

RESEARCH

Open Access



Histone modifications associated with gene expression and genome accessibility are dynamically enriched at *Plasmodium falciparum* regulatory sequences

Jingyi Tang^{1,2†}, Scott A. Chisholm^{3,4†}, Lee M. Yeoh^{4,5,6}, Paul R. Gilson^{7,8}, Anthony T. Papenfuss^{9,10,11,12}, Karen P. Day^{4,5,6}, Michaela Petter^{1,13} and Michael F. Duffy^{1,4,5,6*} 

Abstract

Background: The malaria parasite *Plasmodium falciparum* has an unusually euchromatic genome with poorly conserved positioning of nucleosomes in intergenic sequences and poorly understood mechanisms of gene regulation. Variant histones and histone modifications determine nucleosome stability and recruit *trans* factors, but their combinatorial contribution to gene regulation is unclear.

Results: Here, we show that the histone H3 acetylations H3K18ac and H3K27ac and the variant histone Pf H2A.Z are enriched together at regulatory sites upstream of genes. H3K18ac and H3K27ac together dynamically mark regulatory regions of genes expressed during the asexual life cycle. In contrast, H3K4me1 is depleted in intergenic sequence and dynamically depleted upstream of expressed genes. The temporal pattern of H3K27ac and H3K18ac enrichment indicates that they accumulate during S phase and mitosis and are retained at regulatory sequences until at least G1 phase and after cessation of expression of the cognate genes. We integrated our ChIPseq data with existing datasets to show that in schizont stages H3K18ac, H3K27ac and Pf H2A.Z colocalise with the transcription factor PfAP2-I and the bromodomain protein PfBDP1 and are enriched at stably positioned nucleosomes within regions of exposed DNA at active transcriptional start sites. Using transient transfections we showed that sequences enriched with colocalised H3K18ac, H3K27ac and Pf H2A.Z possess promoter activity in schizont stages, but no enhancer-like activity.

Conclusions: The dynamic H3 acetylations define *P. falciparum* regulatory sequences and contribute to gene activation. These findings expand the knowledge of the chromatin landscape that regulates gene expression in *P. falciparum*.

Keywords: *Plasmodium falciparum*, Histone modifications, Gene regulation

Background

The parasite *Plasmodium falciparum* is the cause of most global malaria mortality. *P. falciparum* differs from other eukaryotes in the unusually high AT content of its genome (90% in intergenic regions) [1], its unique variant histones [2] and the high proportion of its genome that it maintains as euchromatin during its pathogenic, intra-erythrocytic lifecycle [3]. Structural analysis of *P.*

*Correspondence: mduffy@unimelb.edu.au

†Jingyi Tang and Scott A Chisholm contributed equally

¹ Department of Medicine, The University of Melbourne, Royal Melbourne Hospital, Parkville, VIC 3050, Australia

Full list of author information is available at the end of the article



falciparum nucleosomes confirmed the unique nature of *P. falciparum* chromatin [4], which may therefore exert novel mechanisms of control over gene expression. The majority of *P. falciparum* genes are specifically regulated during its haploid, 48-h intra-erythrocytic developmental cycle (IDC) [5, 6] and groups of *Plasmodium* genes are specifically activated by stage-specific transcription factors [7]. In at least one instance a stage-specific transcription factor [8] requires the presence of the *P. falciparum* bromodomain protein 1 (PfBDP1) to activate a subset of erythrocyte invasion genes [9]. PfBDP1 binds acetylated lysines in histones and presumably acts as a transcriptional co-factor. Thus, the many possible combinations between chromatin-binding proteins and specific transcription factors may explain the tightly regulated, dynamic transcription of most *P. falciparum* genes.

Increased nucleosome occupancy can directly hinder transcription and is dependent on nucleosome stability and thus nucleosome composition. Histone modifications can also recruit factors that facilitate or repress transcription. Associations between combinations of variant histones and histone modifications with depletion of nucleosomes at regulatory elements have been shown in yeast and for the human Encyclopedia of DNA Elements (ENCODE) Project. Yeast and human active promoters are enriched in H3K9ac, H3K18ac, H3K27ac, H3K4me3 and the histone variant H2A.Z [10–20], whereas active human enhancer elements are characterised by high levels of H3K4me1, H3K18ac, H3K27ac and H2A.Z [10, 11, 16, 20, 21]. H2A.Z enrichment levels correlate with human enhancer strength [18, 22] and H3K27ac enrichment levels differentiate active from inactive metazoan enhancers [23, 24]. The transcriptional co-activators and histone acetyltransferases p300/CBP responsible for acetylating H3K27 and H3K18 [25] are also enriched at mammalian enhancers [11, 26–28]. Nucleosomal occupancy is relatively depleted around human promoters, but both nucleosomal occupancy and positioning is maintained around the exposed transcription factor binding site at putative human enhancers [29].

Metazoan enhancers can be many kb distant from the cognate gene, but yeast have upstream activating sequences (UAS) that are typically located within 450 bp of the transcriptional start site (TSS) [30]. Analogous to distal metazoan enhancers, UASs bind the essential, transcriptional regulator mediator complex [31], specific transcription factors, and chromatin-modifying complexes [32, 33] to regulate expression of the core promoter. Yeast maintain 2–3 nucleosomes in the 500 bp upstream of the TSS [34, 35]. However, gene activation is associated with depletion of the consistently positioned -1 nucleosome at approximately -250 bp [35] and

in ribosomal protein genes a nucleosome depleted region extends from the start codon to the UAS [33].

Throughout the *P. falciparum* intra-erythrocytic cycle gene expression correlates with intergenic, upstream, dynamic enrichment of several euchromatic histone modifications [3, 36–39]. These include enrichment of H3K9ac, H3K4me3 and Pf H2A.Z flanking the regions of greater chromatin accessibility that define promoter sequences and specific transcription factor binding sites [40, 41].

Multiple regions of increased chromatin accessibility [40, 41] and multiple transcriptional start sites (TSSs) [42, 43] are found within the same *P. falciparum* intergenic regions. *P. falciparum* intergenic sequence is maintained in a nucleosomal array aside from a short, nucleosome depleted region (NDR) of variable length that precedes the well positioned +1 nucleosome at TSSs [43]. Whether the multiple TSSs function in the same cell and whether they mark regulatory elements that have a synergistic or antagonistic effect on the core promoter or are simply promiscuous TSSs is unknown. The average intergenic size of approximately 1500 bp in *P. falciparum* is too short to resolve enhancer and promoter interactions within the same intergenic region using Hi-C approaches [44] and more distant interactions have not been detected outside of the contacts due to clustering of rRNA genes and virulence genes, suggesting distant enhancers do not exist in *P. falciparum* [45–47]. Putative *P. falciparum* enhancer-like elements identified by relative H3K4me1 enrichment [48] did not correlate with putative transcription factor binding sites identified by ATACseq (Assay for Transposase Accessible Chromatin using sequencing) [40]. H3K18ac and H3K27ac are enriched at both metazoan enhancers and promoters so they could be enriched at all regulatory sequences in *P. falciparum*. Indeed H3K27ac is enriched in intergenic sequence of euchromatic genes in *P. falciparum* oocysts and sporozoites, but is not clearly associated with gene expression in these mosquito stages [49].

To get better insight into the complex regulatory network underlying *P. falciparum* gene expression we investigated the genome-wide correlation of promoter and enhancer associated histone modifications with cognate gene expression profiles as well as with published DNA accessibility and transcription factor binding data. We demonstrate that dynamic enrichment of H3K18ac and H3K27ac at Pf H2A.Z containing nucleosomes at TSSs correlates with gene expression levels, whereas H3K4me1 is generally absent from intergenic sequences and thus depleted from these putative regulatory elements. Acetylations of H3K18 and H3K27 accumulate during mitosis and appear to mark regulatory sequences of genes for expression rather than as a consequence of expression.

From transient transfection analysis, we concluded that histone modifications that identify active enhancers and promoters in metazoan and yeast cells do identify *P. falciparum* sequences with schizont stage promoter activity but not enhancer-like elements.

Results

To investigate chromatin structure of the *P. falciparum* genome and how this is associated with regulatory sequences and gene expression, we analysed genomic enrichment of the variant histone H2A.Z and the histone modifications H3K4me1, H3K18ac and H3K27ac. These markers of chromatin structure are all enriched at metazoan enhancers and, with the exception of H3K4me1, at yeast and metazoan promoters [10–24].

Levels of H3K4me1 and Pf H2A.Z remain constant across the intra-erythrocytic development cycle (IDC) but levels of H3K18ac and H3K27ac vary

The antibodies we used to investigate the chromatin marks included a Pf H2A.Z antiserum of confirmed specificity [50] and commercial antibodies to H3K4me1, H3K27ac and H3K18ac that were previously validated for chromatin immunoprecipitation (ChIP) (Antibody Validation Database (<https://compbio.med.harvard.edu/antibodies/>)) [51], (Histone Antibody Specificity Database (<https://www.histoneantibodies.com>)) [52]. We confirmed specificity of these anti-human histone antibodies for *P. falciparum* histones by showing that each antibody bound only the single H3 band in western blots of *P. falciparum* histones (Additional file 1: Fig S1). We also showed their differential binding to dot blots of ten custom-made biotinylated *P. falciparum* peptides and six commercially available human peptides that were largely conserved with *P. falciparum* (Additional file 2: Fig S2). The anti-H3K18ac, anti-H3K27ac and anti-H3K4me1 antibodies all bound to their cognate histone modification with greater than 32-fold sensitivity compared to background, well over the ENCODE standard that signals from the cognate peptide should be at least tenfold higher than those from other peptides [53].

Fluorescence microscopy confirmed the predicted nuclear localisation of all three H3 modifications (Fig. 1a). The diffuse signal from anti-H3K18ac, anti-H3K27ac and anti-H3K4me1 largely colocalised with the DAPI-stained nucleus in ring and trophozoite stages.

The abundance of Pf H2A.Z and the H3 modifications throughout the intra-erythrocytic lifecycle was assessed by western blot using histone H3 as a loading control. The levels of H3K4me1 and H2A.Z were constant throughout the lifecycle. The levels of both H3K18ac and H3K27ac were low in ring- to mid-trophozoite-stages and peaked at late-trophozoite- to early-schizont-stages (32–40 hpi)

prior to dropping at mid- to late-schizont stage (Fig. 1b). This was consistent with previous western blot and mass spectrometry data [54, 55] and correlated with the total transcriptional activity that peaks at late stage in the IDC [56], and indicated that H3K18ac and H3K27ac may contribute to regulation of transcription.

Pf H2A.Z, H3, H3K18ac and H3K27ac are differentially enriched across the genome

ChIPseq of Pf H2A.Z, H3K18ac, H3K27ac, H3K4me1 and H3 was performed in ring, trophozoite and schizont-stage parasites in biological duplicates as per ENCODE guidelines [53, 57, 58], and matched input was also sequenced. Matched RNAseq was performed for each duplicate chromatin sample and trophozoite and schizont-stage analyses were augmented with an additional RNA biological replicate. For both RNAseq and ChIPseq the average \pm standard deviation quality (all more than $86.4 \pm 3.6\%$ reads \geq Q30) and coverage (all more than $94.3 \pm 0.7\%$ of reads mapped to *P. falciparum*, all greater than 27 ± 8 X genome coverage) of the libraries were all acceptable (Additional files 3, 4: Tables S1 and S2) and the biological replicates clustered together by principal component analysis (PCA) (Additional file 5: Fig S3) and were clearly reproducible (Additional file 6: Fig S4). The input clustered with H3K4me1 and H3, whereas H3K18ac, H3K27ac and Pf H2A.Z formed a separate cluster. The GC sequencing bias was controlled for by normalising to input.

In all three lifecycle stages Pf H2A.Z, H3K18ac and H3K27ac were globally enriched upstream of transcriptional start sites and downstream of transcriptional end sites (TESSs), but depleted within transcripts determined from the cognate RNAseq (Fig. 2). The intergenic enrichment of H3K18ac and H3K27ac was greater in trophozoites (30–36 hpi) and schizonts (38–44 hpi) than in ring stages, consistent with the western blot data for whole-cell levels of these modifications that peaked at 32–40 hpi. Conversely, H3 and H3K4me1 were depleted up and downstream of transcripts, but were relatively enriched in the transcribed sequences. This explains H3 and H3K4me1 clustering in PCA with input, which has a slight sequencing bias towards coding sequence.

Pf H2A.Z, H3K18ac, H3K27ac, H3K4me1 and unmodified H3 enrichment vary with gene expression

Ranking genes by expression levels demonstrated that the upstream enrichment of Pf H2A.Z, H3K18ac and H3K27ac positively correlated with gene expression in all three lifecycle stages. Conversely, the abundance of H3 and H3K4me1 upstream of coding sequences inversely correlated with transcript levels in all three lifecycle

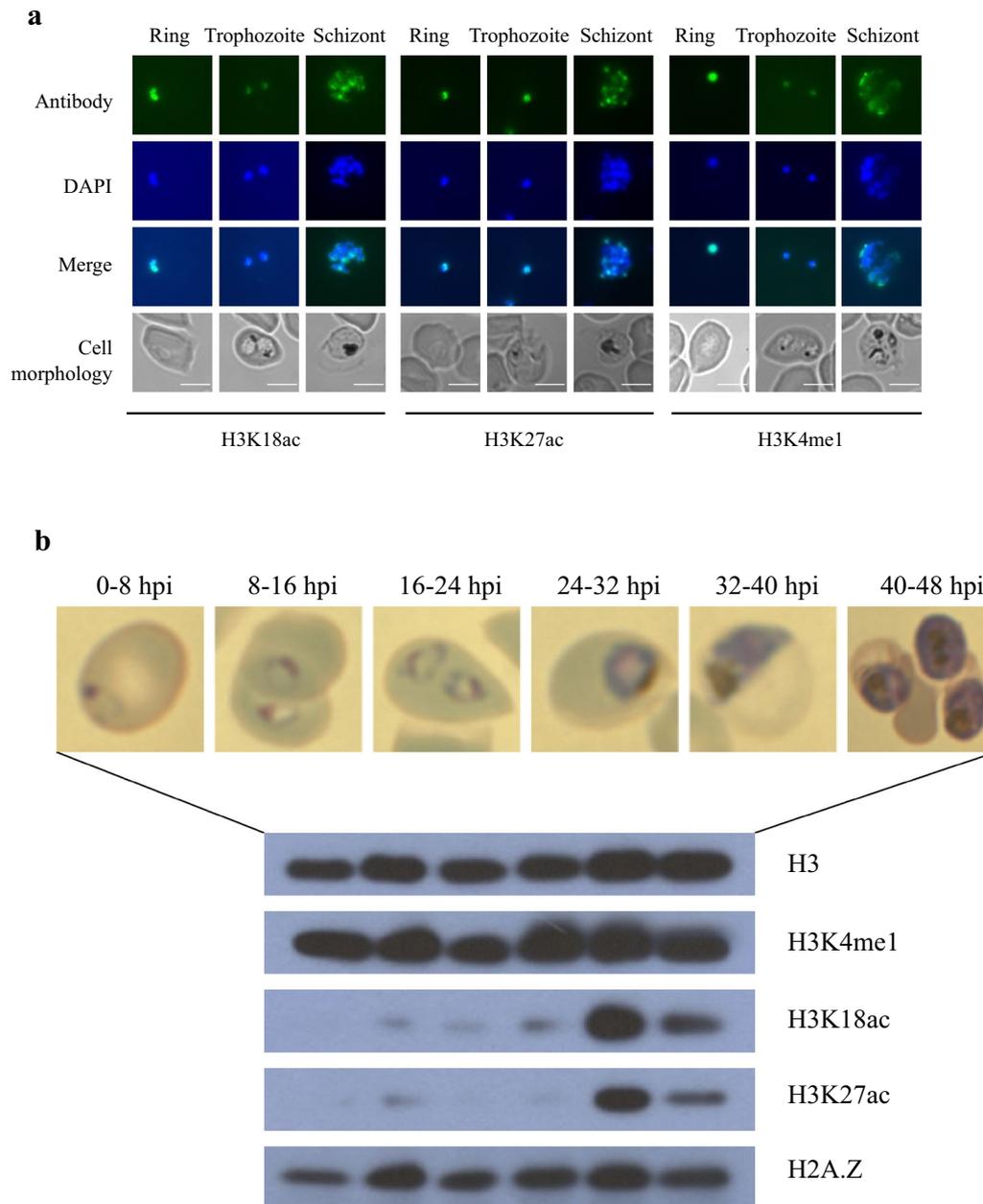


Fig. 1 Distribution of histones and their modifications throughout the intra-erythrocytic developmental cycle. **a** Immunofluorescence assay shows that PfH3K18ac, PfH3K27ac and PfH3K4me1 localise to the nucleus of *P. falciparum* in ring-, trophozoite- and schizont-stages in the IDC. Bound antibody shown in green in the first row was specific for the histone modifications indicated at the bottom of the figure. DNA was stained blue with DAPI in the second row. Bright field microscopy shows the cell morphology in the fourth row. Scale bar is 5 μ m. **b** *P. falciparum* H3K4me1, H3K18ac, H3K27ac and H2A.Z levels at 0–8 hpi, 8–16 hpi, 16–24 hpi, 24–32 hpi, 32–40 hpi and 40–48 hpi during the IDC. Cell morphology at the six time points when the parasite lysates were prepared is indicated on top. H3 levels were shown as the parasite lysate loading control

stages (Fig. 3, Additional file 7: Fig S5) (all Spearman r correlations and p values are provided in the figures).

