What functional genomics has taught us about transcriptional regulation in malaria parasites

Christa G. Toenhake and Richárd Bártfai

Abstract
Malaria parasites are characterized by a complex life cycle that is accompanied by dynamic gene expression patterns. The factors and mechanisms that regulate gene expression in these parasites have been searched for even before the advent of next generation sequencing technologies. Functional genomics approaches have substantially boosted this area of research and have yielded significant insights into the interplay between epigenetic, transcriptional and post-transcriptional mechanisms. Recently, considerable progress has been made in identifying sequence-specific transcription factors and DNA-encoded regulatory elements. Here, we review the insights obtained from these efforts including the characterization of core promoters, the involvement of sequence-specific transcription factors in life cycle progression and the mapping of gene regulatory elements. Furthermore, we discuss recent developments in the field of functional genomics and how they might contribute to further characterization of this complex gene regulatory network.

Key words: Malaria; Plasmodium; transcription factors; regulatory sequences; gene expression

Introduction
Eukaryotic, unicellular parasites of the Plasmodium genus are the causative agents of malaria. During their life cycle, these parasites alternate between a vertebrate and an insect host with multiple, morphologically and functionally distinct stages of development within each host (Figure 2). In the vertebrate host, development and replication occur predominantly within host cells, either a hepatocyte or an erythrocyte. In the mosquito, parasites traverse host cells and reside in extracellular spaces (e.g. midgut lumen or at the basal side of the midgut wall). The larger part of this life cycle is deterministic, including few true cell fate decision events. The decision between continued asexual replication and the formation of gametocytes is typical of all Plasmodium species, while the decision to enter and exit the dormant hypnozoite stage in hepatocytes is made only in a few Plasmodium species (e.g. Plasmodium vivax).

In eukaryotes, gene expression can be regulated at various steps before or after RNA synthesis. Regulatory mechanisms that act on the transcriptional process itself can be formally divided into (i) epigenetic mechanisms that, in simple terms, influence access or recruitment of the transcriptional machinery to the DNA [1–3] and (ii) the activity of sequence-specific transcriptional activators or repressors that interact with cis-regulatory DNA elements [4, 5]. The critical contribution of epigenetic mechanisms to Plasmodium gene expression regulation has been established (e.g. reviewed in this issue and [6–8]), for example in regulating antigenic variation genes and the expression of alternative solute transporters [8, 9]. Furthermore, expansion of heterochromatic domains during gametocytogenesis and mosquito-stage development restricts the expression of stage-specific sets of genes during developmental progression [10]. Finally, higher acetylation levels of histone tails in regulatory regions have been associated with increased gene transcription [11, 12]. On the
other hand, the contribution of DNA regulatory elements to Plasmodium gene expression regulation, and thus the involvement of sequence-specific transcription factors (TFs) has been evident from initial promoter mapping studies (primarily performed in Plasmodium falciparum), which identified DNA regions with transcription enhancing or repressing potential (reviewed in [13, 14]). These regions were, however, limited to only a small set of genes and the identity of the protein or protein complex binding to the regulatory DNA sequence could in most cases not be clarified.

Recent functional genomic approaches have led to the genome-wide characterization of regulatory DNA elements, including transcription start sites (TSSs) [15–17], and have lent support to the essential role of sequence-specific TFs, in particular apicomplexan apetala 2 (ApiAP2) factors [18–34]. In this review, we summarize recent findings on how the delicate interplay between core promoters, regulatory DNA elements and sequence-specific TFs, regulates specific sets of genes during the developmental progression of malaria parasites. Notably, the majority of studies on this topic have been performed in a limited number of Plasmodium species, i.e. human-infecting P. falciparum and rodent-infecting Plasmodium berghei and Plasmodium yoelli and findings cannot be generalized.

Core promoter recognition in the compact Plasmodium genome

At the basis of the transcriptional process lies recognition of the core promoter by general transcription factors that, together with RNA polymerase II, make up the pre-initiation complex (PIC). In most eukaryotes, core promoters have characteristic DNA sequences, nucleosome positioning and/or chromatin marking [1]. Plasmodium parasites contain a haploid set of 14 chromosomes during most of their life cycle. The 20–30 Mb genomes encode 5500–6500 protein-coding genes, which results in approximately a 50:50 ratio of coding versus non-coding DNA and rather short intergenic sequences (on average 1.4–2 Kb intergenic [35]). Notably, the genome of several Plasmodium species is among the most AT-rich nuclear genomes [36]. In particular that of P. falciparum, which has coding and subtelomeric sequences somewhat richer in G and C, but an average 87% AT bases in intergenic regions. Why and how these extremely AT-biased genomes evolved specifically in some Plasmodium species remains unclear. Nevertheless, it is likely that the transcriptional apparatus and other DNA-associated proteins have adapted to this nucleotide bias. For example, the Apicomplexan-specific H2A.Z/H2B.Z double-variant nucleosomes which, at least in P. falciparum, associate with AT-rich intergenic regions [37, 38], that could not be occupied by the canonical histone H2A and H2B containing nucleosomes [39].

The search is still ongoing for Plasmodium DNA- and chromatin-related elements that guide the PIC to the core promoter and signal the site of transcription. P. falciparum TSSs have been mapped to multiple small windows in most promoter regions [15–17], and divergent transcription initiation is highly prevalent [16, 17, 40]. Interestingly, while most genes have multiple TSS windows, some of these initiation sites are more prevalent than others [17]. Furthermore, while in most cases the windows initiate transcription simultaneously, in a small number of promoter regions these windows exhibit differential regulation during intra-erythrocytic development (3.4%) [16].

Several DNA- and chromatin-based features have been associated with TSSs in P. falciparum and may guide core promoter recognition, TSS selection and/or promoter strength. Sequence-based DNA features include a typical TA dinucleotide at position −1.0 [16, 40], which is also observed at mammalian TSSs and may partially reflect the initiator core promoter element [41]. However, whether this resemblance has any functional relevance in the AT-rich P. falciparum genome is questionable. Additional sequence-based features include GC-rich sequence elements at ∼150 bp and ∼210 bp downstream of the TSS [16], polymeric AAAAA- or TTTTTT-stretches within 50 bp upstream of the TSS [42] and a local increase of GC-content around weaker TSSs [16]. Several DNA structural features may also predict TSSs with reasonable accuracy [40] but this study made use of TSS mappings for a limited number of P. falciparum genes, and it is unclear whether the same features hold true for TSSs identified on a genome-wide scale. Lastly, while the classic core promoter element, TATA-box (TATAA), is recognized by the P. falciparum TATA-binding protein in vitro [43], the relevance of this motif for in vivo TSS selection in an AT-rich genome requires experimental validation.

Chromatin-related features associated with P. falciparum TSSs include a well-positioned nucleosome just downstream of the TSS (the so-called ‘+1 nucleosome’) [16, 17] and a nucleosome-depleted region (NDR) directly upstream of the TSS [17]. The NDR is more pronounced for highly expressed genes [16, 17] and shows dynamic nucleosome loss that correlates with transcriptional activity [17]. As nucleosome positioning can be determined by the underlying DNA sequence, these observations may reflect DNA sequence features. In particular, homopolymeric poly(dA:dT) sequences have been reported to be stiff and resistant to nucleosome formation [44] and are enriched next to TSSs in P. falciparum alongside other well-positioned nucleosomes close to the ATG start codon, stop codon and splice sites [17]. However, these tracts only partially correlate with nucleosome positioning, pointing to the involvement of adenosine triphosphate-dependent chromatin remodeling complexes [17]. Finally, typical eukaryotic histone marks associated with the core promoter of model organisms like H3K4me3, H3K9ac and histone variant H2A.Z are also found broadly covering P. falciparum intergenic regions; however, they only moderately correlate with transcriptional output [11, 16].

