specifically, 80 µg of total RNA (at a concentration of 0.4 µg/µL) was incubated at room temperature protected from light for 3 h in the presence of 160 µL of 1 mg/mL solution of EZ-link Biotin-HDPP (Thermo Fisher Scientific). Biotinylated total RNA was precipitated and resuspended in DEPC-treated water to a final concentration of 0.5 µg/mL. Incorporation of 4-TU and biotinylatation was determined by NanoDrop and confirmed using Dynabeads MyOne™ Streptavidin C1 magnetic beads (Life Technologies) at a concentration of 2 µg/µL of RNA. Beads were washed as per manufacturers protocol to remove any RNases and blocked with 16 µg of yeast tRNA (Life Technologies). 4-TU-labeled biotinylated RNA was eluted from the Dynabeads MyOne™ Streptavidin C1 magnetic beads using the available Lombr enclave package for R. The resulting peak times for each gene can be found in Supplementary Data 1. These peak times were used to order genes based on peak time of total abundance and to bin genes based on their peak times of transcription and stabilization.
Rates of transcription for each gene were determined based upon the slope of the calculated expression values from the start of transcription to the adjacent peak. Conversely, rates of decay were calculated by determining the slope of the expression values of unlabeled RNA (stabilized fraction) from the peak to the following trough. Calculated rates of transcription and decay for each gene were then binned into groups based upon the peak timing of nascent transcription or stabilization of that gene (Supplementary Data 2). These groups represent six major morphological stage transitions: early ring (0–10 hpi), mid-ring (11–15 hpi), late ring/early trophozoite (16–21 hpi), mid-trophozoite (22–26 hpi), late trophozoite (27–32 hpi), and schizont (33–48 hpi). The calculated rates are plotted as a box-whisker plot with Tukey’s Biweight distribution and median represented.

Motif and GO-term enrichment analysis. We identified motifs enriched in the 5′ UTRs of genes co-transcribed throughout the IDC using a regulatory element discovery algorithm (FIRE) and these results are represented as a heatmap. The activity profile of each motif was used to identify a refined list of target genes that were both enriched for the motif and similarly transcribed. Gene lists identified by FIRE analysis to be enriched for either the AP2-1 or AP2-2 motif were compared to the previously published target gene listed identified by ChiP-seq.

GO-term enrichment of select gene groups or clusters was carried out using the Analysis Tool at http://www.PlasmidDB.org and terms that are significantly enriched (Fisher’s exact test, Bonferroni corrected p-value ≤ 0.05) are summarized in Supplementary Data 3.

Data availability. DNA microarray data that was generated for this study are available on Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo) under accession number GSE66669 and on fighshare with the identifier https://doi.org/10.6084/m9.figshare.6200792.v1. The authors declare that all other data supporting the findings of this study are available within the article and its Supplementary Information files, or are available from the authors upon request.