The depletion of intergenic H3 was consistent with the correlation between intergenic nucleosome depletion and gene expression in *P. falciparum* [43]. Silent

genes and genes in the lowest tercile by expression had very similar upstream, intergenic profiles of moderate depletion of H3K4me1 and unmodified H3, suggesting that H3K4me1 levels simply mirrored nucleosomal occupancy for these genes. However, genes in the

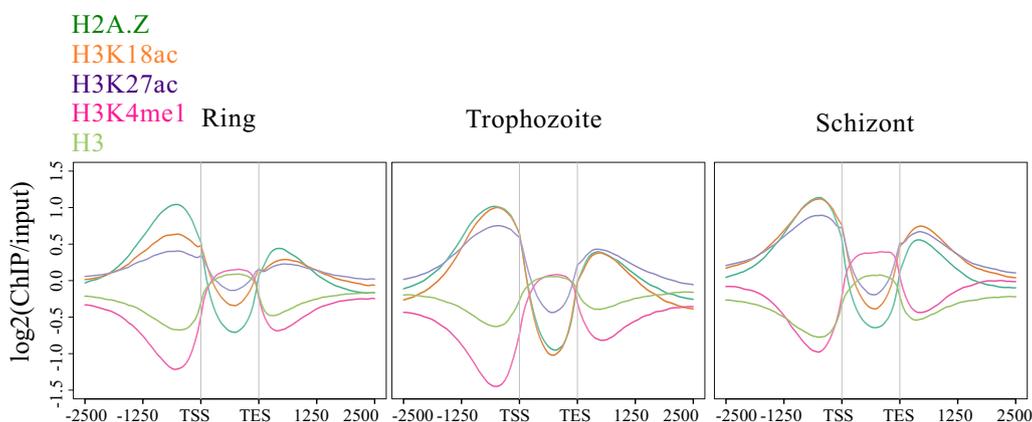


Fig. 2 Enrichment profiles of histones and their modifications relative to transcriptional units. Average \pm SE enrichment profiles (\log_2 ratio of ChIP over stage-matched input) of H2A.Z (dark green), H3K18ac (orange), H3K27ac (purple), H3K4me1 (pink) and H3 (pale green) plotted across all transcripts from the transcription start sites (TSSs) to the transcription end sites (TESs) with an extension of 2500 bp upstream and downstream in ring, trophozoite stage and schizont stages

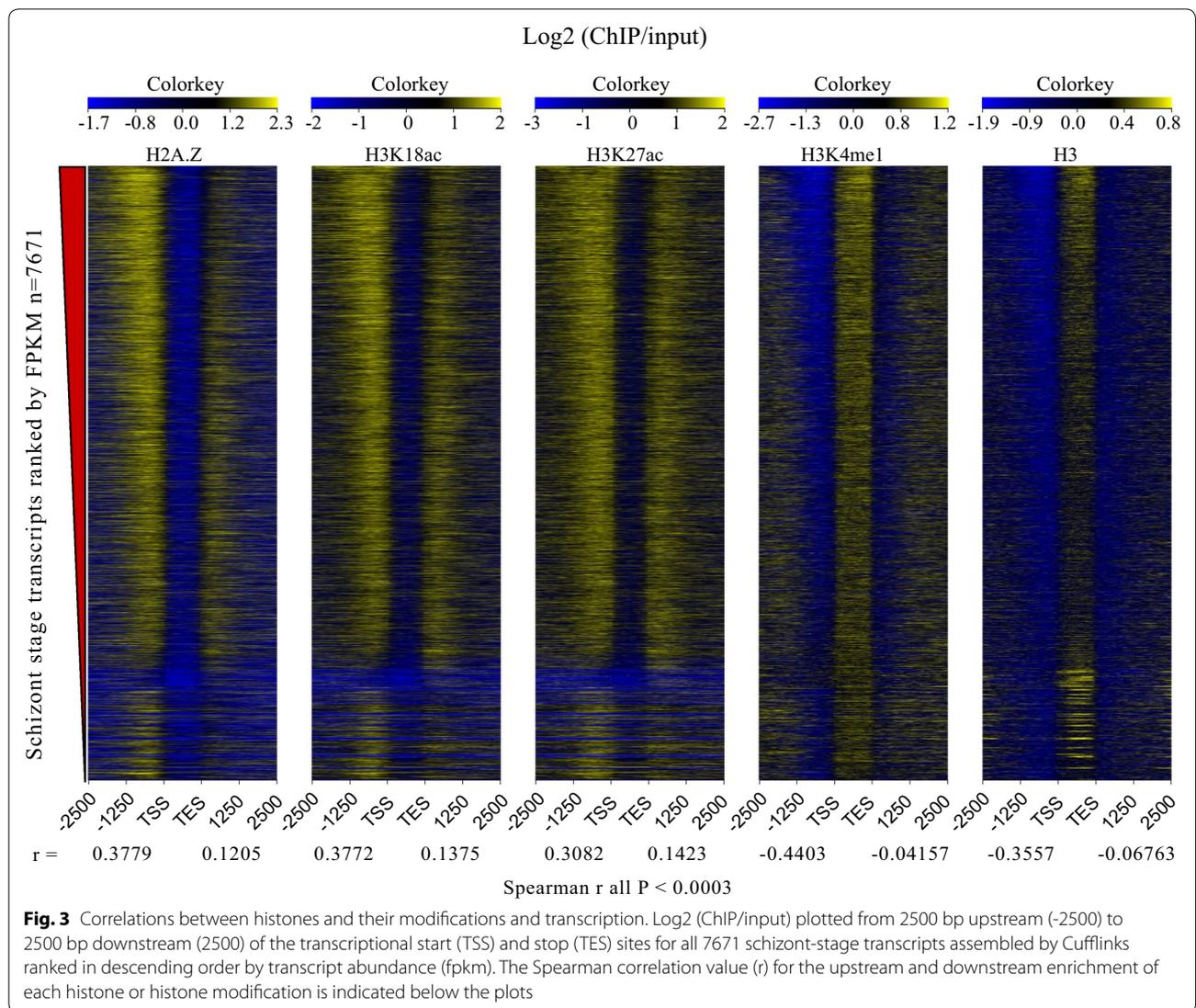
top tercile by expression were markedly depleted of upstream H3K4me1 compared to H3 (Additional file 8: Fig S6 D,E,I,J,N,O), indicating active depletion of H3K4me1 at the promoter of highly expressed genes. This would be consistent with additional methylation of H3K4 and previous reports of H3K4me3 enrichment at regulatory sequences [3, 36, 38].

Although on average intergenic H3K4me1 was depleted, 76 intergenic peaks of H3K4me1 were detected in schizont stages. Bidirectional transcription originated from the H3K4me1 peaks and the peaks also colocalised with relative depletion of H3K4me3 in a published schizont dataset [38] (Fig. 4), which are characteristics of some enhancer sequences [11]. However, H3 levels were increased at these peaks while DNA accessibility was low, as indicated by published ATACseq signals [41] (Fig. 4); both of these properties were consistent with increased nucleosomal occupancy and thus inconsistent with enhancer function. Furthermore, Pf H2A.Z, H3K27ac and H3K18ac were not enriched at the intergenic H3K4me1 peaks and transcription of the closest, downstream genes (median 11.7 rpkm IQR 1.4, 70.5 rpkm) was similar to the average of all the other protein coding genes in the *P. falciparum* genome in schizonts (median 19.9 IQR 3, 72.8 rpkm Mann–Whitney test $p=0.2636$). Interestingly, GC content was very high for intergenic sequence.

To further investigate whether the H3K4me1/H3K4me3 ratio was informative of putative enhancer function, we identified peaks of H3K4me1 calculated relative to published H3K4me3 (Additional file 9: Fig S7A) [38]. This set of peaks included most of the H3K4me1 peaks identified by comparison to input (Additional

file 9: Fig S7C) and had a similar profile (Additional file 9: Fig S7A) for all characteristics described above except that the adjacent downstream genes were expressed at significantly lower levels (median 8.9 IQR 0.7, 43.1 rpkm Mann–Whitney $p=0.0009$) than all of the other protein coding genes in the *P. falciparum* genome in schizonts. Overall, the H3K4me1/input and H3K4me1/H3K4me3 data suggested that intergenic H3K4me1 enrichment was not marking transcription regulating sequences. Indeed, the AT content and chromatin characteristics of these peaks were more similar to coding than intergenic sequence.

A recent study reported that intergenic enrichment of H3K4me1 in *P. falciparum* marked regulatory regions [48]. In this report regions of H3K4me1 enrichment were divided into “strong” and “weak” peaks, but only 259 of the 462 “strong” peaks were longer than one nucleotide. None of the H3K4me1 intergenic peaks we detected overlapped with these 259 previously reported peaks (Additional file 9: Fig S7C) [48]. However, H3K4me1 was also enriched relative to adjacent intergenic sequence in our ChIP data at these 259 sites, albeit at a lower level than at the intergenic H3K4me1 peaks we found de novo, probably reflecting natural variation between the parasites (Additional file 9: Fig S7B). Although not highly correlated (Additional file 9: Fig S7D), the broad chromatin and transcriptional patterns of the published 259 H3K4me1 intergenic peaks (Additional file 9: Fig S7B) were similar to those of the intergenic H3K4me1 peaks discovered de novo in this study (Fig. 4), suggesting similar features were detected in both studies.

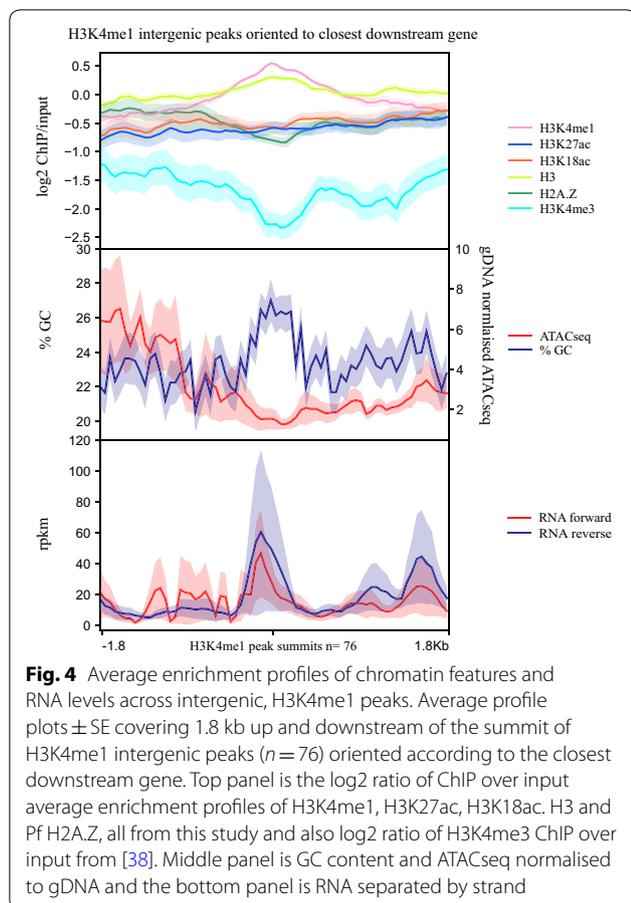


H3K18ac and H3K27ac upstream enrichment dynamically correlates with expression of genes upregulated in mature parasites

P. falciparum genes are dynamically regulated across the IDC. Transcription of highly expressed genes can still vary through the cycle, but even at their minimum level of expression many may still be in a high percentile of genes by transcript level. Thus the correlation between steady-state mRNA levels and upstream enrichment of a histone mark may not reflect a dynamic association with gene expression, as has been noted before for Pf H2A.Z [59]. To assess whether Pf H2A.Z, H3K27ac, H3K18ac and H3K4me1 were dynamically associated with gene expression, their enrichment was examined across the coding sequences and for 2500 bp upstream and downstream of genes that were differentially expressed throughout the IDC. Average enrichment profiles were

plotted for gene sets that were expressed in the top quartile in a particular IDC stage and were expressed at least threefold more in that stage than in another, compared IDC stage (Additional file 10: Table S3). The stage-specific transcription of these genes was confirmed using a published nascent RNA dataset [60] (Additional file 11).

H3 is depleted in non-coding sequence due either to decreased nucleosomal occupancy [61] or to technical artefacts relating to the high AT content of intergenic sequence affecting nucleosome stability during experimental procedures [43]. Although our H3 ChIP protocol cannot distinguish between these alternatives, we did control for the level of histone modifications per nucleosome by normalising to H3 levels. Upstream, intergenic enrichment of all three H3 modifications and H2A.Z were statistically different between stages except for H3K18ac levels in trophozoite stage-expressed genes,



which were not different between trophozoite and schizont stages (Wilcoxon matched-pairs signed rank test, Additional file 12: Table S4). However, the variation in the magnitude of these differences suggested some were not biologically significant. H3K4me1 was depleted upstream of dynamically regulated genes regardless of their expression level (Fig. 5, Additional file 13, 14: Fig S8, S9). Compared to ring stages, H3K18ac and H3K27ac levels clearly increased in trophozoites and schizonts upstream of genes that had peak expression in mature trophozoite and schizont-stage parasites, whereas Pf H2A.Z levels were steady and did not obviously vary with differential expression (gene set 1, Fig. 5 and gene set 3, Additional file 13: Fig S8). Conversely, Pf H2A.Z, H3K18ac and H3K27ac were all enriched upstream of genes that had peak expression in ring stages with no obvious variation compared to trophozoites and schizonts (gene set 2, Fig. 5, and gene set 4, Additional file 13: Fig S8). The dynamic, expression-associated enrichment was also apparent upstream of schizont-expressed genes when compared to trophozoites, but the level of enrichment was already relatively high in trophozoites (gene set 5, Additional file 14: Fig S9). In contrast, the H3K18ac and

H3K27ac enrichment upstream of trophozoite-expressed genes did not obviously decrease in schizonts (gene set 6, Additional file 14: Fig S9).

