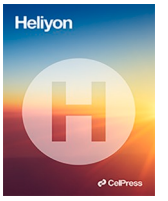
Heliyon 10 (2024) e38500

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Comparative genomics of two protozoans *Dictyostelium discoideum* and *Plasmodium falciparum* reveals conserved as well as distinct regulatory pathways crucial for exploring novel therapeutic targets for Malaria

Shivam Nanda [a], Rajan Pandey [b], Rahila Sardar [b], Ashutosh Panda [c], Aruna Naorem [a,**], Dinesh Gupta [b,***], Pawan Malhotra [c,*]

[a] *Department of Genetics, University of Delhi, South Campus, New Delhi, 110 021, India*
[b] *Translational Bioinformatics Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, 110 067, India*
[c] *Malaria Biology Group, International Centre for Genetic Engineering and Biotechnology, New Delhi, 110 067, India*

ARTICLE INFO

Keywords:
Protozoans
Plasmodium
Dictyostelium
Comparative genomics
Post-translational modulators

ABSTRACT

Plasmodium falciparum, which causes life-threatening cerebral malaria has rapidly gained resistance against most frontline anti-malarial drugs, thereby generating an urgent need to develop novel therapeutic approaches. Conducting in-depth investigations on *Plasmodium* in its native form is challenging, thereby necessitating the requirement of an efficient model system. In line, mounting evidence suggests that *Dictyostelium discoideum* retains both conformational and functional properties of *Plasmodium* proteins, however, the true potential of *Dictyostelium* as a host system is not fully explored. In the present study, we have exploited comparative genomics as a tool to extract, compare, and curate the extensive data available on the organism-specific databases to evaluate if *D. discoideum* can be established as a prime model system for functional characterization of *P. falciparum* genes. Through comprehensive *in silico* analysis, we report that despite the presence of adaptation-specific genes, the two display noteworthy conservation in the housekeeping genes, signaling pathway components, transcription regulators, and post-translational modulators. Furthermore, through orthologue analysis, the known, potential, and novel drug target genes of *P. falciparum* were found to be significantly conserved in *D. discoideum*. Our findings advocate that *D. discoideum* can be employed to express and functionally characterize difficult-to-express *P. falciparum* genes.

1. Introduction

Malaria, an infectious disease is a serious public health problem affecting approximately 249 million individuals worldwide with a global tally of 5,80,000 deaths, specifically in Asian and African subcontinents [1]. Although various *Plasmodium* species such as *P. vivax*, *P. malariae*, *P. ovale,* and *P. knowlesi* have been implicated in malarial etiology, *P. falciparum* is notable for causing cerebral

* Corresponding author.
** Corresponding author.
*** Corresponding author.
E-mail addresses: aruna.naorem@south.du.ac.in (A. Naorem), dinesh@icgeb.res.in (D. Gupta), pawanm@icgeb.res.in (P. Malhotra).

https://doi.org/10.1016/j.heliyon.2024.e38500
Received 8 June 2024; Received in revised form 23 September 2024; Accepted 25 September 2024
Available online 26 September 2024

malaria [2]. The challenge to prevent *P. falciparum* inflicted malarial infections is gaining momentum due to the growing resistance presented by the parasite against most frontline anti-malarial drugs namely, chloroquine, artemisinin, sulphadoxin-pyrimethamine, etc. [3–5]. The surging episodes of drug resistance generate an urgent need to identify both novel drug targets and alternative treatments.

P. falciparum is a parasitic protozoan that exhibits a complex life cycle encompassing two host systems. The parasite undergoes a series of transformations resulting in different stages of its life cycle with its asexual phase taking place in the human host, and the sexual as well as the sporogonic phase occurring in the female *Anopheles* mosquito [6]. Approximately 50 % of the total genes present in *P. falciparum* genome have not been functionally annotated [7,8] primarily because it is challenging to genetically manipulate the organism and its A + T-rich genome renders deriving its proteins in an active form difficult [9]. Several heterologous systems have been used to express and characterize *P. falciparum* genes but with limited success [10]. Recently, *Dictyostelium discoideum* has emerged as a better alternative for conducting such studies [11–14].

D. discoideum, like *P. falciparum,* has an A + T rich genome but unlike the parasite, it is a free-living amoeboid that is abundantly found in soil and moist leaf litter. It displays a life cycle altering between unicellular and multicellular forms which is strictly regulated by the availability of nutrition [15]. It exists as a single-celled haploid organism feeding voraciously on soil microbes representing the vegetative phase until the food source lasts. As the food exhausts, the developmental phase ensues marked by an orchestrated aggregation of starving amoebae transitioning from an independent unicellular form into various multicellular structures [16]. Accredited to this developmental property, ease of culture, and amenability to genetic manipulation, *D. discoideum* has emerged as a prime model organism for studying various biological functions such as cell motility, cytokinesis, signal transduction, phagocytosis, chemotaxis, and cell differentiation [17,18]. Furthermore, evolutionarily conserved genes and signaling pathways with higher organisms have established *D. discoideum* as a promising model system for various biomedical and human diseases-related research including neurodegenerative disorders and host-pathogen interactions [19,20].

Based on the reports that *D. discoideum* retains both conformational and functional properties of *Plasmodium* proteins [11,13], we compared the genomes of *D. discoideum* and that of *P. falciparum* using the data available in public databases, to explore whether *D. discoideum* can be developed as an ideal model system to express *Plasmodium* genes and to carry out the functional analysis of these proteins. The analysis involved a comparison of genome characteristics of the two organisms in interest, comparing gene ontology pertaining to various cellular processes. In addition, an orthologue analysis was performed for genes encoding proteins involved in cell cycle, metabolism, cell motility, autophagy, and post-translational modifications. The findings of the comparative genome analysis of these two protozoans are presented in this study.

2. Materials and methods

2.1. Comparison of the genome characteristics of D. discoideum and P. falciparum

For the comparison of various genome-specific characteristics namely, genome size, chromosome number, number of genes and proteins encoded, etc., appropriate data were retrieved for the AX4 strain of *D. discoideum* from Dictybase (http://dictybase.org/Dicty_Info/genome_statistics.html, accessed on 17th April 2024) and 3D7 strain of *P. falciparum* from PlasmoDB (release 68, accessed on 10th May 2024) along with the genome database of NCBI for the two organisms (https://www.ncbi.nlm.nih.gov/genome, accessed on 17th April 2024) [21,22]. The information retrieved for both organisms was subsequently tabulated and compared. In line, the OrthoMCL database was referred to determine the abundance of proteins corresponding to different orthologous categories for both organisms. Lastly, an OrthoMCL (release 6.21, accessed on May 14, 2024) search was performed to estimate the overall conservation of *P. falciparum* proteins in *D. discoideum*.

2.2. Phylogenetic analysis

For phylogenetic analysis, rRNA gene sequences were taken. These sequences are essential for ribosome function and consist of hypervariable regions that provide species-specific signature sequences and conserved regions that reflect phylogenetic relationships among species. These are highly valuable tools for genomics and are commonly used for phylogenetic analysis [23,24]. The sequences of 16S, 17S, and 18S rRNA for the species of *D. discoideum*, *P. falciparum,* and other organisms from different taxa including *Tetrahymena thermophila* were retrieved from the nucleotide database of NCBI (https://www.ncbi.nlm.nih.gov/nucleotide/, accessed on May 21, 2024). MUSCLE (a program for aligning multiple amino acid and nucleotide sequences, embedded in MEGA X software) was used to generate multiple sequence alignments of the retrieved sequences [25]. MUSCLE alignment parameters included a gap opening penalty (GOP) of −400 and a gap extension penalty (GOE) of 0 with 8 iterations. Other parameters included a minimum diagonal length of 24 and the UPGMB clustering method. The evolutionary history was inferred using the Maximum Likelihood method based on the Tamura-Nei model [26]. The bootstrap consensus tree inferred from 1000 replicates was taken to represent the evolutionary history of the taxa analyzed [27]. Branches corresponding to partitions reproduced in less than 50 % bootstrap replicates were collapsed. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [27]. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Joining and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach and then selecting the topology with superior log likelihood value. The analysis involved 10 nucleotide sequences. All positions with less than 95 % site coverage were eliminated, i.e., fewer than 5 % alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 1517 positions in the final dataset. Evolutionary analyze s were conducted in MEGA X software [28].

2.3. Gene ontology and comparative analysis for genes involved in metabolic pathways

To perform gene ontology, AmiGO's Gene Ontology database search (https://amigo.geneontology.org/amigo, accessed on May 25, 2024) and PANTHER program (Release 17.0, accessed on May 25, 2024) were used to retrieve gene annotations for *D. discoideum* and *P. falciparum*. These databases assign a functional annotation to a protein on the basis of sequence similarity/identity using sequence alignment [21,22,29,30]. To perform a comprehensive comparison of the metabolic pathways of *D. discoideum* and *P. falciparum*, metabolic pathway-specific genes of both organisms were one-to-one compared to identify unique and diverse functional genes using the PANTHER program.

2.4. Annotation of D. discoideum proteins with conserved post-translational modulator domains and orthologue analysis with human malarial parasite P. falciparum

Initially, previously published articles were referred [31] and text-based searches with keywords for various post-translational modulators like "kinase", phosphatase", "peptidase/protease", "acetylase", deacetylase, "methyltransferase", "heat shock proteins" and "PPIases" were performed on Dictybase to identify the various cognate proteins encoded by the *D. discoideum* genome. Next, the complete proteome of *D. discoideum* (AX4 strain) was retrieved from Dictybase [22] and was thoroughly scanned for the presence of conserved post-translational modulator domains. This was performed using the NCBI Conserved Domain Database search (version 3.21–62456 PSSMs, accessed on May 12, 2024) keeping a threshold e-value of 10^{-6} [32]. Additionally, a search for the above-mentioned conserved domains was also performed for each of these protein sequences using InterProScan 5 (https://www.ebi.ac.uk/jdispatcher/pfa/iprscan5, accessed on May 12, 2024) [33]. The results obtained from both the methods (NCBI CDD search and InterProScan 5) were manually compared and verified if the protein sequence harbors identical domain(s) before assigning it to a particular category of post-translational modulator(s). This data generated for *D. discoideum* was subsequently used for orthologue comparison with *P. falciparum* and *Homo sapiens*.