In conclusion, while several associations have been unveiled between DNA/chromatin features and TSSs (summarized in Figure 1), we still do not understand how the Plasmodium general TFs recognize the core promoter, nor how the PIC identifies the TSS to be used. Hence, the functional relevance and causative role of the observed features requires further experimental investigation.

Sequence-specific TFs

Sequence-specific TFs mediate the transcriptional regulation of specific sets of genes through their ability to recognize DNA motifs within gene regulatory elements and either directly or indirectly influence the recruitment and activity of PIC. The most comprehensive list of potential sequence-specific TFs have been drawn up for P. falciparum [18, 45] and consists of 27 ApiAP2s, 12 C2H2-type zinc finger (ZnF-C2H2), eight helix-turn-helix [HTH including high mobility group box 3 protein (HMGB3)], one β-scaffold factor with minor groove contacts [45] as well as the K homology (KH) domain-containing PREB (Pox regulatory element (PRE) binding protein, PF3D7_1011800/PF10_0115 [46] and a homeodomain-like TF (PF3D7_1466200/PF14_0631, Björn Kafasch personal communication). The majority of these genes have syntenic orthologues in other Plasmodium species (Table 1). However, sequence conservation is generally low and gene products might have acquired other functions during evolution.
Currently, the ApiAP2 family is considered as the principal family of TFs in Plasmodium and we therefore discuss it in a separate section. The HTH factor PfMYB1 was among the first TFs studied in Plasmodium [47, 48]. PfMYB1 can bind a putative Myb regulatory element (MRE, wAACnGh) upstream of *P. falciparum* genes and has been associated with the promoter of genes that were downregulated upon Pfmyb1 knockdown [47, 48]. PfMYB2, on the other hand, shows high homology to the pre-messenger RNA (mRNA)-splicing factor CEF1 in *Saccharomyces cerevisiae* [49]. Of the ~12 zinc finger C2H2 domain-containing proteins, only one has been studied. In *P. falciparum*, it is called PfTRZ (telomere repeat-binding ZnF protein) because of its binding to telomeric TT(T/C)AGGG repeats in vitro and in vivo. Interestingly, this same factor also binds to and regulates expression of 5S rDNA loci pointing to an evolutionary relationship with the general transcription factor TFIIA, a combination of functions that is unique among eukaryotes [50]. Of the four Plasmodium HMGB proteins, only one has a configuration linked to sequence-specific binding, while the other three are implicated in chromatin binding [45]. Lastly, PREBP is a KH domain-containing protein. Although this domain is normally linked to RNA-binding and -processing, the *P. falciparum* protein bound dsDNA in a sequence-specific manner and was implicated in the expression of peroxiredoxin genes [46].
Table 1. Overview of candidate sequence-specific TFs in *P. falciparum* and their syntenic orthologues in *P. knowlesi*, *P. vivax*, *P. berghei* and *P. yoelli*, based on [45, 96]. For *P. falciparum*, *P. berghei* and *P. yoelli*, geneIDs are color-coded based on whether the coding sequence (CDS) is mutable (*P. falciparum* [55]), or the gene product is essential in the IDC (*P. berghei* and *P. yoelli* [30, 34, 95]). Green, CDS is mutable or KO could be generated; red, CDS is not mutable and KO could not be generated; blue, KO does not show a phenotype; orange, CDS is mutable but tentative because of small CDS size; # CDS is mutable but KO could not be generated using conventional strategies [23, 32]. * The HTH factor ADA2 (Alteration/Deficiency in Activation 2) is a transcriptional coactivator and part of the GCN5-containing histone acetyltransferase-complex (GCN5, general control of amino acid synthesis 5) [97, 98]. The species in which the TF has been described, is indicated in brackets behind the TF name. † Only C2H2 domain-containing ZnFs are included in this table. Studies that investigated the function of the individual TF, are cited as well. Syntenic orthologues were retrieved from the PlasmoDB database, release 39 [35].

| DNA binding domain | Gene ID | *P. falciparum* 3D7 | Previous gene ID | *P. knowlesi* strain H | *P. vivax* P01 | *P. berghei* ANKA | *P. yoelli* yoelli 17X | TF name | Ref |
|--------------------|---------|---------------------|------------------|-----------------------|----------------|-----------------|-----------------------|---------|-----|
| ApiAP2             | PFD0200c| PKNH_0102100        | PVPO_0103400     | PRANKA_1001800        | PY17X_1003200  | AP2-SP2 (Pb, Py) | [30, 34, 55]         |
|                    | PFD01680| PKNH_1016500        | PVPO_1017200     | PRANKA_1231600        | PY17X_1235000  | AP2-O2 (Pb, Py)  | [30, 34, 55]         |
|                    | PFD0040c| PKNH_1146400        | PVPO_0114800     | PRANKA_0102900        | PY17X_0104500  | SIP2 (Pf)       | [20, 23, 30, 34, 55] |
|                    | PFD01220| PKNH_1139200        | PVPO_1138000     | PRANKA_0109500        | PY17X_0111100  |                 | [30, 34, 55]         |
|                    | PFD02290| PKNH_1127300        | PVPO_1126600     | PRANKA_1121800        | PY17X_1123200  |                 | [30, 33, 34, 55]    |
|                    | PFD07300| PKNH_0215200        | PVPO_0216000     | PRANKA_0214900        | PY17X_0215800  |                 | [25, 30, 55]        |
|                    | PFD01139| PKNH_0937300        | PVPO_0940100     | PRANKA_0909600        | PY17X_0911000  |                 | [30, 34, 55]        |
|                    | PF11_0404| PKNH_1441500        | PVPO_1440600     | not found              | not found      |                 | [55]                 |
|                    | PF12240| PPL1075w             | PKNH_1441700     | PVPO_1440800          | PRANKA_1437500 | AP2-G (Pb, Pf, Py)| [26, 27, 30, 34, 55, 58, 59] |
|                    | PF12260| PPL1085w             | PKNH_1441700     | PVPO_1406100          | PRANKA_1405400 |                 | [30, 34, 55]        |
|                    | PF13052| PKNH_1417900        | PVPO_1418100     | PRANKA_1415700        | PY17X_1417400  | AP2-G3 (Pf)     | [30, 34, 55]        |
|                    | PF13172| PKNH_1405400        | PVPO_1406100     | PRANKA_1405700        | PY17X_1405400  |                 | [30, 34, 55]        |
|                    | PF13290| PKNH_1258500        | PVPO_1211900     | PRANKA_1365000        | PY17X_1361700  |                 | [30, 34, 55]        |
|                    | PF13082| PKNH_1350200        | PVPO_1340800     | PRANKA_1304300        | PY17X_1306700  | AP2-G2 (Pb, Py) | [27, 29, 30, 34, 55] |
|                    | PF14292| PKNH_1329500        | PVPO_1320200     | PRANKA_1015500        | PY17X_1017000  | AP2-O3 (Pb, Py) | [30, 34, 55]        |
|                    | PF14560| PKNH_1223500        | PVPO_1250900     | PRANKA_1319700        | PY17X_1323500  |                 | [30, 34, 55]        |
|                    | PF14640| PKNH_1215100        | PVPO_1240400     | PRANKA_1329800        | PY17X_1334500  | AP2-EXP (Py), AP2-SP (Pb)| [20, 22, 30, 32, 34, 55] |
|                    | PF16655| PKNH_0733100        | PVPO_0731100     | PRANKA_0835200        | PY17X_0835800  |                 | [30, 34, 55]        |
|                    | PF13267| PKNH_1234300        | PVPO_1203900     | PRANKA_1363700        | PY17X_1369400  |                 | [30, 34, 55]        |
|                    | PF0985w| PKNH_0513100        | PVPO_0525800     | PRANKA_0521700        | PY17X_0523100  |                 | [30, 34, 55]        |
|                    | PF06310| PKNH_1138000        | PVPO_1135330     | PRANKA_0112100        | PY17X_0113700  |                 | [30, 34, 55]        |
|                    | PF08010| MAL8P1.153          | not found         | not found              | PRANKA_1228100 |                 | [30, 34, 55]        |
|                    | PF10077| PKNH_0806500        | PVPO_0807400     | PRANKA_1205900        | PY17X_1209100  | AP2-1 (Pf)      | [30, 31, 34, 55]    |
|                    | PF11780| PKNH_0905400        | PVPO_0908600     | PRANKA_0939100        | PY17X_0941600  |                 | [30, 34, 55]        |
|                    | PF11550| PKNH_01613           | PVPO_0191300     | PRANKA_0932300        | PRANKA_0934300 |                 | [30, 34, 55]        |
|                    | PF11341| PKNH_0941000        | PVPO_0943800     | PRANKA_0905900        | PY17X_0907300  | AP2-0 (Pb, Py)  | [28, 28, 30, 34, 55] |
|                    | PF12392| PKNH_1459000        | PVPO_1457300     | PRANKA_1453700        | PY17X_1456200  |                 | [30, 34, 55]        |
|                    | PF14495| PKNH_1232600        | PVPO_1257500     | PRANKA_1313200        | PY17X_1317000  |                 | [34, 55]            |