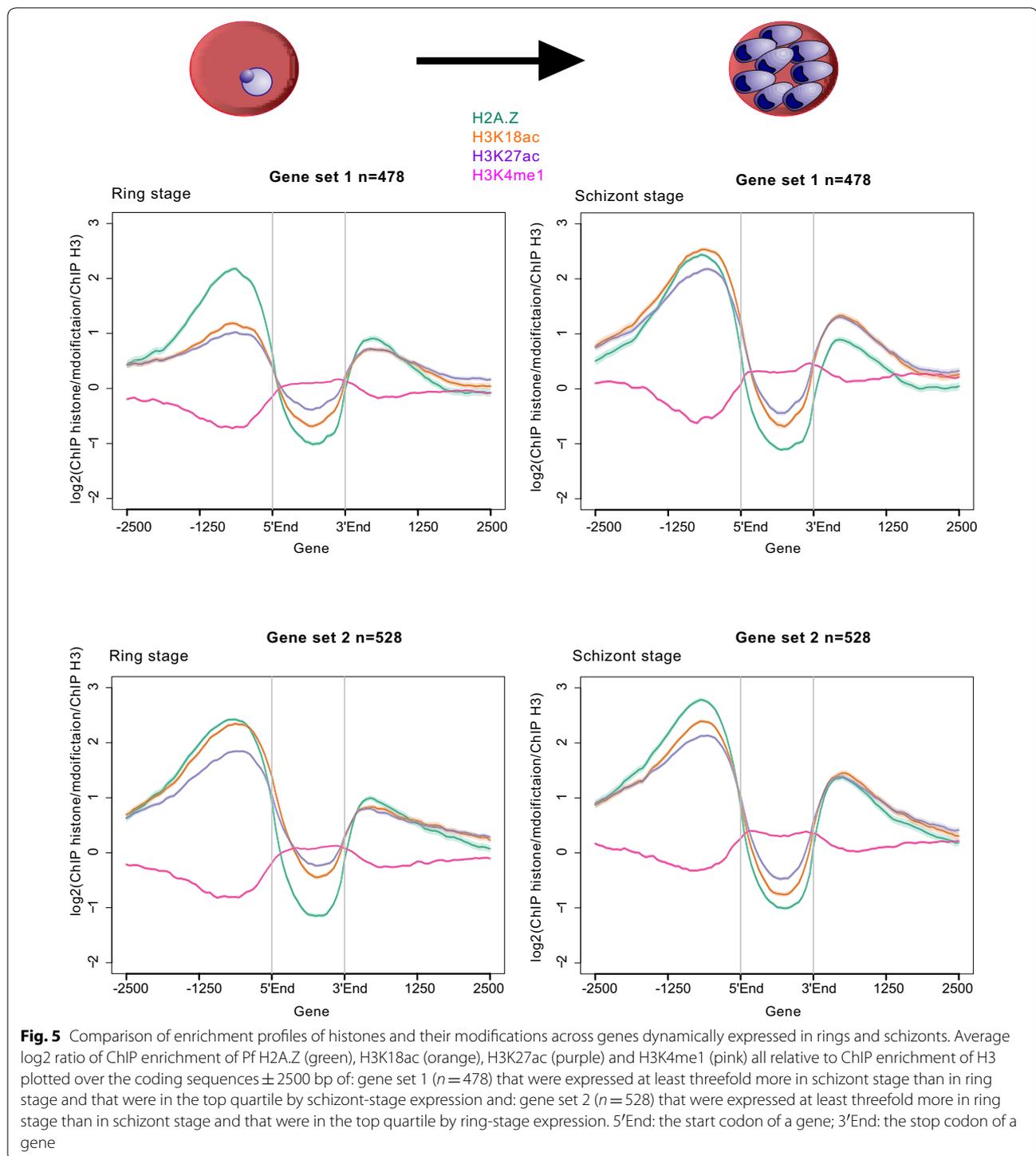
The patterns of H3K18ac and H3K27ac enrichment may indicate that these histone modifications accumulate during mitosis (approximately 27–40 h post-invasion [62]) in nucleosomes upstream of all genes to be expressed until the next round of replication in the following IDC. They were also consistent with western blot evidence of peak H3K18ac and H3K27ac abundance from 32 to 40 h post-invasion (Fig. 1). The modifications appeared to be removed in G1 phase after invasion, but only upstream of those genes that had been expressed and silenced since the commencement of mitosis, i.e. genes expressed in trophozoites and schizonts. Thus, in ring stages, the histone modifications remained only in ring stage-expressed genes. Therefore, we conclude that these histone modifications precede gene expression and may mark genes for subsequent expression, while their removal (by unknown HDACs) occurs after repression of the downstream gene.

Colocalised, intergenic peaks of H3K18ac, H3K27ac and Pf H2A.Z overlap, or are close, to putative regulatory regions

The dynamic, enrichment of H3K27ac and H3K18ac upstream of expressed genes in *P. falciparum* was consistent with their collective marking of regulatory sequences, as they do, together with H2A.Z, in yeast and metazoans. To investigate this possibility, Pf H2A.Z peaks that intersected H3K18ac and H3K27ac peaks were filtered to retain only the 2196 ring stage and 2645 schizont stage, intergenic, colocalised peaks (ICPs). Existing datasets of nucleosomal occupancy (DNaseq of micrococcal nuclease-treated chromatin) [43], DNA accessibility (ATACseq) [40, 41] and transcriptional start sites [42] were examined for associations with possible regulatory sequences defined by ICPs.

The inferred nucleosomes at the Pf H2A.Z summits of the 2645 ICPs in schizont stages (Fig. 6a) overlaid a pronounced peak of nucleosomal occupancy previously mapped by MNaseSeq [43] (Fig. 6b). The nucleosome occupancy at the ICPs was significantly higher than the occupancy at all intergenic regions combined (Fig. 7a). Thus, the nucleosomes defined by the ICPs presumably represented well-positioned nucleosomes because they were conserved between these studies. From matched RNASeq 35% of the 8415 active TSSs intersected with colocalised Pf H2A.Z, H3K18ac and H3K27ac peaks.

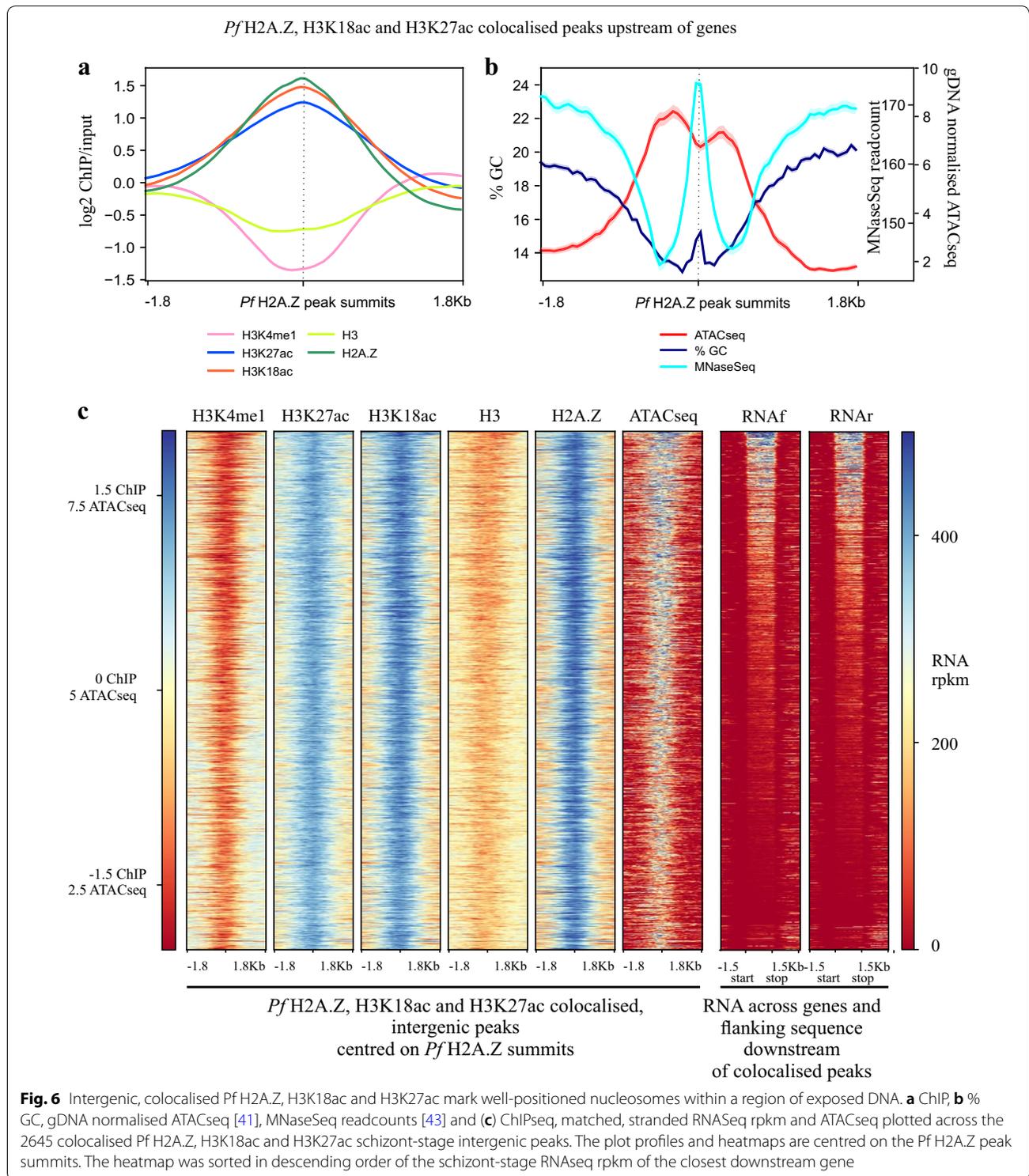
The association between the ICPs and TSSs was confirmed with an orthogonal TSS sequence tag dataset [42]. In both ring stages (10 hpi) and schizonts (42 hpi), the ICPs intersected with more TSS sequence tags [42] than



did other intergenic sequences (Mann–Whitney U test $p < 0.0001$) (Fig. 7b).

The association between the TSSs and the ICPs suggested an association with regulatory sequences. ATAC-seq reveals exposed DNA presumed to be available to

bind transcription factors. The 2645 schizont-stage ICPs overlaid ATACseq peaks [41] and regions of low GC content (Fig. 6b, c). Similarly, nearly all ATACseq peaks [41] were enriched in H3K27ac, H3K18ac and Pf H2A.Z (Additional file 15: Fig S10), the latter having been



shown previously [40]. The *Pf* H2A.Z summits of the ICPs overlaid a distinct dip in ATACseq reads and a distinct increase in GC content (Fig. 6b), probably reflecting reduced accessibility due to the *Pf* H2A.Z containing

nucleosome positioned at the summit peak, and possibly a role for GC in positioning of the *Pf* H2A.Z nucleosome. A slight spike in %GC at the *P. falciparum* TSS has been previously noted [43].

In schizont stages the Pf H2A.Z summits of the ICPs were on average 1.3-fold closer (4400 bp) to schizont-stage ATACseq peaks [40] than were summits of intergenic, non-colocalised Pf H2A.Z peaks (5900 bp) ($p < 0.00005$ Welch's t test with Bonferroni correction) (Fig. 7c). The trophozoite-stage ATACseq peaks were also 1.2-fold closer to schizont ICPs compared to schizont intergenic non-ICPs, but the difference was much less significant ($p = 0.005$) whilst ring-stage ATACseq peaks were not closer to schizont ICPs compared to schizont intergenic non-ICPs. These differences are consistent with a stage-dependent proximity of the ICPs to the stage-matched ATACseq peaks.

To further investigate the associations between Pf H2A.Z, H3K27ac, H3K18ac and a transcription factor (PfAP2-I), its binding sites (ATACseq) and co-activator (PfbBDP1), the well characterised PfAP2-I/PfbBDP1-regulated invasion genes were analysed [8, 9, 41]. Pf H2A.Z, H3K27ac and H3K18ac were all enriched around the subset of schizont-stage ATACseq peaks [41] that colocalise with enrichment of the invasion gene regulators PfAP2-I [8] and PfbBDP1 [9] upstream of invasion genes ($n = 151$) (Fig. 8). The patterns of H3K27ac and H3K18ac enrichment were consistent with enrichment at the nucleosomes flanking the exposed DNA indicated by the ATACseq peak with bound PfAP2-I. However, PfbBDP1 and Pf H2A.Z levels did not dip at the ATACseq peak summit (Fig. 8). Thus, for these genes it seems that Pf H2A.Z containing nucleosomes relatively depleted in the H3 acetylations were distributed across the summit of the ATACseq peak and some of these were bound by PfbBDP1. Interestingly, PfbBDP1 was also more enriched upstream than downstream of the ATACseq summit suggesting that PfbBDP1 additionally binds nucleosomes upstream of the PfAP2-I binding site.

Integrating ChIPseq and RNAseq to identify putative regulatory sequences

To investigate whether ICPs marked promoter or enhancer-like regulatory sequences, the ICPs were filtered by a set of stringent criteria to generate a set of sequences for further analysis (Additional file 16: Fig S11). For schizont stages, the 2645 ICPs were filtered for differential enrichment in

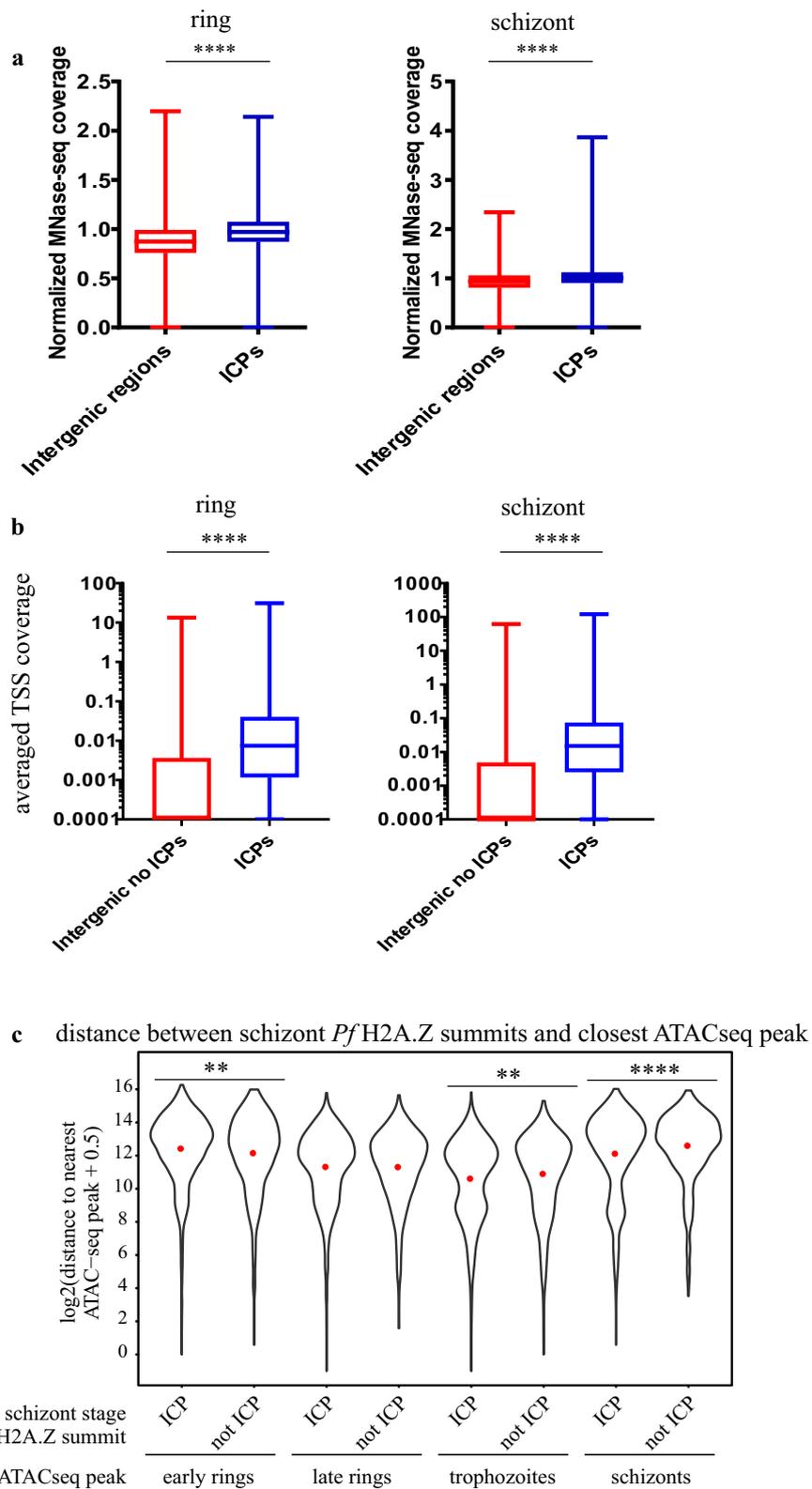
H3K18ac and H3K27ac in schizonts compared to rings ($n = 1371$) and then filtered for the presence of an adjacent downstream gene within 2500 bp that was in the top quartile by expression in schizonts and upregulated at least threefold in schizonts compared to ring stages.

