Proteases in Malaria Parasites - A Phylogenomic Perspective

Hong Cai¹, Rui Kuang², Jianying Gu³* and Yufeng Wang*,¹,⁴

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

Abstract: Malaria continues to be one of the most devastating global health problems due to the high morbidity and mortality it causes in endemic regions. The search for new antimalarial targets is of high priority because of the increasing prevalence of drug resistance in malaria parasites. Malarial proteases constitute a class of promising therapeutic targets as they play important roles in the parasite life cycle and it is possible to design and screen for specific protease inhibitors. In this mini-review, we provide a phylogenomic overview of malarial proteases. An evolutionary perspective on the origin and divergence of these proteases will provide insights into the adaptive mechanisms of parasite growth, development, infection, and pathogenesis.

Received on: May 20, 2011 - Revised on: July 17, 2011 - Accepted on: July 20, 2011

Keywords: Protease, malaria, Plasmodium, phylogenomics, genomics, target, vaccine, systems biology, remote homology detection.

INTRODUCTION

Malaria is one of the most important and persistent global infectious diseases. It is re-emerging as the number one infectious killer, responsible for over one million deaths yearly. The causative agents of malaria are a group of parasites in the Plasmodium genus. Five species, P. falciparum, P. vivax, P. malariae, P. ovale, and P. knowlesi are human pathogens. P. falciparum causes the most deadly form of malaria.

The re-emergence of malaria is largely due to the growing prevalence of parasite populations that show resistance to multiple drug treatment. With the advent of high throughput genomic, transcriptomic, proteomic, metabolomic, and pharmacogenomic technologies, enormous efforts have been focused on the identification and characterization of new and effective antimalarial targets [1-7]. These targets are selected based on several common criteria: (1) They are essential for parasite biology. The disruption of these genes or gene products leads to deleterious effects on parasite growth, development, or invasion. For example, cyclin-dependent protein kinases (CDKs) play indispensible roles in cell cycle progression and signal transduction [8-17]; (2) It is feasible to design or screen for effective pharmacophores or candidate inhibitors. For example, two compounds, chalcones and tryptanthrins, were identified by rational drug design, compound screening and molecular modeling as potent and specific inhibitors for the P. falciparum CDK7 homolog, Pfmrk [18]; (3) The drugs directed at the selected targets should have no or minimal adverse effects on humans. Some of the potential targets such as 1-deoxy-D-xylulose 5-phosphate (DOXP) reductoisomerase [19, 20] and apicoplast gyrase [21] are localized to apicoplast, an organelle uniquely present in Plasmodium parasites and other parasites in the Apicomplexa phylum. These enzymes are crucial for apicoplast metabolism, replication, transcription and translation. Because the apicoplast is of prokaryotic origin, the inhibitors of these targets may have small or no side effects on the host.

Proteases are a class of promising antimalarial targets. They are digestive enzymes that degrade peptide bonds. They have demonstrated roles in parasite nutrition, development, invasion and egress: (1) a cascade of aspartic proteases plasmepsins [22, 23], cysteine proteases falcipains [24, 25] and metalloproteases [26-28] mediate massive degradation of host hemoglobin to release amino acids for parasite nutrition; (2) serine proteases (subtilases) have been implicated in erythrocyte invasion and parasite exit from the host [29-32]; (3) proteases are active mediators for cell cycle regulation and cell signaling [33-35]. Because the mechanisms of enzymatic action for many classes of proteases are known or can be derived from structural modeling or computer-aided drug design, it is possible to design or screen for protease inhibitors. The inhibitor classes for plasmepsins and falcipains have been investigated and evaluated [36-44].

Proteases, in addition to their potential as drug targets, are a prime example of supergene families with complex evolutionary histories involving gene duplication, domain shuffling, and lateral gene transfer. In this paper, we present a phylogenomic survey of malarial proteases. A better
understanding of protease evolution will bring new insights into the genetic basis of adaptive phenotypes such as pathogenesis and virulence.

PHYLGENEOMICS FOR THE IN SILICO PREDICTION OF PROTEASES IN THE PLASMODIUM GENOMES

Phylogenomics is an emerging discipline that combines molecular evolution theory and genomics [45, 46]. One of its direct and most important applications is to make functional predictions for previously uncharacterized proteins. The major hurdle that plagues all genomics-driven efforts in antimarial target identification is the annotation problem [47]. In *Plasmodium* species, sequence similarity can be low, due to mutation, insertion, deletion, shuffling and recombination events, meaning high-confidence alignments between descendant sequences are not feasible and functional assignments are obscured. Genome annotation using traditional alignment-based algorithms has failed to assign functionality to over 60% of the ORFs in *P. falciparum* [48]. Popular methods for building probabilistic alignment models, such as PSI-BLAST [49], hidden Markov models (HMMs) [50], COMPASS [51] and HHSearch [52] show low accuracy and coverage when sequence similarity falls below 30% [53-55].