(Continued)
| DNA binding domain | Gene ID   | P. falciparum 3D7 Previous gene ID | P. knoewlesi strain H | P. vivax PO1 | P. berghei ANKA | P. yoelii yoelii 17X | TF name | Ref |
|-------------------|-----------|------------------------------------|-----------------------|--------------|-----------------|------------------------|---------|-----|
| HTH               | PF3D7_1123000 | PF11_0241 PKNH_0920800 | PVP01_0923600 | PBANKA_0925300 | PY17X_0927300 | [55] |
| PF3D7_1205800     | PF10_0290w  | PKNH_1305500 | PVP01_1304900 | PBANKA_0604500 | PY17X_0607000 | HMG3 | [55] |
| PF3D7_1225200     | PF11_0631   | not found | not found | not found | not found | [55] |
| PF3D7_1215900     | PFL0815w    | PKNH_1436500 | PVP01_1436700 | PBANKA_1432500 | PY17X_1434900 | [55, 95] |
| PF3D7_1225200     | PFL1215c    | PKNH_1444600 | PVP01_1443400 | PBANKA_1440100 | PY17X_1442600 | [55, 95] |
| PF3D7_1309800     | PF13_0054   | PKNH_1410200 | PVP01_1410800 | PBANKA_1408300 | PY17X_1410100 | [55, 95] |
| PF3D7_1315800     | PF13_0088   | PKNH_1415600 | PVP01_1416700 | PBANKA_1414300 | PY17X_1416000 | MYB1 (Pf) | [47, 48, 55] |
| PF3D7_1014600     | PF10_0143   | PKNH_0814700 | PVP01_0814700 | PBANKA_1213000 | PY17X_1216200 | ADA2* | [55, 95, 97, 87] |
| KH-domain         | PF3D7_1011800 | PF11_0115 PKNH_0811600 | PVP01_0811900 | PBANKA_1215200 | PY17X_1213400 | PREBP (Pf) | [46, 55, 95] |
| Znf-C2H2²         | PF3D7_0315800 | PFCO0690c PKNH_0826300 | PVP01_0824300 | PBANKA_0413600 | PY17X_0416400 | [55, 95] |
| PF3D7_0407600     | PFD0375w    | not found | not found | not found | not found | [55, 95] |
| PF3D7_1009400     | PF10_0091   | PKNH_0809100 | PVP01_0809400 | PBANKA_1207700 | PY17X_1210900 | ZNF2 | [55, 95] |
| PF3D7_1209100     | PFL0455c    | not found | not found | not found | not found | [95] |
| PF3D7_1209300     | PFL0465c    | PKNH_1309200 | PVP01_1308400 | PBANKA_0607700 | PY17X_0610200 | TRZ (Pf) | [50, 55, 95] |
| PF3D7_1450400     | PF14_0479   | PKNH_1231700 | PVP01_1256600 | PBANKA_1514100 | PY17X_1317900 | [95] |
| PF3D7_1463900     | PF14_0657   | PKNH_1212300 | PVP01_1237900 | PBANKA_1332100 | PY17X_1336800 | [95] |
| PF3D7_1473900     | PF14_0703   | PKNH_1207200 | PVP01_1223000 | PBANKA_1333100 | PY17X_1347100 | [95] |
| PF3D7_1408500     | PFD0455w    | PKNH_0307900 | PVP01_0307900 | PBANKA_1007400 | PY17X_1008800 | [95] |
| PF3D7_1243100     | PFL2075c    | PKNH_1462400 | PVP01_1460200 | PBANKA_1456500 | PY17X_1459000 | [95] |
| PF3D7_1458600     | PF14_0559   | PKNH_1223200 | PVP01_1248300 | PBANKA_1322300 | PY17X_1326100 | [95, 95] |
| PF3D7_1464400     | PF14_0612   | PKNH_1217100 | PVP01_1242400 | PBANKA_1327800 | PY17X_1332500 | [95] |
| β-scaffold        | PF3D7_0109600 | PFA0470c PKNH_0204200 | PVP01_0205500 | PBANKA_0204000 | PY17X_0205400 | [95] |
What functional genomics has taught us about transcriptional regulation in malaria parasites

Figure 2. Overview of the *Plasmodium* life cycle and the corresponding AP2 factors that are essential for the development of individual stages (boxed), as identified in *P. falciparum* (pink dot), *P. berghei* (light blue dot) and/or *P. yoelli* (dark blue dot). Note that gametocyte stages reflect the morphology of *P. falciparum* gametocytes only. Data were derived from references [19, 22, 23, 25–27, 29–34].

