



## Data in Brief

# *Plasmodium falciparum* epigenome: A distinct dynamic epigenetic regulation of gene expression

Mukul Rawat<sup>1</sup>, Madhvi A. Bhosale<sup>1</sup>, Krishanpal Karmodiya<sup>\*</sup>

Department of Biology, Indian Institute of Science Education and Research, Pune, Maharashtra, India

## ARTICLE INFO

## Article history:

Received 16 November 2015

Accepted 30 November 2015

Available online 1 December 2015

## Keywords:

*Plasmodium falciparum*  
Genome-wide mapping  
Histone modifications  
Chromatin  
Transcription

## ABSTRACT

Histone modification profiles are predictive of gene expression and most of the knowledge gained is acquired through studies done in higher eukaryotes. However, genome-wide studies involving *Plasmodium falciparum*, the causative agent of malaria, have been rather few, at lower resolution (mostly using ChIP-on-chip), and covering limited number of histone modifications. In our recent study [1], we have performed extensive genome-wide analyses of multiple histone modifications including the active (H3K4me2, H3K4me3, H3K9ac, H3K14ac, H3K27ac and H4ac), inactive (H3K9me3 and H3K27me3), elongation (H3K79me3) and regulatory element (H3K4me1) in a stage-specific manner. Furthermore, we used a ligation-based method suitable for sequencing homopolymeric stretches as seen in *P. falciparum* for next-generation sequencing library amplification [2], enabling highly quantitative analysis of the extremely AT-rich *P. falciparum* genome. Our recently published study suggests that transcription regulation by virtue of poised chromatin and differential histone modifications is unique to *P. falciparum* [1]. Here we describe the experiments, quality controls and chromatin immunoprecipitation-sequencing data analysis of our associated study published in Epigenetics and Chromatin [1]. Stage-specific ChIP-sequencing data for histone modifications is submitted to Gene Expression Omnibus (GEO) database under the accession number GSE63369.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Organism/cell line/tissue	<i>Plasmodium falciparum</i>
Sex	Not applicable
Sequencer or array type	AB SOLiD 4 system high-throughput genome sequencer
Data format	Raw and analyzed
Experimental factors	ChIP-sequencing of histone modification using antibody
Experimental features	ChIP-sequencing of <i>Plasmodium falciparum</i> histone modifications in a stage specific manner
Consent	Not applicable
Sample source location	Not applicable

Direct link to deposited data [provide URL below]

The direct link for the ChIP-sequencing data is -<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE63369>

## 1. Experimental design, materials and methods

### 1.1. *Plasmodium falciparum* culture and synchronization

*P. falciparum* strain 3D7 was cultured in RPMI 1640 medium supplemented with 25 mM HEPES, 0.5% AlbuMAX I, 1.77 mM sodium bicarbonate, 100  $\mu$ M hypoxanthine and 12.5  $\mu$ g ml<sup>-1</sup> gentamicin sulfate, at

a pH of 7.2 in 5% CO<sub>2</sub>. Incomplete medium was prepared weekly without sodium bicarbonate and stored at 4 °C. Complete medium was freshly prepared with sodium bicarbonate. Fresh whole blood from a healthy donor (O<sup>+</sup>ve) was used for preparing RBCs and stored at 4 °C for at least 1 day. RBCs were washed thrice with washing medium (complete medium without AlbuMAX I) before use. Subculturing was done every 2 days for 6–8 h before invasion by equally dividing the contents of each flask into two or more flasks and quickly restoring the haematocrit between 1 and 1.5% in the required volume of culture medium [3]. Medium was changed every 24 h. Asynchronous culture with early ring stage (less than 10 h) was synchronized using 5% sorbitol, which was added 10 times the volume of infected RBCs pellet followed by vigorous vortexing for 30 s to rupture mature parasitic forms. Culture was then kept for incubation at 37 °C for 8 min under shaking at 240 rpm. Culture was centrifuged at 250 g for 5 min to get rid of ruptured RBCs. Pellet was washed twice with washing medium and transferred to a flask containing complete medium. Parasitemia was monitored with acridine orange stained thin blood smear. The synchronized culture was harvested at 18, 30 and 40 hpi for chromatin immunoprecipitation.

### 1.2. Screening of histone modification antibody for chromatin immunoprecipitation (ChIP)

Infected RBCs were cross-linked with 1% formaldehyde (Catalogue number 28908, Thermo Scientific), which was directly added to the

<sup>\*</sup> Corresponding author.

E-mail address: [krish@iiserpune.ac.in](mailto:krish@iiserpune.ac.in) (K. Karmodiya).