For orthologue analysis, an OrthoMCL database search was performed which categorizes two proteins as orthologues of each other based on sequence homology. In the present analysis, OrthoMCL pre-set default values for various parameters such as average percent identity (80 percent to 100 percent), average percent homology (80 percent to 100 percent), e-value (−200 to −150), etc., were taken [34,35]. Accordingly, the UniProt IDs of post-translational modulators of *D. discoideum* were fed to the OrthoMCL database and eventually, the *P. falciparum-specific* orthologues were obtained at the end of the search. The resultant list was then compared with the list of post-translational modulators of *P. falciparum* made by referring to the previously published reports, and databases such as PlasmoDB and Kipho (https://bioinfo.icgeb.res.in/kipho/, accessed on May 14, 2024) for further verification [36–39]. Furthermore, to discern the conserved post-translational modulators between *D. discoideum* and *H. sapiens*, an additional OrthoMCL analysis was also conducted. The final output data was compared to ascertain the unique and conserved post-translational modulators amongst all the three organisms of interest.

2.5. Identification of conserved gene(s) and protein(s) participating in various cellular processes, transcription regulation, and potential drug-based targets

To ascertain the presence of orthologues of *P. falciparum* proteins involved in virulence, hemoglobin digestion, heme detoxification, cellular motility, autophagy, etc., in *D. discoideum* both NCBI's BLASTP (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins, performed on May 15, 2024) and OrthoMCL analysis using Uniport IDs of respective *Plasmodium* proteins against *D. discoideum* were carried out. In case, no significant hits were obtained through primary sequence NCBI blast search, secondary structure-based HHpred (https://toolkit.tuebingen.mpg.de/tools/hhpred, performed on May 17, 2024) analysis was performed with default parameters to find orthologues for *P. falciparum* [40]. Further using previously published reports as a reference and CisBP (Database build 1.02, May 25, 2024) depositories [41], we performed a detailed analysis of transcription factors [42–44]. Lastly, to find out the orthologues in *D. discoideum* against various known and potential drug targets in *P. falciparum*, previously published studies were referred to collect reported drug targets which were then fed to the OrthoMCL database for further analysis [45,46].

2.6. Tabulation and graphical representation

The data retrieved through various databases and analyses were systematically entered and tabulated in Microsoft Office Excel 2013 and the graphs depicting the genes and proteins distribution were generated using PRISM GraphPad software (ver8.0.2). A flowchart depicting the methodology and analysis pipeline followed for this study has also been provided (Supplementary Fig. 1).

3. Results

3.1. Genome characteristics of D. discoideum and P. falciparum

To get insights into the features of *P. falciparum* and *D. discoideum* genomes, we compared the genome information of *P. falciparum* and *D. discoideum* using their respective databases; DictyBase (http://dictybase.org/Dicty_Info/genome_statistics.html) and PlasmoDB ver 68 (Table 1). The genome size of *D. discoideum* and *P. falciparum* are 34 Mb and 23.3 Mb respectively, although the number of chromosomes is more than twice in *P. falciparum*, 14 as compared to the 6 present in *D. discoideum*, [22,47]. Both organisms have very

high AT content, 81 % in *P. falciparum* and 77.6 % in *D. discoideum*. The AT content in *P. falciparum* rises to approximately 90 % in the introns and intergenic regions while in *D. discoideum* it can exceed up to 95 % in the intergenic region [48,49]. It is also found that the total number of genes both coding and non-coding, are higher in *D. discoideum* except for snoRNA genes and other ncRNA genes that are comparatively more in *P. falciparum* (Table 1). Recent investigations using RNA-seq data have shown that *P. falciparum* and *D. discoideum* encode 164 and 621 long-non-coding RNA, respectively, and these regulate various molecular functions during development [50,51]. The number of annotated genes is higher in *P. falciparum*, probably because of the extensive work carried out on *P. falciparum*. The information retrieved through organism-specific databases and NCBI's genome database showed 12,739 proteins are encoded by *D. discoideum* genome and 5318 proteins encoded by *P. falciparum* genome. These proteins have been classified into 8758 and 4674 orthologues groups, respectively by OrthoMCL (Supplementary File 1). Further analysis showed 1827 *P. falciparum* proteins have homologues in *D. discoideum* (2029 proteins distributed in 1667 orthologues groups) (Supplementary File 1 and Supplementary Figure 2A). Of note, *P. falciparum* has a higher number of proteins having functional repeats and amino acid homo polymer repeat regions or low complexity regions in comparison with *D. discoideum* which also harbors a large number of such proteins, however, has a comparatively smaller number of pseudogenes (Table 1) [52–56].

3.2. Phylogenetic analysis shows that D. discoideum and P. falciparum are evolutionarily distinct

To analyze how *D. discoideum* and *P. falciparum* are phylogenetically linked, we performed evolutionary phylogenetic analysis using 16S/17S/18S rRNA gene sequences for different species of *Dictyostelium, Plasmodium,* and several other organisms including *T. thermophila,* which is a free-living ciliate that can switch from communalistic to pathogenic mode of survival under different conditions [57]. The genomes of *D. discoideum* and *P. falciparum* are AT-rich while the genomes of *P. vivax* and *P. knowlesi* have high GC content. The phylogenetic analysis showed that distinct clusters were formed by *Plasmodium* spp. and *Dictyostelium* spp, separated by *T. thermophila*, which shows both the characteristics. This indicates that comparatively small differences lie among *Plasmodium* species despite having differences in their genome composition (Fig. 1). However, as expected, it was observed that *Plasmodium* spp., *Dictyostelium* spp. and other protozoan spp. formed distinct clusters suggestive of the significant changes that had occurred in the genomes of protozoans and metazoans.

3.3. Comparative gene ontology and metabolic pathway analysis reveals the presence of both conserved and distinct proteins

Gene ontology data retrieved from AmiGO database (AmiGO 2) resulted in the identification of 5447 *P. falciparum* genes and 9435 *D. discoideum* genes with annotated biological processes and subcellular localization (Fig. 2). Gene ontology analysis shows a similar pattern for conserved genes involved in housekeeping functions as the ratio of the number of genes involved in these cellular processes is in accordance with the genome size of the two organisms. However, a significant change is seen in the ratio or the number of genes involved in the cellular processes related to the adaptive pathways required by these two organisms for their survival in their respective niches (Fig. 2B). For example, in cellular component analysis, the number of proteins involved in cellular homeostasis, cell periphery and membrane localization were found to be 4 times higher in *D. discoideum* as compared to *P. falciparum,* whereas apical proteins, rhoptry proteins and cell surface proteins were comparatively much higher in *P. falciparum* (Fig. 2A). It appears that the differences in these proteins reflect their respective ecology and adaptation for survival.

Likewise, the number of genes involved in various biological processes was seen to be much higher in *D. discoideum* than in *P. falciparum* based on molecular functional analysis using AmiGO (source: AmiGO gene ontology database version AmiGO 2). However, in some biological processes, this correlation between genome size and the number of proteins does not hold true (Fig. 2B). For instance, proteins involved in cell adhesion regulation and stimulus responses were higher in number *in P. falciparum* than in *D. discoideum.* On the contrary, it was observed that *D. discoideum* has more proteins involved in detoxification than in *P. falciparum.*

Table 1
Comparison between *P. falciparum* (PlasmoDB version 68) and *D. discoideum* (dictybase).

Features	*P. falciparum* (3D7 strain)	*D. discoideum* (AX4 strain)
Genome size	23.3 Mb	34 Mb
Number of Chromosomes	14	6
% GC content	19	22.4
% AT content	81	77.6
Total genes	5720	12947
Number of protein-coding genes	5389	12257
rRNA genes	65	200
tRNA genes	78	418
snRNA genes	5	18
snoRNA genes	41	19
Other ncRNA genes	55	35
GO annotations	5447	9435
Repeat regions	0.44	0.11
Scaffolds	15	41
Pfam Domains	6300	14238
Number of Pseudogenes	158	650

Table 2
Transcription factors families in the CisBP database (Database build 1.02).

Transcription Factors Family	Organisms			
	P. falciparum (3D7 strain)	*P. vivax* (P01 strain)	*D. discoideum* (AX4 strain)	*D. purpureum* (DpAX1 strain)
AP2	26	27	–	–
Myb/SANT	7	5	28	32
Sox	3	4	5	6
Arid/Bright	1	1	3	2
C2H2 ZF	2	3	8	6
CBF/NF-Y	1	1	–	–
CSD	1	–	–	–
Ets	1	–	–	–
mTERE	1	1	–	–
p53	1	–	–	–
TBP	3	3	1	1
GATA	–	–	23	35
Homeodomain	–	–	13	21
E2F	–	–	7	2
MADS box	–	–	4	2
Forkhead	–	–	17	1
HSF	–	–	1	1
NAC/NAM	–	–	1	1
bZIP	–	–	19	14
Ndt80/PhoG	–	–	2	3
NFX	–	–	1	1
Nuclear receptor	–	–	1	–
RWP-RF	–	–	5	2
Zinc cluster	–	–	2	2
AT hook	–	–	1	–
WRKY	–	–	1	2
ABF 1	–	–	2	–
TCR/Cxc	–	–	1	1

Note: "-" indicates "no orthologues" found.