**ApiAP2 factors and Plasmodium life cycle progression**

The *Plasmodium* AP2 family, the principal TF family in apicomplexan parasites, is homologous to the plant *apetala 2*/ethylene response factor TF family [18]. Due to its absence in humans, it has been proposed as a potential anti-malarial drug target [20, 21, 27] and several members may mediate drug resistance in *Plasmodium* [51]. The 60-aa globular AP2 domain has a conserved core consisting of three $\beta$-sheets, which make base-specific contacts with the DNA and a stabilizing C-terminal $\alpha$-helix [21]. A cysteine residue between the first two $\beta$-sheets can facilitate dimerization [21], thereby providing a means for the proposed combinatorial mode of regulation [24, 52]. *Plasmodium* AP2 proteins have one to three AP2 domains per protein, and some encode the accessory C-terminal domain, the function of which is not yet understood (ACDC, AP2-coincident C-terminal domain; PFAM id: PF14733 [53]). In vitro, the different AP2 domains show diverse DNA-binding preferences [24], although it is important to note that not all AP2 domains have to be involved in DNA-binding in vivo. PfAP2-I, for example, requires only its third AP2 domain to interact with the DNA, at least in blood stage parasites [31]. The opposite is true for PfSIP2 (SPE2-interacting protein, PF3D7_0604100/PFF0200c) where both AP2 domains are required for binding bona fide SPE2 elements (subtelomeric var promoter element 2), at least in vitro [23]. Additionally, post-translational modifications can alter the DNA-binding ability of AP2 domains in vivo. For example, lysine acetylation, which is prevalent among *P. falciparum* AP2 proteins, diminishes the DNA-binding ability of PfAP2-I [54]. To decipher their contribution to *Plasmodium* life cycle progression, these DNA-binding proteins have been studied on a per gene basis [19, 22, 23, 25–29, 31–33], as well as, more recently, in systematic knock-out (KO) screens [30, 34]. Several AP2s have now been assigned names based on the stage of the developmental defect in the respective KO line (Figure 2, Table 1).

Almost half of the 27–28 AP2 factors are essential to the intra-erythrocytic developmental cycle (IDC). The number of so-called ‘essential’ AP2s varies slightly between species—11 in *P. falciparum*, 12 in *P. yoelli* and 14 in *P. berghei* (Table 1).
[30, 34, 55]—and does not show a perfect overlap, especially between *P. falciparum* and the rodent species. Whether this is due to technical factors (e.g., culturing conditions, KO strategy used) or whether it points to a considerable evolutionary rewiring of the transcription regulation program in the IDC, is unclear. As their essential nature in the IDC requires the use of techniques other than phenotyping the respective KO line, four of them have been studied in detail in *P. falciparum*: *Pfsip2*, *Pfap2-1*, *Pfap2tel* and *Pfap2-exp*. *PSIP2* recognizes SPE2 repeats within *ups8*-var promoters and in telomere-associated sequences and has been implicated in *ups8*-type var gene silencing and heterochromatin biology in *P. falciparum* [23]. *PfAP2-1* has a critical role in the expression of invasion-related genes in merozoites [31]. Besides silencing [10, 26, 60, 61], and this epigenetic control is key to understanding heterochromatin protein 1 (HP1)-mediated epigenetic control [26]. Interestingly, in all species studied so far, involving two other AP2 factors in gametocyte development, *PfAP2-1* also targets promoters of nucleosome- and chromatin-related genes (including seven *Pfap2* genes), cell-cycle-related genes, and genes associated with vesicle transport and host-cell remodeling [31, 56]. In addition, two *P. falciparum* studies have suggested roles for AP2 factors in the IDC that were not obvious from rodent studies. The orthologue of *AP2-SP* in *P. falciparum* is essential for intra-erythrocytic development and is called *PfAP2-EXP* because of its involvement in the proper expression of multi-gene family genes [32]. Similarly, the orthologue of *AP2-SP3* in *P. falciparum* has been named *PfAP2Tel* as it recognises telomere GCCGGT(C/G) repeats in *in vitro* and *in vivo*, pointing to a role in telomere maintenance [33]. If this role were conserved in rodent-infecting species, an accompanying growth defect would be expected in the corresponding KO line, but this was never observed [30, 34]. Besides these ‘essential’ AP2s, the deletion of several *Pbap2* genes (*Pbap2-g2, Pbap2-o, Pbap2-sp, Pbap2-l*), while lethal for the development of mosquito- or liver-stage parasites, causes transcriptional deregulation and a growth delay in asexual stages [30]. Thus, it is clear that a considerable number of AP2s are involved in *Plasmodium* blood-stage development, due to the difficulty of manipulating genes essential for the IDC, as well as the use of different techniques in different species, it is difficult to determine to what extent functions are conserved. Furthermore, alternative strategies are needed to determine the exact targets of the essential AP2 factors and their contribution to *Plasmodium* blood-stage development.

The differentiation to gametocytes represents the sole developmental switch in the predominantly deterministic life cycle of *Plasmodium* parasites. The epigenetic, transcriptional and post-transcriptional factors involved in this pathway have recently been reviewed elsewhere [57]. Despite the existence of two possible routes for gametocyte commitment [58, 59] and despite different trajectories of gametocytogenesis among *Plasmodium* species [57], extensive forward and reverse genetic approaches have identified a conserved AP2 as the principal regulator of this switch: AP2-G (PBANKA_1437500 [27] and PF3D7_1222400 [26]). Interestingly, in all *Plasmodium* species studied so far, *ap2-g* is the only single-locus heterochromatic gene that is under heterochromatin protein 1 (HP1)-mediated epigenetic silencing [10, 26, 60, 61], and this epigenetic control is key to the regulated expression of *ap2-g* and gametocyte conversion [62]. In the current model, *PGDV1* interacts with *PHHP1* and this binding is associated with *PHHP1* eviction from the *Pfap2-g* locus by a hitherto unknown mechanism [62]. However, as homologues of *Pgdv1* are so far only detected in primate malarials and the avian parasite *Plasmodium gallinaceum* [63], different control mechanisms are likely to operate in the other species. Besides AP2-G, there is considerable evidence for the involvement of two other AP2s in gametocyte development, at least in rodent-infecting malaria species. *Pb(b/y)AP2-G2* is regarded as a general transcriptional repressor, and during gametocytogenesis should release the repression of gametocyte-specific genes, start the repression of asexual genes and maintain the repression of genes specific to other stages [29, 30, 34]. *PyAP2-G3* has been implicated in gametocyte commitment in *P. yoelli* [34] and *P. falciparum* [64]. Studies of *P. yoelli* suggest that it acts upstream of *PyAP-G*, since deletion of *Pyap2-g3* reduced *Pyap2-g* expression but not vice versa. Additionally, *PyAP2-G3* is highly abundant in the cytoplasm, indicating that it might relay commitment signals to the nucleus [34]. Besides these factors identified by KO studies, several other AP2s have been implicated in gametocytogenesis. Using single-cell transcriptomics, it was possible to detect transcripts that were significantly upregulated shortly after *Pfap2-g* expression in *pfa2-g*-NI54 parasites, including those for the AP2 factors *PF3D7_1222400* and *PF3D7_1139300* [65]. In addition, Brancucci et al. [66] showed that the serum component lysophosphatidylcholine (lysoPC) inhibited *P. falciparum* gametocyte differentiation in vitro and detected significant upregulation of seven *Pfap2* transcripts besides *Pfap2-g* in the absence of LysoPC (PF3D7_0516800 (orthologue of *Pfap2-o*), PF3D7_0613800, PF3D7_0802100, *Pfap2-g2/PF3D7_1007700, PF3D7_1222400, PF3D7_1239200 and PF3D7_1456000 (not essential in *P. yoelli* or *P. berghei*)). *PF3D7_1222400* is an interesting candidate as it was found in both studies and is unique to primate-infecting *Plasmodium* species. As the majority of the other genes were essential for the IDC of rodent-infecting *Plasmodium* species, alternative methods are required to decipher their contribution to gametocytogenesis. It should also be noted that, although the majority of *apiap2* genes show differential mRNA expression in *P. falciparum* male and female gametocytes [67], thus far, no single AP2 factor could be associated with sex-specific gene expression in the KO screens [30, 34].