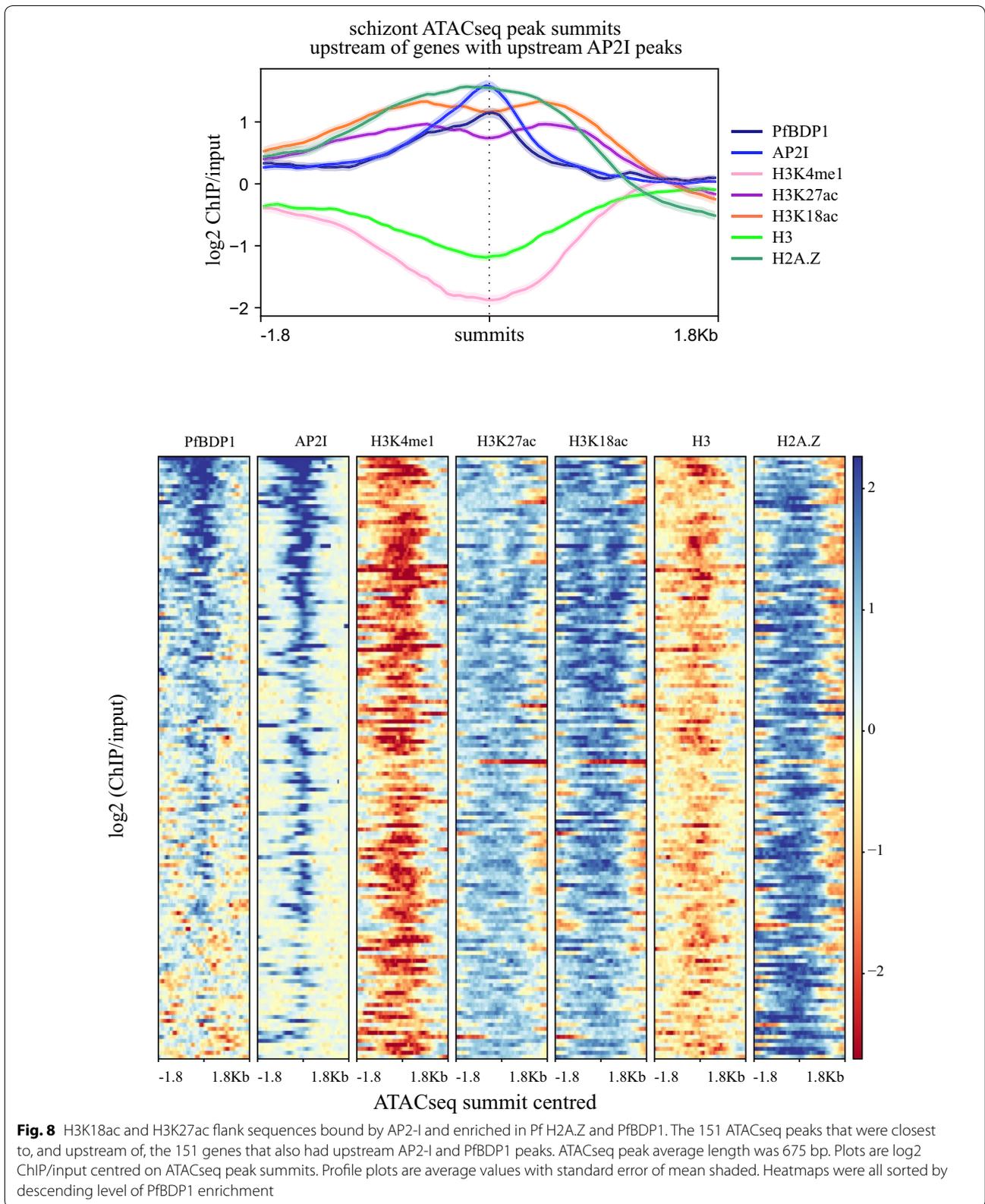
To identify potential promoter and enhancer-like sequence pairs marked by the ICPs, we further filtered the set of genes downstream of the filtered peaks for the presence of two sets of upstream, differentially enriched, colocalised ChIPseq peaks that were both closer to the downstream gene than any other gene. We hypothesised that the gene-proximal upstream ICP was a putative promoter whilst the gene-distal upstream ICP was a putative enhancer-like sequence. This generated a set of 30 differentially expressed genes that each had tandem upstream putative, regulatory sequences with differential enrichment of H3K18ac and H3K27ac and constitutive enrichment of Pf H2A.Z (Fig. 9a). A similar approach was used to identify 24 ring stage-expressed genes with tandem sets of upstream ICPs (Fig. 9b). However, the filtering for differential H3K18ac and H3K27ac enrichment was omitted for the ring-stage peaks because enrichment levels of these modifications are maintained throughout the IDC upstream of genes that are upregulated in ring stages (Fig. 5 gene set 2).

In both ring and schizont stages it was unlikely that the putative enhancer-like sequence exerted a positive regulatory effect upon the adjacent upstream gene because they were repressed or transcribed at low levels compared to the genes adjacent and downstream of the tandem ICPs (median, [IQR] rpkm of upstream/downstream genes in ring stages: 13 [3, 134]/606 [171, 978] and schizont stages: 41 [4, 224]/251 [134, 10537]; Wilcoxon test comparison between rpkm of upstream and downstream genes in rings $p = 0.0002$ and schizonts $p = 0.005$). In both stages, the putative promoters, but not enhancers, intersected with more TSS tags [42] than did intergenic sequences without ICPs (Kruskal-Wallis test both $p < 0.0001$, Dunn's multiple comparisons test all $p < 0.0003$). The putative promoters also intersected more sense strand TSS tags than antisense (Dunn's multiple comparisons test all $p < 0.0001$), but the putative enhancers did not (Fig. 9c). This showed

(See figure on next page.)

Fig. 7 Intergenic, colocalised peaks of Pf H2A.Z, H3K18ac and H3K27ac compared to other intergenic regions have increased nucleosomal occupancy and increased density of TSSs and are closer to regions of exposed DNA. **a** Per site normalised *P. falciparum* MNase-seq read coverage [43] from ring stages (15 h post-invasion) and schizont stages (40 h post-invasion) plotted across all 5567 intergenic regions and across the 2197 ring stage and 2645 schizont stage intergenic, colocalised H2A.Z, H3K18ac and H3K27ac ChIP peaks. Whiskers are minimum and maximum values. Mann Whitney U test **** $p < 0.0001$. **b** Average, stranded, *P. falciparum* TSS tag coverage at 10 and 42 h post-invasion [42] at ICPs and at all other intergenic regions. (Mann-Whitney U test **** $p < 0.0001$). **c** Distance between ATACseq peaks [40] in early rings, late rings, trophozoites or schizonts and schizont-stage Pf H2A.Z summits of colocalised, intergenic peaks (ICP) or schizont-stage Pf H2A.Z, intergenic summits that did not colocalise with H3K18ac or H3K27ac peaks (not ICP), t test **** $p < 0.0001$, ** $p < 0.01$





that the putative promoters were driving sense transcription, consistent with promoter function, and that the putative enhancers did not have canonical promoter activity.

Gene-proximal colocalised peaks of Pf H2A.Z, H3K18ac and H3K27ac mark sequences that have promoter activity

To investigate whether the putative promoter and enhancer pairs had regulatory activity, they were cloned upstream of a nanoluciferase reporter gene and used to transiently transfect *P. falciparum* along with a control plasmid expressing firefly luciferase. Transfected parasites were assayed for chemiluminescence normalised to transfected plasmid copy number and parasite genome number by quantitative PCR (qPCR) of matched DNA samples. We used qPCR for normalisation because firefly luciferase activity was too low for accurate detection. Luminescence was determined for parasites expressing nanoluciferase driven by (1) each of three schizont- and three ring-stage putative promoters, selected randomly from the 54 putative promoter sequences; (2) the hsp86 promoter as a positive control; (3) an AT content-matched *P. falciparum* intergenic sequence that lacked any chromatin modifications indicative of regulatory activity and that was located upstream of a non-expressed gene to control for non-sequence specific effects and (4) a no DNA control (Additional file 17: Table S5). Each of the putative promoter sequences also had the cognate, putative enhancer sequence cloned upstream and these were analysed for enhanced luminescence expression compared to plasmids containing the cognate promoter with a scrambled version of the putative enhancer from PF3D7_1362000 cloned upstream to control for non-sequence specific effects.

Two of the schizont-stage putative promoters had greater promoter activity than the AT content-matched control sequence (ANOVA $p=0.0205$, Sidak's multiple comparisons test PF3D7_1362000 $p=0.0196$, PF3D7_0920700 $p=0.0228$) and the trend to greater activity approached significance for the third schizont-stage putative promoter (PF3D7_1460600 $p=0.0908$), but none of the three ring-stage putative promoters had greater activity than the AT content-matched control

sequence (Fig. 10). Curiously the hsp86 control promoter sequence had no greater activity than the AT content-matched control sequence, indicating that the control sequence itself had moderate promoter activity and that the gene-proximal ICPs had identified strong schizont-stage promoters. None of the three ring stage nor three schizont-stage putative enhancers affected activity of the cognate, putative promoter. We concluded that the distal ICPs in the paired sets of ICPs did not correspond to enhancer-like elements.

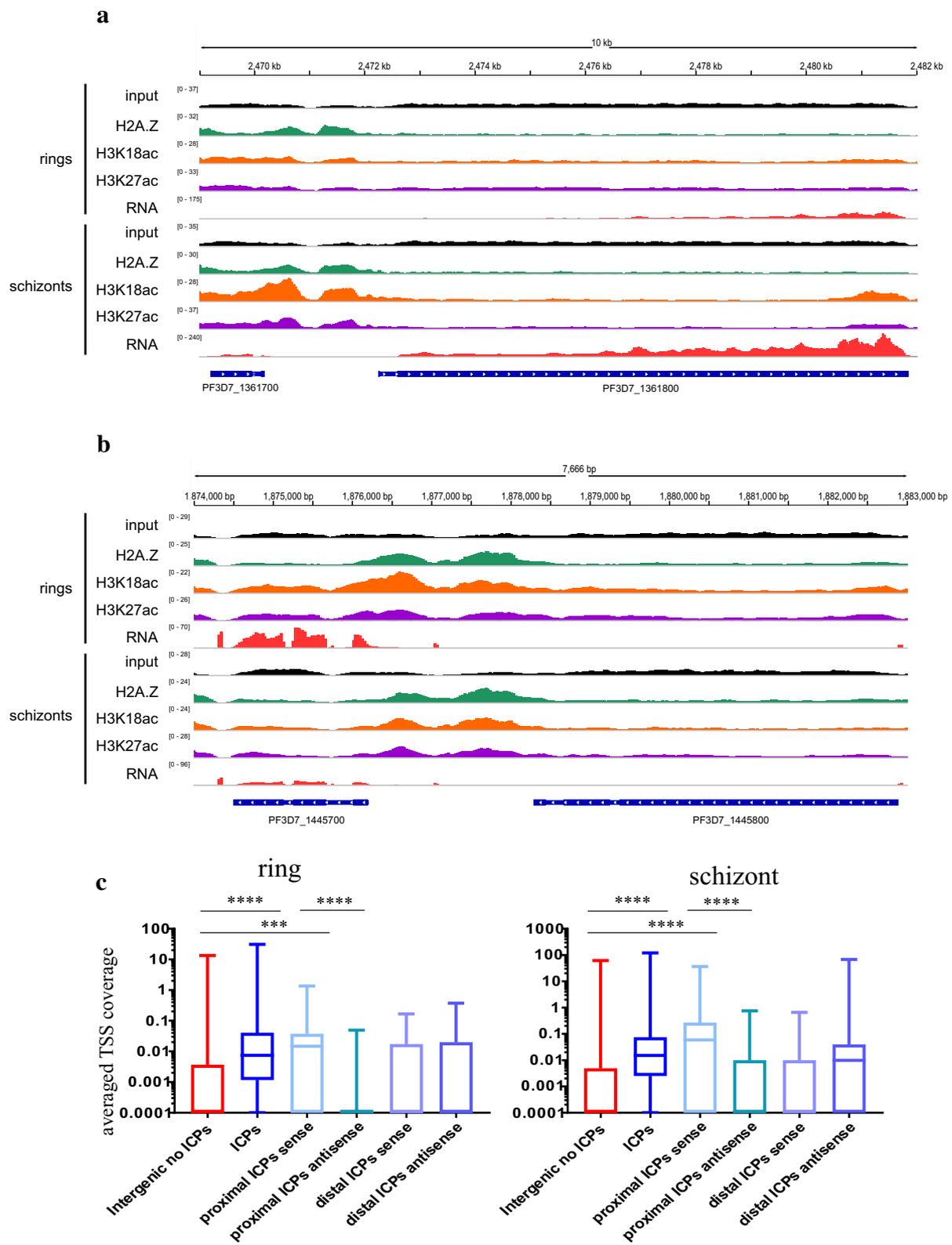
Discussion

Throughout the *P. falciparum* intra-erythrocytic developmental cycle Pf H2A.Z, H3K18ac and H3K27ac were enriched in intergenic sequence. A model is proposed in Fig. 11 that describes the patterns of enrichment of Pf H2A.Z, H3K18ac, H3K27ac and H3K4me1 throughout the intra-erythrocytic developmental cycle and their associations with gene expression. As previously reported, the intergenic enrichment of Pf H2A.Z correlated with steady-state expression levels, but Pf H2A.Z was not dynamically enriched at expressed genes [59]. In other eukaryotes, H2A.Z is required to mark promoters that are subsequently activated by other chromatin dynamics [63, 64]. H3K18ac and H3K27ac were also enriched at, and upstream of, the TSS but their intergenic enrichment dynamically correlated with gene expression. H3K18ac and H3K27ac thus marked active promoters and possibly additional regulatory elements, consistent with evidence from other eukaryotes [10–12, 20, 23–25, 65].

H3K18ac and H3K27ac appeared to accumulate at intergenic sites during asynchronous S phase and mitosis and were removed during the subsequent G1 phase following expression of the cognate, downstream gene (Fig. 11). Thus, these marks at a promoter preceded expression of the cognate gene. Increased accessibility of promoters has been previously noted to also precede their activity [40]. Maximum abundance of H3K27ac during mitosis and enrichment of H3K27ac during mitosis upstream of genes expressed at any stage during the following cell cycle has also been described in human cells [66]. H3K27ac enrichment at human

(See figure on next page.)