This can be explained by the fact that *D. discoideum*, a free-living soil organism, cohabits in soil with many soil microorganisms secreting various toxins for example DNA-damaging agents, etc., and the proteins of the organism involved in the detoxification process detoxify and counter the hostile and fluctuating surrounding environments in soil. On the other hand, *P. falciparum* having a parasitic life cycle, may use host proteins in the detoxification process, in addition to its own proteins whenever required. Some of the published reports suggest that *P. falciparum* imports human Peroxiredoxin-2 and Peroxiredoin-6 proteins into its cytosol, from the erythrocytes for peroxide detoxification. Inhibition of these host-specific enzymes has been lethal for the parasite [58,59]. Thus, the observed pattern reflects the adaptation of the organisms to their respective environment.

Classification using the Panther program categorizes *D. discoideum* and *P. falciparum* proteins into 23 proteins classes with the maximum number clustered into unclassified protein classes, 58.8 % and 53.3 %, respectively (Fig. 3B). The analysis showed that a maximum number of proteins belonged to metabolite interconversion enzyme, followed by protein modifying enzymes and RNA metabolism proteins (Fig. 3A and B). Interestingly, classification of proteins into different groups reveals comparable percent distribution as the overall number of proteins encoded by the organism in accordance with its genome size with an exception of some groups such as RNA metabolism proteins, translational proteins and protein-binding activity modulators.

Similarly, metabolic pathway analysis showed conservation of the enzymes involved in cellular housekeeping pathways such as glycolysis, ATP synthesis, heme biosynthesis, ubiquitin-proteasome pathway, DNA replication and cell cycle, etc., between the two organisms. However, significant difference was observed in the number of annotated proteins for cholesterol biosynthesis, FS-7-associated surface antigen (FAS) signaling (required for programmed cell death/apoptosis) and Fibroblast growth factor (FGF) signaling, etc., in these two genomes (Supplementary File 2).

3.4. Comparative analysis of post translational modulators indicates significant number of conserved proteins

To gain insights into the regulatory proteins present in these two genomes, comparative analysis for the post-translational modulators related to these two genomes was performed. Although reports on post-translational modulators in *P. falciparum* genome are available, but no report on a comprehensive analysis for *D. discoideum* post-translation modulators is documented yet. The search resulted in finding various proteins involved in protein folding, modification and processing pathways such as chaperones, PPIases, kinases, phosphatases, proteases, methyltransferases, acetyltransferases and deacetylases, etc. of *D. discoideum.* (Fig. 4 and Supplementary Tables 1–5). Further, comparative orthologue analysis of post-translational modulators between *D. discoideum* and *P. falciparum* genomes revealed expansion for kinases, phosphatases, proteases, methyltransferases, acetyltransferases and deacetylases in *D. discoideum* genome (Fig. 4). The analysis revealed a similar distribution pattern for kinases in different families but showed great expansion of Protein tyrosine phosphatase (PTP) superfamily in *D. discoideum* as compared to *P. falciparum* (Supplementary

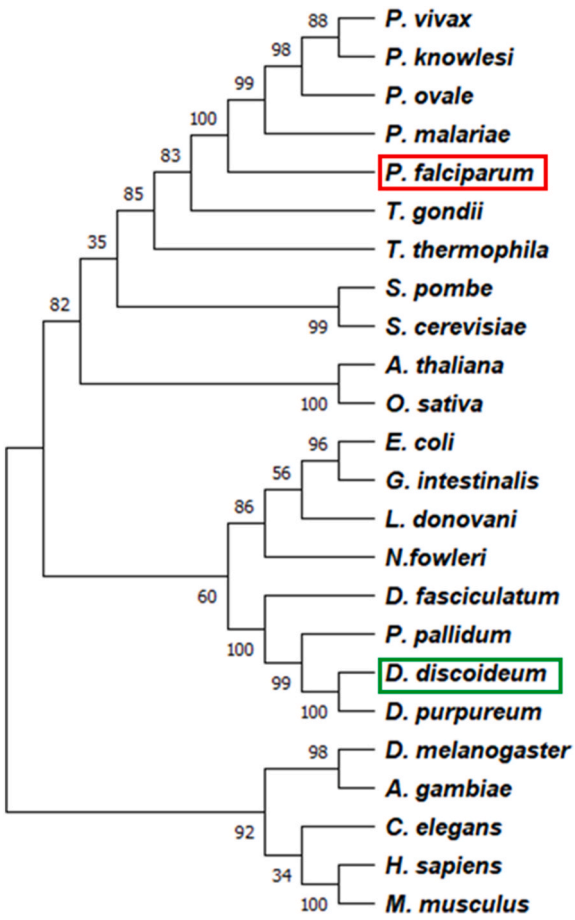


Fig. 1. Phylogenetic tree generation and evolutionary relationship among *Plasmodium* and *Dictyostelium* spp. using 16S/17S and 18S rRNA gene sequences and MEGA X software. The Maximum-Likelihood method with 1000 bootstraps was used to perform the phylogenetic analysis. The numerical values show the number of times a sequence occupied that place while the phylogenetic tree was constructed 1000 times (bootstrap value) by the program. The values are shown out of 100 instead of 1000.

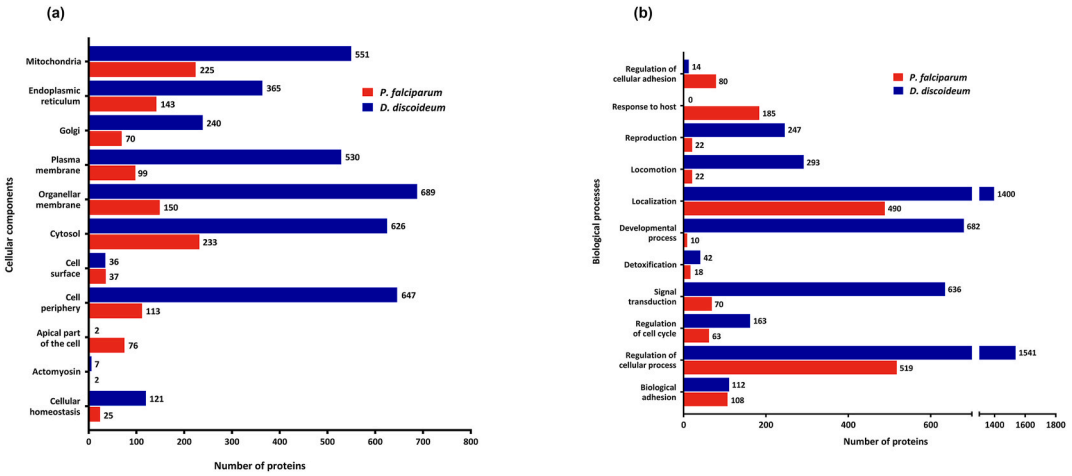


Fig. 2. Comparative gene ontology (GO) analysis for *P. falciparum* and *D. discoideum* (Amigo database) (A) Cellular localization (B) Biological processes in which identified proteins are involved. Red colored bars represent data for *P. falciparum*, whereas blue colored bars represent data for *D. discoideum*. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

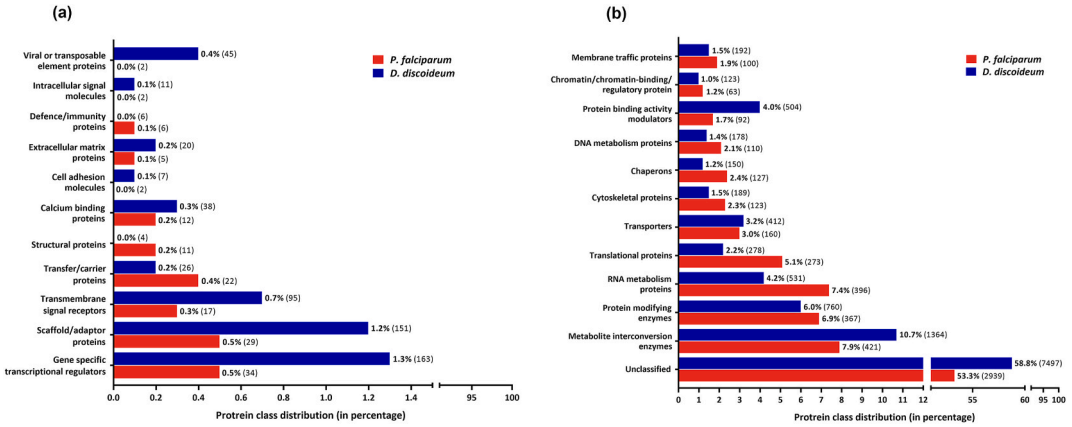


Fig. 3. Comparative protein class distribution for *P. falciparum* and *D. discoideum*. Red colored bars represent data for *P. falciparum*, whereas blue colored bars represent data for *D. discoideum*. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 2). Additionally, sequence based orthologue analysis through the OrthoMCL database showed that *D. discoideum* and *P. falciparum* have 19, 44 and 64 conserved orthologue groups for kinases, phosphatases and proteases, respectively (Supplementary Tables 1, 2, 3 and Supplementary Figures 2-B, C and D). Similarly, *D. discoideum* and *P. falciparum* have 3 and 9 conserved orthologous groups for methyltransferases and (acetyl/deacetyl) transferases (Supplementary Tables 4 and 5). The number of proteins belonging to studied post-translational modulators for identified conserved orthologue groups is higher in *D. discoideum* than in *P. falciparum* but an expansion of proteins was found to be in few orthologue groups in *D. discoideum* (Supplementary Tables 1–5). It was also noted that these proteins found in *D. discoideum* showed a considerable degree of conservation with *H. sapiens* proteins.