Only a handful of proteases had been discovered and characterized prior to the completion of genome sequencing for *P. falciparum* [48]. Using a comparative genomic approach, we predicted that a total of 92 protease homologs were present in *P. falciparum* genome, and at least 88 of them were expressed at the mRNA level by microarray and RT-PCR assays [56]. Subsequent data mining on the parasite proteome revealed that 67 of these predicted proteases were expressed at the protein level at least in one stage of the life cycle [57]. Recently we extended our study to other sibling species of malaria parasites, including *P. vivax* [58], which is the most widely distributed human malaria parasite, and three rodent species *P. berghei*, *P. chabaudi*, and *P. yoelii yoelii* [59, 60], which serve as the animal models for human malaria. In addition to traditional BLAST searches, we adopted a novel support vector machine (SVM)-based, supervised machine learning approach to tackle the remote homology problem. The underlying principle for remote homology detection lies in the domain of phylogenomics: these algorithms are designed to capture subtle similarities between the unknown proteins and the annotated proteins based on the evolutionarily conserved characteristics of the genes/proteins. A SVM classifier is a function that separates the training data into two classes and also maximizes the geometric margin between them in a feature space. Unlike most alignment-based algorithms which build models only with positive sequences, SVMs also use negative sequences (proteins outside the protein family) to learn the difference between the two classes. The SVM approach discovered several putative proteases that were not detectable by PSI-BLAST. For example, one putative PPPDE protease (PF09400c) is a member of a novel family with a papain-like fold. This family was postulated to play a role in deubiquitination and cell cycle regulation [61]. The total number of predicted proteases in *P. falciparum* was increased from 92 to 123.

The degradome of five malaria parasite species is comprised of 115-137 putative proteases in five distinct catalytic classes (aspartic, cysteine, metallo, serine and threonine) (See Table 2 and Table 3 in [47]), which account for 0.9-2.3% of the open reading frames (ORFs) in the genome. They form 37 protein families based on their evolutionary relationship and structure conservation, according to the MEROPS protease classification system [62], and 29 of these families are commonly shared in five species. These proteases are important players in metabolism, cell cycle regulation, invasion, stress response, transcriptional regulation, signal transduction, and trafficking. A number of these proteases are becoming targets for functional characterization and rational inhibitor design [43, 63-69].

PHYLGENEOMICS FOR FUNCTIONAL CHARACTERIZATION OF MALARIAL PROTEASES

Phylogenomic analysis provides a cost-effective means to examine the evolutionary profiles of genes and gene products for functional prediction or characterization. The procedure often involves homology identification, multiple sequence alignment, phylogenetic reconstruction, inference of function, evolutionary analysis of orthology and paralogy, and identification of lateral gene transfer events [70]. This approach is particularly useful for the studies of protein families and protein superfamilies such as kinases and transporters [71-74]. In the domain of protease researches, phylogenomics has contributed to, for example, the classification and reconstruction of evolutionary diversification of serine proteases in fungi [75], the evolutionary profiling of cystatins, which comprise a superfamily of cysteine protease inhibitors [76], and the development of a statistical framework that was able to detect site-specific functional divergence in the caspase family of cysteine proteases [77]. The annotation of the predicted malarial proteases was essentially based on phylogenomic analysis [56], which revealed an array of novel proteases that could potentially be important for parasite-specific functions. A single copy of calpain was identified in the five surveyed *Plasmodium* genomes. It contains active site residues (C1035-N-1371-H1391 in *P. falciparum*) that are conserved in known or characterized calpains. Although calpains are well-known modulators for signal transduction, differentiation, cell motility, cell cycle regulation and cell-cell communication from bacteria to humans, its physiological role in parasite biology is yet to be defined. Nevertheless, partial knockdown assays indicated that the malarial calpain is crucial for the optimal growth of the parasite and cell cycle progression [78]. Interestingly, phylogenetic analysis revealed that the malarial calpain belongs to a clade of calcium-independent calpains, a lineage restricted to alveolates; the divergence from major human calpains makes it a possible drug target.