Fig. 9 Tandem, intergenic colocalised peaks upstream of expressed genes have characteristics of regulatory sequences. Mapped reads normalised for mapped library size from concatenated replicates of input DNA and ChIP of Pf H2A.Z, H3K18ac and H3K27ac and from matched RNAseq for schizont and ring stages. Mapped reads are plotted over the preceding gene, the upstream intergenic sequence and a gene (a) in the forward orientation and expressed in schizont stages and (b) in the reverse orientation and expressed in ring stages. c As per Fig. 7b) Average, stranded, *P. falciparum* TSS tag coverage at 10 and 42 h post-invasion [42] at ICPs, all other intergenic regions and gene-proximal and distal ICPs from tandem pairs of ICPs. For this comparison the proximal and distal ICPs were defined as ± 50 bp from the colocalized H2A.Z peak summit. Kruskal–Wallis test showed significant variation across all categories in both rings and schizonts ($p < 0.0001$), shown are P values for post hoc comparisons (Dunn's multiple comparisons test) **** $p < 0.0001$, ** $p < 0.001$



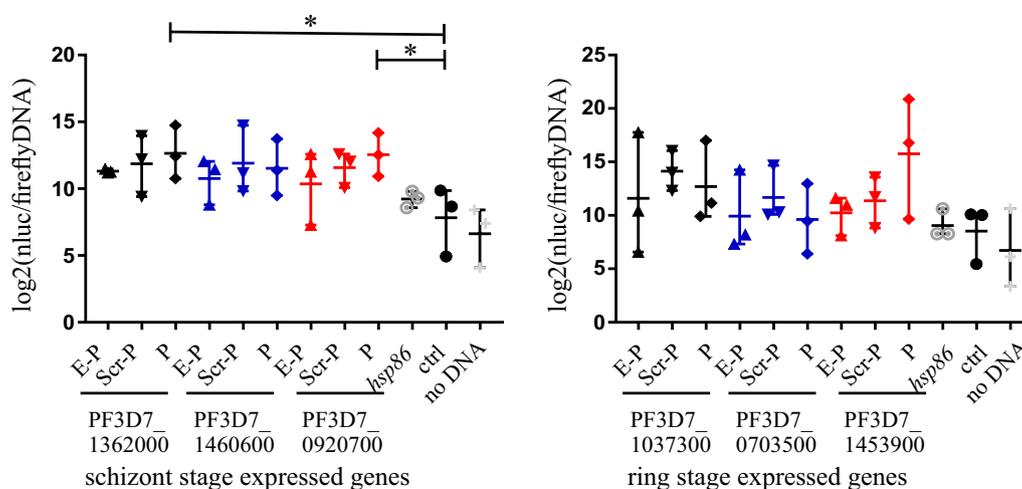


Fig. 10 Transient transfections to detect whether intergenic colocalised peaks have gene regulatory activity. Shown are Log₂ transformed nanoluciferase luminescence normalised to levels of firefly luciferase DNA and parasite genome copy number detected by qPCR. Each transfected plasmid encoded nanoluciferase with its transcription driven by upstream sequences consisting of: the putative enhancer upstream of the putative promoter (E-P); a scrambled version of the PF3D7_1362000 putative enhancer upstream of the putative promoter (Scr-P); the putative promoter alone (P); the *hsp86* promoter (*hsp86*); an AT-matched intergenic control sequence that had no regulatory effect on neighbouring genes during the asexual lifecycle (ctrl) and a mock transfection control (no DNA). ANOVA with post hoc Sidak's multiple comparisons test of the AT-matched control sequence to each of the putative promoter-alone constructs and the *hsp86* promoter alone. **p* < 0.05. Bars are the means, whiskers are the ranges

regulatory sequences in mitosis is also predictive of a transcriptional burst in human cells at the mitosis/G1 transition, leading to the proposition that H3K27ac may be important for cell cycle progression [67].

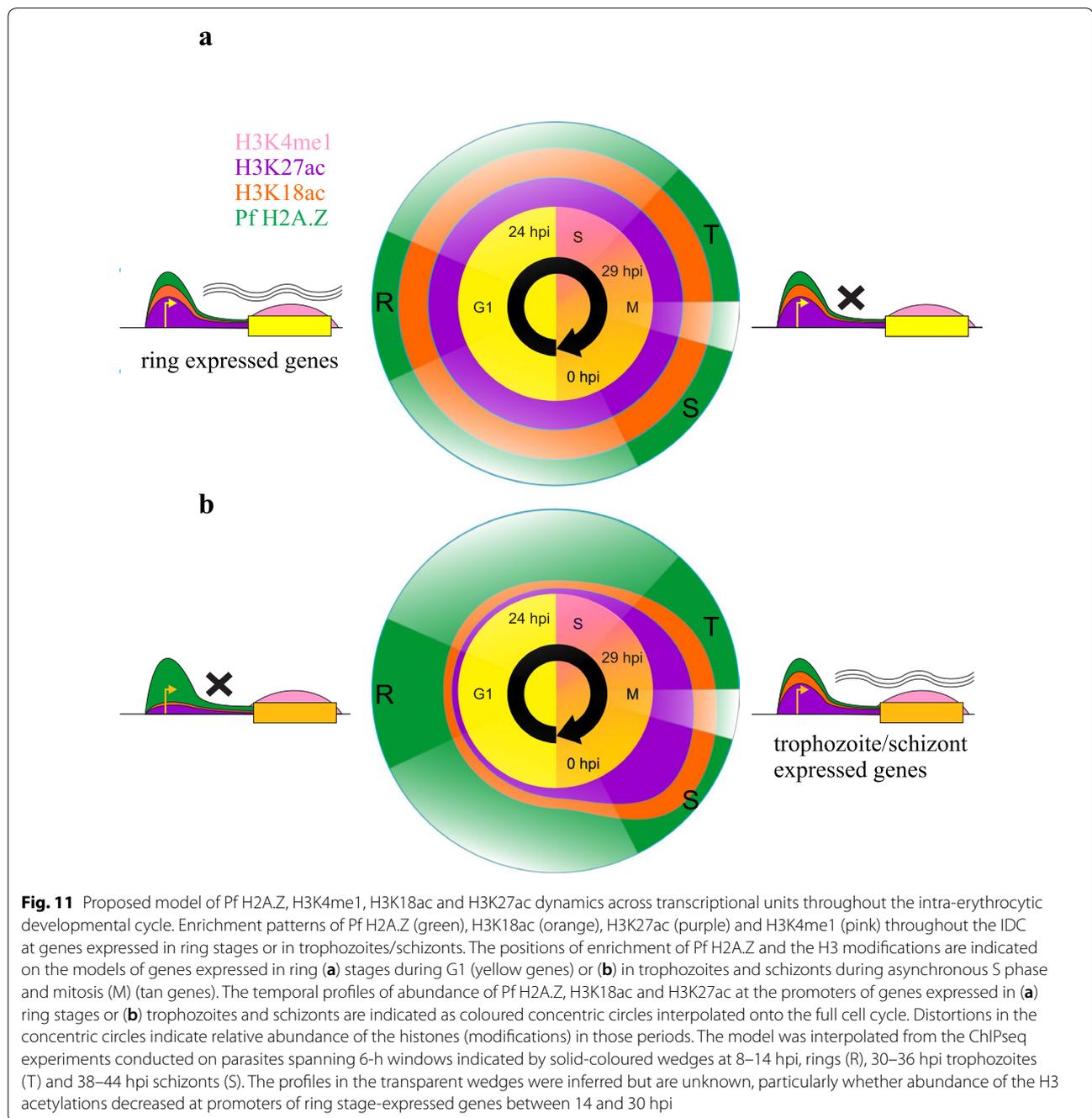
We had observed this pattern of H3K27ac enrichment before in a limited set of *P. falciparum* genes including active *var* genes [37]. The genome-wide pattern revealed here by ChIPseq does not exclude a role for H3K18ac and H3K27ac in epigenetic memory for *var* gene regulation, but the apparent persistence of these marks throughout the cell cycle is more probably a consequence of the genome-wide timing of deposition and removal of these histone modifications.

The low complexity of the *P. falciparum* intergenic sequence with approximately 90% AT content has hindered attempts to identify gene regulatory sequences. Studies of *P. falciparum* genes using episomes with various truncations of upstream regions characterised core promoters around the TSS, but also identified additional sequences starting between 300 and 1400 bp upstream of the TSSs that were required for efficient promoter activity [68–73], that conferred stage-specific expression [72, 73], that bound nuclear *trans* factors [69, 71] and that were functional regardless of orientation, a classical eukaryotic enhancer characteristic [69–72]. These repeated demonstrations of *cis* regulatory elements that act upon a core promoter are convincing evidence of enhancer or UAS-like elements that bind

specific transcription factors to regulate gene expression in *P. falciparum*.

As the enrichment levels of intergenic H3K18ac and H3K27ac were highly correlated and dynamically associated with downstream gene expression, we attempted to use these modifications to define potential regulatory sequences in the intergenic regions upstream of genes. This approach could not detect enhancer-like sequences that were distant from genes or in gene-bodies, although Hi-C studies suggest the former are unlikely to function in asexual blood stage *P. falciparum* [45, 46]. Throughout eukaryotes H2A.Z is also enriched around regulatory sequences so colocalised peaks of Pf H2A.Z, H3K18ac and H3K27ac (ICPs) were analysed for characteristics of regulatory sequences. The coverage of published TSSs [42] by the ICPs was consistent with them marking regulatory sequences. In *P. falciparum*, intergenic nucleosomal occupancy peaks at the well positioned +1 nucleosome that overlies the TSS, this is immediately downstream of a short nucleosome depleted region that is itself downstream of a region of above average occupancy of poorly positioned nucleosomes [43]. The ICPs overlaid a sharply defined MNaseSeq peak representing a nucleosome with a position conserved between the parasites analysed in these orthogonal studies, presumably equivalent to the strongly positioned +1 nucleosome at the TSS [43].

The observed stage-dependent proximity of ICPs to ATACseq peaks and the patterns of histone modification



enrichment at ATACseq peaks suggest that these ICPs demarcate regulatory sequences that bind transcription factors. The example of the AP2-I/PfBDP1 association with ICPs clearly demonstrated enrichment of these histone modifications at nucleosomes flanking the bound transcription factor at the ATACseq peaks. PfBDP1 was recently shown to have promiscuous affinity for the Pf H2A.Z acetylations [74] that alone, or in combination with other, untested histone acetylations, may be

actually binding PfBDP1 explaining the relative depletion of H3K18ac and H3K27ac at the summits of the PfBDP1/AP2-I peaks.

We further filtered ICPs by successive, stringent criteria to generate a small set of putative, proximal promoter and distal enhancer-like regulatory sequence elements. Our attempt to infer the presence of regulatory sequences from the presence of expression-associated ICPs was clearly corroborated by the promoter activity of

sequences underlying the schizont-stage, gene-proximal ICPs in transient transfection experiments. However, the ring stage sequences and the schizont-stage putative enhancers did not have regulatory effects on the episomal reporter. We concluded that the ICPs we analysed did not define sequences with enhancer function that was detectable using our transient transfection system. It is probable that ATACseq datasets contain the identities of enhancer or UAS-like elements in *P. falciparum*. However, the large number of ATACseq peaks detected in those studies [40, 41] means that intersection with other orthogonal datasets is required to filter false positives and identify the enigmatic, regulatory elements that lie outside the core promoter. The ChIPseq datasets we provide here would be a valuable contribution to such a meta-analysis.

The intergenic depletion and coding sequence enrichment of H3K4me1 was consistent with previous reports from *P. falciparum* [39], yeast [13, 75], and humans [18]. However H3K4me1 was not enriched upstream of the TSS as it is in yeast [13, 75, 76], nor at nucleosomes flanking the TSS and at enhancers as it is in humans [16, 18, 21]. Intergenic levels of H3K4me1 inversely correlated with downstream gene expression consistent with the association in yeast between H3K4me1 around the TSS and gene repression [65]. However, upstream H3K4me1 depletion did not vary with dynamic gene expression indicating that H3K4me1 was not marking regulatory sequences. In *P. falciparum* H3K4me3 is largely not dynamically enriched at active promoters, but instead levels increase throughout the lifecycle with the greatest intergenic enrichment in trophozoites and schizonts at the most highly expressed genes [38]. This pattern inversely mirrors the profile of intergenic H3K4me1, so we concluded that the accentuated depletion of H3K4me1 at highly expressed promoters was likely a consequence of additional methylation of H3K4 to H3K4me3 and that H3K4me1 was not specifically a feature of regulatory sequences in *P. falciparum*.

A recent study reported that intergenic enrichment of H3K4me1 in *P. falciparum* marked regulatory regions [48]. However, the intergenic sequences that were enriched in H3K4me1 in that study had chromatin characteristics more like coding sequence than regulatory sequence and they did not lie upstream of highly expressed genes, making it unlikely that they had enhancer function. These sequences may be analogous to the GC-rich (upstream) uORFs, previously implicated in post-transcriptional regulation of *var* genes [77, 78], although none of the H3K4me1 peaks lay upstream of a *var* gene.

Conclusions

In conclusion, we show that the histone modifications H3K18ac and H3K27ac mark promoters preceding and during gene expression, that they demarcate regulatory sequences defined by ATACseq and that together with colocalised *Pf* H2AZ they identify schizont stage sequences with promoter activity. In contrast, H3K4me1 did not appear to be associated with intergenic sequences exerting a regulatory effect, but the small number of intergenic H3K4me1 peaks identified did have a curious chromatin and GC composition suggestive of coding sequence, which may warrant further investigation. These findings contribute to the understanding of chromatin structure at the poorly defined gene regulatory elements of *P. falciparum* and how the contribution of histone modifications to gene regulation may be influenced by the *P. falciparum* cell cycle.

Materials and methods

Parasite culture

The *P. falciparum* 3D7 clone was cultured [79] using a modified method. Briefly, parasites were grown in RPMI-HEPES supplemented with 0.5% AlbuMAX[®] II Lipid-Rich BSA (Life technologies, 11,021), 0.2% sodium bicarbonate and O+ red blood cells at 3% hematocrit. Parasites were incubated at 37 °C in 5% CO₂, 1% O₂ and 94% N₂. Ring-stage parasites were synchronised to an approximately 4-h window by two incubations with 5% D-sorbitol [80] separated by a 12-h interval.

Antibodies

Primary antibodies employed in dot blot, immunofluorescence assay, western blot and ChIP assays in this study were rabbit anti-Pf H2A.Z and pre-immune rabbit serum [50], rabbit anti-H3 (Abcam, Ab1791), rabbit anti-H3K4me1 (Abcam, Ab8895), rabbit anti-H3K27ac (Abcam, Ab4729), rabbit anti-H3K18ac (Abcam, Ab1191) and rabbit control IgG (Abcam, Ab46540). Secondary antibody used for dot blot and western blot was goat anti-rabbit IgG antibody conjugated with horseradish peroxidase (HRP) (Invitrogen, A16110). Secondary antibody for immunofluorescence assay was Alexa Fluor[®] 488 anti-rabbit antibody (life technologies).

Western blots

Parasites at 5–10% parasitemia were harvested at 8-h intervals, erythrocytes were lysed in 0.075% saponin and parasite pellets then dissolved in 2 × reducing Laemmli buffer (125 mM Tris HCl pH 6.8, 20% glycerol, 4% SDS, 0.005% bromophenol blue, 5% beta-mercaptoethanol) at a concentration of 5 × 10⁶ infected erythrocytes/μl. Reduced parasite lysates were denatured at 95 °C for 5 min and approximately 1 × 10⁷ infected erythrocytes

were separated by SDS-PAGE on NuPAGE® Novex® 4–12% Bis-Tris protein gel (Life technologies, NP0323BOX) in 1 × NuPAGE® MES SDS running buffer (Life technologies, NP0002) at 150 V for 70 min. Proteins were transferred to an Amersham HYBOND-ECL membrane (GE Healthcare, 45-000-929) in cold transfer buffer (192 mM glycine, 25 mM Tris, 3.5 mM SDS, 20% methanol) at 400 mA for 1 h and successful transfer was verified by Ponceau S (0.1% Ponceau S, 5% acetic acid) staining. The membranes were incubated with primary and secondary antibodies and bound antibodies were detected using Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore, WBKLS0500).

Dot blot

Custom-made, biotinylated *P. falciparum* peptides PfH3 unmodified 1 (residues 1–21), PfH3 unmodified 2 (residues 21–43), PfH3 unmodified 3 (residues 47–65), PfH3K4me1 (residues 1–21), PfH3K9ac, PfH3K14ac, PfH3K18ac, PfH3K27ac, PfH3K23,27ac and PfH3K56ac and biotinylated human peptides HuH4 (residues 1–21) (Epigentek, R-1007-100), HuH4K20me1 (Epigentek, R-1054-100), HuH4K20me3 (Epigentek, R-1056-100) and nonbiotinylated human peptides HuH3K4me3 (Abcam, Ab1342), HuH3K36me3 (Abcam, Ab1785) and HuH3K9me3 (Abcam, Ab1773) were used for testing antibody specificity. All *P. falciparum* peptides and nonbiotinylated human peptides were used at 0.25 mg/ml except for HuH4K20me1 and HuH4K20me3 which were used at 0.07 mg/ml while HuH4 was 0.065 mg/ml. All peptides were then tested at further dilutions of 1:4, 1:8, 1:16 and 1:32. One microlitre of each dilution was loaded on an Amersham Hybond™ HYBOND-ECL membrane (GETM Healthcare, 45-000-929). One membrane was blocked with 5% w/v skim milk powder in TBS-T, washed, incubated with streptavidin–HRP diluted 1:50,000 in TBS-T, washed, and the membrane bound streptavidin–HRP was detected using Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore, WBKLS0500). The other membranes were blocked with 5% w/v skim milk powder in in TBS-T, each was incubated with a primary antibody, washed, incubated with secondary antibody conjugated to HRP, washed again then bound antibodies were detected using Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore, WBKLS0500).