Additionally, an orthologue analysis was performed on proteins with chaperone activities; heat shock proteins (HSPs) and peptidyl prolyl *cis/trans* isomerases (PPIases). HSPs function as molecular chaperones and are essential for survival under normal and stress conditions. They play an important role in maintaining cellular proteostasis by integrating the fundamental process of protein folding and degradation. Some of the heat shock proteins of *D. discoideum* were found to be conserved in both *P. falciparum* and *H. sapiens* (Table 3a, 3b). On the other hand, Peptidyl-prolyl *cis/trans* isomerases which form a chaperone superfamily mainly consist of cyclophilins, FK-506 binding proteins (FKBPs), parvulins and Protein-ser/thr-phosphatase2A activators (PTPA). These enzymes catalyze the *cis/trans* isomerization of peptide bonds preceding proline in proteins, thereby regulating protein folding, activation, degradation and localization of several protein substrates. We also observed that based on protein sequence similarity, some of the cyclophilins and FKBPs found in *D. discoideum* were also found to be conserved in *P. falciparum*. It is notable that high degree of conservation among all PPIases in *D. discoideum* and those in *H. sapiens* was also observed (Table 4).

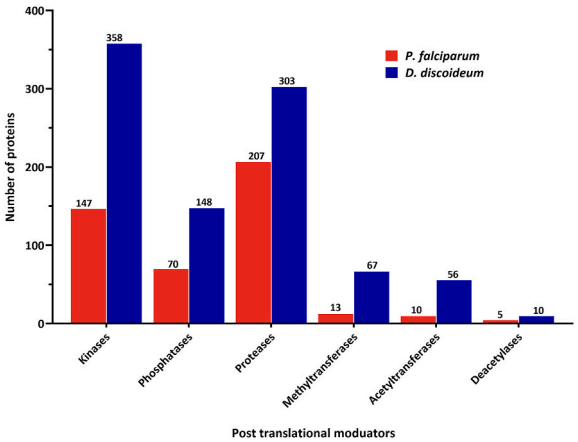


Fig. 4. Comparative genome-wide distribution of post-translational modulators in *P. falciparum* and *D. discoideum*. Red colored bars represent data for *P. falciparum*, whereas blue colored bars represent data for *D. discoideum*. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 3a
D. discoideum heat shock proteins and their orthologues in *P. falciparum* and *H. sapiens*.

D. discoideum (AX4 strain) ID	*D. discoideum* proteins	*P. falciparum* (3D7 strain) ID	*H. sapiens* ID	OrthoMCL ID
Q9NKX1	Glucose regulated protein 94	Q25869	P08238 P07900-1	OG6_100204
P54651	Hsp90	Q25869	P08238 P07900-1	OG6_100204
Q55ED4	NF-KappaB activating protein	O15739	Q8N5F7	OG6_103179
Q6TMK3	Heat shock protein 70	Q8IC01	B4DYH1	OG6_100680
Q557E0	Heat shock protein 70 E	Q07615	P11021	OG6_100083
Q54BE0	Hsp70H2	K7NTP5	P11021	OG6_100083
Q810H7	mHSP70	Q8II24	P38646	OG6_100423
Q54EV3	Hsp101	QN9FG9	Q9H078-1	OG6_100223
P36415	hspB	P12794	Q53FA3 P54652 P08107 P34931 P11021 P17066 P11142-1	OG6_100083

Note: All proteins have been denoted with their UniProt IDs.

Table 3b
D. discoideum heat shock proteins having no orthologues in *P. falciparum* and *H. sapiens*.

D. discoideum (AX4 strain) ID	*D. discoideum* Protein	OrthoMCL ID
Q76NU5	HspG12	OG6_114203
Q8MPA5	HspG7	OG6_114203
Q550E9	HspG8	OG6_114203
Q8MPA7	HspG5	OG6_114203
Q86KF5	HspG9	OG6_114203
Q550E5	HspG10	OG6_114203
Q550E3	HspG11	OG6_114203
Q55GH8	HspL	OG6_100183
Q552K9	HspG2	OG6_114203
Q86I25	HspG1	OG6_114203
Q54MR6	HspN	OG6_592477
Q54IG2	Hsp83	OG6_114203
Q1ZXL6	HspF-2	OG6_591334
Q54WT6	HspJ	OG6_590801
Q54W43	HspM	OG6_114203
Q551F8	HspG3	OG6_114203
P54658	HspC	OG6_590870
Q54VP4	Hsp48	OG6_100183
Q54I91	HspI	OG6_129811
C7G080	Hsp19	OG6_100183
Q551G1	HspG4	OG6_114203
Q8MMN1	HspG6	OG6_114203
Q54QE9	Hsp69	OG6_174950
Q86H60	HspK	OG6_165654

Note: All proteins have been denoted with their UniProt IDs.

3.5. Orthologue analysis showed the conservation of proteins involved in cell motility and autophagy pathways with distinct proteins involved in P. falciparum pathogenicity

Regulation at the level of transcription and post-translational modifications plays a vital role in many cellular processes such as cell survival, cell motility, development, proliferation, etc. To check whether the proteins in some of these cellular pathways are present in both *D. discoideum* and *P. falciparum*, a search using BLASTP was performed to identify orthologues in the genomes of these organisms. The analysis showed that the proteins which are involved in regulating cell cycle and cell motility, for instance, Cyclin-dependent kinases (Cdks) and phosphatases, Cdc25, etc., seemed to have significant conservation (Supplementary Tables 1, 2 and 6). Actin and Tubulin involved in cellular motility also showed a high degree of similarity (At 96 % query coverage these proteins show 80 % and 68 % identity respectively, data not shown). These results indicate that proteins essential for vital cellular activities are significantly conserved between these organisms suggesting probable conservation in the signaling pathways.

It is interesting to note that orthologous proteins involved in hemoglobin degradation and heme detoxification pathways in *P. falciparum* were absent in *D. discoideum* (Supplementary File 3). Similarly, genes implicated in pathogenicity and virulence for *P. falciparum*, such as Erythrocyte membrane protein 1 (PfEMP1) [60,61], Erythrocyte binding antigens (EBA-140, EBA-175 and

Table 4
D. discoideum PPIases and their orthologueues in *P. falciparum* and *H. sapiens*.

D. discoideum (AX4 strain) ID	*D. discoideum* PPIases	*P. falciparum* (3D7 strain) ID	*H. sapiens* ID	OrthoMCL ID
Q54SM3	PpiA	Q76NN7	P23284 P62937 P30405	OG6_100528
Q9UA41	Cyclophilin D	Q76NN7	P23284 P62937 P30405	OG6_100528
Q55F01	Ppwd1	Q8I402	Q96BP3	OG6_102016
Q9NI62	Cyclophilin E	Q8I3I0	Q9Y3C6	OG6_102285
Q86J17	PpiH	Q27716	O43447	OG6_103291
B0G146	Ppil 2	Q8I2K8	Q13556-1	OG6_102815
Q9TW32	Cyclophilin B	Q27745	P23284 P62937 P30405	OG6_100844
Q54E95	Ppil3	Q8IIK3	Q9H2H8-1	OG6_102733
Q54QI6	FKBP4	Q8I4V8	Q13451 P62942 Q02790 P68106-1	OG6_100299
Q554J3	FKBP3	Q8I4V8	Q13451 P62942 Q02790 P68106-1	OG6_100299
Q55643	Cyclophilin-type PPIase	Q8ILM0	Q6UX04-1	OG6_102804
Q54SR7	FKBP-type PPIase	–	P26885 Q9NWM8	OG6_101705
Q54Z53	Pin4	–	Q9Y237-1 Q9Y237-2	OG6_101694
Q55EZ0	PinA	–	Q13526	OG6_100390
Q54NB6	FKBP-type PPIase	–	–	OG6_104296
Q54LG6	FKBP1	–	–	OG6_174979
B0G119	FKBP-type PPIase domain containing protein	–	–	OG6_104296
Q86IX8	Cyclophilin-type PPIase	–	–	OG6_124560
Q54G21	FKBP-type PPIase	–	Q5T1M5-1	OG6_106619
Q54Y27	FKBP6	–	O75344	OG6_107945
Q54CU3	Cyclophilin-type PPIase	–	Q8WUA2	OG6_102469
Q54QS2	PpiD	–	Q08752	OG6_103831
P34137	PtpA1-2/PtpA1-1	–	P23468-1 P10586-1	OG6_100623

Note: "-" indicates "no orthologues" found.
All proteins have been denoted with their UniProt IDs.

EBA-180) [6], Merozoite surface proteins (MSPs) [62] and Apical membrane protein (AMA1) [63,64] were absent in *D. discoideum* substantiating the existence of organism-specific pathways depending on their growth niche.

Further analysis on autophagy proteins using OrthoMCL revealed sequence conservation for a few proteins, namely, ATG3, ATG7, ATG8, ATG12, and ATG18 in these two protozoans. (Supplementary Table 6). Furthermore, it is worth mentioning that autophagy proteins ATG1 and ATG4 of these organisms cannot be considered as orthologues on the basis of their sequence yet they displayed structural similarity through HHpred analysis (data not shown). (Supplementary File 3). However, no significant conservation was detected in sequence identity and secondary structure for the autophagy protein ATG9 in these two protozoa. These findings suggested that conservation of protein function was achieved by retaining protein structure despite the diversification at amino acid sequence in the autophagy pathway proteins.

3.6. Comparative analysis of transcription regulators shows conserved as well as distinct families

Analysis of transcription factors (TFs) revealed the presence of conserved members of Zinc finger, Helix-turn-helix, Myb, Cumin/JMC and HMG families in both *D. discoideum* and *P. falciparum* wherein the number of TFs was found to be higher in *D. discoideum*. A recent report on *P. falciparum* transcription factors has shown 96 transcription factors distributed among 19 families which accounts for a 2-fold higher number as reported in the CisBP database, suggesting a new search method may provide additional TFs in the genome of *D. discoideum* [29]. AP2 family, a major TF family present in *P. falciparum,* has been shown to regulate cell cycle transitions and in controlling different development stages of its parasitic life cycle [65–67]. Similarly, transcription factors belonging to the bZIP family help in the regulation of expression of genes responsible for driving cell proliferation and differentiation in *D. discoideum* [68,69] and the transcription factors of bZIP family do no find orthologues in *P. falciparum* but the AP2 transcription factors help in the regulation of same cellular functions. Furthermore, orthologues of the AP2 transcription factor family along with CBF, CSD, Ets, and mTERE were not found in *D. discoideum* (Table 2). Also, many families of TFs present in *D. discoideum* were found to be unique,

suggesting these unique TFs may play a vital role in regulating processes specific to these two protozoans' life cycle and ecology. However, a family of TFs whose members are found to be present in both the organisms, might regulate different cellular processes. For example, the transcription factors of Myb/SANT family are found in both *D. discoideum* and *P. falciparum*. Sequence analysis of the transcription factors of this family, belonging to both the organisms using SMART protein tool [70] revealed the presence of low-complexity regions and SANT domain (data not shown). In *D. discoideum*, myb transcription factors help in the regulation of genes involved in stalk cell differentiation process [71]. While in *P. falciparum*, myb transcription factors such as *PfMyb*1, is needed for intra-erythrocytic growth of the parasite and it also regulates the expression of key genes involved in cell cycle regulation [72].