Multiple copies of metacaspases were identified in *Plasmodium* genomes [34, 47, 79], suggesting the existence of apoptosis or a similar signaling cascade in malaria parasites. Apoptosis-inducer and the administration of antimalarial drug chloroquine were shown to lead to DNA fragmentation and mitochondrial membrane potential disruption, indicative of the onset of programmed cell death.
in the parasite [80]. Phylogenetic analysis shows that this family may be generated by at least one gene duplication event: metacaspase-3 may represent an ancestral form, while metacaspase-1 and metacaspase-2 are more closely related to each other (Fig. 1). Metacaspase-1, in particular, contains the typical catalytic domain and the active site residues (histidine and cysteine dyad) that are essential for proteolytic function. In addition, it harbors a caspase recruitment domain (CARD) for apoptosis-related signaling [80]. The discovery of metacaspases has triggered the search for proteins/regulators in the parasite apoptotic network [81], which may represent a major stress-response system for parasites to survive under drug treatment and host immune challenges. Another example of potentially important proteases in malaria parasites is the signal peptide peptidase (SPP). SPP is an active player in regulated intramembrane proteolysis (RIP), which initiates signal transduction via processing the transmembrane segments of the substrates. One single copy of SPP is found in *P. falciparum*, *P. vivax*, *P. yoelii yoelii*, and *P. chabaudi*, and two copies are present in *P. berghei*. The putative active sites Tyr–Asp (YD) and Gly–Leu–Gly–Asp (GLGD) motifs which are universally conserved in SPPs are present in all the malarial SPPs (Fig. 2A). Phylogenetic analysis shows that the malarial SPPs form a distinct clade that is distantly related to the SPPs found in other animals from the fruit fly to the human (Fig. 2B). Data mining of the protein-protein interaction network revealed that *P. falciparum* SPP (PfSPP) is a highly connected protein with 54 association partners. Fig. (3) shows a schematic association map of PfSPP with representative partners with relatively high statistical support. Each association between a pair of proteins has a confidence score (S) ranging from 0.15 to 0.999 that was inferred from the evidence used to establish the association [82], including the interolog comparisons, which is rooted on phylogenomic inference of network associations among evolutionarily related organisms, yeast-2-hybrid (Y2H) assays, Gene Ontology (GO) classification for biological functions, cellular processes and sublocalizations, structural configuration, co-expression and co-occurrence patterns, and so on [83, 84]. Phylogenomic inference predicted that PfSPP is associated with a putative ER lumen protein retaining receptor (ERD2) [85] and a secretory protein Sec61 [86], both of which are components in the parasite translocation machinery required for the uptake of nutrients and expulsion of wastes: the SPP homolog was found to be co-expressed with the ERD2 and Sec61 homologs in three model organisms: *Arabidopsis thaliana*, *Caenorhabditis elegans*, and *Drosophila melanogaster*. PfAPP is also associated with signal peptidase, translation initiation and elongation factors, splicing factor, peptide chain release factor, and a variety of enzymes, suggesting that it is involved in transport, translation, posttranslational regulation and metabolism. It has recently been considered as a promising drug target since gene disruption assays indicate it is essential for parasite growth and merozoite invasion [87, 88].

**PHYLOGENOMICS FOR ASSESSING THE SUITABILITY OF MALARIAL PROTEASES AS DRUG TARGETS**

Phylogenomic analysis can reveal the complex evolutionary history of malarial proteases; their origin and...
Fig. (2). (A) The alignment of the catalytic region of the signal peptide peptidases (SPPs) in six *Plasmodium* species and other representative species. The putative active sites Tyr–Asp (YD) and Gly–Leu–Gly–Asp (GLGD) motifs are highlighted. (B) The phylogenetic tree of SPPs, inferred using the neighbor-joining method based on the amino acid sequences with Poisson corrected distance.
relatedness to the host help researchers to assess their suitability as potential antimalarial targets. Jean et al. [89] and Coombs et al. [22] conducted elegant phylogenomic analyses on plasmepsins, a group of aspartic proteases in the pepsin (A1) family. Ten plasmepsins have been identified in *Plasmodium falciparum*, namely PM I-X; PM III is also known as histo-aspartic protease (HAP). They are divided into two classes (Fig. 4): (1) PM I-IV are all intronless with one single exon; They are located in adjacent positions on Chromosome 14, and are likely to be generated by tandem gene duplications. (2) PM V-X form a large clade which may represent an ancestral type of plasmepsins; Introns are present in PM VI, VII, VIII, and IX. The most extreme case is seen in PM VI, which contains 15 exons. An evolutionary model suggested that lateral gene transfer, exon shuffling and intron loss events may lead to the diverse types of plasmepsins in the parasite genome that are required for effective hemoglobin digestion [22, 89]. Inhibition of hemoglobin digestion causes starvation of the parasite and accumulation of intermediates toxic to parasites [23, 90]. Recently, two neutral aminopeptidases, M1 alanyl aminopeptidase (PfM1AAP, MAL13P1.56) and M17 leucine aminopeptidase (PfM17LAP, PF14_0439) have been characterized. Both enzymes are required for the late stage of hemoglobin digestion, which releases free amino acids for parasite nutrition and development inside the human red blood cell. Parasites were not viable *in vivo* and *in vitro* with the treatment of aminopeptidase inhibitors [91]. Because only single copy of PfM17LAP is present in *P. falciparum* genome, its disruption cannot be compensated for by any homolog. The X-ray structure of PfM17LAP has been resolved, opening a promising avenue for rational drug design [92].