Immunofluorescence analysis

Smears of asynchronous parasite cultures were prepared on glass slides and air-dried overnight. Cells were fixed and permeabilised by a mixture of ice cold 10% methanol/90% acetone for 5 min at room temperature and air-dried. Cells

were rehydrated in PBS, stained with primary antibodies, washed with PBS, stained with secondary antibody and DAPI (Sigma-Aldrich, D9542), washed, mounted in Pro-Long® Gold Antifade Reagent (Life technologies, P36934) and left to cure overnight. Imaging was performed using a Zeiss Fluorescence Upright Microscope (ZEISS Axioskop 2 mot plus) with the 100 × oil objective.

Cross-linked chromatin immunoprecipitation

Chromatin was isolated from ring-stage (8–14 h post-invasion), trophozoite-stage (30–36 hpi) and schizont-stage (38–44 hpi) parasites as previously described [50]. Briefly, infected erythrocytes were cross-linked with 1% formaldehyde at 37 °C for 10 min and quenched with 125 mM glycine. Erythrocytes were lysed with saponin and parasites were incubated for 30 min on ice in 10 mM HEPES pH 8.0, 10 mM KCl, 0.1 mM EDTA pH 8.0, 0.1 mM EGTA pH 8.0, 1 mM DL-dithiothreitol (DTT) and 1 × EDTA free protease inhibitors. IGE-PAL® CA-630 (Sigma-aldrich, I8896) was then added to a final concentration of 0.25% before lysing parasites with a Dounce tissue grinder (Sigma-aldrich, D8938-1SET, Pestle B). The nuclei were pelleted at 21,000 × g for 10 min at 4 °C and nuclei from 2 × 10⁹ ring-stage parasites, 1 × 10⁹ trophozoite-stage parasites or 4 × 10⁸ schizont-stage parasites were resuspended in 200 μl SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris pH 8.1, 1 × EDTA free protease inhibitors). Chromatin was sheared into 200–600 bp fragments by sonication in a Bioruptor UCD-200 (Diagenode) for 28 min (High, 30 s on, 30 s off) and diluted 1:10 in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris–HCl pH 8.1, 150 mM NaCl).

For each immunoprecipitation, diluted chromatin was pre-cleared using protein A/G beads (GE Healthcare Life Sciences, 17-5280-01&17-0618-01) at a beads-to-chromatin volume ratio of 1:200 for 1 h at 4 °C. Approximately 470 μl pre-cleared chromatin from 5 × 10⁸ ring-stage or 2.5 × 10⁸ trophozoite-stage or 1 × 10⁸ schizont-stage parasites were immunoprecipitated overnight at 4 °C with 8 μg antibody and 30 μl prewashed protein A/G beads. 1/10th of the volume of immunoprecipitated chromatin was also incubated overnight at 4 °C for use as the input control. Beads were washed once with low-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8.1, 150 mM NaCl), once with high-salt immune complex wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris–HCl pH 8.1, 500 mM NaCl), once with LiCl immune complex wash buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris–HCl pH 8.1) and twice with TE buffer (10 mM Tris–HCl pH 8.1, 1 mM EDTA pH

8.0) and co-immune-precipitated DNA was then eluted twice with freshly made elution buffer (1% SDS, 0.1 M NaHCO₃). Reversal of cross-linking was performed by incubating the eluate at 45 °C overnight with 500 mM NaCl. DNA was purified using the MinElute PCR purification kit (Qiagen, cat 28,006). All ChIP experiments were performed in three biological replicates.

ChIPed DNA library preparation for sequencing

ChIPed DNA libraries were prepared using the NEBNext[®] ultra DNA library prep kit for Illumina[®] (New England BioLabs, E7370). 6 ng input or chromatin-immunoprecipitated DNA was used and adaptor-ligated DNA was purified without size selection using Agencourt AMPure XP beads (Beckman coulter, A63881). The PCR amplification step was performed using 1 unit of high-fidelity KAPA HiFi DNA polymerase (KAPA biosystems, KK2101), 0.3 mM dNTPs, 0.5 μM NEBNext Universal PCR Primers for Illumina and 0.5 μM NEBNext Index Primer for Illumina (NEBNext[®] Multiplex Oligos for Illumina[®] Index Primers Set 1 (E7335) and Set 2 (E7500)) in 1 × KAPA HiFi buffer containing tetramethylammonium chloride for 12 PCR cycles [81]. The PCR conditions were 98 °C for 2 min for the initial denaturation followed by 12 cycles of 98 °C for 15 s and 65 °C for 2 min and a final extension at 65 °C for 5 min. Library clean-up was performed using 1 × AMPure beads and DNA libraries were eluted with 10 mM Tris–HCl pH 7.5.

RNA library preparation for sequencing

Stage-matched RNA from at least two biological replicates was extracted, gDNA digested and RNASeq libraries constructed as described previously [82]. Briefly, the aqueous phase from cells dissolved in TRIzol[®] reagent (Life Technologies, 15596018) was diluted in 70% ethanol and purified using RNeasy Mini columns (Qiagen, cat 74104). DNase digestion with the RNase-Free DNase Set (Qiagen, cat 79254) was repeated until qPCR of the DNased RNA gave a Ct greater than 34 for the *P. falciparum* *hsp70* gene. Total RNA (2.5 μg) was depleted of haemoglobin mRNA using the GLOBINclear[™] Kit (Life Technologies, AM1980) and mRNA then purified using magnetic oligo d(T) beads (NEBNext[®] Poly(A) mRNA Magnetic Isolation Module, New England BioLabs, E7490S). Indexed mRNA libraries were prepared using the NEBNext[®] Ultra[™] Directional RNA Library Prep Kit for Illumina[®] (New England BioLabs, E7420S). PCR enrichment was performed as per the ChIPseq method described above. The PCR conditions were 98 °C for 1 min for the initial denaturation followed by 12 cycles of 98 °C for 10 s and 65 °C for 1 min and a final extension at 65 °C for 5 min. Library clean-up was performed using

0.8 × AMPure beads and RNA libraries were eluted with nuclease-free water.

High-throughput sequencing

Input and chromatin-immunoprecipitated DNA libraries and mRNA libraries were subjected to quantification and quality check for purity and fragment size range using the Agilent Bioanalyzer and Qubit Fluorometer analyses prior to sequencing by the HiSeq2500 (Illumina) at the Melbourne Translational Genomics Platform. 126-bp paired-end reads were generated.

Bioinformatic analyses

Bioinformatic analysis was performed on Ubuntu 14.04.5 LTS (GNU/Linux 3.13.0–105-generic x86_64). Quality control of all raw sequencing data was performed using FastQC (version 0.11.2, <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Adaptor sequences were trimmed and reads were quality filtered using Trim Galore! (version 0.3.7, www.bioinformatics.babraham.ac.uk/projects/trim_galore/). ChIP-seq and RNA-seq reads from biological replicates from each stage mapped to *P. falciparum* 3D7 genome were visualised using IGV [83]. Data range for each track was normalised to the corresponding library size. Replicate reproducibility was checked using MultiBamSummary, PlotCorrelation and PlotPCA tools from DeepTools [84] with default settings.

RNAseq

RNA library sequencing reads were mapped to the *P. falciparum* 3D7 genome assembly using Tophat2 [85]. The resulting BAM files of biological replicates were concatenated using Samtools [86]. Transcript assembly was performed using Cufflinks on concatenated BAM files [87]. Transcripts or genes were ranked by isoform-level or gene-level expression values presented as fragments per kilobase of exon per million reads mapped (FPKM) in each parasite stage during the IDC. In each stage, genes with Cufflinks IDs were divided into four groups by expression levels, i.e. three terciles and silent genes. Any Cufflinks gene ID that could not be assigned to any *P. falciparum* 3D7 gene or was assigned to multiple genes was omitted from the analysis. Differential gene expression analysis between any two stages of the IDC was performed using Cuffdiff [87].

ChIPseq

Input DNA and chromatin-immunoprecipitated DNA library sequencing reads were mapped to the *P. falciparum* 3D7 genome (PlasmoDB v12) using Subread [88]. The resulting BAM files of biological replicates were concatenated using Samtools [86]. H2A.Z, H3K18ac, H3K27ac and H3K4me1 peaks from each stage were

called from concatenated BAM files using MACS2 ($q=0.01$) [89]. NGSplot [90] was used to plot log2ChIP/ (input or H3 ChIP) across specified genomic regions as averaged line plots and heatmaps.

Schizont and ring-specific candidate enhancers and promoters and their cognate genes were identified using BEDTools [91]. Briefly, to identify schizont-specific candidate regulatory elements, schizont-stage colocalised, intergenic H2A.Z, H3K18ac and H3K27ac peaks were detected using 'bedtools intersect' and peak boundaries were defined using H2A.Z peaks. H3K18ac and H3K27ac ChIP-seq peaks that were differentially enriched in schizonts versus rings were identified using csaw FDR < 0.05 [92] and intersected with the schizont-stage colocalised, intergenic H2A.Z, H3K18ac and H3K27ac peaks using 'bedtools intersect'. Peaks were then filtered for the presence of a gene within 2500 bp downstream that was expressed at least threefold more in schizont stages than in ring stages and that were in the top quartile by schizont-stage expression (gene set 1 from Fig. 5) using 'bedtools window'. Finally, the closest differentially expressed genes to two upstream colocalised, intergenic, H3K18ac and H3K27ac peaks with differential stage enrichment were identified using 'bedtools closest'. The closer peak to the gene was the putative promoter while the further peak was the putative enhancer.

Comparisons with published datasets: nucleosome occupancy, transcription start sites, DNA accessibility, PfBDP1 and PfAP2-I

P. falciparum MNase-seq read coverage normalised per site in BEDGRAPH format at time points 15 h post-invasion and 40 h post-invasion was retrieved from the Gene Expression Omnibus GSE66185 [43]. Coordinates of total 5590 genes from *P. falciparum* 3D7 v12 were retrieved from PlasmoDB (<https://plasmodb.org>). Coordinates of total 5567 intergenic regions were obtained using "bedtools complement". The average score of BEDGRAPH records that intersected with *P. falciparum* intergenic regions and colocalised H2A.Z, H3K18ac and H3K27ac peaks was determined using "bedtools map". Mann-Whitney U test was performed using GraphPad Prism version 6.

Stranded *P. falciparum* TSS tag coverage at 10 and 42 h post-invasion were retrieved from the Gene Expression Omnibus GSE68982 [42]. 2196 ring- and 2943 schizont-stage intergenic, colocalised peaks were identified as described above. For this comparison the peaks were defined as ± 50 bp from the colocalised H2A.Z peak summit. Mann Whitney U test or Kruskal-Wallis test and Dunn's multiple comparisons test were performed using GraphPad Prism version 6.

ATACseq data from *P. falciparum*-3D7-t40 replicate 1 GSM2789028 was mapped to the *P. falciparum* 3D7 strain genome release 12 version 3 (PlasmoDB) using BWA-MEM [93] as per [41] and peaks were called using MACS2 [89]. The same schizont-stage ATACseq data normalised to gDNA was converted from the supplied bedgraph file ncbi accession GSM2789028 to bigwig format using bedGraphToBigWig (UCSC) [94] in bioconda [95]. The resulting bigwig file was used in subsequent graphical visualisations in deepTools [84].

The schizont-stage MNase-seq T40A sample from SRX885818: GSM1616491: comprised 5 paired-end fastq libraries SRR1813391 to SRR1813395 [43]. Adapters and low-quality ends (phred < 20) were trimmed using trim_galore (cutadapt) (version 0.3.7, www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and reads were then mapped to the *P. falciparum* 3D7 strain genome release 12 version 3 (PlasmoDB) by BWA-MEM v 0.7.17 [93] and filtered using samtools [86] for alignments with quality > 30 and to remove alignments that were a PCR or optical duplicate or were a secondary or supplementary alignment (flags 256, 1024, 2048) as per [43]. The bam files were concatenated and used to make bigwig files for analysis in deepTools [84].

The schizont-stage AP2I ChIPseq and input data [8] from GEO: GSE80293 comprising 3 replicate schizont-stage AP2I ChIP and 3 matched input fastq files were trimmed of adapters and low-quality ends (phred < 20) using trim_galore (cutadapt) (version 0.3.7, www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and then reads were mapped to the *P. falciparum* 3D7 strain genome release 12 version 3 (<https://plasmodb.org>) by BWA-MEM v 0.7.17 [93] and filtered using samtools [86] for alignments with quality > 30 and to remove alignments that were a PCR or optical duplicate or were a secondary or supplementary alignment (flags 256, 1024, 2048) as per [8]. The bam files were concatenated and used to make bigwig files for analysis in deepTools [84]. The duplicate schizont-stage PfBDP1 ChIPseq and input data were processed as per [9] to generate bam files and MACS2 [89] called peaks.

In order to compare ICP peaks with chromatin accessibility across asexual stages, ATACseq data were acquired and mean peak positions derived from the start and end of the reported peaks [40]. The distance between each H2AZ peak summit and the nearest ATAC-seq peak was calculated for each of the ATAC-seq stages reported. These distances were plotted for ICP versus non-ICP Pf H2AZ peak summits [40] (Fig. 7c), and compared pairwise using Welch's t-test with Bonferroni correction.

Transient transfections

Plasmid constructs

To create the control plasmids for nanoluciferase expression, the sequence encoding nanoluciferase was cloned into the pPf86 plasmid [96] using the *NcoI* and *XbaI* restriction sites to create the plasmid pNluc. An AT content-matched *P. falciparum* intergenic sequence that lacked any chromatin modifications indicative of regulatory activity and that was located upstream of a non-expressed gene was used to control for non-sequence specific effects (Additional file 17: Table S5). The control sequence was amplified from *P. falciparum* 3D7 genomic DNA and cloned into *MluI* and *XhoI* restriction sites in pNluc, directly upstream of the hsp86 promoter to generate the pNR3 control plasmid.

The putative promoter and enhancer sequences were amplified from *P. falciparum* 3D7 genomic DNA (Additional file 17: Table S5). Candidate promoter sequences were cloned into *XhoI* and *NcoI* restriction sites of the pNluc plasmid using Gibson assembly. The corresponding enhancer sequences were cloned into these plasmids using the *MluI* and *XhoI* restriction sites. In parallel, the scrambled enhancer sequence was designed by scrambling the putative enhancer sequence for PF3D7_1362000 using the Sequence Manipulation Suite (https://www.bioinformatics.org/sms2/shuffle_dna.html), and synthesising the sequence with GeneArt (Life Technologies), and also cloned into the *MluI* and *XhoI* restriction sites of the promoter-containing pNluc constructs. The promoter-only constructs were made by digesting promoter-containing constructs using *MluI* and *XhoI* restriction enzymes, removing overhangs using Klenow fragment and ligating to form the final plasmid.

Transient transfections

5 µg of plasmids containing test sequences driving Nluc, along with 15 µg of pPf86 [96] were ethanol precipitated and resuspended in 10 µL of sterile TE buffer. Schizont-stage parasites were purified from a 67% (v/v) Percoll cushion and resuspended in 100 µL supplemented Nucleofector™ solution (Lonza) containing the plasmid DNA. The parasites were transfected with the Amaxa™ 4D-Nucleofector™ System (Lonza), using program FP158. The transfected parasites were immediately cultured in medium containing fresh erythrocytes and harvested 6–8 days after transfection. The parasites transfected with putative ring-stage regulatory elements were harvested at ring stage and the parasites transfected with putative schizont-stage regulatory elements were harvested at schizont stage.

Luciferase measurement

Parasites were treated with 0.1% (w/v) saponin in PBS and washed in PBS. Freed parasites were added to white 96-well plates and treated with NanoDLR™ Stop & Glo reagent, to measure nanoluciferase activity. In addition, the levels of firefly luciferase DNA and the number of parasite genomes were measured by quantitative PCR using SYBR™ Green PCR Master Mix (ThermoFisher). The level of nanoluciferase activity was calculated relative to the amount of firefly luciferase DNA per parasite genome to determine the expression of nanoluciferase per plasmid per parasite for all transfections.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s13072-020-00365-5>.