3.7. Orthologue analysis of known, potential and novel drug target genes identified in P. falciparum showed significant conservation in D. discoideum

Some novel and potential drug target genes have been identified in *P. falciparum*. Performing an orthologue analysis for already documented, potential and novel drug target genes using OrthoMCL showed that many of these genes find their orthologues in *D. discoideum* and a few are specific to *P. falciparum* (Supplementary Table 7).

In summary, the study revealed the presence of both the conserved and unique proteins in the studied AT-rich protozoan spp. for vital and organism-specific biological processes respectively.

4. Discussion

Plasmodium falciparum is responsible for causing cerebral malaria, the most severe form of the disease [2]. This parasite has gained widespread resistance against most frontline antimalarial drugs, so, there is an urgent need to search for novel drug targets as well as potent pharmacological compounds [3–5]. The biological pathways which play a crucial role in its pathogenesis and in imparting resistance against anti-malarial drugs are still enigmatic [73]. The lack of functional annotation of many of the *P. falciparum* genes [7,8] and difficulty in acquiring its proteins in active and soluble form for downstream processing [9] have made the requirement of a heterologous expression system essential. Recently, utilizing *D. discoideum* as an expression host, some of the *P. falciparum* proteins were successfully expressed for their functional characterization [11–14]. Furthermore, *P. falciparum* chloroquine resistance transporter was characterized using vesicles derived from *D. discoideum* as a substitute for hematin-containing digestive vacuoles from erythrocytes infected with *P. falciparum* permitting easy assessments of drug accumulation, pH and membrane potential, which are otherwise difficult to measure [12].

Attempts to express *Plasmodium* proteins in various other heterologous systems such as *Escherichia coli, Saccharomyces cerevisiae, Baculovirus, Xenopus laevis*, etc., were partially successful due to the aggregation of expressed proteins, formation of truncated proteins and proteins with undesired conformations [10]. In earlier studies, *T. thermophila* has also been used as a model to express genes of certain parasites, including *P. falciparum* [74,75]. Nevertheless, *D. discoideum* possesses certain advantages over *T. thermophila*, as a heterologous model system to express *P. falciparum* genes. In case of *T. thermophila*, to achieve a cell density of 2.2×10^7 cells/ml, a bioreactor is needed [75], whereas in *D. discoideum*, the same or even a higher cell density can be achieved with simple laboratory culturing conditions [76]. Both prokaryotes and eukaryotes use three codons, namely UAA, UAG and UGA, one at a time, to terminate the process of protein synthesis [77] but in case of *T. thermophile*, UAA and UAG are not used as stop codons. These codons code for the amino acid glutamine in this organism [75]. So, the genes having UAA and UAG as the stop codons would have to be codon optimized before expression in *T. thermophila*. On the other hand, *D. discoideum* uses all three codons (UAA, UAG and UGA) as stop codons, so codon optimization is not required prior to expression [78].

In the present study, we compared the genomes of *D. discoideum* (~34 Mb) and that of *P. falciparum* (~23.3 Mb) and our observations point towards developing *D. discoideum* as a model system to study the functions of *P. falciparum* genes/proteins. The rationale for the study is based on the fact that both these organisms belong to the protozoan family and their genomes are AT-rich. Comparative genomic analysis revealed that the genomes of *D. discoideum* and *P. falciparum* code for genus-specific proteins, which regulate specific biological processes essential for their adaptation to hosts and to the environment. For example, the proteins involved in hemoglobin degradation and invasion to specific host cells are specifically present in *P. falciparum* [79] but absent in *D. discoideum*. However, the overall number of genes involved in cell movement is found to be higher in *D. discoideum* implying more genes are required to maintain the complexity of cell movement and its regulation during hunt for food in naturally fluctuating environments and also in various cellular processes such as phagocytosis, cell division and development [80].

The proteins involved in several metabolic pathways, including enzymes necessary for glucose metabolism, purine and pyrimidine biosynthesis, DNA replication, ATP generation and cell division cycle are highly conserved in both of these organisms. Even the proteins associated with the autophagy pathway are also conserved in both. For example, the ATG8a and ATG8b proteins of *D. discoideum* exhibit sequence similarity with the ATG8 protein of *P. falciparum* (Supplementary Table 6). Significantly, key kinases, phosphatases, proteases, acetyl and methyl transferases required for post-translational regulation of different cellular activities along with essential membrane transporters found in *P. falciparum* [81] are conserved in these two organisms [82,83]. It can be surmised that the conservation of these post-translational modulators between the two organisms is expected to aid in preserving the structural conformation and function of *P. falciparum* proteins in *D. discoideum* [84].

Another interesting aspect of the present study is that the single amino acid homo polymers and functional domain repeats are found to be of common occurrence in both *P. falciparum* and *D. discoideum*. Proteins containing such low complexity regions are known to perform specialized functions in stress response, development, transcription, organelle biogenesis, transport and protein-protein interactions [85]. *D. discoideum* has the ability to keep such low complexity repeat-containing proteins in soluble form even if the

length of such repeats exceeds the pathogenic range for humans [86]. This property of *D. discoideum* could be exploited to express and characterize certain essential genes of *P. falciparum* such as the ones coding for ABC transporters, heat shock proteins and PPIases which also contain such repeats and are known to be essential for the survival of the parasite [87].

This study is an *in silico* analysis based on available curated genomes of the organisms studied here. The annotations of these genomes may slightly change over time, as further annotations and experiments are performed. Despite that, it is envisaged that the inference of the study shall remain largely the same. Our study reveals that *D. discoideum* is a promising and congenial system for the expression of *Plasmodium* genes/proteins, this organism could be further exploited to study the most important issues pertaining to *P. falciparum*. Several proteins in *D. discoideum* that are classified as peptidyl-prolyl cis/trans isomerases (PPIases) and heat shock proteins have corresponding orthologues in *P. falciparum*. It would be worthwhile to investigate if some of the conserved PPIases or heat shock proteins could also be potential targets for developing novel anti-malarial drug molecules. Furthermore, many of the known, potential and novel drug target genes identified in *P. falciparum*, have their orthologues in *D. discoideum*. This information could be utilized in performing complementation studies. In addition to this, functional characterization of potential drug target genes of *P. falciparum* can be attempted in *D. discoideum*, even if they do not find their orthologues in this organism.

5. Conclusions

The study reveals that the genomes of *D. discoideum* and *P. falciparum* code for genus-specific proteins which regulate biological processes specific to both the organisms. However, conservation is seen in the genes/proteins involved in various cellular processes such as cell division, cell motility, regulation of gene expression and autophagy. Both the organisms possess similar post-translational modification pathways and the proteins involved in these pathways. *D. discoideum* also harbors orthologues of potential and novel drug target genes/proteins of *P. falciparum*. The data thus suggest that *D. discoideum* can be a suitable host for expressing and characterizing *Plasmodium* genes, whereas similar attempts made using other systems like *E. coli, S. cerevisiae* etc. did not give desired results. In summary, *D. discoideum* can be exploited further for both understanding the parasitic disease and for developing potential drug targets in particular for *P. falciparum*.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this article and its supplementary information files.

Funding

Translational Bioinformatics Group at ICGEB (New Delhi) led by DG is supported by Department of Biotechnology, Government of India, India grants BT/PR40151/BTIS/137/5/2021 & BT/IC-06/003/91-Flagship Program. Malaria Biology Group at ICGEB, New Delhi, led by PM is supported by the Department of Biotechnology (DBT), Government of India, India grant BT/IC-06/003/91-Flagship Program and Welcome Team Science Grant (WTA/24/0006) as well as by the JC Bose grant (DST/20/015) provided by the Department of Science and Technology, Science and Engineering Research Board, Government of India. Gene Regulation Group at the Department of Genetics, University of Delhi, India, led by AN is supported by the Science and Engineering Research Board, Department of Science and Technology, Government of India, India, EMR/2016/002994-EMR and EEQ/2022/000330 projects.

CRediT authorship contribution statement

Shivam Nanda: Writing – original draft, Methodology, Investigation, Data curation. **Rajan Pandey:** Writing – review & editing, Methodology, Formal analysis, Data curation. **Rahila Sardar:** Formal analysis, Data curation. **Ashutosh Panda:** Writing – review & editing, Supervision, Formal analysis. **Aruna Naorem:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis. **Dinesh Gupta:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis, Conceptualization. **Pawan Malhotra:** Writing – review & editing, Supervision, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

SN acknowledges University Grants Commission, India for research fellowship. SN also sincerely thanks Himanshu Mishra, Vipul Yadav, Khanchuila Shingnaisui, Prerna Aggarwal and Somya Nanda for their valuable suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e38500.