Gene duplication and lateral gene transfer are implicated in the evolution of other protease families such as subtilases and falcipains. Three copies of subtilases are found in *P. falciparum* genome and 2-5 copies are present in the other *Plasmodium* genomes. Subtilases are required for parasite invasion and egress from the human host [30, 93-95]. Evolutionarily, they are probably acquired via lateral gene transfer from a bacteria origin where subtilsins are commonly found. No statistically significant subtilase homologs are found in the humans, which is a desirable feature for drug targets. Similarly, falcipains are crucial for parasite biology; they may have dual roles in both hemoglobin digestion and host cell egress [30, 96-98]. They are evolutionarily closely related to the papains found in viruses and fungi. Lineage-specific expansion is evident in the evolution of rhomboid proteases (the S54 serine protease family) in *Plasmodium*: eight copies are present in *P. falciparum*, and 5-8 copies are present in the other species. They are the central players in regulated intramembrane proteolysis (RIP) and have been implicated in parasite development, invasion, cell signaling and pathogenesis [35, 99-101]. Different rhomboid proteases may have specific substrates preferences; their potential substrates include various adhesins and surface antigens [33]. Phylogenetic analysis revealed that they are closely related to the rhomboid homologs present in other apicomplexan parasites including *Toxoplasma gondii*, *Eimeria tenella*, and *Toxoplasma gondii*.
Cryptosporidium spp. and Theileria spp [102]. There is only weak sequence similarity between the malarial rhomboids and a mitochondrial rhomboid protease, PARL, in the human genome. Targeting rhomboids and their associated signaling pathways therefore may be a novel therapeutic strategy. Other proteases with potentially important functions, such as calpain, metacaspases, and signal peptidase I, are also phylogenetically divergent from the host lineage.

Phylogenomic analysis also revealed a group of putative malarial proteases that are destined to parasite-specific organelles of prokaryotic origin. The top target organelle for drug development is the apicoplast. It is essential for parasite life cycle, as inhibitors for apicoplast metabolism and replication resulted in the death of parasite [103, 104]. Using two independent algorithms, PATS (an artificial neural network algorithm) [105] and PlasmoAP [106], which combines signal peptide prediction [107] and rule-based classification. Ralph et al. [104] identified more than 540 genes in *P. falciparum* that are targeted to the apicoplast. Twenty-one of these genes encode putative proteases (Table 1). Because the apicoplast is derived from an ancient endosymbiosis in which the eukaryotic ancestor engulfed a red alga with a solitary chloroplast, its disruption does not cause significant interference with the host functions. The cyanobacterial heritage of the apicoplast enhances the potential of these apicoplast-targeted proteases as drug targets. Notably, five putative proteases from the ClpP endopeptidase family (S14) are predicted to be localized to the apicoplast (Table 1). They are the central players in the parasite heat shock response system, a key system for...

**Fig. (4).** The phylogenetic tree of plasmpepsins, inferred using the neighbor-joining method based on the amino acid sequences with Poisson corrected distance. PM I-IV contain one single exon without any introns. PM VI contains 15 exons, PM VIII contains 13 exons, PM VII and PM IX contain eight exons, and PM V and PM X contain one exon.
Proteases in Malaria Parasites

Table 1. Putative Proteases that are Predicted to Localize to the Apicoplast in P. falciparum

| Catalytic Type | Protease Family | P. falciparum protease ID | Annotation |
|---------------|----------------|--------------------------|------------|
| Aspartic      | A1 (pepsin family) | PF13_0133 | Plasmepsin V |
|               | A22 (presenilin family) | PF14_0625 | Plasmepsin VIII |
|               | C1 (papain family) | PF14_0543 | signal peptide peptidase |
| Metallo       | M1 (aminopeptidase N family) | MAL13P1.56 | M1-family alanyl aminopeptidase AMPN |
|               | M16 (pirimysin family) | PF14_0382 | Pirrmysin |
|               | M17 (leucyl aminopeptidase family) | PF14_0439 | M17 leucyl aminopeptidase |
|               | M24 (methionyl aminopeptidase 1) | MAL8P1.140 | putative methionine aminopeptidase 1c |
|               | M41 (PspH endopeptidase family) | PF14_0616 | putative ATP-dependent protease la |
| Serine        | S8 (subtilisin family) | PFE0355c | putative subtilisin-like protease 3 |
|               | S14 (ClpP endopeptidase family) | PFC0310c | ATP-dependent Clp protease proteolytic subunit |
|               | S33 (prolyl aminopeptidase) | PF14_0348 | ATP-dependent Clp protease proteolytic subunit |
|               | S34 (Rhomboide family) | PF14_0063 | putative ClpB |
|               | S35 (Rhomboid family) | PF11_0175 | Clp protease, heat shock protein 101 |
|               | S36 (prolyl aminopeptidase) | PF08_0063 | putative aminopeptidase |
|               | S37 (prolyl aminopeptidase) | PF14_0015 | putative aminopeptidase |
|               | U48 (prenyl protease 2 family) | PF13_0312 | rhomboid protease ROM3 |
|               | U49 (prenyl protease 2 family) | PF0660c | putative protease |

CONCLUSIONS

The phylogenomic approach has played an important role in the identification and in silico characterization of proteases in malaria parasites, providing a promising and largely uncharacterized set of targets for wet lab functional characterization and drug design. An evolutionary perspective on the origin and divergence of these proteases provides insights into the adaptive mechanisms of parasite growth, development, infection, and pathogenesis.
ACKNOWLEDGEMENTS

This work is supported by NIH grants GM081068, SC1AI080579, and AI067543 to YW, and the PSC-CUNY Research Award PSRREG-39-497 to JG. YW is also supported by NIH grant RR013646. We thank the Computational Biology Initiative at UTSA for providing computational support. 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, or the National Institutes of Health.

