Predicting and exploring network components involved in pathogenesis in the malaria parasite via novel subnetwork alignments

Hong Cai†, Timothy G Lilburn‡, Changjin Hong†, Jianying Gu⁴, Rui Kuang³*, Yufeng Wang¹,⁵*

From The International Conference on Intelligent Biology and Medicine (ICIBM) 2014
San Antonio, TX, USA. 04-06 December 2014

Abstract

Background: Malaria is a major health threat, affecting over 40% of the world’s population. The latest report released by the World Health Organization estimated about 207 million cases of malaria infection, and about 627,000 deaths in 2012 alone. During the past decade, new therapeutic targets have been identified and are at various stages of characterization, thanks to the emerging omics-based technologies. However, the mechanism of malaria pathogenesis remains largely unknown. In this paper, we employ a novel neighborhood subnetwork alignment approach to identify network components that are potentially involved in pathogenesis.

Results: Our module-based subnetwork alignment approach identified 24 functional homologs of pathogenesis-related proteins in the malaria parasite P. falciparum, using the protein-protein interaction networks in Escherichia coli as references. Eighteen out of these 24 proteins are associated with 418 other proteins that are related to DNA replication, transcriptional regulation, translation, signaling, metabolism, cell cycle regulation, as well as cytoadherence and entry to the host.

Conclusions: The subnetwork alignments and subsequent protein-protein association network mining predicted a group of malarial proteins that may be involved in parasite development and parasite-host interaction, opening a new systems-level view of parasite pathogenesis and virulence.

Background

Malaria is a severe global infectious disease. The World Health Organization estimated 207 million clinical cases, and 627,000 deaths due to malaria in 2012. Malaria infection was reported in 97 countries in 2013, representing a major public health concern. The vast majority of the deaths occurred in children under age 5 in sub-Saharan areas; it is estimated that a child dies from malaria every minute.

The causative agents of malaria are a group of parasites from the Genus Plasmodium. Among five human malaria parasites, Plasmodium falciparum causes the most severe form of malaria. It is notoriously difficult to study malaria biology as the parasite has a complex life cycle involving a mosquito vector and a human host. Within a human host, the parasite has a relatively dormant liver stage and an active red blood cell (RBC) stage. Clinical symptoms such as high fever, chills, headache, profuse sweats, fatigue, nausea, and vomiting are manifested at the blood stage, the stage that most drugs target against.

The effectiveness of antimalarial drugs, however, has been constantly challenged during the past decades due to the fast evolution of parasites that are resistant to multiple lines of drugs. This underscores an urgent need to search for novel drug targets. New lines of antimalarial targets have been identified [1-5] and are at various stages of functional and pharmacochemical characterization, thanks to the development of customized omic-based high
throughput technologies, including genome sequencing and annotation [6-11], microarray [12-23], RNA-Seq [21,22,24], mass spectrometry [4,5,25-32], yeast two-hybrid protein-protein interaction assays [33-36], and large-scale compound screening [37]. Moreover, these amazing genomic, transcriptomic, proteomic, metabolomics [38-42], and interactomic data are enabling a new systems biology paradigm in malaria research; the complex and dynamic cellular processes such as pathogenesis and virulence have recently begun to be unveiled.

By nature, pathogenicity is an emergent property, that is, it is the result of the concerted effects of many genes. The suggestion by Mu et al. that differences in drug resistance in *P. falciparum* parasites from different continents were due to “overlapping, but not identical, sets of genes, including many encoding unknown proteins” [43] and the observation that the costs of maintaining antibiotic resistance in bacteria can be offset or lost thanks to changes at other loci not directly involved in antibiotic resistance [44] are both strongly suggestive of the reality of systems-level adaptation and of the existence of systems producing the virulence phenotype [45]. However, unlike bacterial systems, whose pathogenesis mechanism is much better characterized, to date, malaria pathogenesis, i.e., the development of the disease condition, remains largely unknown. A critical barrier is our limited knowledge of the cellular components, cellular events and reactions and other pathologic mechanisms associated with the disease development.