References

[1] World malaria report, n.d. https://www.who.int/teams/global-malaria-programme/reports/world-malaria-report-2023, 2023. (Accessed 23 February 2024).
[2] S. Sato, Plasmodium-a brief introduction to the parasites causing human malaria and their basic biology, J. Physiol. Anthropol. 40 (2021) 1, https://doi.org/10.1186/s40101-020-00251-9.
[3] A.M. Dondorp, F. Nosten, P. Yi, D. Das, A.P. Phyo, J. Tarning, K.M. Lwin, F. Ariey, W. Hanpithakpong, S.J. Lee, P. Ringwald, K. Silamut, M. Imwong, K. Chotivanich, P. Lim, T. Herdman, S.S. An, S. Yeung, P. Singhasivanon, N.P.J. Day, N. Lindegardh, D. Socheat, N.J. White, Artemisinin resistance in Plasmodium falciparum malaria, N. Engl. J. Med. 361 (2009) 455–467, https://doi.org/10.1056/NEJMoa0808859.
[4] P.K. Mohapatra, A. Prakash, K. Taison, K. Negmu, A.C. Gohain, N.S. Namchoom, D. Wange, D.R. Bhattacharyya, B.K. Goswami, B.K. Borgohain, J. Mahanta, Evaluation of chloroquine (CQ) and sulphadoxine/pyrimethamine (SP) therapy in uncomplicated falciparum malaria in Indo-Myanmar border areas, Trop. Med. Int. Health 10 (2005) 478–483, https://doi.org/10.1111/j.1365-3156.2005.01401.x.
[5] C. Wongsrichanalai, A.L. Pickard, W.H. Wernsdorfer, S.R. Meshnick, Epidemiology of drug-resistant malaria, Lancet Infect. Dis. 2 (2002) 209–218, https://doi.org/10.1016/S1473-3099(02)00239-6.
[6] V. Soulard, H. Bosson-Vanga, A. Lorthiois, C. Roucher, J.-F. Franetich, G. Zanghi, M. Bordessoulles, M. Tefit, M. Thellier, S. Morosan, G. Le Naour, F. Capron, H. Suemizu, G. Snounou, A. Moreno-Sabater, D. Mazier, Plasmodium falciparum full life cycle and Plasmodium ovale liver stages in humanized mice, Nat. Commun. 6 (2015) 7690, https://doi.org/10.1038/ncomms8690.
[7] L. Florens, M.P. Washburn, J.D. Raine, R.M. Anthony, M. Grainger, J.D. Haynes, J.K. Moch, N. Muster, J.B. Sacci, D.L. Tabb, A.A. Witney, D. Wolters, Y. Wu, M. J. Gardner, A.A. Holder, R.E. Sinden, J.R. Yates, D.J. Carucci, A proteomic view of the Plasmodium falciparum life cycle, Nature 419 (2002) 520–526, https://doi.org/10.1038/nature01107.
[8] M.J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R.W. Hyman, J.M. Carlton, A. Pain, K.E. Nelson, S. Bowman, I.T. Paulsen, K. James, J.A. Eisen, K. Rutherford, S.L. Salzberg, A. Craig, S. Kyes, M.-S. Chan, V. Nene, S.J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A.B. Vaidya, D.M.A. Martin, A.H. Fairlamb, M.J. Fraunholz, D.S. Roos, S.A. Ralph, G.I. McFadden, L.M. Cummings, G.M. Subramanian, C. Mungall, J.C. Venter, D.J. Carucci, S.L. Hoffman, C. Newbold, R.W. Davis, C.M. Fraser, B. Barrell, Genome sequence of the human malaria parasite Plasmodium falciparum, Nature 419 (2002) 498–511, https://doi.org/10.1038/nature01097.
[9] V. Muralidharan, A. Oksman, P. Pal, S. Lindquist, D.E. Goldberg, Plasmodium falciparum heat shock protein 110 stabilizes the asparagine repeat-rich parasite proteome during malarial fevers, Nat. Commun. 3 (2012) 1310, https://doi.org/10.1038/ncomms2306.
[10] L.-M. Birkholtz, G. Blatch, T.L. Coetzer, H.C. Hoppe, E. Human, E.J. Morris, Z. Ngcete, L. Oldfield, R. Roth, A. Shonhai, L. Stephens, A.I. Louw, Heterologous expression of plasmodial proteins for structural studies and functional annotation, Malar. J. 7 (2008) 197, https://doi.org/10.1186/1475-2875-7-197.
[11] N. Fasel, C. Begdadi-Rais, M. Bernard, C. Bron, G. Corradin, C.D. Reymond, Dictyostelium discoideum as an expression host for the circumsporozoite protein of Plasmodium falciparum, Gene 111 (1992) 157–163, https://doi.org/10.1016/0378-1119(92)90683-g.
[12] J. Papakrivos, J.M. Sá, T.E. Wellems, Functional characterization of the Plasmodium falciparum chloroquine-resistance transporter (PfCRT) in transformed Dictyostelium discoideum vesicles, PLoS One 7 (2012) e39569, https://doi.org/10.1371/journal.pone.0039569.
[13] C.D. Reymond, C. Beghdadi-Rais, M. Roggero, E.A. Duarte, C. Desponds, M. Bernard, D. Groux, H. Matile, C. Bron, G. Corradin, N.J. Fasel, Anchoring of an immunogenic Plasmodium falciparum circumsporozoite protein on the surface of Dictyostelium discoideum, J. Biol. Chem. 270 (1995) 12941–12947, https://doi.org/10.1074/jbc.270.21.12941.
[14] J.M. Sá, M.M. Yamamoto, C. Fernandez-Becerra, M.F. de Azevedo, J. Papakrivos, B. Naudé, T.E. Wellems, H.A. del Portillo, Expression and function of *pvcrt-o*, a *Plasmodium vivax* ortholog of *pfcrt*, in Plasmodium falciparum and Dictyostelium discoideum, Molecular and Biochemical Parasitology 150 (2006) 219–228, https://doi.org/10.1016/j.molbiopara.2006.08.006.
[15] P. Schaap, Evolutionary crossroads in developmental biology: Dictyostelium discoideum, Development 138 (2011) 387–396, https://doi.org/10.1242/dev.048934.
[16] J.D. Dunn, C. Bosmani, C. Barisch, L. Raykov, L.H. Lefrançois, E. Cardenal-Muñoz, A.T. López-Jiménez, T. Soldati, Eat prey, live: Dictyostelium discoideum as a model for cell-autonomous defenses, Front. Immunol. 8 (2018) 1906, https://doi.org/10.3389/fimmu.2017.01906.
[17] S. Bozzaro, The past, present and future of Dictyostelium as a model system, Int. J. Dev. Biol. 63 (2019) 321–331, https://doi.org/10.1387/ijdb.190128sb.
[18] C.J. Pears, J.D. Gross, Microbe Profile: Dictyostelium discoideum: model system for development, chemotaxis and biomedical research, Microbiology (Read.) 167 (2021), https://doi.org/10.1099/mic.0.001040.
[19] J. Martín-González, J.-F. Montero-Bullón, J. Lacal, Dictyostelium discoideum as a non-mammalian biomedical model, Microb. Biotechnol. 14 (2021) 111–125, https://doi.org/10.1111/1751-7915.13692.
[20] C.L. Storey, R.S.B. Williams, P.R. Fisher, S.J. Annesley, Dictyostelium discoideum: a model system for neurological disorders, Cells 11 (2022) 463, https://doi.org/10.3390/cells11030463.
[21] S. Bozzaro, The model organism Dictyostelium discoideum, Methods Mol. Biol. 983 (2013) 17–37, https://doi.org/10.1007/978-1-62703-302-2_2.
[22] A. Bahl, B. Bunk, R.L. Coppel, J. Crabtree, S.J. Diskin, M.J. Fraunholz, G.R. Grant, D. Gupta, R. Huestis, J.C. Kissinger, P. Labo, L. Li, S.K. McWeeney, A. J. Milgram, D.S. Roos, J. Shug, C.J.S. Jnr, PlasmoDB: the Plasmodium genome resource. An integrated database providing tools for accessing, analyzing and mapping expression and sequence data (both finished and unfinished), Nucleic Acids Res. 30 (2002) 87–90.
[23] D.J. Lane, B. Pace, G.J. Olsen, D.A. Stahl, M.L. Sogin, N.R. Pace, Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyze s, Proc Natl Acad Sci U S A 82 (1985) 6955–6959, https://doi.org/10.1073/pnas.82.20.6955.
[24] S.G. Tringe, P. Hugenholtz, A renaissance for the pioneering 16S rRNA gene, Curr. Opin. Microbiol. 11 (2008) 442–446, https://doi.org/10.1016/j.mib.2008.09.011.
[25] R.C. Edgar, MUSCLE: a multiple sequence alignment method with reduced time and space complexity, BMC Bioinf. 5 (2004) 113, https://doi.org/10.1186/1471-2105-5-113.
[26] K. Tamura, M. Nei, Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees, Mol. Biol. Evol. 10 (1993) 512–526, https://doi.org/10.1093/oxfordjournals.molbev.a040023.
[27] J. Felsenstein, Confidence limits on phylogenies: an approach using the bootstrap, Evolution 39 (1985) 783–791, https://doi.org/10.1111/j.1558-5646.1985.tb00420.x.

[28] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, Mega X: molecular evolutionary genetics analysis across computing platforms, Mol. Biol. Evol. 35 (2018) 1547–1549, https://doi.org/10.1093/molbev/msy096.

[29] H. Attrill, P. Gaudet, R.P. Huntley, R.C. Lovering, S.R. Engel, S. Poux, K.M. Van Auken, G. Georghiou, M.C. Chibucos, T.Z. Berardini, V. Wood, H. Drabkin, P. Fey, P. Garmiri, M.A. Harris, T. Sawford, L. Reiser, R. Tauber, S. Toro, Gene Ontology Consortium, Annotation of gene product function from high-throughput studies using the Gene Ontology, Database 2019 (2019), https://doi.org/10.1093/database/baz007 baz007.

[30] S. Carbon, A. Ireland, C.J. Mungall, S. Shu, B. Marshall, S. Lewis, AmiGO. Hub, Web Presence Working Group, AmiGO: online access to ontology and annotation data, Bioinformatics 25 (2009) 288–289, https://doi.org/10.1093/bioinformatics/btn615.