KO studies have demonstrated the essential role of five AP2 factors for successful mosquito infection, at least in rodent-infecting *Plasmodium* species (*AP2-O through AP2-O5*) [19, 28, 30, 34]. Of these, the function of *PbAP2-O* has been most extensively studied. This AP2 has a central role in ookinete gene expression, and without *Pbap2-o*, development of *P. berghei* parasites halts in zygote-to-ookinete transition. In wild-type parasites, the *Pbap2-o* transcript is translationally repressed in development of zygote inhibited (DOZI)–containing ribonucleoprotein complexes in female gametocytes [68, 69] and, after being translated, it induces the expression of many ookinete-specific genes [19, 28]. Deletion of each of the other four *ap2* genes in *P. berghei* or *P. yoelli* halts parasite development somewhere along the differentiation path towards mature oocysts or settling oocysts at the basal side of the mosquito midgut epithelium [30, 34]. *P(b/y)AP2-O3* and -O4 are required for proper ookinete maturation and early oocyst formation, respectively. The *Pbap2-o3* KO showed an upregulation of male-specific transcripts in gametocytes in the absence of the expected shift in male:female gametocyte ratios in blood films and might therefore be required for the repression of male-specific genes. In addition, an upregulation of translationally repressed transcripts was detected in the ookinete cultures of this KO, suggesting a role for *PbAP2-O3* in translational repression [30] (although these hypotheses remain to be tested). While KOs for *ap2-o−o5* and −o4 showed the same morphological phenotype in *P. berghei* and *P. yoelli*, deletion of *ap2-o2* did not [19, 30, 34]. In *P. berghei*, *ap2-o2* deletion greatly reduced in vivo zygote-to-ookinete conversion efficiency and this line formed only a few oocysts in vivo that were unable to sporulate. In *P. yoelli*, on the other hand, *ap2-o2* deletion did not affect ookinete maturation.
What functional genomics has taught us about transcriptional regulation in malaria parasites | 297

but did affect oocyst formation and greatly reduced sporozoite numbers in midgut oocysts and salivary glands. It is interesting that even in these closely related species, homologous AP2 factors appear to have different functions. Lastly, PyAP2-05 was found to be essential for ookinete motility. Obviously, additional experiments are needed to elucidate the role, target genes and reciprocal interactions of these five factors during development of functional ookinetes.

Finally, at least three further AP2 factors are required for the formation of mature sporozoites in P. berghei and P. yoelli (PbAP2-SP, -SP2, -SP3), each functioning at a different stage in development [22, 30, 34]. While AP2-SP was the first AP2 to be studied in detail [22], its genome-wide target genes and mode of action have not yet been elucidated. In addition, as mentioned above, besides its transcription-activating function in sporozoites, recent findings indicate that this factor might have transcriptional-repressive properties in the IDC of P. berghei [30] and might play a role in the expression of multi-gene family genes in P. falciparum [32]. Finally, at least one AP2 (PbAP2-L) is essential for the development of mature liver-stage schizonts in P. berghei [25]. Recent transcriptome analyses of liver-stage schizonts and hypnozoites of P. vivax [70] and Plasmodium cynomolgi [71, 72] attempted to identify additional AP2s involved in liver-stage development of these species including quiescent hypnozoite forms. However, the resulting gene lists show very little overlap among the three studies, indicating that further experimental validation is desired to characterize the role of AP2s in liver-stage development in general and in hypnozoite formation in particular.

Taken together, the above studies support the essential role of the ApiAP2 family in Plasmodium life cycle progression. The various developmental stages seem to rely critically on one or several AP2 family members but further studies are needed to investigate the functional contribution of the AP2 in question, beyond the stage at which it is essential.

Genome-wide mapping of gene regulatory elements in the Plasmodium genome

In order to modulate the transcriptional process, TFs need to bind cis-regulatory DNA elements. TF binding sites (TFBSs) are the core of larger gene regulatory elements like enhancers or promoters and play a central role in molecular gene regulatory networks [4]. The preferred approach to identify TFBSs relies on chromatin immunoprecipitation sequencing (ChIP-seq) in combination with transcriptome analyses of KO lines. However, ChIP-seq of TFs, in particular, has proven to be difficult in Plasmodium and has only been performed for a few factors [23, 28, 29, 31, 33]. Therefore, alternative approaches to identify potential regulatory DNA elements and associate those with TFBSs, TFs and potential target genes, have been employed. These can roughly be divided into two strategies that are complementary to each other. One approach starts with finding the TFBS for each TF using in vitro sequence preferences of individual DNA-binding domains. To this end, protein-binding microarray (PBM) experiments have been performed to predict DNA-binding preferences for the AP2 domains of 20 P. falciparum AP2 proteins [24]. These predictions have been validated and refined by additional biochemical, transcriptomic and/or ChIP experiments for PfSIP2 [23], PfAP2-G [26] and PfAP2-1 [31] in P. falciparum as well as for PbAP2-G2 [29] and PbAP2-O [28] in P. berghei, demonstrating the quality of the PBM predictions. However, in vivo binding is unlikely to occur at every genomic occurrence of the in vitro predicted motif, and not all AP2 domains are necessarily required for DNA binding. In addition, it is well appreciated that DNA-binding preferences can be influenced by the flanking DNA [24] and the interaction with epigenetic reader proteins [31, 56]. Lastly, while AP2 proteins are largely conserved across Plasmodium species [18], this does not exclude the possibility of altered DNA-binding preferences or alternative use of orthologous TFs in related species [24, 73]. Hence, while in vitro DNA-sequence preferences provide a valuable starting point for target site prediction, additional experiments are needed to validate and refine these in vitro predictions. The alternative strategy assumes that regulatory regions of co-expressed and/or functionally related genes share conserved DNA-sequences that serve as TFBS for a particular TF, operational at that stage. The G-box element, upstream of Plasmodium heat shock genes, was identified using this reasoning [74]. Subsequent genome-wide transcriptome analyses, either stand-alone or in combination with chromatin landscape profiling data sets, allowed motif predictions to be made on a larger scale using various bioinformatics approaches [52, 56, 75–82]. However, in many cases, the predicted motifs were not functionally tested and the TFs recognizing the predicted motifs were not identified.

As mentioned above, an additional degree of refinement that is useful for transferring PBM-derived motifs to regulatory gene elements, as well as for de novo bioinformatics motif prediction strategies, can be gained by taking the chromatin landscape into account. TF binding disrupts the local nucleosome structure and/or prevents local occupancy by nucleosomes, thereby creating a relative ‘open’ chromatin region. Genome-wide indexing of these open chromatin regions can be achieved using exonuclease deoxyribonuclease 1 (DNase I-seq [83, 84]) or Tn5 transposition (assay for transposase-accessible chromatin using sequencing; ATAC-seq [85]) followed by sequencing of the purified fragments. Alternatively, these NDRs can be purified from the pellet of nucleosomal DNA after crosslinking and quantified by DNA sequencing (formaldehyde-assisted isolation of regulatory elements; FAIRE-seq [86]). Each of these methods is affected by some level of bias, either enzymatic or due to cross-linking, and appropriate controls should be used [87]. Both FAIRE-seq and ATAC-seq have been applied to the P. falciparum genome during blood-stage development. Of these two techniques, ATAC-seq appears to provide a higher resolution [56, 80, 82, 88]. In general, the relative level of enrichment or accessibility correlated positively with the transcript abundance of the closest downstream gene [56, 82, 88], and the differential expression of clonally variant genes across P. falciparum strains could be explained by the presence of distinct accessibility patterns [82]. These observations not only support the contribution of transcriptional control to global IDC gene expression but also suggest that, during the IDC, transcriptional activating events prevail over repressive ones. Furthermore, there is an indication that the majority of regulatory elements appear to locate close to their target gene, as is also observed in ChIP-seq studies [28, 29, 31, 56, 82]. Whether the negatively correlating events, although few in number, reflected the activity of candidate repressors like AP2-G2, AP2-O or AP2-SP [30] was not investigated. As these accessible regions represent candidate in vivo TF binding events, they can be used, either alone or in combination with other data sets, to make more informed motif predictions, identify collaborative binding events and construct global gene regulatory networks. For example, in-depth motif analyses could substantiate and refine PBM-predicted motifs and suggest candidate novel motifs in P. falciparum [56]. To
identify TFs binding to such novel DNA elements, DNA pulldown coupled with quantitative proteomics can be used [89]. Interestingly, mainly AP2 factors were identified using such an approach as the respective binding partners of several de novo motifs [56], confirming their status as major *P. falciparum* TF family.