Additional file 1: Fig S1. Specificity of ChIPseq antibodies for *P. falciparum* histones. Western blot of *P. falciparum* whole-cell lysate probed with the antibodies used in this study and with antibody to H4K12ac as an additional control showing the presence of multiple histones on the western blot.

Additional file 2: Fig S2. Specificity of ChIPseq antibodies for *P. falciparum* histone modifications. Dot blot of biotinylated histone peptides probed with rabbit antibody to a) H3K18ac (Abcam, Ab1191), b) rabbit antibody to H3K27ac (Abcam, Ab4729), c and f) rabbit antibody to H3K4me1 (Abcam, Ab8895) and d and e) streptavidin as a loading control. Peptides indicated on the left were diluted from right to left as shown above the figure. Starting concentrations (neat) were 0.25 mg/ml for all *P. falciparum* (Pf) peptides and Human (Hu) H3K4me3, H3K36me3 and H3K9me3 peptides; 0.07 mg/ml for Human H4K20me1 and H4K20me3; 0.065 mg/ml for Human H4. Human H3K4me3, H3K36me3 and H3K9me3 were not biotinylated and their loading was confirmed by ponceau red staining shown for neat peptides in lanes to the left. *P. falciparum* H3 unmodified 1, 2 and 3 peptides represented residues 1–21, 21–43 and 47–65, respectively. A minor cross reaction of anti-H3K18ac antibody with a *P. falciparum* histone H3 unmodified peptide (residues 47–65) was observed, but the antibody was at least 32-fold more sensitive for H3K18ac than H3 (residues 47–65). The rabbit anti-H3K27ac antibody strongly bound both the *P. falciparum* H3K27ac peptide and the H3K23acK27ac peptide. The anti-H3K27ac antibody also bound the *P. falciparum* H3K9ac peptide but by densitometry it generated no more than 1/32 of the signal of H3K9ac.

Additional file 3: Table S1. Quality summary of ChIPseq data.

Additional file 4: Table S2. Quality summary of RNAseq data.

Additional file 5: Fig S3. Reproducibility of ChIPseq. PCA plots and associated scree plots for all ChIPseq and input replicates for ring stages, trophozoite stages and schizont stages.

Additional file 6: Fig S4. Individual sequencing tracks. Read coverage for ChIP, input and RNA sequencing replicates across 78 kb within chromosome 14.

Additional file 7: Fig S5. Correlations between expression level and histones and their modifications across transcriptional units. Log₂ (ChIP/input) plotted from 2500 bp upstream (–2500) to 2500 bp downstream (2500) of the transcriptional start (TSS) and stop (TES) sites for 7612 ring stage transcripts and 7711 trophozoite stage transcripts assembled by Cufflinks ranked in descending order by transcript abundance (fpkm). The Spearman r correlation value (r) for the upstream and downstream enrichment of each histone or histone modification is indicated below the plots.

Additional file 8: Fig S6. Average enrichment profiles of histones and their modifications across genes binned by expression level. The average profile of H2A.Z, H3K18ac, H3K27ac, H3K4me1 and H3 over the coding

sequences ± 2500 bp for genes grouped into terciles by expression level (top green, middle orange, bottom purple) and for silent genes (pink). The enrichment profile of each chromatin mark was presented as the log₂ ratio over the input control. In rings, trophozoites and schizonts gene expression categories were, respectively, the top ($n = 1636, 1603, 1613$), middle ($n = 1495, 1539, 1519$) and bottom ($n = 1580, 1356, 1386$) terciles, and silent genes ($n = 476, 530, 602$). 5'End: the start codon of a gene; 3'End: the stop codon of a gene.

Additional file 9: Fig S7. A) Average enrichment profiles of chromatin features and RNA levels from this study across the summits ± 1.8 kb of intergenic, peaks of H3K4me1 ChIP from this study relative to H3K4me3 from [38] B) Average enrichment profiles of chromatin features and RNA levels from this study across the intergenic peaks of H3K4me1 identified by Ubhe et al [48]. Average profile plots $\pm SE$ covering 1.5 kb up and downstream of published, "strong" H3K4me1 intergenic peaks longer than 1 bp ($n = 259$) [48] oriented according to the closest downstream gene. Top panel is the log₂ ratio of ChIP over input average enrichment profiles of H3K4me1, H3K27ac, H3K18ac, H3 and Pf H2A.Z, middle panel is GC content and ATACseq normalised to gDNA and the bottom panel is RNA separated by strand. C) The number of intergenic peaks and their intersections from the H3K4me1/input from this study, the H3K4me1 from this study over the H3K4me3 from [38], and the H3K4me1 enriched intergenic regions from [48]. D) Spearman correlations for the single schizont H3K4me1 ChIP sample and matched input (Ubhe) analysed in [48] and for our two schizont replicates (S1 and S2) of H3K4me1 ChIP and matched input.

Additional file 10: Table S3. Number of highly expressed genes differentially expressed between stages.

Additional file 11: Comparison with nascent RNA.

Additional file 12: Table S4. Comparison between lifecycle stages of ChIP enrichment upstream of highly expressed genes differentially expressed between stages.

Additional file 13: Fig S8. Comparison of enrichment profiles of histones and their modifications across genes dynamically expressed in rings and trophozoites. Average log₂ ratio of ChIP enrichment of Pf H2A.Z (green), H3K18ac (orange), H3K27ac (purple) and H3K4me1 (pink) all relative to ChIP enrichment of H3 plotted over the coding sequences ± 2500 bp of: gene set 3 ($n = 299$) that were expressed at least threefold more in trophozoite stage than in ring stage and that were in the top quartile by trophozoite-stage expression and: gene set 4 ($n = 455$) that were expressed at least threefold more in ring stage than in trophozoite stage and that were in the top quartile by ring-stage expression. 5'End: the start codon of a gene; 3'End: the stop codon of a gene.

Additional file 14: Fig S9. Comparison of enrichment profiles of histones and their modifications across genes dynamically expressed in trophozoites and schizonts. Average log₂ ratio of ChIP enrichment of Pf H2A.Z (green), H3K18ac (orange), H3K27ac (purple) and H3K4me1 (pink) all relative to ChIP enrichment of H3 plotted over the coding sequences ± 2500 bp of: gene set 5 ($n = 433$) that were expressed at least threefold more in schizont stage than in trophozoite stage and that were in the top quartile by schizont-stage expression and: gene set 6 ($n = 173$) that were expressed at least threefold more in trophozoite stage than in schizont stage and that were in the top quartile by trophozoite-stage expression. 5'End: the start codon of a gene; 3'End: the stop codon of a gene.

Additional file 15: Fig S10. Enrichment of histones and their modifications across ATACseq peak summits. Log₂ ChIP H3K4me1 (this study)/H3K4me3 [38], Log₂ ChIP/input of H3K4me1, H3, H3K18ac, H3K27ac and Pf H2A.Z and ATACseq coverage normalised to gDNA [41] plotted across all schizont stage ATACseq peak summits ± 1800 bp. Average profile plots are shown on top, heatmaps below are all ranked in descending order of ATACseq coverage.

Additional file 16: Fig S11. Flow chart for the bioinformatic strategy used to identify candidate proximal and distal regulatory sequences upstream of schizont stage-expressed genes.

Additional file 17: Table S5. Coordinates of tandem ICPs tested for regulatory activity.

Acknowledgements

The authors thank Prof Graham V Brown for useful discussions and input during the study and helpful comments on the manuscript. We thank Dr Troy Attard for manufacturing the H3K4me1 peptide at the Bio21 Institute.

Author contributions

JT designed, performed and analysed experiments, performed data analysis and was a major contributor to writing the MS. SAC performed and analysed experiments, performed data analysis and contributed to writing the MS. LMY analysed data and contributed to writing the MS. PRG contributed to study design, provided reagents and contributed to writing the MS. ATP contributed to study design and contributed to writing the MS. KPD provided reagents and contributed to writing the MS. MP contributed to study design, analysed data and was a major contributor to writing the MS. MFD designed the study, performed data analysis and was a major contributor to writing the MS. All authors read and approved the final manuscript.

Funding

This work was supported by the Australian Research Council grant number DP110100483 and the National Health and Medical Research Council grant number APP1128975. The funding bodies had no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the NCBI Sequence Read Archive BioProject PRJNA612099, study SRP252482, BioSamples SAMN14361305 to SAMN14361313, accessions SRR11292854 to SRR11292897, Gene Expression Omnibus accession number GSE64691 [9], Gene Expression Omnibus accession number GSE104075 [41], NCBI Sequence Read Archive, accession number GSE80293 [8], Gene Expression Omnibus accession number GSE66185 [43], Gene Expression Omnibus accession number GSE68982 [42], Gene Expression Omnibus accession number GSE109599 [40].

Ethics approval and consent to participate

This work used human blood products for the in vitro culture of *P. falciparum*. This work was approved by the Human Research Ethics Committee of The University of Melbourne approval number 1442753.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Medicine, The University of Melbourne, Royal Melbourne Hospital, Parkville, VIC 3050, Australia. ² School of Medicine, Faculty of Health, Deakin University, Geelong Waurn Ponds Campus, Waurn Ponds, VIC 3216, Australia. ³ School of BioSciences, The University of Melbourne, Parkville, VIC 3052, Australia. ⁴ Bio21 Institute, Parkville, VIC 3052, Australia. ⁵ Peter Doherty Institute, Melbourne, VIC 3000, Australia. ⁶ Department of Microbiology and Immunology, The University of Melbourne, Victoria 3000, Australia. ⁷ Macfarlane Burnet Institute for Medical Research and Public Health, Melbourne, VIC 3004, Australia. ⁸ Monash University, Melbourne, VIC 3800, Australia. ⁹ Bioinformatics Division, Walter and Eliza Hall Institute of Medical Research, Parkville, VIC, Australia. ¹⁰ Department of Mathematics and Statistics, University of Melbourne, Victoria, Australia. ¹¹ Department of Medical Biology, University of Melbourne, Parkville, VIC, Australia. ¹² Sir Peter MacCallum, Department of Oncology, University of Melbourne, Parkville, VIC, Australia. ¹³ Erlangen University, 91054 Erlangen, Germany.

Received: 27 March 2020 Accepted: 29 September 2020
Published online: 23 November 2020

References

- Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, et al. Genome sequence of the human malaria parasite *Plasmodium*

- falciiparum*. *Nature*. 2002;419:498–511. <https://doi.org/10.1038/nature01097>.
2. Miao J, Fan Q, Cui L, Li J. The malaria parasite *Plasmodium falciparum* histones: organization, expression, and acetylation. *Gene*. 2006;369:53–65.
 3. Salcedo-Amaya AM, van Driel MA, Alako BT, Trelle MB, van den Elzen AM, Cohen AM, et al. Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A*. 2009;106:9655–60. <https://doi.org/10.1073/pnas.0902515106>.
 4. Silberhorn E, Schwartz U, Loffler P, Schmitz S, Symelka A, de Koning-Ward T, et al. *Plasmodium falciparum* nucleosomes exhibit reduced stability and lost sequence dependent nucleosome positioning. *PLoS Pathog*. 2016;12:e1006080. <https://doi.org/10.1371/journal.ppat.1006080>.
 5. Bozdech Z, Llinas M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol*. 2003;1:85–100.
 6. Reid AJ, Talman AM, Bennett HM, Gomes AR, Sanders MJ, Illingworth CJR, et al. Single-cell RNA-seq reveals hidden transcriptional variation in malaria parasites. *Elife*. 2018. <https://doi.org/10.7554/eLife.33105>.
 7. Modrzynska K, Pfander C, Chappell L, Yu L, Suarez C, Dundas K, 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. <https://doi.org/10.1016/j.chom.2016.12.003>.
 8. Santos JM, Josling G, Ross P, Joshi P, Orchard L, Campbell T, et al. Red blood cell invasion by the malaria parasite is coordinated by the PfAP2-1 transcription factor. *Cell Host Microbe*. 2017;21:731–41. <https://doi.org/10.1016/j.chom.2017.05.006>.
 9. Josling GA, Petteer M, Oehring SC, Gupta AP, Dietz O, Wilson DW, et al. A *Plasmodium falciparum* bromodomain protein regulates invasion gene expression. *Cell Host Microbe*. 2015;17:741–51. <https://doi.org/10.1016/j.chom.2015.05.009>.
 10. Wang Z, Zang C, Rosenfeld JA, Schones DE, Barski A, Cuddapah S, et al. Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet*. 2008;40:897–903. <https://doi.org/10.1038/ng.154>.
 11. Heintzman ND, Hon GC, Hawkins RD, Kheradpour P, Stark A, Harp LF, et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*. 2009;459:108–12. <https://doi.org/10.1038/nature07829>.
 12. Kurdistani SK, Tavazoie S, Grunstein M. Mapping global histone acetylation patterns to gene expression. *Cell*. 2004;117:721–33. <https://doi.org/10.1016/j.cell.2004.05.023>.
 13. Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, et al. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*. 2005;122:517–27. <https://doi.org/10.1016/j.cell.2005.06.026>.
 14. Bernstein BE, Humphrey EL, Erlich RL, Schneider R, Bouman P, Liu JS, et al. Methylation of histone H3 Lys 4 in coding regions of active genes. *Proc Natl Acad Sci U S A*. 2002;99:8695–700. <https://doi.org/10.1073/pnas.082249499>.
 15. Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, et al. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell*. 2005;120:169–81. <https://doi.org/10.1016/j.cell.2005.01.001>.
 16. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, et al. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet*. 2007;39:311–8. <https://doi.org/10.1038/ng1966>.
 17. Koch CM, Andrews RM, Flicek P, Dillon SC, Karaoz U, Clelland GK, et al. The landscape of histone modifications across 1% of the human genome in five human cell lines. *Genome Res*. 2007;17:691–707. <https://doi.org/10.1101/gr.5704207>.
 18. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. *Cell*. 2007;129:823–37.
 19. Roh TY, Cuddapah S, Zhao K. Active chromatin domains are defined by acetylation islands revealed by genome-wide mapping. *Genes Dev*. 2005;19:542–52. <https://doi.org/10.1101/gad.1272505>.
 20. Encode Project Consortium, Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, et al. An integrated encyclopedia of DNA elements in the human genome. *Nature*. 2012;489:57–74. <https://doi.org/10.1038/nature11247>.
 21. Encode Project Consortium, Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*. 2007;447:799–816. <https://doi.org/10.1038/nature05874>.
 22. Ernst J, Kheradpour P, Mikkelsen TS, Shoresh N, Ward LD, Epstein CB, et al. Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature*. 2011;473:43–9. <https://doi.org/10.1038/nature09906>.
 23. Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A*. 2010;107:21931–6. <https://doi.org/10.1073/pnas.1016071107>.
 24. Bonn S, Zinzen RP, Girardot C, Gustafson EH, Perez-Gonzalez A, Delhomme N, et al. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat Genet*. 2012;44:148–56. <https://doi.org/10.1038/ng.1064>.
 25. Jin Q, Yu LR, Wang L, Zhang Z, Kasper LH, Lee JE, et al. Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J*. 2011;30:249–62. <https://doi.org/10.1038/emboj.2010.318>.
 26. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, Wu J, et al. Widespread transcription at neuronal activity-regulated enhancers. *Nature*. 2010;465:182–7. <https://doi.org/10.1038/nature09033>.
 27. Rada-Iglesias A, Bajpai R, Prescott S, Brugmann SA, Swigut T, Wysocka J. Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell*. 2012;11:633–48. <https://doi.org/10.1016/j.stem.2012.07.006>.
 28. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, Holt A, et al. ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature*. 2009;457:854–8. <https://doi.org/10.1038/nature07730>.
 29. Wang J, Zhuang J, Iyer S, Lin X, Whitfield TW, Greven MC, et al. Sequence features and chromatin structure around the genomic regions bound by 119 human transcription factors. *Genome Res*. 2012;22:1798–812. <https://doi.org/10.1101/gr.139105.112>.
 30. Kristiansson E, Thorsen M, Tamas MJ, Nerman O. Evolutionary forces act on promoter length: identification of enriched cis-regulatory elements. *Mol Biol Evol*. 2009;26:1299–307. <https://doi.org/10.1093/molbev/msp040>.
 31. Grunberg S, Henikoff S, Hahn S, Zentner GE. Mediator binding to UASs is broadly uncoupled from transcription and cooperative with TFIID recruitment to promoters. *EMBO J*. 2016;35:2435–2446. <https://doi.org/10.15252/embj.201695020>.
 32. Eriksson PR, Ganguli D, Nagarajavel V, Clark DJ. Regulation of histone gene expression in budding yeast. *Genetics*. 2012;191:7–20. <https://doi.org/10.1534/genetics.112.140145>.
 33. Reja R, Vinayachandran V, Ghosh S, Pugh BF. Molecular mechanisms of ribosomal protein gene coregulation. *Genes Dev*. 2015;29:1942–54. <https://doi.org/10.1101/gad.268896.115>.
 34. Brogaard K, Xi L, Wang JP, Widom J. A map of nucleosome positions in yeast at base-pair resolution. *Nature*. 2012;486:496–501. <https://doi.org/10.1038/nature11142>.
 35. Schep AN, Buenostro JD, Denny SK, Schwartz K, Sherlock G, Greenleaf WJ. Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Res*. 2015;25:1757–70. <https://doi.org/10.1101/gr.192294.115>.
 36. Gupta AP, Chin WH, Zhu L, Mok S, Luah YH, Lim EH, et al. Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite *Plasmodium falciparum*. *PLoS Pathog*. 2013;9:e1003170. <https://doi.org/10.1371/journal.ppat.1003170>.
 37. Duffy MF, Tang J, Sumardy F, Nguyen HH, Selvarajah SA, Josling GA, et al. Activation and clustering of a *Plasmodium falciparum* var gene are affected by subtelomeric sequences. *FEBS J*. 2017;284:237–57. <https://doi.org/10.1111/febs.13967>.
 38. Bartfai R, Hoeijmakers WA, Salcedo-Amaya AM, Smits AH, Janssen-Megens E, Kaan A, 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. <https://doi.org/10.1371/journal.ppat.1001223>.
 39. Karmodiya K, Pradhan SJ, Joshi B, Jangid R, Reddy PC, Galande S. A comprehensive epigenome map of *Plasmodium falciparum* reveals unique mechanisms of transcriptional regulation and identifies H3K36me2 as a