[31] J.M. Goldberg, G. Manning, A. Liu, P. Fey, K.E. Pilcher, Y. Xu, J.L. Smith, The Dictyostelium kinome—analysis of the protein kinases from a simple model organism, PLoS Genet. 2 (2006) e38, https://doi.org/10.1371/journal.pgen.0020038.

[32] S. Lu, J. Wang, F. Chitsaz, M.K. Derbyshire, R.C. Geer, N.R. Gonzales, M. Gwadz, D.I. Hurwitz, G.H. Marchler, J.S. Song, N. Thanki, R.A. Yamashita, M. Yang, D. Zhang, C. Zheng, C.J. Lanczycki, A. Marchler-Bauer, CDD/SPARCLE: the conserved domain database in 2020, Nucleic Acids Res. 48 (2020) D265–D268, https://doi.org/10.1093/nar/gkz991.

[33] P. Jones, D. Binns, H.-Y. Chang, M. Fraser, W. Li, C. McAnulla, H. McWilliam, J. Maslen, A. Mitchell, G. Nuka, S. Pesseat, A.F. Quinn, A. Sangrador-Vegas, M. Scheremetjew, S.-Y. Yong, R. Lopez, S. Hunter, InterProScan 5: genome-scale protein function classification, Bioinformatics 30 (2014) 1236–1240, https://doi.org/10.1093/bioinformatics/btu031.

[34] S. Fischer, B.P. Brunk, F. Chen, X. Gao, O.S. Harb, J.B. Iodice, D. Shanmugam, D.S. Roos, C.J. Stoeckert, Using OrthoMCL to assign proteins to OrthoMCL-DB groups or to cluster proteomes into new ortholog groups, Curr Protoc Bioinformatics Chapter 6 (6.12.1–6.12.19) (2011), https://doi.org/10.1002/0471250953.bi0612s35.

[35] L. Li, C.J. Stoeckert, D.S. Roos, OrthoMCL: identification of ortholog groups for eukaryotic genomes, Genome Res. 13 (2003) 2178–2189, https://doi.org/10.1101/gr.1224503.

[36] C. Doerig, J.C. Rayner, A. Scherf, A.B. Tobin, Post-translational protein modifications in malaria parasites, Nat. Rev. Microbiol. 13 (2015) 160–172, https://doi.org/10.1038/nrmicro3402.

[37] R. Pandey, P. Kumar, D. Gupta, KiPho: malaria parasite kinome and phosphatome portal, Database 2017 (2017), https://doi.org/10.1093/database/bax063 bax063.

[38] R. Pandey, A. Mohmmed, C. Pierrot, J. Khalife, P. Malhotra, D. Gupta, Genome wide in silico analysis of Plasmodium falciparum phosphatome, BMC Genom. 15 (2014) 1024, https://doi.org/10.1186/1471-2164-15-1024.

[39] R. Sardar, A. Kaushik, R. Pandey, A. Mohmmed, S. Ali, D. Gupta, ApicoTFdb: the comprehensive web repository of apicomplexan transcription factors and transcription-associated co-factors, Database 2019 (2019), https://doi.org/10.1093/database/baz094 baz094.

[40] L. Zimmermann, A. Stephens, S.-Z. Nam, D. Rau, J. Kübler, M. Lozajic, F. Gabler, J. Söding, A.N. Lupas, V. Alva, A completely reimplemented MPI Bioinformatics toolkit with a new HHpred server at its core, J. Mol. Biol. 430 (2018) 2237–2243, https://doi.org/10.1016/j.jmb.2017.12.007.

[41] M.T. Weirauch, A. Yang, M. Albu, A.G. Cote, A. Montenegro-Montero, P. Drewe, H.S. Najafabadi, S.A. Lambert, I. Mann, K. Cook, H. Zheng, A. Goity, H. van Bakel, J.-C. Lozano, M. Galli, M.G. Lewsey, E. Huang, T. Mukherjee, X. Chen, J.S. Reece-Hoyes, S. Govindarajan, G. Shaulsky, A.J.M. Walhout, F.-Y. Bouget, G. Ratsch, L.F. Larrondo, J.R. Ecker, T.R. Hughes, Determination and inference of eukaryotic transcription factor sequence specificity, Cell 158 (2014) 1431–1443, https://doi.org/10.1016/j.cell.2014.08.009.

[42] E. Bischoff, C. Vaquero, In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of Plasmodium falciparum, BMC Genom. 11 (2010) 34, https://doi.org/10.1186/1471-2164-11-34.

[43] R. Sardar, A. Kaushik, R. Pandey, A. Mohmmed, S. Ali, D. Gupta, ApicoTFdb: the comprehensive web repository of apicomplexan transcription factors and transcription-associated co-factors, Database 2019 (2019), https://doi.org/10.1093/database/baz094 baz094.

[44] R. Sucgang, A. Kuo, X. Tian, W. Salerno, A. Parikh, C.L. Feasley, E. Dalin, H. Tu, E. Huang, K. Barry, E. Lindquist, H. Shapiro, D. Bruce, J. Schmutz, A. Salamov, P. Fey, P. Gaudet, C. Anjard, M.M. Babu, S. Basu, Y. Bushmanova, H. van der Wel, M. Katoh-Kurasawa, C. Dinh, P.M. Coutinho, T. Saito, M. Elias, P. Schaap, R. R. Kay, B. Henrissat, L. Eichinger, F. Rivero, N.H. Putnam, C.M. West, W.F. Loomis, R.L. Chisholm, G. Shaulsky, J.E. Strassmann, D.C. Queller, A. Kuspa, I. V. Grigoriev, Comparative genomics of the social amoebae Dictyostelium discoideum and Dictyostelium purpureum, Genome Biol. 12 (2011), https://doi.org/10.1186/gb-2011-12-2-r20.

[45] P.M. Cheuka, P. Njaria, G. Mayoka, E. Funjika, Emerging drug targets for antimalarial drug discovery: validation and insights into molecular mechanisms of function, J. Med. Chem. 67 (2024) 838–863, https://doi.org/10.1021/acs.jmedchem.3c01828.

[46] A. Kone, M. van de Vegte-Bolmer, R. Siebelink-Stoter, G.-J. van Gemert, A. Dara, H. Niangaly, A. Luty, O.K. Doumbo, R. Sauerwein, A.A. Djimde, Sulfadoxine-pyrimethamine impairs Plasmodium falciparum gametocyte infectivity and Anopheles mosquito survival, Int. J. Parasitol. 40 (2010) 1221–1228, https://doi.org/10.1016/j.ijpara.2010.05.004.

[47] S. Basu, P. Fey, Y. Pandit, R. Dodson, W.A. Kibbe, R.L. Chisholm, dictyBase 2013: integrating multiple Dictyostelid species, Nucleic Acids Res. 41 (2013) D676–D683, https://doi.org/10.1093/nar/gks1064.

[48] M.J. Gardner, N. Hall, E. Fung, O. White, M. Berriman, R.W. Hyman, J.M. Carlton, A. Pain, K.E. Nelson, S. Bowman, I.T. Paulsen, K. James, J.A. Eisen, K. Rutherford, S.L. Salzberg, A. Craig, S. Kyes, M.-S. Chan, V. Nene, S.J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A.B. Vaidya, D.M.A. Martin, A.H. Fairlamb, M.J. Fraunholz, D.S. Roos, S.A. Ralph, G.I. McFadden, L.M. Cummings, G.M. Subramanian, C. Mungall, J.C. Venter, D.J. Carucci, S.L. Hoffman, C. Newbold, R.W. Davis, C.M. Fraser, B. Barrell, Genome sequence of the human malaria parasite Plasmodium falciparum, Nature 419 (2002) 498–511, https://doi.org/10.1038/nature01097.

[49] G. Glöckner, K. Szafranski, T. Winckler, T. Dingermann, M.A. Quail, E. Cox, L. Eichinger, A.A. Noegel, A. Rosenthal, The complex repeats of Dictyostelium discoideum, Genome Res. 11 (2001) 585–594, https://doi.org/10.1101/gr.162201.

[50] R.D. Rosengarten, B. Santhanam, J. Kokosar, G. Shaulsky, The long noncoding RNA transcriptome of Dictyostelium discoideum, Development, G3 (Bethesda) 7 (2017) 387–398, https://doi.org/10.1534/g3.116.037150.

[51] K. Simantov, M. Goyal, R. Dzikowski, Emerging biology of noncoding RNAs in malaria parasites, PLoS Pathog. 18 (2022) e1010600, https://doi.org/10.1371/journal.ppat.1010600.

[52] A. Aspegren, A. Hinas, P. Larsson, A. Larsson, F. Söderbom, Novel non-coding RNAs in Dictyostelium discoideum and their expression during development, Nucleic Acids Res. 32 (2004) 4646–4656, https://doi.org/10.1093/nar/gkh804.

[53] U. Böhme, T.D. Otto, M. Sanders, C.I. Newbold, M. Berriman, Progression of the canonical reference malaria parasite genome from 2002-2019, Wellcome Open Res 4 (2019) 58, https://doi.org/10.12688/wellcomeopenres.15194.2.

[54] L. Chappell, P. Ross, L. Orchard, T.J. Russell, T.D. Otto, M. Berriman, J.C. Rayner, M. Llinás, Refining the transcriptome of the human malaria parasite Plasmodium falciparum using amplification-free RNA-seq, BMC Genom. 21 (2020) 395, https://doi.org/10.1186/s12864-020-06787-5.

[55] P. Gaudet, P. Fey, S. Basu, Y.A. Bushmanova, R. Dodson, K.A. Sheppard, E.M. Just, W.A. Kibbe, R.L. Chisholm, dictyBase update 2011: web 2.0 functionality and the initial steps towards a genome portal for the Amoebozoa, Nucleic Acids Res. 39 (2011) D620–D624, https://doi.org/10.1093/nar/gkq1103.