Clearly, chromatin accessibility-based approaches combined with computational modeling have great potential in our quest to understand the regulatory network that dictates *Plasmodium* gene expression. Until now, the discovery of regulatory elements has mainly been limited to *P. falciparum* IDIC parasites; however, the low-input requirements of ATAC-seq provide the means to study gene regulation in other *Plasmodium* species and stages as well.

**Conclusion and outlook**

The combined use of systematic KO screens, transcriptomics, chromatin accessibility mappings and computational approaches has established the involvement of transcriptional regulation in Plasmodium life cycle progression. Thus far, the apicomplexan AP2 family emerges as the main TF family in *Plasmodium* parasites, although other putative and yet unidentified DNA-binding proteins might also contribute. The exact modes of action of the different AP2 proteins should be subject to further investigation including their proposed combinatorial mode of action [30, 52] and pleiotropic functionality [30]. Interaction proteomic studies and ChIP-re-ChIP-type approaches are needed to reveal the proposed interactions between different DNA-binding as well as epigenetic reader proteins. Similarly, stage-specific interactions between TFs and/or stage-specific post-translational modifications (PTM) could help to explain how certain TFs can perform activating as well as suppressive functions at different stages. Such interaction and PTM-profiling studies might actually be the first step in drawing up the signaling cascades that control TF function. Additionally, they may shed light on the question of whether (and how) environmental changes provoke transcriptional responses that may contribute to the adaptation of parasites.

Similarly, there remains a lot to be learned about both core promoters and TFBSSs. While chromatin accessibility-based approaches and ChIP-seq substantially improved the identification of putative TFBSS sites, we are not yet able to pinpoint target sequences confidently for the majority of TFs across parasite stages. This will likely require systematic mapping of the in vivo binding patterns of large numbers of individual TFs at different stages, followed by CRISPR-Cas9 mutagenesis of the endogenous regulatory elements, whether TSSs or TFBSSs. These, in combination with other genome-wide data sets, could also reveal the features involved in targeted transcription factors beyond the DNA sequence, be that DNA structure, DNA modifications or non-coding RNAs.

Eventually, the gene regulatory network of *Plasmodium* species will be built by combining evidence from different methodologies across life cycle stages and *Plasmodium* species. Low-input methodologies [90] and the advent of single-cell (multi-)omics approaches [91–93], together with interaction proteomic analyses, provide exciting opportunities, not only to study the characteristics of the network in stages and species that cannot be cultured in vitro, but also to incorporate epigenetic and post-transcriptional regulatory elements in order to model its heterogeneity and stochastic nature [94].

---

**Key Points**

Recent functional genomics studies establish a central role for transcriptional mechanisms in *Plasmodium* gene expression regulation.

Transcription initiation in *P. falciparum* is rather promiscuous and core promoters are not clearly delineated.

Apetala 2 (AP2) domain-containing proteins are the main transcription factor family in *Plasmodium* parasites.

The majority of identified gene regulatory elements can be associated with an immediate downstream target gene.

**References**

1. Haberle V, Stark A. Eukaryotic core promoters and the functional basis of transcription initiation. Nat Rev Mol Cell Biol 2018;19:621–57.
2. Slattery M, Zhou T, Yang L, et al. Absence of a simple code: how transcription factors read the genome. Trends Biochem Sci 2014;39:381–99.
3. Levo M, Segal E. In pursuit of design principles of regulatory sequences. Nat Rev Genet 2014;15:453–68.
4. Doane AS, Elemento O. Regulatory elements in molecular networks. Wiley Interdiscip Rev Syst Biol Med 2017;9:e1374–83.
5. Clayton CE. Life without transcriptional control? From fly to man and back again. EMBO J 2002;21:1881–8.
6. Duffy MF, Selvarajah SA, Josling GA, et al. Epigenetic regulation of the plasmodium falciparum genome. Brief Funct Genomics 2014;13:203–16.
7. Kirchner S, Power BJ, Waters AP. Recent advances in malaria genomics and epigenomics. Genome Med 2016;8:17.
8. Cortes A, Deitsch KW. Malaria epigenetics. Cold Spring Harb Perspect Med 2017;7:23.
9. Duraisingh MT, Skillman KM. Epigenetic variation and regulation in malaria parasites. Annu Rev Microbiol 2018;72:355–75.
10. Fraschka SA, Filarsky M, Hoo R, et al. Comparative heterochromatin profiling reveals unique and unique epigenome signatures linked to adaptation and development of malaria parasites. Cell Host Microbe 2018;23:407–20.
11. Bartfai R, Hoeijmakers WA, Salcedo-Amaya AM, et al. H2A.Z demarcates intergenic regions of the plasmodium falciparum epigenome that are dynamically marked by H3K9ac and H3K4me3. PLoS Pathog 2010;6:e1001223.
12. Gupta AP, Chin WH, Zhu L, et al. Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite Plasmodium falciparum. PLoS Pathog 2013;9:e1003170.
13. Horrocks P, Decherling K, Lanzer M. Control of gene expression in Plasmodium falciparum. Mol Biochem Parasitol 1998;95:171–81.
14. Horrocks P, Wong E, Russell K, et al. Control of gene expression in Plasmodium falciparum—ten years on. Mol Biochem Parasitol 2009;164:9–25.
15. Watanabe J, Sasaki M, Suzuki Y, et al. Analysis of transcriptomes of human malaria parasite Plasmodium falciparum using full-length enriched library: identification of novel genes and diverse transcription start sites of messenger RNAs. Gene 2002;291:105–13.
16. Adjalley SH, Chabbert CD, Klaus B, et al. Landscape and dynamics of transcription initiation in the malaria parasite Plasmodium falciparum. Cell Rep 2016;14:2463–75.
What functional genomics has taught us about transcriptional regulation in malaria parasites

17. Kensche PR, Hoeijmakers WA, Toenhake CG, et al. The nucleosome landscape of Plasmodium falciparum reveals chromatin architecture and dynamics of regulatory sequences. Nucleic Acids Res 2016;44:2110–24.

18. Balaji S, Babu MM, Iyer LM, et al. 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 2005;33:3994–4006.

19. Yuda M, Iwanaga S, Shigenobu S, et al. Identification of a transcription factor in the mosquito-invasive stage of malaria parasites. Mol Microbiol 2009;71:1402–14.

20. De Silva EK, Gehrke AR, Olszewski K, et al. Specific DNA-binding by apicomplexan AP2 transcription factors. Proc Natl Acad Sci U S A 2008;105:8393–8.