<sup>1</sup> These authors contributed equally to this work.

culture medium drop-wise in chemical-hood and mixed by rotating for 10 min at room temperature. Formaldehyde fixed cells were quenched with 150 mM glycine for 10 min at room temperature. Infected RBCs were washed twice with  $1 \times$  PIC and 1 mM PMSF in cold PBS. Resultant pellet was dissolved in swelling buffer (25 mM Tris pH 7.9, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, 0.1% NP40, 1 mM DTT, 0.5 mM PMSF,  $1 \times$  PIC) for nuclei isolation. Nuclei were isolated by dounce homogenization using loose piston (B). Isolated nuclei were lysed and sonicated in sonication buffer (10 mM Tris-HCl pH 7.5, 200 mM NaCl, 1% SDS, 4% NP-40, 1 mM PMSF) to obtain an average chromatin size of 200–400 bp. Chromatin was pre-cleared using 50  $\mu$ l of a 50% protein A Sepharose (GE healthcare) slurry for 1 h at 4 °C with gentle inverting. Immunoprecipitations were carried out in 1 ml of IP buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton-X 100). Three micrograms of antibody was used per 20  $\mu$ g purified chromatin. 10% Input chromatin was obtained after pre-clearing by de-crosslinking and purified using the Qiaquick column (Qiagen) according to the manufacturer's instructions. Immunoprecipitations were carried out with inverting at 4 °C for 14–16 h. The samples were then incubated with 50  $\mu$ l of a 50% Protein A Sepharose slurry for 3 h at 4 °C with gentle inverting. IP samples were reverse-crosslinked and the DNA was purified using a Qiaquick column (Qiagen). Specificity of ChIP was determined by quantitative PCR for the known histone modification enriched genomic region and an arbitrarily chosen control genomic region. A serial dilution of input sample was performed to calculate the % input enrichment. Samples were processed for ChIP-sequencing if the enrichment was observed more than 1% of input and 5 fold to control genomic region (Fig. 1). Primers used for cold genomic region (Forward 5'-AACGTTAAATTTGAATCCGAGA-3', Reverse 5'-AATCTCCGAGACCGGAAT-3'), Pf11\_0468 (Forward 5'-TGTG CACATGGGAATTTCA-3', Reverse 5'-CTCTCAATAGCATCTCTTCATT-3'), PF10\_0287 (Forward 5'-CCATGAACTGCGACGTCTAC-3', Reverse 5'-AAAAATCCCTTAAAAAGATGAGTGA-3'), and PF13\_0303 (Forward 5'-CAACCATCGTTCCTTGACCT-3' Reverse 5'-GTAACCGTCCGTGTGCTTTA-3'). We found this method reproducible as the normalization is performed with respect to the control genomic region from the same experimental condition.

### 1.3. RNA extraction and strand-specific RT-PCR

Synchronized culture was harvested at 18 (rings), 30 (trophozoites) and 48 (schizonts) hpi. Parasites were isolated by saponin (8 mg/ml in PBS) lysis at 37 °C for 15 min. Total RNA was extracted from isolated parasites by adding pre-warmed TRIzol to the pellet and incubated at room temperature for 5 min.  $0.2 \times$  TRIzol volumes of chloroform was added and vigorous shaking followed by 2–3 min incubation at room temperature. The samples were centrifuged for 30 min to collect upper layer and  $0.5 \times$  TRIzol volume of Isopropanol added to precipitate RNA. RNA was treated with DNaseI (Ambion) as described in the manufacturer's protocol followed by phenol-chloroform extraction.

RT-PCR was performed using 500 ng total RNA. For each gene, reverse transcription (RT) was performed using a forward and reverse primer. After first strand cDNA synthesis PCR was performed using sense and anti-sense primers.

## 2. Data analysis

### 2.1. *P. falciparum* ChIP-seq data processing and quality control measures

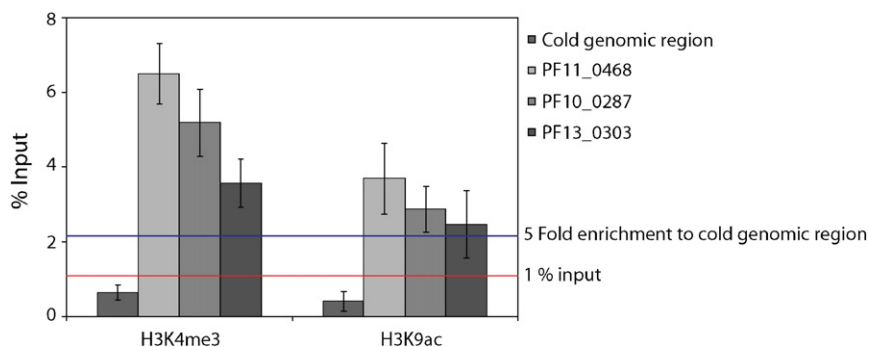
Reference genome (PlasmoDB-9.3\_P.falciparum 3D7) was converted to color-space using bowtie-build. ChIP-seq data were mapped using the Bowtie allowing two mismatches to generate SAM files. The parameters “-5 5 -3 5 -a -m 1 -C” were used to select for only uniquely mapping reads. Bed files generated using sam2bed tool (<https://bedpos.readthedocs.org/en/latest>).