To unveil previously unknown proteins that are potentially involved in pathogenesis in *P. falciparum*, we developed a module-based subnetwork alignment approach. Traditional homology transfer using a protein-protein network alignment approach [46-54] does not appear feasible to predict functional orthologs in the malaria parasite, because of the remote homology between *P. falciparum* and other known model organisms: BLAST-based sequence similarity search was unable to make functional annotation for over 60% of the genes in *P. falciparum* [11]. Our neighborhood subnetwork alignment algorithm [55] could potentially circumvent this limitation by searching for the similarities between functional modules. This algorithm, which was custom-developed for malaria research, was used to predict novel transcriptional regulators and cell cycle regulatory proteins in *P. falciparum* [55,56]. In this paper, we extended the subnetwork alignment approach to identify the proteins related to pathogenesis and explore their potential functional roles.

**Results and discussion**

**Twenty-four (24) novel proteins were predicted to be associated with pathogenesis in malaria parasite**

Malaria pathogenesis is an emergent property that involves dynamic regulation of parasite metabolism, parasite signaling, and host-parasite interactions [57]. Gene Ontology using Term GO:0009405 (pathogenesis) predicted 95 *P. falciparum* genes (Additional File 1), not surprisingly, including 78 members of the *Plasmodium falciparum* erythrocyte membrane protein (PfEMP1) family. PfEMP1 is encoded by the highly polymorphic Var gene, and is one of the largest and most widely-studied protein families in *P. falciparum*. PfEMP1 is believed to play an important role in parasite pathogenesis and immune evasion [58]; it contributes to antigenic variation and cytoadherence, and the expression of specific PfEMP1 variants is associated with severe malaria in children [59].

Given the observed complexity of malaria pathogenesis, one might envisage that more network components have yet to be identified. Using a neighborhood subnetwork alignment algorithm, we predicted that 24 novel proteins in *P. falciparum* were functional orthologs of proteins known to be involved in pathogenesis in *E. coli* (Additional File 2). *E. coli* was chosen because its pathogenesis and virulence mechanisms are among the best-characterized of all known infectious agents. Seven of these predicted functional orthologs were annotated as “conserved *Plasmodium* proteins with unknown function”, six of which do not have any predicted interactors in malarial protein-protein association networks (Table 1). Seventeen of these 24 predicted functional orthologs may be involved in various biological processes from post-transcriptional regulation, protein translation, protein ubiquitination and modification, oxidation-reduction process, signaling, metabolism, to histone modification. Notably, one functional ortholog (PF3D7_0300100) is a variant of PfEMP1, and it is also consistent with the GO prediction that it may be associated with pathogenesis (Additional File 1).

**The set of functional orthologs is involved in important biological processes that are related to pathogenesis and virulence**

We performed network mining on the 24 predicted pathogenesis-related functional orthologs. Eighteen proteins are associated with 418 other proteins with high confidence score (>0.7). The network centralization parameter, a measure of the distribution of network density [60], is 0.0548, indicating that the network is decentralized. The degree of connectivity of these 18 proteins ranges from 180 to one and the degree distribution offers a good fit to a power law distribution (Table 1). This type of distribution indicates that most nodes are connected to a few others and a few nodes are connected to many others (Figure 1). The network heterogeneity coefficient assesses the relative number of hubs in a network; in this network it is relatively high (1.455), which is in accord with the observed power law distribution.

Table 2 shows representative functional categories of the proteins that are associated with predicted...
pathogenesis-related *P. falciparum* proteins. They are active players in crucial cellular processes in parasite development, growth, and response to host and environmental stresses. In the following sections, we discuss critical network components and their potential associations with pathogenesis and virulence.