ABBREVIATIONS

CARD = Caspase recruitment domain
CDK = Cyclin-dependent protein kinase
DOXP = 1-Deoxy-D-xylulose 5-phosphate
ER = Endoplasmic reticulum
GO = Gene Ontology
HAP = Histo-aspartic protease
HSP = Heat shock protein
HMM = Hidden Markov models
ORF = Open reading frame
RIP = Regulated intramembrane proteolysis
SPP = Signal peptide peptidase
SVM = Support vector machine
Y2H = Yeast 2-hybrid

REFERENCES

[1] Dvorin, J. D.; Martyn, D. C.; Patel, S. D.; Grimley, J. S.; Collins, C. R.; Hopp, C. S.; Bright, A. T.; Westenberger, S.; Winzeler, E.; Blackman, M. J.; Baker, D. A.; Wandless, T. J.; Duraisingham, M. T. A plant-like kinase in Plasmodium falciparum regulates parasite egress from erythrocytes. Science, 2010, 328(5980), 910-912.
[2] Ecker, A.; Lakshmanan, V.; Sinis, P.; Coppens, I.; Fidock, D. A. Evidence That Mutant PICRT Facilitates the Transmission to Mosquitoes of Chloroquine-Treated Plasmodium Gametocytes. J. Infect. Dis., 2011, 203(2), 228-236.
[3] Jana, S.; Falwill, J. Novel molecular targets for antimalarial chemotherapy. Int. J. Antimicrob. Agents, 2007, 30(1), 4-10.
[4] Joet, T.; Eckstein-Ludwig, U.; Morin, C.; Krishna, S. Validation of the hexose transporter of Plasmodium falciparum as a novel drug target. Proc. Natl. Acad. Sci. USA, 2003, 100(13), 7476-7479.
[5] Johnson, D. J.; Fidock, D. A.; Mungthin, M.; Lakshmanan, V.; Sidhu, A. B.; Bray, P. G.; Ward, S. A. Evidence for a central role of PfCRT in circumventing Plasmodium falciparum resistance to diverse antimalarial agents. Mol. Cell, 2004, 15(6), 867-877.
[6] Kirk, K.; Howitt, S. M.; Broer, S.; Sulston, J. E.; Downie, M. J. Purine uptake in Plasmodium: transport versus metabolism. Trends Parasitol, 2009, 25(6), 246-249.
[7] Patel, A. P.; Staines, H. M.; Krishna, S. New antimalarial targets: the example of glucose transport. Travel Med. Infect. Dis., 2008, 6(1-2), 58-66.
[8] Rangarajan, R.; Rea, A.; Henry, N.; Madame, P.; Parzy, D.; Nizey, M. P.; Doering, C.; Sultan, A. Pbcrcr-1, the Plasmodium berghei orthologue of P. falciparum cdc-2 related kinase 1 (Pfckr-1), is essential for completion of the intraerythrocytic asexual cycle. Exp. Parasitol., 2006, 112(3), 202-207.
[9] Dorin, D.; Sembli, J. P.; Poulett, P.; Alano, P.; Goldring, J. P.; Whittle, C.; Patterson, S.; Chakrabarti, D.; Doering, C. PIPK7, an atypical MEK-related protein kinase, reflects the absence of classical three-component MAPK pathways in the human malaria parasite Plasmodium falciparum. Mol. Microbiol., 2005, 55(1), 184-196.
[10] Merckx, A.; Le Roch, K.; Nizey, M. P.; Dorin, D.; Alano, P.; Gutierrez, G. J.; Neberda, A. R.; Goldring, D.; Whittle, C.; Patterson, S.; Chakrabarti, D.; Doering, C. Identification and initial characterization of three novel cyclin-related proteins of the human malaria parasite Plasmodium falciparum. J. Biol. Chem., 2003, 278(41), 39839-39850.
[11] Morgan, D. O. Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu. Rev. Cell Dev. Biol., 1997, 13, 261-291.
[12] Grant, K. M. Targeting the cell cycle in the pursuit of novel chemotherapies against parasitic protozoa. Curr. Pharm. Des., 2008, 14(9), 917-924.
[13] Koyama, F. C.; Chakrabarti, D.; Garcia, C. R. Molecular machinery of signal transduction and cell cycle regulation in Plasmodium. Mol. Biochem. Parasitol., 2009, 165(1), 1-7.