- global mark of gene suppression. *Epigenetics Chromatin*. 2015;8:32. <https://doi.org/10.1186/s13072-015-0029-1>.
40. Ruiz JL, Tena JJ, Bancells C, Cortes A, Gomez-Skarmeta JL, Gomez-Diaz E. Characterization of the accessible genome in the human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Res*. 2018;46:9414–31. <https://doi.org/10.1093/nar/gky643>.
 41. Toenhake CG, Frasncka SA, Vijayabaskar MS, Westhead DR, van Heeringen SJ, Bartfai R. Chromatin accessibility-based characterization of the gene regulatory network underlying *Plasmodium falciparum* blood-stage development. *Cell Host Microbe*. 2018;23(557–569):e559. <https://doi.org/10.1016/j.chom.2018.03.007>.
 42. Adjalley SH, Chabbert CD, Klaus B, Pelechano V, Steinmetz LM. Landscape and dynamics of transcription initiation in the malaria parasite *Plasmodium falciparum*. *Cell Rep*. 2016;14:2463–75. <https://doi.org/10.1016/j.celrep.2016.02.025>.
 43. Kenschke PR, Hoeijmakers WA, Toenhake CG, Bras M, Chappell L, Berriman M, et al. The nucleosome landscape of *Plasmodium falciparum* reveals chromatin architecture and dynamics of regulatory sequences. *Nucleic Acids Res*. 2016;44:2110–244. <https://doi.org/10.1093/nar/gkv1214>.
 44. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, Lee AY, et al. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature*. 2013;503:290–4. <https://doi.org/10.1038/nature12644>.
 45. Lemieux JE, Kyes SA, Otto TD, Feller AI, Eastman RT, Pinches RA, et al. Genome-wide profiling of chromosome interactions in *Plasmodium falciparum* characterizes nuclear architecture and reconfigurations associated with antigenic variation. *Mol Microbiol*. 2013;90:519–37. <https://doi.org/10.1111/mmi.12381>.
 46. Ay F, Bunnik EM, Varoquaux N, Bol SM, Prudhomme J, Vert JP, et al. Three-dimensional modeling of the *P. falciparum* genome during the erythrocytic cycle reveals a strong connection between genome architecture and gene expression. *Genome Res*. 2014;24:974–88. <https://doi.org/10.1101/gr.169417.113>.
 47. Bunnik EM, Cook KB, Varoquaux N, Batugedara G, Prudhomme J, Cort A, et al. Changes in genome organization of parasite-specific gene families during the *Plasmodium* transmission stages. *Nat Commun*. 2018;9:1910. <https://doi.org/10.1038/s41467-018-04295-5>.
 48. Ubhe S, Rawat M, Verma S, Anamika K, Karmodiya K. Genome-wide identification of novel intergenic enhancer-like elements: implications in the regulation of transcription in *Plasmodium falciparum*. *BMC Genomics*. 2017;18:656. <https://doi.org/10.1186/s12864-017-4052-4>.
 49. Gomez-Diaz E, Yerbanga RS, Lefevre T, Cohuet A, Rowley MJ, Ouedraogo JB, et al. Epigenetic regulation of *Plasmodium falciparum* clonally variant gene expression during development in *Anopheles gambiae*. *Sci Rep*. 2017;7:40655. <https://doi.org/10.1038/srep40655>.
 50. Petter M, Lee CC, Byrne TJ, Boysen KE, Volz J, Ralph SA, et al. Expression of *P. falciparum* var genes involves exchange of the histone variant H2A.Z at the promoter. *PLoS Pathog*. 2011;7:e1001292. <https://doi.org/10.1371/journal.ppat.1001292>.
 51. Egelhofer TA, Minoda A, Klugman S, Lee K, Kolasinska-Zwiercz P, Alekseyenko AA, et al. An assessment of histone-modification antibody quality. *Nat Struct Mol Biol*. 2011;18:91–3. <https://doi.org/10.1038/nsmb.1972>.
 52. Rothbart SB, Dickson BM, Raab JR, Grzybowski AT, Krajewski K, Guo AH, et al. An interactive database for the assessment of histone antibody specificity. *Mol Cell*. 2015;59:502–11. <https://doi.org/10.1016/j.molcel.2015.06.022>.
 53. Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res*. 2012;22:1813–31. <https://doi.org/10.1101/gr.136184.111>.
 54. Coetzee N, Sidoli S, van Bijljon R, Painter H, Llinas M, Garcia BA, et al. Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual *Plasmodium falciparum* parasites. *Sci Rep*. 2017;7:607. <https://doi.org/10.1038/s41598-017-00687-7>.
 55. Saraf A, Cervantes S, Bunnik EM, Pons N, Sardu ME, Chung DW, et al. Dynamic and combinatorial landscape of histone modifications during the intraerythrocytic developmental cycle of the malaria parasite. *J Proteome Res*. 2016;15:2787–801. <https://doi.org/10.1021/acs.jproteome.6b00366>.
 56. Sims JS, Militello KT, Sims PA, Patel VP, Kasper JM, Wirth DF. Patterns of gene-specific and total transcriptional activity during the *Plasmodium falciparum* intraerythrocytic developmental cycle. *Eukaryot Cell*. 2009;8:327–38. <https://doi.org/10.1128/EC.00340-08>.
 57. Rozowsky J, Euskirchen G, Auerbach RK, Zhang ZD, Gibson T, Bjornson R, et al. PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nat Biotechnol*. 2009;27:66–75. <https://doi.org/10.1038/nbt.1518>.
 58. ENCODE. ENCODE Experimental Guidelines for ENCODE3 ChIP-seq 2017. <https://www.encodeproject.org/about/experiment-guidelines/>.
 59. Hoeijmakers WAM, Salcedo-Amaya AM, Smits AH, Francois KJ, Trecek M, Gilberger TW, 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. <https://doi.org/10.1111/mmi.12151>.
 60. Painter HJ, Chung NC, Sebastian A, Albert I, Storey JD, Llinas M. Genome-wide real-time in vivo transcriptional dynamics during *Plasmodium falciparum* blood-stage development. *Nat Commun*. 2018;9:2656. <https://doi.org/10.1038/s41467-018-04966-3>.
 61. Westenberger SJ, Cui L, Dharia N, Winzler E. Genome-wide nucleosome mapping of *Plasmodium falciparum* reveals histone-rich coding and histone-poor intergenic regions and chromatin remodeling of core and subtelomeric genes. *BMC Genomics*. 2009;10:610. <https://doi.org/10.1186/1471-2164-10-610>.
 62. Arnot DE, Ronander E, Bengtsson DC. The progression of the intraerythrocytic cell cycle of *Plasmodium falciparum* and the role of the centriolar plaques in asynchronous mitotic division during schizogony. *Int J Parasitol*. 2011;41:71–80. <https://doi.org/10.1016/j.ijpara.2010.07.012>.
 63. Li B, Pattenden SG, Lee D, Gutierrez J, Chen J, Seidel C, et al. Preferential occupancy of histone variant H2A.Z at inactive promoters influences local histone modifications and chromatin remodeling. *Proc Natl Acad Sci U S A*. 2005;102:18385–900. <https://doi.org/10.1073/pnas.0507975102>.
 64. Chen P, Zhao J, Wang Y, Wang M, Long H, Liang D, et al. H3.3 actively marks enhancers and primes gene transcription via opening higher-ordered chromatin. *Genes Dev*. 2013;27:2109–24. <https://doi.org/10.1101/gad.222174.113>.
 65. Cui XJ, Li H, Liu GQ. Combinatorial patterns of histone modifications in *Saccharomyces cerevisiae*. *Yeast*. 2011;28:683–91. <https://doi.org/10.1002/yea.1896>.
 66. Liu Y, Chen S, Wang S, Soares F, Fischer M, Meng F, et al. Transcriptional landscape of the human cell cycle. *Proc Natl Acad Sci U S A*. 2017. <https://doi.org/10.1073/pnas.1617636114>.
 67. Hsiung CCS, Bartman CR, Huang P, Ginart P, Stonestrom AJ, Keller CA, et al. A hyperactive transcriptional state marks genome reactivation at the mitosis–G1 transition. *Genes Dev*. 2016;30:1423–39. <https://doi.org/10.1101/gad.280859>.
 68. Horrocks P, Kilbey BJ. Physical and functional mapping of the transcriptional start sites of *Plasmodium falciparum* proliferating cell nuclear antigen. *Mol Biochem Parasitol*. 1996;82:207–15.
 69. Horrocks P, Lanzer M. Mutational analysis identifies a five base pair cis-acting sequence essential for GBP130 promoter activity in *Plasmodium falciparum*. *Mol Biochem Parasitol*. 1999;99:77–87.
 70. Crabb BS, Cowman AF. Characterization of promoters and stable transfection by homologous and nonhomologous recombination in *Plasmodium falciparum*. *Proc Natl Acad Sci USA*. 1996;93:7289–94.
 71. Osta M, Gannoun-Zaki L, Bonnefoy S, Roy C, Vial HJ. A 24 bp cis-acting element essential for the transcriptional activity of *Plasmodium falciparum* CDP-diacylglycerol synthase gene promoter. *Mol Biochem Parasitol*. 2002;121:87–98.
 72. Komaki-Yasuda K, Okuwaki M, Kano S, Nagata K, Kawazu S. 5' sequence and chromatin modification-dependent gene expression in *Plasmodium falciparum* erythrocytic stage. *Mol Biochem Parasitol*. 2008;162:40–51. <https://doi.org/10.1016/j.molbiopara.2008.07.002>.
 73. Lopez-Estrano C, Gopalakrishnan AM, Semblat JP, Fergus MR, Mazier D, Haldar K. An enhancer-like region regulates *hrp3* promoter stage-specific gene expression in the human malaria parasite *Plasmodium falciparum*. *Biochim Biophys Acta*. 2007;1769:506–13. <https://doi.org/10.1016/j.bbaexp.2007.04.009>.
 74. Hoeijmakers WAM, Miao J, Schmidt S, Toenhake CG, Shrestha S, Venhuizen J, et al. Epigenetic reader complexes of the human malaria parasite *Plasmodium falciparum*. *Nucleic Acids Res*. 2019;47:11574–88. <https://doi.org/10.1093/nar/gkz1044>.
 75. Filletton F, Chuffart F, Nagarajan M, Bottin-Duplus H, Yvert G. The complex pattern of epigenomic variation between natural yeast strains at

- single-nucleosome resolution. *Epigenetics Chromatin*. 2015;8:26. <https://doi.org/10.1186/s13072-015-0019-3>.
76. Huang F, Ramakrishnan S, Pokhrel S, Pflueger C, Parnell TJ, Kasten MM, et al. Interaction of the Jhd2 Histone H3 Lys-4 demethylase with chromatin is controlled by histone H2A surfaces and restricted by H2B ubiquitination. *J Biol Chem*. 2015;290:28760–77. <https://doi.org/10.1074/jbc.M115.693085>.
 77. Amulic B, Salanti A, Lavstsen T, Nielsen MA, Deitsch KW. An upstream open reading frame controls translation of *var2csa*, a gene implicated in placental malaria. *PLoS Pathog*. 2009;5:e1000256.
 78. Brancucci NM, Witmer K, Schmid C, Voss TS. A var gene upstream element controls protein synthesis at the level of translation initiation in *Plasmodium falciparum*. *PLoS ONE*. 2014;9:e100183. <https://doi.org/10.1371/journal.pone.0100183>.
 79. Trager W, Jensen JB. Human malaria parasites in continuous culture. *Science*. 1976;193:673–5.
 80. Lambros C, Vanderberg JP. Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. *J Parasitol*. 1979;65:418–20.
 81. Oyola SO, Otto TD, Gu Y, Maslen G, Manske M, Campino S, et al. Optimizing Illumina next-generation sequencing library preparation for extremely AT-biased genomes. *BMC Genomics*. 2012;13:1. <https://doi.org/10.1186/1471-2164-13-1>.
 82. Tonkin-Hill GQ, Trianty L, Noviyanti R, Nguyen HHT, Sebayang BF, Lampah DA, et al. The *Plasmodium falciparum* transcriptome in severe malaria reveals altered expression of genes involved in important processes including surface antigen-encoding var genes. *PLoS Biol*. 2018;16:e2004328. <https://doi.org/10.1371/journal.pbio.2004328>.
 83. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol*. 2011;29:24–6. <https://doi.org/10.1038/nbt.1754>.
 84. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res*. 2014;42:W187–191. <https://doi.org/10.1093/nar/gku365>.
 85. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14:R36. <https://doi.org/10.1186/gb-2013-14-4-r36>.
 86. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25:2078–9. <https://doi.org/10.1093/bioinformatics/btp352>.
 87. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc*. 2012;7:562–78. <https://doi.org/10.1038/nprot.2012.016>.
 88. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*. 2013;41:e108. <https://doi.org/10.1093/nar/gkt214>.
 89. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol*. 2008;9:R137. <https://doi.org/10.1186/gb-2008-9-9-r137>.
 90. Shen L, Shao N, Liu X, Nestler E. ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*. 2014;15:284.
 91. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841–2. <https://doi.org/10.1093/bioinformatics/btq033>.
 92. Lun AT, Smyth GK. csaw: a Bioconductor package for differential binding analysis of ChIP-seq data using sliding windows. *Nucleic Acids Res*. 2016;44:e45. <https://doi.org/10.1093/nar/gkv1191>.
 93. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv. 2013;1303.3997 [q-bio.GN].
 94. Kent WJ, Zweig AS, Barber G, Hinrichs AS, Karolchik D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics*. 2010;26:2204–7. <https://doi.org/10.1093/bioinformatics/btq351>.
 95. Gruning B, Dale R, Sjodin A, Chapman BA, Rowe J, Tomkins-Tinch CH, et al. Bioconda: sustainable and comprehensive software distribution for the life sciences. *Nat Methods*. 2018;15:475–6. <https://doi.org/10.1038/s41592-018-0046-7>.
 96. Militello KT, Wirth DF. A new reporter gene for transient transfection of *Plasmodium falciparum*. *Parasitol Res*. 2003;89:154–7. <https://doi.org/10.1007/s00436-002-0721-5>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