Received: 14 February 2018 Accepted: 31 May 2018
Published online: 09 July 2018

References
1. WHO. World Malaria Report 2016 (World Health Organization, Geneva, 2016)
2. Gardner, M. J. et al. Genome sequence of the human malaria parasite Plasmodium falciparum. Nature 419, 486–511 (2002).
3. Bannister, L. H., Hopkins, J. M., Fowler, R. E., Krishna, S. & Mitchell, G. H. A brief illustrated guide to the ultrastructure of Plasmodium falciparum asexual blood stages. Parasitol. Today 16, 427–433 (2000).
4. Josling, G. A. & Llinas, M. Sexual development in Plasmodium parasites: knowing when it’s time to commit. Nat. Rev. Microbiol. 13, 573–586 (2015).
5. Le Roch, K. G. et al. Discovery of gene function by expression profiling of the malaria parasite life cycle. Science 301, 1503–1508 (2003).
6. Tarun, A. S. et al. A combined transcriptome and proteome survey of malaria parasite liver stages. Proc. Natl. Acad. Sci. U.S.A. 105, 305–310 (2008).
7. Lasonder, E. et al. Integrated transcriptomic and proteomic analyses of P. falciparum gametocytes: molecular insight into sex-specific processes and translational repression. Nucleic Acids Res. 44, 6087–6101 (2016).
8. Young, J. A. et al. The Plasmodium falciparum sexual development transcriptome: a microarray analysis using ontology-based pattern identification. Mol. Biochem. Parasitol. 143, 67–79 (2005).
9. Lopez-Barragan, M. J. et al. Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum. BMC Genomics 12, 587 (2011).
10. Bozdech, Z. et al. The transcriptome of the intraerythrocytic developmental cycle of Plasmodium falciparum. PLoS Biol. 1, E5 (2003).
11. Llinas, M., Bozdech, Z., Wong, E. D., Adai, A. T. & DeRisi, J. L. Comparative whole genome transcriptome analysis of three Plasmodium falciparum strains. Nucleic Acids Res. 34, 1166–1173 (2006).
12. Painter, H. J., Campbell, T. L. & Llinas, M. The Apicomplexan AP2 family: integral factors regulating Plasmodium development. Mol. Biochem. Parasitol. 176, 1–7 (2011).
13. Hughes, K. R., Philip, N., Starnes, G. L., Taylor, S. & Waters, A. P. From cradle to grave: RNA biology in malaria parasites. Wiley Interdiscip. Rev. RNA 1, 287–303 (2010).
14. Balaji, S., Babu, M. M., Iyer, I. M. & Aravind, L. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. Nucleic Acids Res. 33, 3994–4006 (2005).
15. Horrocks, P., Wong, E., Russell, K. & Emes, R. D. Control of gene expression in Plasmodium falciparum—ten years on. Mol. Biochem. Parasitol. 164, 9–25 (2009).
46. Cleary, M. D., Meiering, C. D., Jan, E., Guymon, R. & Boothroyd, J. C. Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay. Nat. Biotechnol. 23, 232–237 (2005).

47. Gay, L., Karalis, K. V., Miller, M. R., Doe, C. Q. & Stankunas, K. Applying thiouracil tagging to mouse transcriptome analysis. Nat. Protoc. 9, 410–420 (2014).

48. Kunzmann, M. et al. Microarray analysis of newly synthesized RNA in cells and animals. Proc. Natl. Acad. Sci. U.S.A. 104, 6164–6169 (2007).

49. Windhager, L. et al. Ultra short and progressive 4sU-tagging reveals key characteristics of RNA processing at nucleotide resolution. Genome Res. 22, 2031–2042 (2012).

50. Dolken, L. et al. High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. RNA 14, 1959–1972 (2008).

51. Hyde, J. E. Targeting purine and pyrimidine metabolism in human apicomplexan parasites. Curr. Drug Targets 8, 31–47 (2007).

52. Nkumah, L. J. et al. Efficient site-specific integration in Plasmodium falciparum chromosomes mediated by mycobacteriophage Bba1 integrase. Nat. Methods 3, 615–621 (2006).

53. Sangurdekar, D. P. Ruts: R normalization and inference of time series data v1.2.0. (The R Foundation, 2014).

54. Hardin, J., Mitani, A., Hicks, L. & VanKoten, B. A robust measure of correlation between two genes on a microarray. BMC Bioinformatics 8, 220 (2007).

55. Tukey, J. Data analysis, computation, and mathematics. Q. Appl. Math. 30, 51–65 (1972).

56. AURRECOECHA, C. et al. PlasmoDB: a functional genomic database for malaria parasites. Nucleic Acids Res. 37, D539–D543 (2009).

57. Preiser, P. R. et al. Recombination associated with replication of malarial mitochondrion DNA. EMBO J. 15, 684–693 (1996).

58. Williamson, D. H. et al. The plastid DNA of the malaria parasite Plasmodium falciparum is replicated by two mechanisms. Mol. Microbiol. 45, 533–542 (2002).

59. Foth, B. J. et al. Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum. Science 299, 705–708 (2003).

60. Chan, J. A., Foecken, J. F. B. & Beeson, J. G. Surface antigens of Plasmodium falciparum-infected erythrocytes as immune targets and malaria vaccine candidates. Cell. Mol. Life Sci. 71, 3633–3657 (2014).

61. Howitt, C. A. et al. Clonally variant gene families in Plasmodium falciparum share a common activation factor. Mol. Microbiol. 73, 1171–1185 (2009).

62. Bovire-Graeils, N. et al. Transcriptional variation in the malaria parasite Plasmodium falciparum. Genome Res. 22, 925–938 (2012).

63. Counihan, N. A., Kalanon, M., Coppel, R. L. & de Koning-Ward, T. F. Plasmodium rhoptry proteins: why order is important. Trends Parasitol. 29, 228–236 (2013).