**1. Tight associations between translation, protein turnover, proteolysis, signaling, and stress response**

The predicted functional ortholog with the highest connectivity (180) is a putative 60S ribosomal protein L40/UBI (PF3D7_1365900) (Figure 1). This protein is unique as it is a fusion protein of two domains: a C-terminal ribosomal L40e family domain and an N-terminal ubiquitin domain. It therefore plays dual functions: On one hand, it is a central member of the ribosomal protein complex in *P. falciparum*, which includes small and large subunit ribosomal proteins of varying sizes. This complex is associated with several translation initiation and elongation factors. On the other hand, L40/UBI is implicated in the ubiquitin-proteasome system [61]. UPS is an essential mechanism in *P. falciparum* for degrading misfolded or unneeded proteins [62]. UPS is becoming a promising target for antimalarial development due to its critical roles in cell cycle regulation and stress response and the relative low toxicity of its inhibitors [63]. Notably, malarial L40/UBI may be involved in other processes that are currently not fully appreciated, such as (1) signaling. L40/UBI and PfRACK (*P. falciparum* receptor for activated C kinase homolog, PF3D7_0826700) are associated with a high confidence score (0.908), suggested by the co-expression of their homologs in *Caenorhabditis elegans*, *Drosophila melanogaster*, and humans, and protein-protein interaction found by affinity capture-MS assay in humans. The evidence that PfRACK could directly inhibit inositol 1,4,5-trisphosphate receptor-mediated calcium-signaling in mammalian host cells is indicative of a pathogenesis mechanism acting through the disruption of host activities [64]. (2) Genome stability. L40/UBI is associated with two putative proliferating cell nuclear antigens (PCNAs), auxiliary factors of DNA polymerase. Ubiquitin, along with SUMO (small ubiquitin-related modifier), were shown to effectively regulate DNA damage recognition and repair in yeast [65]. A similar mechanism, if it exists in the malaria parasite, would promote the accuracy of replication and timely repair, both of which are crucial for parasite survival within the host.

### Table 1 Novel *Plasmodium falciparum* proteins that were predicted to be functional orthologs of pathogenesis-related proteins in *E. coli*

| PlasmoDB Accession Number | Annotation                                      | Degree of Connectivity |
|---------------------------|-------------------------------------------------|-----------------------|
| PF3D7_1365900             | 60S ribosomal protein L40/UBI, putative          | 180                   |
| PF3D7_1241800             | DEAD/DEAH box ATP-dependent RNA helicase, putative | 98                    |
| PF3D7_0306800             | T-complex protein beta subunit, putative         | 43                    |
| PF3D7_0823300             | histone acetyltransferase GCNS (GCNS)            | 20                    |
| PF3D7_1318700             | conserved Plasmodium protein, unknown function   | 19                    |
| PF3D7_1212800             | iron-sulfur subunit of succinate dehydrogenase   | 18                    |
| PF3D7_0208600             | ribosome-recycling factor, putative (RRF)        | 16                    |
| PF3D7_1416900             | prefoldin subunit 2, putative                   | 15                    |
| PF3D7_0731100             | Plasmodium exported protein (PHigTc), unknown function (GEXP11) | 14                   |
| PF3D7_1013500             | phosphoinositide-specific phospholipase C (PI-PLC) | 11                   |
| PF3D7_0810600             | RNA helicase, putative                          | 9                     |
| PF3D7_1337200             | 1-deoxy-D-xylulose 5-phosphate synthase          | 7                     |
| PF3D7_1148800             | Plasmodium exported protein (hyp11), unknown function | 6            |
| PF3D7_1008600             | inositol polyphosphate kinase, putative (IPK1)  | 6                     |
| PF3D7_0919000             | nucleosome assembly protein (NAPS)               | 6                     |
| PF3D7_1418000             | ubiquitin fusion degradation protein (UFD1, putative (UFD1)) | 5            |
| PF3D7_0300100             | erythrocyte membrane protein 1, PEM1 (VAR)      | 5                     |
| PF3D7_1342400             | casein kinase II beta chain (CK2beta2)          | 1                     |
| PF3D7_1318700             | conserved Plasmodium protein, unknown function   | 0                     |
| PF3D7_1364200             | conserved Plasmodium protein, unknown function   | 0                     |
| PF3D7_1229300             | conserved Plasmodium protein, unknown function   | 0                     |
| PF3D7_1451200             | conserved Plasmodium protein, unknown function   | 0                     |
| PF3D7_0310900             | conserved Plasmodium protein, unknown function   | 0                     |
| PF3D7_1008100             | conserved Plasmodium protein, unknown function   | 0                     |
2. RNA metabolism and transcriptional regulation