[14] Kozlov, S.; Waters, N. C.; Chavchich, M. Leveraging cell cycle analysis in anticancer drug discovery to identify novel plasmodial drug targets. Infect. Disord. Drug Targets, 2010, 10(3), 165-190.
[15] Holland, Z.; Prudent, R.; Reiser, J. B.; Cochet, C.; Doering, C. Functional analysis of protein kinase CK2 of the human malaria parasite Plasmodium falciparum. Eukaryot. Cell, 2009, 8(3), 388-397.
[16] Merckx, A.; Echalier, A.; Langford, K.; Sicard, A.; Langsley, G.; Jouve, J.; Doering, C.; Noble, J.; Baccietti, J. Structures of P. falciparum protein kinase 7 identify an activation motif and leads for inhibitor design. Structure, 2008, 16(2), 228-238.
[17] Waters, N. C.; Geyer, J. A. Cyclin-dependent protein kinases as therapeutic drug targets for antimarial drug development. Expert. Opin. Ther. Targets, 2003, 7(1), 7-17.
[18] Keenan, S. M.; Geyer, J. A.; Welsh, W. J.; Prige, S. T.; Waters, N. C. Rational inhibitor design and iterative screening in the identification of selective plasmodial cyclin dependent kinase inhibitors. Comb. Chem. High Throughput Screen., 2005, 8(1), 27-38.
[19] Jomaa, H.; Wiesner, J.; Sanderbrand, S.; Altincicek, B.; Weidemeyer, C.; Hintz, M.; Turbachova, I.; Eberl, M.; Zeidler, J.; Lichtenhalter, H. K.; Soldati, D.; Beck, E. Inhibitors of the non-ovalsolute pathway of isopenoid biosynthesis as antimalarial drugs. Science, 1999, 285(5433), 1573-1576.
[20] Wiesner, J.; Jomaa, H. Isopenoid biosynthesis of the apicoplast as drug target. Curr. Drug Targets, 2007, 8(1), 3-13.
[21] Dah, E. L.; Rosenthal, P. J. Apicoplast translation, transcription and genome replication: targets for antimalarial antibiotics. Trends Parasitol., 2008, 24(6), 279-284.
[22] Coombs, G. H.; Goldberg, D. E.; Klemba, M.; Berry, C.; Kay, J.; Mottram, J. C. Aspartic proteases of Plasmodium falciparum and other parasitic protozoa as drug targets. Trends Parasitol., 2001, 17(11), 532-542.
[23] Goldberg, D. E. Hemoglobin degradation. Curr. Top. Microbiol. Immunol., 2005, 295, 275-291.
[24] Rosenthal, P. J. Hydrolysis of erythrocyte proteins by proteases of malaria parasites. Curr. Opin. Hematol., 2002, 9(2), 140-145.
[25] Rosenthal, P. J. Cysteine proteases of malaria parasites. Int. J. Parasitol., 2004, 34(13-14), 1489-1499.
[26] Eggleston, K. H.; Duffin, K. L.; Goldberg, D. E. Identification and characterization of falcinsin, a metalloproteinase involved in hemoglobin catabolism within the malaria parasite Plasmodium falciparum. J. Biol. Chem., 1999, 274(45), 32411-32417.
[27] Murata, C. E.; Goldberg, D. E. Plasmodium falciparum falcinsin: a metalloprotease with dual specificity. J. Biol. Chem., 2003, 278(39), 38022-38028.
[28] Murata, C. E.; Goldberg, D. E. Plasmodium falciparum falcinsin: an unprocessed food vacuole enzyme. Mol. Biochem. Parasitol., 2003, 129(1), 123-126.
[29] Blackman, M. J. Proteases involved in erythrocyte invasion by the malaria parasite: function and potential as chemotherapeutic targets. Curr. Drug Targets, 2000, 1(1), 59-83.
[30] Blackman, M. J. Malarial proteases and host cell egress: an 'emerging' cascade. Cell Microbiol., 2008, 10(10), 1925-1934.
[31] Blackman, M. J.; Fujioka, H.; Stafford, W. H.; Sajid, M.; Clough, B.; Fleck, S. L.; Aikawa, M.; Grainger, M.; Hackett, F. A subtilisin-like protein in secretory organelles of Plasmodium falciparum merozoites. J. Biol. Chem., 1998, 273(36), 23398-23409.
Proteases in Malaria Parasites