21. Lindner SE, De Silva EK, Keck JL, et al. Structural determinants of DNA binding by a P. falciparum ApiAP2 transcriptional regulator. J Mol Biol 2010;395:558–67.

22. Yuda M, Iwanaga S, Shigenobu S, et al. Transcription factor AP2-Sp and its target genes in malarial sporozoites. Mol Microbiol 2010;75:854–63.

23. Flueck C, Bartfai R, Niederwieser I, et al. A major role for the Plasmodium falciparum ApiAP2 protein PSIP2 in chromosome end biology. PLoS Pathog 2010;6:e1000784.

24. Campbell TL, De Silva EK, Olszewski KL, et al. Identification and genome-wide prediction of DNA binding specificities for the ApiAP2 family of regulators from the malaria parasite. PLoS Pathog 2010;6:15.

25. Iwanaga S, Kaneko I, Kato T, et al. Identification of an AP2-family protein that is critical for malaria liver stage development. PLoS One 2012;7:e47557.

26. Kafsack BFC, Rovira-Graells N, Clark TG, et al. A transcriptional switch underlies commitment to sexual development in malaria parasites. Nature 2014;507:248.

27. Sinha A, Hughes KR, Modrzynska KK, et al. A cascade of DNA-binding proteins for sexual commitment and development in Plasmodium. Nature 2014;507:253.

28. Kaneko I, Iwanaga S, Kato T, et al. Genome-wide identification of the target genes of AP2-O, a Plasmodium AP2-family transcription factor. PLoS Pathog 2015;11:e1004905.

29. Yuda M, Iwanaga S, Kaneko I, et al. Global transcriptional repression: an initial and essential step for Plasmodium sexual development. Proc Natl Acad Sci U S A 2015;112:12824–9.

30. Modrzynska K, Pfander C, Chappell L, et al. A knockout screen of ApiAP2 genes reveals networks of interacting transcriptional regulators controlling the Plasmodium life cycle. Cell Host Microbe 2017;21:11–22.

31. Santos JM, Josifin G, Ross P, et al. Red blood cell invasion by the malaria parasite is coordinated by the PfAP2-I transcription factor. Cell Host Microbe 2017;21:731.

32. Martins RM, Macpherson CR, Clae A, et al. An ApiAP2 member regulates expression of clonally variant genes of the human malaria parasite Plasmodium falciparum. Sci Rep 2017;7:14042.

33. Sierra-Miranda M, Vembar SS, Delgadillo DM, et al. PfAP2Tel, harbouring a non-canonical DNA-binding AP2 domain, binds to Plasmodium falciparum telomeres. Cell Microbiol 2017;19:e12742–53.

34. Zhang C, Li Z, Cui H, et al. Systematic CRISPR-Cas9-mediated modifications of Plasmodium yoelii ApiAP2 genes reveal functional insights into parasite development. MBio 2017;8:e01986–17.

35. Aurrecoechea C, Brestelli J, Brunk BP, et al. PlasmidDB: a functional genomic database for malaria parasites. Nucleic Acids Res 2009;37:D539–43.

36. Dechering KJ, Cuelenaere K, Konings RN, et al. Distinct frequency-distributions of homopolymeric DNA tracts in different genomes. Nucleic Acids Res 1998;26:4056–62.

37. Hoeijmakers WA, Salcedo-Amaya AM, Smits AH, et al. H2A.Z/H2B.Z double-variant nucleosomes inhabit the AT-rich promoter regions of the Plasmodium falciparum genome. Mol Microbiol 2013;87:1061–73.

38. Petter M, Selvarajah SA, Lee CC, et al. H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy var gene promoters in the malaria parasite Plasmodium falciparum. Mol Microbiol 2013;87:1167–82.

39. Silberhorn E, Schwartz U, Loffler P, et al. Plasmodium falciparum nucleosomes exhibit reduced stability and lost sequence-dependent nucleosome positioning. PLoS Pathog 2016;12:e1006080.

40. Brick K, Watanebe J, Pizzi E. Core promoters are predicted by their distinct physicochemical properties in the genome of Plasmodium falciparum. Genome Biol 2008;9:16.

41. Carninci P, Sandelin A, Lenhard B, et al. Genome-wide analysis of mammalian promoter architecture and evolution. Nat Genet 2006;38:626.

42. Russell K, Cheng CH, Bizzaro JW, et al. Homopolymer tract organization in the human malarial parasite Plasmodium falciparum and related apicomplexan parasites. BMC Genomics 2014;15:848.

43. Ruvalcaba-Salazar OK, del Carmen Ramirez-Estudillo M, Montiel-Condado D, et al. Recombinant and native Plasmodium falciparum TATA-binding-protein binds to a specific TATA box element in promoter regions. Mol Biochem Parasitol 2005;140:183–96.

44. Struhl K, Segal E. Determinants of nucleosome positioning. Nat Struct Mol Biol 2013;20:267–73.

45. Bischoff E, Vaquero C. In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of Plasmodium falciparum. BMC Genomics 2010;11:34.

46. Komaki-Yasuda K, Okuwa M, Nagata K, et al. Identification of a novel and unique transcription factor in the malaria parasite Plasmodium falciparum. BMC Genomics 2010;11:34.

47. Bischoff E, Vaquero C. In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of Plasmodium falciparum. BMC Genomics 2010;11:34.

48. Boschet C, Gissot M, Briquet S, et al. Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 Plasmodium falciparum clones. Mol Biochem Parasitol 2004;138:159–63.

49. Gissot M, Briquet S, Refour P, et al. PfMyb1, a Plasmodium falciparum transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation. J Mol Biol 2005;346:29–42.

50. Burns CG, Ohi R, Krainer AR, et al. Evidence that Myb-related CDC5 proteins are required for pre-mRNA splicing. Proc Natl Acad Sci U S A 1999;96:13789–94.

51. Bertsch NL, Toenhake CG, Zou A, et al. Malaria parasites possess a telomere repeat-binding protein that shares ancestry with transcription factor IIIA. Nat Microbiol 2017;2:17033.

52. Cowell AN, Istvan ES, Lukens AK, et al. Mapping the malaria parasite druggable genome by using in vitro evolution and chemogenomics. Science 2018;359:191–9.

53. de Vries H, Huyten MA. Combinatorial gene regulation in Plasmodium falciparum. Trends Genet 2006;22:73–8.
62. Filarsky M, Fraschka SA, Niederwieser I, Lopez-Rubio JJ, Mancio-Silva L, Scherf A. Genome-wide analysis of the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. Science 2018;360:eaap7847.

53. El-Gebali S, Mistry J, Bateman A, et al. The Pfam protein families database in 2019. Nucleic Acids Res 2018;47:D427–D432.

54. Cobbold SA, Santos JM, Ochoa A, et al. Proteome-wide analysis reveals widespread lysine acetylation of major protein complexes in the malaria parasite. Sci Rep 2016;6:14.

55. Zhang M, Wang C, Otto TD, et al. Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. Science 2018;360:eaap7847.

56. Toenhake CG, Fraschka SA, Vijayabaskar MS, et al. Chromatin accessibility-based characterization of the gene regulatory network underlying Plasmodium falciparum blood-stage development. Cell Host Microbe 2018;23:557–69.

57. Josling GA, Williamson KC, Llinàs M. Regulation of sexual commitment and gametocytogenesis in malaria parasites. Annu Rev Microbiol 2018;72:501–19.