[56] A. Hinas, P. Larsson, L. Avesson, L.A. Kirsebom, A. Virtanen, F. Söderbom, Identification of the major spliceosomal RNAs in Dictyostelium discoideum reveals developmentally regulated U2 variants and polyadenylated snRNAs, Eukaryot. Cell 5 (2006) 924–934, https://doi.org/10.1128/EC.00065-06.

[57] M. Pimenta Leibowitz, R. Ariav, D. Zilberg, Environmental and physiological conditions affecting Tetrahymena sp. infection in guppies, Poecilia reticulata Peters, J. Fish. Dis. 28 (2005) 539–547, https://doi.org/10.1111/j.1365-2761.2005.00658.x.

[58] S. Koncarevic, P. Rohrbach, M. Deponte, G. Krohne, J.H. Prieto, J. Yates, S. Rahlfs, K. Becker, The malarial parasite Plasmodium falciparum imports the human protein peroxiredoxin 2 for peroxide detoxification, Proc Natl Acad Sci U S A 106 (2009) 13323–13328, https://doi.org/10.1073/pnas.0905387106.

[59] M.P. Wagner, P. Formaglio, O. Gorgette, J.M. Dziekan, C. Huon, I. Berneburg, S. Rahlfs, J.-C. Barale, S.I. Feinstein, A.B. Fisher, D. Ménard, Z. Bozdech, R. Amino, L. Touqui, C.E. Chitnis, Human peroxiredoxin 6 is essential for malaria parasites and provides a host-based drug target, Cell Rep. 39 (2022) 110923, https://doi.org/10.1016/j.celrep.2022.110923.

[60] B.S. Crabb, A.F. Cowman, Plasmodium falciparum virulence determinants unveiled, Genome Biol. 3 (2002), https://doi.org/10.1186/gb-2002-3-11-reviews1031 reviews1031.1-reviews1031.4.
[61] J.D. Smith, The role of PfEMP1 adhesion domain classification in Plasmodium falciparum pathogenesis research, Mol. Biochem. Parasitol. 195 (2014) 82–87, https://doi.org/10.1016/j.molbiopara.2014.07.006.
[62] J.G. Beeson, D.R. Drew, M.J. Boyle, G. Feng, F.J.I. Fowkes, J.S. Richards, Merozoite surface proteins in red blood cell invasion, immunity and vaccines against malaria, FEMS Microbiol. Rev. 40 (2016) 343–372, https://doi.org/10.1093/femsre/fuw001.
[63] R.R. Soares, C.F. Cunha, R. Ferraz-Nogueira, A. Marins-Dos-Santos, R.N. Rodrigues-da-Silva, I. da Silva Soares, J. da Costa Lima-Junior, A.L. Bertho, M. U. Ferreira, K.K.G. Scopel, Apical membrane protein 1-specific antibody profile and temporal changes in peripheral blood B-cell populations in Plasmodium vivax malaria, Parasite Immunol. 41 (2019) e12662, https://doi.org/10.1111/pim.12662.
[64] T. Triglia, J. Healer, S.R. Caruana, A.N. Hodder, R.F. Anders, B.S. Crabb, A.F. Cowman, Apical membrane antigen 1 plays a central role in erythrocyte invasion by Plasmodium species, Mol. Microbiol. 38 (2000) 706–718, https://doi.org/10.1046/j.1365-2958.2000.02175.x.
[65] M.D. Jeninga, J.E. Quinn, M. Petter, ApiAP2 transcription factors in apicomplexan parasites, Pathogens 8 (2019) 47, https://doi.org/10.3390/pathogens8020047.
[66] I. Kaneko, S. Iwanaga, T. Kato, I. Kobayashi, M. Yuda, Genome-wide identification of the target genes of AP2-O, a Plasmodium AP2-family transcription factor, PLoS Pathog. 11 (2015) e1004905, https://doi.org/10.1371/journal.ppat.1004905.
[67] K. Modrzynska, C. Pfander, L. Chappell, L. Yu, C. Suarez, K. Dundas, A.R. Gomes, D. Goulding, J.C. Rayner, J. Choudhary, O. Billker, A knockout screen of ApiAP2 genes reveals networks of interacting transcriptional regulators controlling the Plasmodium life cycle, Cell Host Microbe 21 (2017) 11–22, https://doi.org/10.1016/j.chom.2016.12.003.
[68] J.E. Phillips, E. Huang, G. Shaulsky, R.H. Gomer, The putative bZIP transcripton factor BzpN slows proliferation and functions in the regulation of cell density by autocrine signals in Dictyostelium, PLoS One 6 (2011) e21765, https://doi.org/10.1371/journal.pone.0021765.
[69] N.V. Zhukovskaya, M. Fukuzawa, Y. Yamada, T. Araki, J.G. Williams, The Dictyostelium bZIP transcription factor DimB regulates prestalk-specific gene expression, Development 133 (2006) 439–448, https://doi.org/10.1242/dev.02190.
[70] J. Schultz, R.R. Copley, T. Doerks, C.P. Ponting, P. Bork, SMART: a web-based tool for the study of genetically mobile domains, Nucleic Acids Res. 28 (2000) 231–234, https://doi.org/10.1093/nar/28.1.231.
[71] M. Tsujioka, N. Zhukovskaya, Y. Yamada, M. Fukuzawa, S. Ross, J.G. Williams, Dictyostelium myb transcription factors function at culmination as activators of ancillary stalk differentiation, Eukaryot. Cell 6 (2007) 568–570, https://doi.org/10.1128/EC.00373-06.
[72] M. Gissot, S. Briquet, P. Refour, C. Boschet, C. Vaquero, PfMyb1, a Plasmodium falciparum transcription factor, is required for intra-erythrocytic growth and controls key genes for cell cycle regulation, J. Mol. Biol. 346 (2005) 29–42, https://doi.org/10.1016/j.jmb.2004.11.045.
[73] L. Anton, D.W. Cobb, C.-M. Ho, Structural parasitology of the malaria parasite Plasmodium falciparum, Trends Biochem. Sci. 47 (2022) 149–159, https://doi.org/10.1016/j.tibs.2021.10.006.
[74] D.S. Peterson, Y. Gao, K. Asokan, J. Gaertig, The circumsporozoite protein of Plasmodium falciparum is expressed and localized to the cell surface in the free-living ciliate Tetrahymena thermophila, Mol. Biochem. Parasitol. 122 (2002) 119–126, https://doi.org/10.1016/s0166-6851(02)00079-8.
[75] J. Gaertig, Y. Gao, T. Tishgarten, T.G. Clark, H.W. Dickerson, Surface display of a parasite antigen in the ciliate Tetrahymena thermophila, Nat. Biotechnol. 17 (1999) 462–465, https://doi.org/10.1038/8638.
[76] P. Fey, A.S. Kowal, P. Gaudet, K.E. Pilcher, R.L. Chisholm, Protocols for growth and development of Dictyostelium discoideum, Nat. Protoc. 2 (2007) 1307–1316, https://doi.org/10.1038/nprot.2007.178.
[77] A.T. Ho, L.D. Hurst, Stop codon usage as a window into genome evolution: mutation, selection, biased gene conversion and the TAG paradox, Genome Biol Evol 14 (2022) evac115, https://doi.org/10.1093/gbe/evac115.
[78] Codon preference in Dictyostelium discoideum. - PMC, (n.d.). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC338318/(accessed August 9, 2024).
[79] B. Elsworth, C.D. Keroack, M.T. Duraisingh, Elucidating host cell uptake by malaria parasites, Trends Parasitol. 35 (2019) 333–335, https://doi.org/10.1016/j.pt.2019.03.005.
[80] W.F. Loomis, Cell signaling during development of Dictyostelium, Dev. Biol. 391 (2014) 1–16, https://doi.org/10.1016/j.ydbio.2014.04.001.
[81] J. Weiner, T.W.A. Kooij, Phylogenetic profiles of all membrane transport proteins of the malaria parasite highlight new drug targets, Microb Cell 3 (2016) 511–521, https://doi.org/10.15698/mic2016.10.534.
[82] N. Dissmeyer, A. Schnittger, The age of protein kinases, Methods Mol. Biol. 779 (2011) 7–52, https://doi.org/10.1007/978-1-61779-264-9_2.
[83] S.S. Taylor, A.P. Kornev, Protein kinases: evolution of dynamic regulatory proteins, Trends Biochem. Sci. 36 (2011) 65–77, https://doi.org/10.1016/j.tibs.2010.09.006.
[84] Q. Zhong, X. Xiao, Y. Qiu, Z. Xu, C. Chen, B. Chong, X. Zhao, S. Hai, S. Li, Z. An, L. Dai, Protein posttranslational modifications in health and diseases: functions, regulatory mechanisms, and therapeutic implications, MedComm 4 (2020) (2023) e261, https://doi.org/10.1002/mco2.261.
[85] K. Kastano, P. Mier, M.A. Andrade-Navarro, The role of low complexity regions in protein interaction modes: an illustration in huntingtin, Int. J. Mol. Sci. 22 (2021) 1727, https://doi.org/10.3390/ijms22041727.
[86] S. Santarriaga, A. Petersen, K. Ndukwe, A. Brandt, N. Gerges, J. Bruns Scaglione, K.M. Scaglione, The social amoeba Dictyostelium discoideum is highly resistant to polyglutamine aggregation, J. Biol. Chem. 290 (2015) 25571–25578, https://doi.org/10.1074/jbc.M115.676247.
[87] J. Okombo, A.I. Abdi, S.M. Kiara, L. Mwai, L. Pole, C.J. Sutherland, A. Nzila, L.I. Ochola-Oyier, Repeat polymorphisms in the low-complexity regions of Plasmodium falciparum ABC transporters and associations with in vitro antimalarial responses, Antimicrob. Agents Chemother. 57 (2013) 6196–6204, https://doi.org/10.1128/AAC.01465-13.