### 2.2. Analysis of ChIP-sequencing data

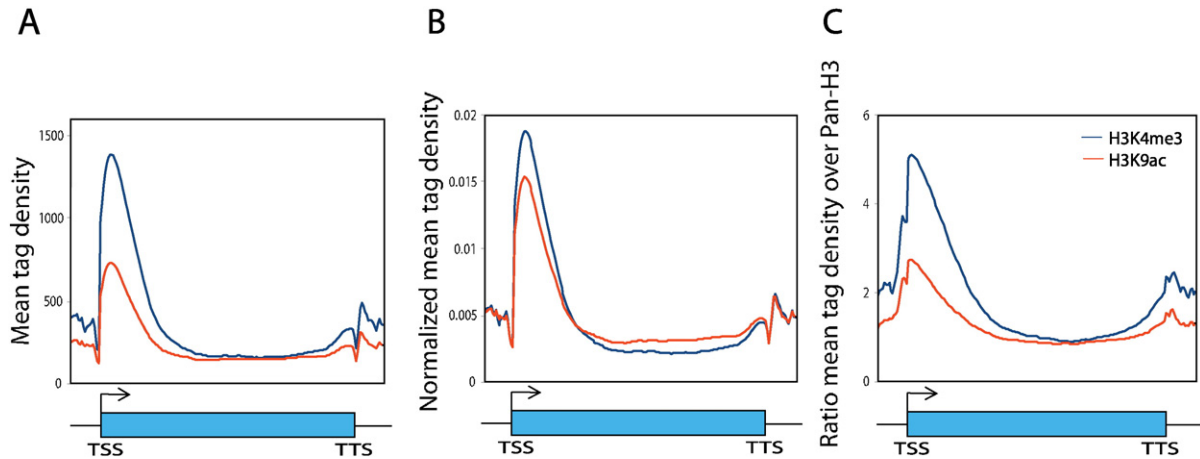
Our interest was to generate the distribution profile of histone modifications over the gene body in *P. falciparum* to generate a comprehensive map. However, the transcription start sites (TSSs) and transcription termination sites (TTSs) were not known in *P. falciparum*. To determine the TSSs and TTSs, we pooled previously reported RNA-seq data covering 8 samples collected at different time points of erythrocytic cycle stage [4]. Pfalciparum3D7\_Genome\_v9.3 genome index was built by bowtie2-build (version 2.2.3) and all raw reads from RNA-seq results from different stages were aligned using TopHat (v2.0.13) to indexed genome [5] using parameters “-min-intron-length 10 -max-intron-length 10,000”. In the alignment process a total of 89,645,433 (76.6%) reads were mapped out of 117,030,592 reads. Assembly of the transcripts from aligned reads was carried out by cufflinks (v2.2.1) program using PlasmoDB-9.3\_Pfalciparum3D7.gff as GTF reference. Cufflinks program was ran with following parameters, minimum intron length set as 10 nt and a maximum intron length as 10,000 nt and assembly was guided by GFF file (PlasmoDB-9.3\_Pfalciparum3D7.gff) to annotate the transcripts [5]. GTF file given by Cufflink was merged with reference GTF file (PlasmoDB-9.3\_Pfalciparum3D7.gtf) using Cuffmerge function (part of Tuxedo suite). Merged GTF was converted to GFF by 'gffread' utility available in Tuxedo suite [5] and a gene co-ordinate GFF file was created using AWK script (awk '/gene\_name/{print\$0}'./merged.gff > ./genes.gff). Genes coding for rRNA and tRNA were removed from this file. This modified file was used to merge overlapping gene features using mergeBed tool from BEDTools. From this Bed files TSSs and TTSs were extracted and used for further analysis.

### 2.3. Average profile calculations of histone modifications on TSSs and TTSs

Profile of histone modifications is calculated using the program seqMiner [6]. Histone modification files (.bed) were used to generate



**Fig. 1.** Assessment of H3K4me3 and H3K9ac ChIP by ChIP-qPCR. Serial dilution of input DNA was performed to make the standard curve to determine the relative concentration for each primer pair. Fold enrichment is calculated over an arbitrarily chosen control genomic region. Error bars represent the standard deviation for three technical replicates. Samples were processed for ChIP-sequencing if the enrichment was observed more than 1% of input and 5 fold to control genomic region.