The predicted pathogenesis-related functional orthologs with the second largest connectivity (98) is a putative DEAD/DEAH box ATP-dependent RNA helicase (PF3D7_1241800) (Figure 1). The main function of this family of RNA helicases is to unwind RNA [66]. Helicase is considered to be a potential antimalarial target due to its essentiality for parasite life cycle [67]. Fifty-one DEAD/DEAH helicase homologs were identified in P. falciparum by genomic analysis [68]. Our network mining indicated that PF3D7_1241800 is broadly involved in RNA metabolism; it is associated with (1) several RNA polymerase subunits and a number of other helicases required for transcription, (2) nucleolar GTP-binding proteins, putative GTPases, and BRIX-domain containing proteins for ribosomal biogenesis and maturation, (3) at least three putative U3 snoRNP associated proteins for pre-rRNA processing, (4) a ribosomal subunit export protein for RNA transport, and (5) an eukaryotic initiation factor required for translation. In addition to RNA metabolism,
PF3D7_1241800 is associated with a putative transcription factor (PF3D7_0206600). This transcription factor is a member of the general transcription factor TFIIS family with a characteristic zinc ribbon conformation. The accurate information flow of basal transcriptional regulation and RNA metabolism is essential for parasite development.

3. Chromatin remodeling, epigenetic regulation, and antigenic variation

It is becoming clear that chromatin remodeling/histone modification controls genome-wide gene expression in *P. falciparum* and epigenetic regulation may be a neglected major contributing factor to antigenic variation and evasion of the human immune system during parasite pathogenesis [69-71]. Our network analysis identified a subnetwork module centered on PfGCN5 (PF3D7_0823300), a bona fide histone acetyltransferase. PfGCN5 is associated with 20 other proteins (Figure 2), including histones H2A, H3, and H4, DNA helicase, three histone deacetylases, a putative chromodomain-helicase-DNA-binding protein, and a putative chromatin assembly protein ASF1, suggesting its role in chromatin remodeling. In addition, PfGCN5 may be involved in transcriptional regulation: it is associated with the transcriptional activator ADA2 with a confidence score of 0.955 [72]; in yeast, GCN5, ADA2, and ADA3 interact and form a catalytic core for histone acetyltransferase (HAT) activity. PfGCN5 may also have protein-protein interactions with a general transcription initiation factor TFIID and a putative transcriptional regulator. Interestingly, PfGCN5 is predicted to be associated with a transcriptional regulatory protein Sir2a (PF3D7_1328800) (Figure 2). Sir2a and its paralog Sir2b are implicated in epigenetic regulation of PfEMP1 antigenic variation, thus controlling the pathogenicity of malaria [73,74].

PfGCN5 is also indirectly associated with three predicted pathogenesis-related functional orthologs (labelled in yellow in Figure 2): (1) an NAPS protein (PF3D7_0919000) involved in nucleosome assembly, (2) a putative ribosome-recycling factor (RRF, PF3D7_0208600) that is essential for recycling of organellar ribosomes during protein translation, and (3) a putative inositol polyphosphate kinase (IPK1, PF3D7_1008000) that may be involved in signaling, chromatin remodeling, and mRNA export.

Unlike the PfGCN5-centered chromatin-remodeling module, very little is known about another subnetwork centered on two *Plasmodium* exported proteins hyp11.
(PF3D7_1148800) and GEXP11 (PF3D7_0731100) (Figure 3). The function of these two proteins is unclear, and the majority of their interactors are hypothetical proteins as well; the evidence of association is mainly established based on co-expression profiles and does not necessarily suggest direct functional association.

**Conclusions**

Using a neighborhood subnetwork alignment algorithm, we identified network components associated with 24 potential pathogenesis-related proteins that may be involved in malaria pathogenesis. The proteins play roles in parasite information processing, signaling, epigenetic regulation, and entry to the host, all processes that are related to pathogenesis and virulence. A better understanding of the network components containing these proteins creates a new list of potential rational targets for antimalarial intervention.