[52] Soding, J. Protein homology detection by HMM-HMM comparison. Bioinformatics, 2005, 21(7), 951-960.

[53] Chubb, D.; Jefferys, B. R.; Sternberg, M. J.; Kelley, L. A. Sequencing delivers diminishing returns for homology detection: implications for mapping the protein universe. Bioinformatics, 2010, 26(21), 2664-2671.

[54] Qi, Y.; Sadreyev, R. I.; Wang, Y.; Kim, B. H.; Grishin, N. V. A comprehensive system for evaluation of remote sequence similarity detection. BMC Bioinformatics, 2007, 8, 314.

[55] Reid, A. J.; Yeats, C.; Orengo, C. A. Methods of remote homology detection can be combined to increase coverage by 10% in the midnight zone. Bioinformatics, 2007, 23(18), 2353-2360.

[56] Liu, W. Y. M.; Wang, X. Y.; Liu, X. Y. Data-mining approaches reveal hidden families of proteases in the genome of malaria parasite. Genome Res., 2003, 13(4), 601-616.

[57] Wang, Y.; Wu, Y. Computer assisted searches for drug targets with emphasis on malaria proteases and their inhibitors. Curr. Drug Targets Infect. Disord., 2004, 4(1), 25-40.

[58] Carlton, J. M.; Adams, J. H.; Silva, J. C.; Bidwell, S. L.; Lorenzi, H.; Caler, E.; Crabtree, J.; Angiolii, S. V.; Merino, E. F.; Amedeo, P.; Cheng, Q.; Coulson, R. M.; Crabb, B. S.; Del Portillo, H. A.; Essien, K.; Feldblumy, T. V.; Fernandez-Becerra, C.; Gilson, P. R.; Gueye, A. H.; Guo, X.; Kangala, S.; Kooi, T. W.; Korsnickyz, M.; Meyer, E. V.; Nene, V.; Paulsen, I. H.; White, O. R.; Sali, A.; Rosenberg, S. Q.; Sargeant, J. T.; Salzberg, S. L.; Stoeckert, C. J.; Sullivan, S. A.; Yamamoto, M. M.; Hoffman, S. L.; Wotruba, J. R.; Gardner, M. J.; Galinski, M. R.; Barnwell, J. W.; Fraser-Liggett, C. M. Comparative genomics of the neglected human malaria parasite Plasmodium vivax. Nature, 2008, 455(7214), 757-763.

[59] Carlton, J.; Silva, J.; Hall, N. The genome of model malaria parasites, and comparative genomics. Curr. Issues Mol. Biol., 2005, 7(1), 23-37.

[60] Carlton, J. M.; Angiolii, S. V.; Suh, B. B.; Kooi, T. W.; Pertea, M.; Silva, J. C.; Ermolaeva, M. D.; Allen, J. E.; Selengut, J. D.; Koo, H. L.; Peterson, J. D.; Pop, M.; Kossack, D. S.; Shumway, M. F.; Bidwell, S. L.; Shalom, J. S.; van Aken, S. E.; Riedmuller, S. B.; Feldblumy, T. V.; Cho, J. K.; Quackenbush, J.; Selegal, M.; Shaps, A.; Crouse, K. M.; Florens, L.; Yates, J. R.; Raine, J. D.; Sinden, R. E.; Harris, M. A.; Cunningham, D. A.; Preiser, P. S.; Bergman, L. W.; Vaidya, A. B.; Van Lin, L. H.; Janse, C. J.; Waters, A. P.; Smith, H. O.; White, O. R.; Salzberg, S. L.; Venter, J. C.; Fraser, C. M.; Hoffman, S. L.; Gardner, M. J.; Carucci, D. J. Genome sequence and comparative analysis of the model rodent malaria parasite Plasmodium yoelii yoelii. Nature, 2002, 419(6906), 512-519.

[61] Iyer, L. M.; Koonin, E. V.; Aravind, L. Novel predicted peptidases with a potential role in the ubiquitin signaling pathway. Cell Cycle, 2004, 3(11), 1440-1450.

[62] Rawlings, N. D. A large and accurate collection of peptidase cleavages in the MEROPS database. Database (Oxford), 2009, 2009, ba015.

[63] Boddey, J. A.; Hodder, A. N.; Gunther, S.; Gilson, P. R.; Patsiouras, H.; Kapp, E. A.; Pearce, J. A.; de Koning-Ward, T. F.; Simpson, R. J.; Crabb, B. S.; Cowman, A. F. An aspartyl protease directs malaria effector proteins to the host cell. Nature, 2010, 463(7281), 627-31.

[64] Macgowan, S.; Seal, C. A.; Birru, W. A.; Cardacivos-Tay, T.; Stack, C. M.; Lownther, J.; Skinner-Adams, T.; Mucha, A.; Kafarski, P.; Grembecka, J.; Trenholme, K. R.; Bucks, A. M.; Gardiner, D. L.; Dalton, J. P.; Whistcock, J. C. Structure of the Plasmodium falciparum M17 aminopeptidase and significance for the design of drugs targeting the neutral exopeptidases. Proc. Natl. Acad. Sci. USA, 2010, 107(6), 2449-2454.
Russo, I.; Oksman, A.; Vaupel, B.; Goldberg, D. E. A calpain 426
Wu, Y.; Wang, X.; Liu, X.; Wang, Y. Data-mining approaches 426
von Mering, C.; Jensen, L. J.; Snel, B.; Hooper, S. D.; Krupp, M.; 426
Meslin, B.; Barnadas, C.; Boni, V.; Latour, C.; De Monbrison, F.; 426
Sjolander, K. Phylogenomic inference of protein molecular 426
Jung, K. H.; Cao, P.; Seo, Y. S.; Dardick, C.; Ronald, P. C. The 426
Hu, G.; Leger, R. J. A phylogenomic approach to reconstructing 426
Wang, Y. F.; Gu, X. Functional divergence in the caspase gene 426
Pradel, G. Malaria proteases mediate inside-out egress of 426
Rice Kinase Phylogenomics Database: a guide for systematic 426
Kaiser, K.; Picot, S. Features of apoptosis in Plasmodium 426
Kirkard, J.; Pascall, J. C.; Brown, K. D.; Soldati, D. 426
Srinivasan, P.; Coppens, I.; Jacobs-Lorena, M. Distinct roles of 426
Dowse, T. J.; Pascall, J. C.; Brown, K. D.; Soldati, D. 426
Dowse, T. J.; Soldati, D. Rhomboid-like proteins in Apicomplexa: 426
Sijwali, P. S.; Koo, J.; Singh, N.; Rosenthal, P. J. Gene disruptions 426
Shenai, B. R.; Sijwali, P. S.; Singh, A.; Rosenthal, P. J. 426
Jean, L.; Hackett, F.; Martin, S. R.; Blackman, M. J. Functional 426
Jean, L.; Withers-Martinez, C.; Hackett, F.; Blackman, M. J. 426
Brown, W. C.; Azad, A. F.; Sobral, B. W. Phylogenomics reveals a 426
Gardiner, D. L.; Dalton, J. P.; Whisstock, J. C. Structure of the 426
Stack, C. M.; Lowther, J.; Skinner-Adams, T.; Mucha, A.; 426
Kafarski, P.; Goldberg, D. E. Four plasmepsins are active in the Plasmodium 426
Goldberg, D. E. A calpain 426
Gardiner, D. L.; Dalton, J. P.; Whisstock, J. C. Structure of the 426
Li, X.; Chen, H.; Oh, S. S.; Chishti, A. H. A Presenilin-like protease associated with Plasmodium falciparum micronemes is involved in erythrocyte invasion. Mol. Biochem. Parasitol., 2008, 158(1), 22-31.