**Fig. 2.** Profiles of H3K4me3 and H3K9ac over averaged *Plasmodium falciparum* genes. (A) Mean tag density and (B) Normalized mean tag density of gene profile of H3K4me3 and H3K9ac histone modifications over 500 expressed *P. falciparum* genes. To normalize the tag density between histone modifications, reads in each bin are divided by total number of reads for a histone modification. (C) Profiles of H3K4me3 and H3K9ac normalized with pan-H3. Gene unit including transcription start site (TSS) and transcription termination site (TTS) is shown below.

average gene profiles. Gene profile was divided into 100 bins (80 inside the gene body and 10 each on 5' and 3' ends). ChIP-sequencing tag density values were calculated as described earlier [1,7] for each dataset and each bin. To normalize the tag density between histone modifications, reads in each bin are divided by total number of reads for a histone modification (Fig. 2A and B). The profiles of histone modification are also normalized by dividing the reads in each bin with the histone H3 ChIP-sequencing (nucleosome) reads to nullify the uneven nucleosome distribution in *P. falciparum* (Fig. 2C).

#### 2.4. Correlation of histone modifications with gene expression

We categorized 5200 *P. falciparum* genes into 52 groups based on increasing expression levels. Histone modifications are found to be distributed throughout the gene unit in *P. falciparum* and therefore enrichment of histone modifications is calculated over the entire gene unit [7,8]. On each group, histone modification enrichment is summed by adding the mean tag density using R function for the gene unit. Whisker plots were plotted using R boxplot function with default parameters for each category for all the histone modifications. Pearson correlation was calculated between the median of gene expression and histone modification for each category.

### 3. Discussion

Here we describe information about *P. falciparum* culturing, chromatin immunoprecipitation controls and data analysis. The dataset comprises ChIP-sequencing of multiple histone modifications in a stage specific manner in *P. falciparum*. The dataset, which is recently published [1], first time visualized histone modification profiles at the TSSs and TTSs as opposed to translation start sites and translation termination sites [4,9]. Integrative analysis of datasets we have generated reveals distinct modes of epigenetic and transcriptional regulation in *P. falciparum*, which can influence parasite virulence and pathogenicity [1]. In addition, the datasets presented here serves as a framework for further investigation of histone modifications to better understand the epigenetic machinery in *Plasmodium falciparum* and other related parasites.

### Acknowledgments

We thank Prof. Sanjeev Galande and Saurabh J. Pradhan for performing ChIP-sequencing, Dr. P.C. Reddy for TSSs and TTSs prediction analysis and critical reading of the manuscript. This work is supported by grants under the DST-INSPIRE Faculty Award (IFA-13 LSBM-53) to KK.

### References

- [1] K. Karmodiya, S.J. Pradhan, B. Joshi, R. Jangid, P.C. Reddy, S. Galande, 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* 8 (2015) 32.
- [2] W.A. Hoeijmakers, R. Bártfai, K.J. François, H.G. Stunnenberg, Linear amplification for deep sequencing. *Nat. Protoc.* 6 (7) (2011) 1026–1036.
- [3] A. Radfar, D. Mendez, C. Moneriz, M. Linares, P. Marín-García, A. Puyet, A. Diez, J.M. Bautista, Synchronous culture of *Plasmodium falciparum* at high parasitemia levels. *Nat. Protoc.* 4 (2009) 1828–1844.
- [4] R. Bártfai, W.A.M. Hoeijmakers, A.M. Salcedo-Amaya, A.H. Smits, E. Janssen-Megens, A. Kaan, M. Treeck, T. Gilberger, K. François, H.G. Stunnenberg, H2A.Z demarcates intergenic regions of the *Plasmodium falciparum* epigenome that are dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog.* 6 (2010), e1001223.
- [5] C. Trapnell, A. Roberts, L. Goff, G. Pertea, D. Kim, D.R. Kelley, H. Pimentel, S.L. Salzberg, J.L. Rinn, L. Pachter, Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7 (2012) 562–578.
- [6] T. Ye, A. Krebs, M.A. Choukallah, C. Keime, F. Plewniak, I. Davidson, L. Tora, seqMINER: an integrated ChIP-seq data interpretation platform. *Nucleic Acids Res.* 39 (2011), e35.
- [7] A.R. Krebs, K. Karmodiya, M. Lindahl-Allen, K. Struhl, L. Tora, SAGA and ATAC histone acetyl transferase complexes regulate distinct sets of genes and ATAC defines a class of p300-independent enhancers. *Mol. Cell* 44 (2011) 410–423.
- [8] K. Karmodiya, A. Krebs, M. Oulad-Abdelghani, H. Kimura, L. Tora, H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics* 13 (2012) 424.
- [9] A.M. Salcedo-Amaya, M.A. van Driel, B.T. Alako, M.B. Trelle, A.M. van den Elzen, A.M. Cohen, E.M. Janssen-Megens, M. van de Vegte-Bolmer, R.R. Selzer, A.L. Iniguez, R.D. Green, R.W. Sauerwein, O.N. Jensen, H.G. Stunnenberg, Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 9655–9660.