**Methods**

**Querying subnetworks using neighborhood alignments**

We have framed the prediction of functional orthologs in *P. falciparum* as a subnetwork querying problem. Network alignment algorithms search the network of the organism of interest for subnetworks that are similar to subnetworks in other, better-understood organisms [75,76]. The network we used to search the “target” network is a well-studied functional module from a model
organism. Such queries can find similar modules in the less studied target organism, permitting us to link information about biological functions across organisms [77]. We have used this method in previous studies to predict hitherto unrecognized transcriptional regulators and cell cycle regulators playing important roles in the parasite life cycle [55,56]. We adopted the same algorithm to identify proteins involved in pathogenesis.

We first mapped a set of proteins related to pathogenesis (GO: 0009405) from E. coli onto its own PPI network. A set of “neighbors” was selected for each pathogenesis protein in E. coli, creating a set of subnetworks, which, by inference, form a network of subnetworks in the query network. Similarly, each P. falciparum protein was mapped into its own PPI network, and a subnetwork of neighbors was constructed. Subnetwork size was controlled by including only proteins that are k hops from the central protein in each case and k was chosen such that the neighbor size was under 500, unless the central protein had more than 500 directly connected neighbors.

After deducing the neighborhood subnetworks for both the query E. coli pathogenesis-related proteins and all P. falciparum proteins, the E. coli subnetworks were combinatorially aligned against the P. falciparum subnetworks. When good alignments were found (see below), the central protein of the best-aligned P. falciparum subnetwork was labeled a functional

---

**Figure 3** A subnetwork showing the proteins associated with two putative Plasmodium exported proteins (PF3D7_0731100 and PF3D7_1148800). Node size is proportional to the degree of the connectivity of the node. Nodes are colored according to their functional classification in the eggNOG database [80]. The visualization is as for Figure 1.
ortholog of the cognate protein at the center of the relevant pathogenesis subnetwork in *E. coli*.

In order to analyze and score the degree of similarity between the *P. falciparum* and *E. coli* neighborhood subnetworks a shortest-path graph kernel was used to measure the similarity between two labeled networks, and a numerical score for each alignment was assigned [78]. To optimize the graph kernel, only paths between the central protein and other subnetwork proteins are counted. Each shortest path through the central protein can be considered a chain of cellular activities, and the path defines the dynamic function of this protein. Given a subnetwork *S*<sub>p</sub> with central protein *p* and a query subnetwork *S*<sub>q</sub> with central protein *q*, the shortest path similarity function is defined as follows,

$$K(S_q, S_p) = \frac{1}{\|S_q\| + \|S_p\|} \prod_{(i1, i2) \in S_q} B((i1, i2), S_p)$$

where

$$B((i1, i2), S_p) = \max_{(j1, j2) \in S_p} \frac{2E(i1, j1)E(i2, j2)}{dist(i1, i2) + dist(j1, j2)}.$$  

$$E(x, y) = \exp\left(-\frac{Eval(x, y)}{\sigma}\right)$$

with the normalization parameter $\sigma = 10$ measures the sequence similarity between proteins $x$ and $y$ based on the E-value of the sequence alignment, and $dist(x, y)$ is the length of the shortest path connecting proteins $x$ and $y$ in the PPI subnetwork. The computation was done on a -log 10 scale. The method outlined here takes each pair of proteins $(i1, i2)$ from one subnetwork and seeks the maximum ratio of sequence similarity with respect to the closeness (shortest path through the central protein) of the networks, in order to identify proteins $(j1, j2)$ in the target subnetwork. Using this algorithm, a subnetwork alignment score is obtained by collecting the shortest paths between two neighborhood subnetworks, getting an alignment score for each pair of proteins, and totaling all of the alignment values. Hence, by quantifying the sequence similarity and network similarity and evaluating the role of the central protein in the query network we can summarize the functional coherence, and distance between two central proteins, as a numerical score.

**Network data and network analysis**

We downloaded protein-protein interaction data for *E. coli* from the IntAct database [79], and protein-protein association data for *P. falciparum* from the STRING database [45]. STRING uses numerous data types, including sequence similarity estimates, pathway analysis, chromosome synteny, genome organization, and phylogenetic reconstruction, as well as literature text mining to estimate association confidence scores (S), ranging from 0.15 to 0.999. The data are integrated using a Bayesian approach and the scores approximate the likelihood of finding the pairs of proteins in the same pathway. Cytoscape 3.1 was used for network visualization [59]. EggNOG database was used for functional classification of the network nodes [48]. NetworkAnalyzer was used to compute topological parameters and centrality measures of the cellular networks [49].