Jean, L.; Long, M.; Young, J.; Pery, P.; Tomley, F. Aspartyl protease genes from apicomplexan parasites: evidence for evolution of the gene structure. Trends Parasitol., 2001, 17(10), 491-498.

Banerjee, R.; Liu, J.; Beatty, W.; Pelosof, L.; Klemba, M.; Goldberg, D. E. Four plasmepsins are active in the Plasmodium falciparum food vacuole, including a protease with an active-site histidine. Proc. Natl. Acad. Sci. USA, 2002, 99(20), 990-995.

Skinner-Adams, T. S.; Stack, C. M.; Trenholme, K. R.; Brown, C. L.; Grembecka, J.; Lowther, J.; Mucha, A.; Drag, M.; Kafarski, P.; McGowan, S.; Whistho...
[103] Ralph, S. A.; Foth, B. J.; Hall, N.; McFadden, G. I. Evolutionary pressures on apicoplast transit peptides. *Mol. Biol. Evol.*, 2004, 21(12), 2183-2194.

[104] Ralph, S. A.; van Dooren, G. G.; Waller, R. F.; Crawford, M. J.; Fraunholz, M. J.; Foth, B. J.; Tonkin, C. J.; Roos, D. S.; McFadden, G. I. Tropical infectious diseases: metabolic maps and functions of the Plasmodium falciparum apicoplast. *Nat. Rev. Microbiol.*, 2004, 2(3), 203-216.

[105] Zuegge, J.; Ralph, S.; Schnuiker, M.; McFadden, G. I.; Schneider, G. Deciphering apicoplast targeting signals--feature extraction from nuclear-encoded precursors of Plasmodium falciparum apicoplast proteins. *Gene*, 2001, 280(1-2), 19-26.

[106] Foth, B. J.; Ralph, S. A.; Tonkin, C. J.; Struck, N. S.; Fraunholz, M.; Roos, D. S.; Cowman, A. F.; McFadden, G. I. Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum. *Science*, 2003, 299(5607), 705-708.

[107] Nielsen, H.; Engelbrecht, J.; Brunak, S.; von Heijne, G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.*, 1997, 10(1), 1-6.

[108] El Bakkouri, M.; Pow, A.; Mulichak, A.; Cheung, K. L.; Artz, J. D.; Amani, M.; Fell, S.; de Koning-Ward, T. F.; Goodman, C. D.; McFadden, G. I.; Ortega, J.; Hui, R.; Houry, W. A. The Clp chaperones and proteases of the human malaria parasite Plasmodium falciparum. *J. Mol. Biol.*, 2010, 404(3), 456-477.

[109] Rathore, S.; Sinha, D.; Asad, M.; Bottcher, T.; Afrin, F.; Chauhan, V. S.; Gupta, D.; Sieber, S. A.; Mohmed, A. A cyanobacterial serine protease of *Plasmodium falciparum* is targeted to the apicoplast and plays an important role in its growth and development. *Mol. Microbiol.*, 2010, 77(4), 873-890.

[110] Bottcher, T.; Sieber, S. A. Beta-lactones as specific inhibitors of ClpP attenuate the production of extracellular virulence factors of *Staphylococcus aureus*. *J. Am. Chem. Soc.*, 2008, 130(44), 14400-1.

[111] Saitou, N.; Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, 1987, 4(4), 406-425.

[112] Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol. Biol. Evol.*, 2011.

[113] Smoot, M. E.; Ono, K.; Ruscheinski, J.; Wang, P. L.; Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*, 2010, 27(3), 431-432.