58. Kent RS, Modrzynska K, Cameron R, et al. Inducible developmental reprogramming redefines commitment to sexual development in the malaria parasite Plasmodium berghei. Nat Microbiol 2018;3:1206–13.

59. Bancells C, Lorà-Batlle O, Poran A, et al. Revisiting the initial steps of sexual development in the malaria parasite Plasmodium falciparum. Nat Microbiol 2018;4:144–54.

60. Fluéck C, Bartfai R, Volz J, et al. Plasmodium falciparum heterochromatin protein 1 marks genomic loci linked to heterochromatinization of exported virulence factors. PLoS Pathog 2009;5:e1000569.

61. Lopez-Rubio JJ, Mancio-Silva L, Scherf A. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. Cell Host Microbe 2009;5:179–90.

62. Filarsky M, Fraschka SA, Niederwieser I, et al. GDV1 induces sexual commitment of malaria parasites by antagonizing HP1-dependent gene silencing. Science 2018;359:1259–63.

63. Eksi S, Morahan BJ, Haile Y, et al. Plasmodium falciparum gametocyte development I (Pfdv1) and gametocytogenesis early gene identification and commitment to sexual development. PLoS Pathog 2012;8:e1002964.

64. Ikadai H, Shaw Saliba K, Kanzok SM, et al. Transposon mutagenesis identifies genes essential for Plasmodium falciparum gametocytogenesis. Proc Natl Acad Sci USA 2013;110:E1675–84.

65. Poran A, Notzel C, Aly O, et al. Single-cell RNA sequencing reveals a signature of sexual commitment in malaria parasites. Nature 2017;551:95–9.

66. Brancucci NMB, Gerdt JP, Wang C, et al. Lysophosphatidylcholine regulates sexual stage differentiation in the human malaria parasite Plasmodium falciparum. Cell 2017;171:1532–44.

67. Lasonder E, Rijpma SR, van Schaik BC, et al. Integrated transcriptomic and proteomic analyses of P. falciparum gametocytes: molecular insight into sex-specific processes and translational repression. Nucleic Acids Res 2016;44:6087–101.

68. Mair GR, Braks JA, Garver LS, et al. Regulation of sexual development of Plasmodium by translational repression. Science 2006;313:667–9.

69. Mair GR, Lasonder E, Garver LS, et al. Universal features of post-transcriptional gene regulation are critical for Plasmodium zygote development. PLoS Pathog 2010;6:e1000767.

70. Gural N, Mancio-Silva L, Miller AB, et al. In vitro culture, drug sensitivity, and transcriptome of Plasmodium vivax hypnozoites. Cell Host Microbe 2018;23:395–406.

71. Cubi R, Vembar SS, Biton A, et al. Laser capture microdissection enables transcriptomic analysis of dividing and quiescent liver stages of Plasmodium relapsing species. Cell Microbiol 2017;19:e12735.

72. Voorberg-van der Wel A, Roma G, Gupta DK, et al. A comparative transcriptomic analysis of replicating and dormant liver stages of the relapsing malaria parasite Plasmodium cynomolgi. Elife 2017;6:e29605–26.

73. Oberstaller J, Pumplava Y, Schieler A, et al. The Crytosporidium parvum ApiAP2 gene family: insights into the evolution of apicomplexan AP2 regulatory systems. Nucleic Acids Res 2014;42:8271–84.

74. Militello KT, Dodge M, Bethke L, et al. Identification of regulatory elements in the Plasmodium falciparum genome. Mol Biochem Parasitol 2004;134:75–88.

75. Elemento O, Slonim N, Tava zoie S. A universal framework for regulatory element discovery across all genomes and data types. Mol Cell 2007;28:337–50.

76. Gunasekera AM, Myrick A, Militello KT, et al. Regulatory motifs uncovered among gene expression clusters in Plasmodium falciparum. Mol Biochem Parasitol 2007;153:19–30.

77. Wu J, Sieglaflf DH, Gervin J, et al. Discovering regulatory motifs in the Plasmodium genome using comparative genomics. Bioinformatics 2008;24:1843–9.

78. Young JA, Johnson JR, Benner C, et al. In silico discovery of transcription regulatory elements in Plasmodium falciparum. BMC Genomics 2008;9:70.

79. Jurgenelaitre R, Dijkstra TM, Kocken CH, et al. Gene regulation in the intraerythrocytic cycle of Plasmodium falciparum. Bioinformatics 2009;25:1484–91.

80. Harris EY, Ponts N, Le Roch KG, et al. Chromatin-driven de novo discovery of DNA binding motifs in the human malaria parasite. BMC Genomics 2011;12:601.

81. Russell K, Emes R, Horrocks P. Triaging informative cis-regulatory elements for the combinatorial control of temporal gene expression during Plasmodium falciparum intraerythrocytic development. Parasit Vectors 2015;8:81.

82. Ruiz JL, Tena JJ, Bancells C, et al. Characterization of the accessible genome in the human malaria parasite Plasmodium falciparum. Nucleic Acids Res 2018;46:9414–31.

83. Hesselberth JR, Chen X, Zhang Z, et al. Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. Nat Methods 2009;6:283–9.

84. Nep S, Vierstra J, Stegachis AB, et al. An expansive human regulatory lexicon encoded in transcription factor footprints. Nature 2012;489:83–90.

85. Buenrostro JD, Giresi PG, Zaba LC, et al. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat Methods 2013;10:1213–8.

86. Giresi PG, Kim J, McDaniel RM, et al. FAIRE (formaldehyde-assisted isolation of regulatory elements) isolates active regulatory elements from human chromatin. Genome Res 2007;17:877–85.

87. Meyer CA, Liu XS. Identifying and mitigating bias in next-generation sequencing methods for chromatin biology. Nat Rev Genet 2014;15:709–21.

88. Ponts N, Harris EY, Frudhomme J, et al. Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res 2010;20:228–38.
89. Makowski MM, Willems E, Fang J, et al. An interaction proteomics survey of transcription factor binding at recurrent TERT promoter mutations. Proteomics 2016;16:417–26.
90. Greenleaf WJ. Assaying the epigenome in limited numbers of cells. Methods 2015;72:51–6.
91. Peterson VM, Zhang KX, Kumar N, et al. Multiplexed quantification of proteins and transcripts in single cells. Nat Biotechnol 2017;35:936–9.
92. Stoeckius M, Hafemeister C, Stephenson W, et al. Simultaneous epitope and transcriptome measurement in single cells. Nat Methods 2017;14:865–8.
93. Packer J, Trapnell C. Single-cell multi-omics: an engine for new quantitative models of gene regulation. Trends Genet 2018;34:653–65.
94. Fiers M, Minnoye L, Aibar S, et al. Mapping gene regulatory networks from single-cell omics data. Brief Funct Genomics 2018;17:246–54.
95. Bushell E, Gomes AR, Sanderson T, et al. Functional profiling of a Plasmodium genome reveals an abundance of essential genes. Cell 2017;170:260–72.
96. Painter HJ, Campbell TL, Llinas M. The apicomplexan AP2 family: integral factors regulating Plasmodium development. Mol Biochem Parasitol 2011;176:1–7.
97. Fan Q, An L, Cui L. PfADA2, a Plasmodium falciparum homologue of the transcriptional coactivator ADA2 and its in vivo association with the histone acetyltransferase PfGCNS. Gene 2004;336:251–61.
98. Fan Q, An L, Cui L. Plasmodium falciparum histone acetyltransferase, a yeast GCN5 homologue involved in chromatin remodeling. Eukaryot Cell 2004;3:264–76.