**List of abbreviations used:**

GO: Gene Ontology  
HAT: histone acetyltransferase  
IPK: inositol polyphosphate kinase  
PCNA: proliferating cell nuclear antigen  
PfRACK: *Plasmodium falciparum* receptor for activated C kinase  
PP1: protein-protein interaction  
RBC: red blood cell  
RRF: ribosome-recycling factor  
SUMO: small ubiquitin-related modifier  
UPS: ubiquitin-proteasome system

**Additional material**

Additional file 1: *P. falciparum* proteins that are annotated to be associate with Gene Ontology (GO) term GO0009405 (pathogenesis).

Additional file 2: Predicted functional orthologs involved in pathogenesis in *P. falciparum*. The query genome is *P. falciparum*, and the target genome is *E. coli*. GO: Gene Ontology. BP: Biological Process. MF: Molecular Function. CC: Cellular Component.

**Competing interests**

The authors declare that they have no competing interests.

**Authors’ contributions**

YW and RK conceived and designed the study. All authors performed bioinformatics data analysis and drafted the manuscript. All authors read and approved the final manuscript.

**Acknowledgements**

We thank PlasmoDB for providing access to malaria omic data. This work is supported by NIH grants GM108086, GM081068 and A080579 to YW. KR and CH are supported by University of Minnesota Grant-in-Aid of Research, Artistry and Scholarship. We thank the Computational Biology Initiative at UTSA for providing computational support. This work received computational support from Computational Systems Biology Core, funded by the National Institute on Minority Health and Health Disparities (G12MD007591) from the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences, National Institute of Allergy and Infectious Diseases, National Institute on Minority Health and Health Disparities, or the National Institutes of Health.

**Declarations**

The publication costs for this article were funded by the corresponding author YW.

This article has been published as part of BMC Systems Biology Volume 9 Supplement 4, 2015: Selected articles from The International Conference on
Intelligent Biology and Medicine (ICIBM) 2014: Systems Biology. The full contents of the supplement are available online at http://www.biomedcentral.com/bmcystbio/supplements/9/54.

Authors’ details
1 Department of Biology, University of Texas at San Antonio, San Antonio, TX 78249, USA. 2 Novozymes NA, Durham, NC 27709, USA. 3 Department of Computer Science and Engineering, University of Minnesota Twin Cities, Minneapolis, MN 55455, USA. 4 Department of Biology, College of Staten Island, City University of New York, Staten Island, NY 10314, USA. 5 South Texas Center for Emerging Infectious Diseases, University of Texas at San Antonio, San Antonio, TX 78249, USA.