64. Cowman, A. F., Tonkin, C. J., Tham, W. H. & Duraisingham, M. T. The molecular basis of erythrocyte invasion by malaria parasites. Cell Host Microbe 22, 232–245 (2017).

65. Sherling, E. S. et al. The Plasmodium falciparum rhoptry protein RhopH3 plays essential roles in host cell invasion and nutrient uptake. eLife 6, e23239 (2017).

66. Counihan, N. A. et al. Plasmodium falciparum parasites deploy RhopH2 into the host erythrocyte to obtain nutrients, grow and replicate. eLife 6, e23217 (2017).

67. Ghosh, S. et al. The Plasmodium rhoptry associated protein complex is important for parasitophorous vacuole membrane structure and intraerythrocytic parasite growth. Cell. Microbiol. 19, e12733 (2017).

68. Beeson, J. G. et al. Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria. FEBS Microbiol. Rev. 40, 343–372 (2016).

69. Elemento, O., Slonim, N. & Tavazoie, S. A universal framework for regulatory element discovery across all genomes and data types. Mol. Cell 28, 337–350 (2007).

70. Campbell, T. L., De Silva, E. K., Olsewski, K. L., Elemento, O. & Linas, M. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. PLoS Pathog. 6, e1001165 (2010).

71. Santos, J. M. et al. Red blood cell invasion by the malaria parasite is coordinated by the PAP2-I transcription factor. Cell Host Microbe 21, 731–741 (2017).

72. Yuda, M., Iwanaga, S., Kaneko, I. & Kato, T. Global transcriptional repression: an initial and essential step for Plasmodium sexual development. Proc. Natl. Acad. Sci. U.S.A. 112, 12824–12829 (2015).

73. Balza, A. G. et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell 46, 674–690 (2012).

74. Castello, A. et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. Cell 149, 1393–1406 (2012).

75. Hafner, M. et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141, 129–141 (2010).

76. Maier, A. G., Braks, J. A., Waters, A. P. & Cowman, A. F. Negative selection using yeast cytosine deaminase/uracil phosphoribosyl transferase in Plasmodium falciparum for targeted gene deletion by double crossover recombination. Mol. Biochem. Parasitol. 150, 118–121 (2006).

77. Trager, W. & Jensen, B. Human malaria parasites in continuous culture. Science 193, 673–675 (1976).

78. Fidock, D. A. & Wellens, T. E. Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil. Proc. Natl. Acad. Sci. U.S.A. 94, 10931–10936 (1997).

79. Tonkin et al. Localization of organellar proteins in Plasmodium falciparum using a novel set of transfection vectors and a new immunofluorescence fixation method. Mol. Biochem. Parasitol. 137, 13–21 (2004).

80. Haynes, J. D. & Moeh, J. K. Automated synchronization of Plasmodium falciparum parasites by culture in a temperature-cycling incubator. Methods Mol. Med. 72, 489–497 (2002).

81. Glynn, E. F., Chen, J. & Mushegian, A. R. Detecting periodic patterns in unevenly spaced gene expression time series using Lomb-Scargle periodograms. Bioinformatics 22, 310–316 (2006).

Acknowledgments

This work was funded through support from NIH R21 AI133375. A.S. is supported through the J. Lloyd and Dorothy Foehr Huck Institutes of the Life Sciences at Penn State University. We would like to thank many people who have contributed to this project including Simon Cobbold, Bjorn F.C. Kafsack, Jess O’Hara, Lindsey Orchard, Yoanna Pompalova, Joana Santos, April Williams, and the Genomics Core Facility at Princeton University and its staff, especially Donna Storton and Jessica Buckles Wiggins. We also thank Omar Harb, Brian Brunk, and Shon Cadre at http://www.PlasmoDB.org for making this data publicly available ahead of publication. Lastly, we thank Erik Allman, Gabrielle Josling, Scott Lindner, and Timothy Russell for critical reading of the manuscript.

Author contributions

H.J.P. and M.L. designed the experiments. H.J.P. performed experiments. H.J.P. and N.C.C. analyzed the data. H.J.P. and M.L. wrote the manuscript and generated figures. N.C.C. and J.D.S. created and implemented the statistical model used for data analysis. A.S. and L.A. contributed to data analysis and public availability for hosting the data. M.L. and J.D.S. provided critical comments and project oversight.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-018-04966-3.

Competing interests: The authors declare no competing interests.

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

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