Published: 11 June 2015

References

1. Wu Y, Wang X, Liu X, Wang Y: Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite. Genome research 2003, 13(4):401-416.
2. Wiesner J, Jomaa H: Comparative genomics of the neglecte human malaria parasite Plasmodium yoelii yoelii. Nature 2002, 419(6906):512-519.
3. Hall N, Karras M, Riddle CA, Carlton JM, Kojov-T, Bermann M, Flores L, Janssen CS, Pain A, Christophides GK, et al: A comprehensive survey of the Plasmodium life cycle by genomic, transcriptomic, and proteomic analyses. Science 2005, 307(5706):82-86.
4. Pain A, Bohme U, Berry AE, Mungall K, Finn RD, Jackson AP, Mourier T, Astru M, Blackman MJ, et al: Plant-like kinase in Plasmodium falciparum regulates parasite egress from erythrocytes. Science 2010, 329(5980):910-912.
5. Solyakov L, Halbert J, Alam MM, Semblat JP, Dorin-Semblat D, Reininger L, Palacpac NM, Honma H, Yagi M, Cai et al: Isoprenoid biosynthesis of the apicoplast as drug target. Curr Drug Targets 2007, 8(1):1-3.
6. Dvorin JD, Martyn DC, Patel SD, Grimley JS, Collins CR, Hopp CS, Bright AT, Westenberger S, Wimeler E, Blackman MJ, et al: A Plasmodium falciparum Fe(II)-chisent-EST collection providing clues to chisent specific gene structure and polymorphism. BMC genomics 2009, 10:235.
7. Carlton JM, Angiuoli SV, Suh BB, Kooij TW, Pertea M, Silva JC, Silva-J, Acharya P, Kumar R, Tatu U: Dharma NV, Bright AT, Westenberger SJ, Barnes SW, Batalov S, Carucci DJ, et al: Discovery of gene function by expression profiling of the malaria parasite life cycle. Science 2003, 301(5639):1503-1508.
8. Dharia NV, Bozdech Z, Wong ED, Adai A, DiRienzo JL: Comparative whole genome transcriptome analysis of three Plasmodium falciparum strains. Nucleic acids research 2006, 34(1):1166-1173.
9. Lopez-Barragan MU, Lemieux J, Quinones M, Williamson KC, Molina-Cruz A, Cui K, Barillas-Mury C, Zhao K, Su XZ: Directional gene expression and antisenic transcripts in sexual and asexual stages of Plasmodium falciparum. BMC genomics 2011, 12:587.
10. Watanabe J, Watanabe J, Awataguchi H, Nakamura H, Kato I, Sugano S: Comparative genomics of the Plasmodium falciparum life cycle. Nature 2003, 427(6974):498-511.
11. Floret I, Porcel BM, Guillaume E, Da Silva C, Artigueneau F, Marechal E, Brehelin L, Gasquel O, Chaveaux S, Winkler P, et al: Plasmodium falciparum FcRn-schizont-EST collection providing clues to schizont specific gene structure and polymorphism. BMC genomics 2009, 10:235.
12. Le Roch KG, Zhou Y, Blair PL, Granger M, Moch JK, Haynes JD, De La Vega P, Holder AA, Batalov S, Carucci DJ, et al: Use of high-density tiling microarrays to identify mutations globally and elucidate mechanisms of drug resistance in Plasmodium falciparum. Genome biology 2009, 10(2):R21.
13. Dharia NV, Bozdech Z, Wong ED, Adai A, DiRienzo JL: Computational modeling of the Plasmodium falciparum interactome reveals protein function on a genome-wide scale. Genome research 2006, 16(4):542-549.
14. LaCount DJ, Shoefeld WF, Fields S: Selection of yeast strains with enhanced expression of Plasmodium falciparum proteins. Molecular and biochroemical parasitology 2009, 163(2):119-122.
35. LaCount DJ, Vignali M, Chetter R, Phansalkar A, Bell R, Hesselberth JR, Schoenfeld LW, Ota I, Sahasrabudhe S, Kurschner C, et al. A protein interaction network of the malaria parasite Plasmodium falciparum. Nature 2005, 438(7064):103-107.

36. Suthram S, Stetter T, Ideker T. The Plasmodium falciparum protein network diverges from those of other eukaryotes. Nature 2005, 438(7064):108-112.

37. Yuan J, Cheng KC, Jackson RL, Huang R, Pattardilokrat S, Liu A, Guha R, Fidock DA, Inglese J, Wellens TE, et al. Chemical genomics profiling for anti-malarial therapies, response signatures, and molecular targets. Science 2011, 333(6043):724-729.

38. Ginsburg H. Progress in in silico functional genomics: the malaria Metabolic Pathways database. Trends in parasitology 2006, 22(6):238-240.

39. Ginsburg H, Tilley L. Plasmodium falciparum metabolic pathways (MPMP) project upgraded with a database of subcellular locations of gene products. Trends in parasitology 2011.

40. Lakshmanan V, Rhee KY, Daily JP. Metabolic and biochemical parasitology 2011, 17(2):104-111.

41. Plata G, Hsiao TL, Olzewski KW, Linhas M, Vrkup D. Reconstruction and flux-balance analysis of the Plasmodium falciparum metabolic network. Mol Syst Biol 2010, 6:408.

42. Yeh J, Hanekamp T, Tsoka S, Karp PD, Altman RB. Computational analysis of Plasmodium falciparum metabolism: organizing genomic information to facilitate drug discovery. Genome research 2004, 14(5):917-924.

43. Mu JB, Ferdig MT, Feng XR, Joy DA, Duan JH, Furuya T, Subramanian G, Aravind L, Cooper RA, Wootton JC, et al. Multiple transporters associated with malaria parasite responses to chloroquine and quinine. Molecular microbiology 2003, 49(4):977-989.

44. Schnag SJ, Perrot V, Levin BR. Progress in in silico functional genomics: the malaria Metabolic Pathways database. Trends in parasitology 2006, 22(6):238-240.

45. Yuan J, Cheng KC, Johnson RL, Huang R, Pattaradilokrat S, Liu A, Guha R, Fidock DA, Inglese J, Wellens TE, et al. Chemical genomics profiling for anti-malarial therapies, response signatures, and molecular targets. Science 2011, 333(6043):724-729.

46. Ansena Y, Ramirez F, Schellhorn SE, Lenguater A, Albrecht M. Computing topological parameters of biological networks. Bioinformatics 2008, 24(2):282-284.

47. Henhoe A, Cechanover A. The ubiquitin system. Annual review of biochemistry 1998, 67:425-479.

48. Komander D. The emerging complexity of protein ubiquitination. Biochemical Society transactions 2009, 37(Pt 5):937-953.

49. Chung DW, Le Roch KG. Targeting the Plasmodium ubiquitin/proteasome system with anti-malarial compounds: promises for the future. Infectious disorders drug targets 2010, 10(3):158-164.

50. Waddell H, Araya MJ, Nathanson MH, Garcia CR. The plasmodium receptor for activated C kinase protein inhibits Ca(2+) signaling in mammalian cells. Biochemical and biophysical research communications 2009, 389(4):586-592.

51. Stelter P, Ulrich HD. Control of spontaneous and damage-induced mutagenesis by SUMO and ubiquitin conjugation. Nature 2003, 425(6945):188-191.

52. de la Cruz J, Kessler D, Linder P. Unwinding RNA in Saccharomyces cerevisiae. DEAD-box proteins and related families. Trends in biochemical sciences 1999, 24(2):192-198.

53. Pradhan A, Tuteja R. Bipolar, Dual Plasmodium falciparum helicase 45 expressed in the intraerythrocythrophic developmental cycle is required for parasite growth. Journal of molecular biology 2007, 373(2):268-281.

54. Tuteja R. Genome wide identification of Plasmodium falciparum helicases: a comparison with human host. Cell cycle 2010, 9(1):104-120.

55. Gupta AP, Chi W, Zhu L, Mao S, Luah YH, Lim LH, Bozdech Z. Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite Plasmodium falciparum. PLoS pathogens 2013, 9(2):e1003170.

56. Cui L, Mao J. Chromatin-mediated epigenetic regulation in the malaria parasite Plasmodium falciparum. Europcaricin cell 2010, 9(8):1138-1149.

57. Ponts N, Harris EY, Prudhomme J, Wick I, Eckhardt-Ludika C, Hicks GR, Hardiman G, Lonardi S, Le Roch KG. Nucleosome landscape and control of transcription in the human malaria parasite. Genome research 2010, 20(2):228-238.

58. Durasisingh MT, Voss TS, Ralph SA, Hommel M, Duffy MF, Silva LM, Scherf A, Ivens A, et al. sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. PLoS biology 2009, 7(4):e1000848.

59. Haris AP, Chi W, Zhu L, Mao S, Luah YH, Lim LH, Bozdech Z. Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite Plasmodium falciparum. European journal of medical chemistry 2004, 39(2):268-276.

60. Tonkin CJ, Carret CK, Durasating M, Voss TS, Ralph SA, Hommel M, Duffy MF, Silva LM, Scherf A, Ivens A, et al. sir2 paralogues cooperate to regulate virulence genes and antigenic variation in Plasmodium falciparum. PLoS biology 2009, 7(4):e1000848.
orthologous groups, species and functional annotations. *Nucleic acids research* 2010, **38**(Database):D190-195.

81. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, et al: The COG database: an updated version includes eukaryotes. *BMC bioinformatics* 2003, 4:.

82. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, et al: The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic acids research* 2010, **39**(Database):D561-568.

83. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, et al: STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic acids research* 2013, **41**(Database):D808-815.

Submitted to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit