

1     **PHIST Proteins: At the Center of Host Cell Remodeling**

2

3     Jan D. Warncke<sup>a,b</sup>, Ioannis Vakonakis<sup>c</sup>, Hans-Peter Beck<sup>a,b\*</sup>

4

5     <sup>a</sup> Swiss Tropical and Public Health Institute, Basel, Switzerland

6     <sup>b</sup> University of Basel, Basel, Switzerland

7     <sup>c</sup> Department of Biochemistry, University of Oxford, Oxford, UK

8

9     \* corresponding author: Hans-Peter Beck, Swiss TPH, Socinstrasse 57, CH-4051

10     Basel, Switzerland, phone +41 61 284 8116, fax: +41 61 284 8101, hans-

11     peter.beck@unibas.ch

12

## 13 Table of Contents

14

15	INTRODUCTION .....	4
16	Malaria .....	4
17	Protein Export .....	5
18	PHIST PROTEINS.....	6
19	Methodology for Review .....	6
20	The PHIST Protein Family .....	6
21	Comparison of PHIST and PRESAN Domains.....	8
22	PHISTs in Other <i>Plasmodium</i> Species .....	9
23	The PHIST Subfamilies.....	10
24	PHISTa .....	10
25	PHISTb .....	11
26	PHISTc .....	13
27	Other PHISTs .....	14
28	PHIST Gene Expression.....	14
29	PF3D7_0532400 (LyMP) .....	16
30	PF3D7_0936800 .....	17
31	PF3D7_0731100 (PfPTP2).....	18
32	PF3D7_0402000 .....	19
33	PF3D7_0936600 (GEXP5) .....	19
34	PF3D7_0102200 (RESA).....	20
35	PVX_093680 (CVC-81 <sub>95</sub> ).....	22
36	<i>phist</i> Gene Knockout Study .....	23
37	Structure and Function.....	23
38	CONCLUSION.....	24

## 39 SUMMARY

40 During the asexual cycle, *P. falciparum* extensively remodels the human erythrocyte  
41 to make it a suitable host cell. A large number of exported proteins facilitate this  
42 remodeling process which causes erythrocytes to become more rigid, cytoadherent,  
43 and permeable for nutrients and metabolic products. Among the exported proteins a  
44 family of 89 proteins, called *Plasmodium* helical interspersed subtelomeric protein  
45 family (PHIST) has been identified. While also found in other *Plasmodium* species,  
46 the PHIST family is greatly expanded in *P. falciparum*. Although a decade has  
47 passed since their first description, to date most PHIST proteins remain  
48 uncharacterized, are of unknown function and localization within the host cell, and  
49 little data exist on their interactions with other host or parasite proteins. However,  
50 over the past few years, PHIST proteins have been mentioned in the literature at an  
51 increasing rate owing to their presence at various localizations within the infected  
52 erythrocyte. Expression of PHIST proteins has been implicated in molecular and  
53 cellular processes such as surface display of PfEMP1, gametocytogenesis, changes  
54 in cell rigidity and also in cerebral and pregnancy-associated malaria. Thus, we  
55 conclude that PHIST proteins are central to host cell remodeling but despite their  
56 obvious importance in pathology PHISTs seem to be understudied. Here we review  
57 current knowledge, shed light on the definition of PHISTs, and discuss these proteins  
58 with respect to their localization and probable function. We take into consideration  
59 interaction studies, microarray analyses, or data from natural infected patient blood  
60 samples to combine all available information on this protein family.

## 61 INTRODUCTION

### 62 Malaria

63 Malaria is an infectious disease caused by the protozoan parasite *Plasmodium* and is  
64 transmitted by the female *Anopheles* mosquito to humans during its blood meal. Of  
65 the five *Plasmodium* species causing human malaria, two are of major public health  
66 interest: *P. falciparum* causes the most severe form of malaria whilst *P. vivax* is the  
67 most wide-spread *Plasmodium* species (1, 2). The success of *P. vivax* is due to the  
68 presence of undetectable hypnozoites which represent dormant stages in the liver  
69 and pose a huge problem for malaria control. With half the human population at risk,  
70 roughly 200 million malaria cases, and estimated 438'000 deaths in 2015, malaria  
71 still remains a major threat to public health (3).

72 *Plasmodium* has a complex life cycle alternating between the arthropod vector and  
73 its vertebrate host (Figure 1) (1). During the bite of an infected female *Anopheles*  
74 mosquito, sporozoites are injected into dermal tissue and transported through the  
75 bloodstream to the liver. Here, the sporozoites penetrate and invade hepatocytes  
76 which is followed by several rounds of asexual replication. Subsequently, thousands  
77 of merozoites are released into the bloodstream through so-called merozoites,  
78 starting the blood stage cycle (4, 5).

79 Once in the blood stream, merozoites invade erythrocytes and multiply during an  
80 approximate 48 hour intraerythrocytic development cycle. At the end of the cycle  
81 around 16 to 32 merozoites are released and re-invade new erythrocytes, starting the  
82 cycle anew. Once a merozoite has invaded the erythrocyte it extensively refurbishes  
83 and remodels the host cell. This remodeling of the cell assures the parasite ion  
84 homeostasis within the cell, allows nutrient up-take and is accompanied by changes  
85 in host cell membrane structure and rigidity. All this is achieved by the export of a  
86 large number of proteins from the parasite into the host cell leading to dramatic

87 changes of the infected erythrocytes. Since such a remodeled erythrocyte would be  
88 eliminated in the spleen, the parasite also exports the major virulence factor PfEMP1  
89 that conveys binding of the infected cells to endothelial receptors such as CD36,  
90 ICAM, or EPCR (6, 7). PfEMP1 is displayed on the surface of the infected red blood  
91 cell (iRBC) and is considered to be the key factor for morbidity and mortality. Overall,  
92 these refurbishing processes are considered to be responsible for most symptoms  
93 and the pathology of malaria (1, 8). Hence, recently, research focusing on exported  
94 proteins and processes involved in this host cell remodeling has attracted increased  
95 attention.

96

## 97 **Protein Export**

98 Because exported proteins of *P. falciparum* play such a central role in remodeling the  
99 infected red blood cells, the description of the pentameric amino acid motif,  
100 RxLxE/Q/D, called *Plasmodium* Export Element (PEXEL) (9, 10) was a breakthrough  
101 allowing the prediction of approximately 400 exported proteins. The PEXEL motif was  
102 recently expanded into a more relaxed PEXEL motif (RxLxxE) (11) or a non-  
103 canonical motif (K/HxL/lxE/Q/D) (12), resulting in a total number of more than 460  
104 proteins predicted to be exported in *P. falciparum*. This 'exportome' provides the  
105 basis to study exported proteins and their involvement in pathology of the remodeled  
106 infected erythrocyte much more specifically. Besides the PEXEL defined exportome  
107 there are additional exported proteins that lack a known or discernible export  
108 element, referred to as PEXEL negative proteins (PNEPs) (13, 14). Although the  
109 definitive number of proteins exported into the erythrocyte is unknown, it is without  
110 doubt that they play important roles in pathology through the significant changes they  
111 induce in the infected erythrocyte and that they are major contributors to disease.  
112 Figure 2 shows a schematic representation of differences in cytoskeleton

organization between the uninfected and the infected erythrocyte and the localization of exported proteins during the asexual replicative blood stage of *P. falciparum*.

## **PHIST PROTEINS**

### **Methodology for Review**

We have focused on *P. falciparum* and included all PHIST proteins and genes that were either identified by Sargeant and colleagues (15) or were later identified as PHISTs by Frech and Chen (16). We did not attempt to identify new PHISTs, but only compiled already published data. Gene IDs in PlasmoDB have changed over time, therefore some PHISTs have two or even three gene IDs and various publications use different gene IDs although referring to the same protein or gene. We scanned also supplementary data for information on PHISTs which in some instances was only found there and would not be available through standard literature search. We compiled any available information on each member of the PHIST family by using all gene IDs or names used now or previously. The complete list of PHIST proteins or genes found in the literature and reviewed here is shown in Table 1 and represents an in-depth description of this important protein family.

### **The PHIST Protein Family**

Within the large number of exported proteins Sargeant and colleagues (15) identified a new protein family which they termed *Plasmodium* helical interspersed sub-telomeric (PHIST) protein family. This protein family is characterized by a conserved domain of approximately 150 amino acids predicted to form four consecutive alpha helices (determined by Fugue (17)). Some members of this family comprise little more than an export signal sequence, PEXEL motif and the PHIST domain, whereas other members are substantially elongated and include additional domains, such as a

DnaJ domain (Figure 3). Based on the presence and position of several conserved tryptophan residues within the PHIST domain, the PHIST protein family has been further divided into three subgroups: PHISTa, PHISTb, and PHISTc (15). With more recent data included, PlasmoDB has annotated additional PHIST domains in proteins resulting in a total of 89 currently known PHIST proteins in *P. falciparum* (15, 16). Some of the newly added PHIST proteins were not grouped into the already existing subgroups, but were classified as PHISTa-like or simply as PHIST proteins. Thus, PlasmoDB annotations give the impression that there are more than three PHIST subgroups. However, a phylogenetic tree based on a multiple sequence alignment shows that PHISTa-like and PHIST proteins cluster with the PHISTa subgroup and the PHISTb-DnaJ proteins group among the PHISTb subgroup, giving rise to three distinct PHIST subgroups (Figure 3).

When the protein family was first described, Sargeant and colleagues used the presence and position of conserved tryptophan residues in the amino acid sequence of the PHIST domain to distinguish between the three subgroups (15). Comparison of multiple sequence alignments for each subgroup in which PHISTb-DnaJ and PHISTa-like/PHIST proteins were treated as individual subgroups, reveals a unique pattern of the conserved tryptophan residues for each of the subgroups (Figures 4 and 5A). For each subgroup a different positional pattern of conserved tryptophan residues was found, with PHISTa and PHISTa-like/PHIST proteins possessing only two conserved tryptophan residues within the PHIST domain whilst the remaining subgroups have four. There is only a slight variation in the position of these residues between the PHISTb and PHISTb-DnaJ subgroups, and between the PHISTa and PHISTa-like/PHIST subgroups, respectively. We therefore will use the original three subgroups introduced by Sargeant and colleagues, unless otherwise stated.

Table 1 lists all 89 proteins which we identified as PHIST proteins in *P. falciparum*, of which 64 contain a classical PEXEL motif and are thus predicted to be exported (11, 15). Although PlasmoDB lists 19 *phist* genes as pseudogenes in the reference strain 3D7, some have been found present as proteins or transcripts in other studies. Thus, PHISTs represent a substantial group of exported proteins comprising approximately 14% of all PEXEL proteins or nearly 2% of the complete *P. falciparum* proteome. Despite their potentially important role in host cell remodeling and (in)direct involvement in pathogenicity, most PHIST proteins remain completely uncharacterized, rendering it difficult to assign specific functions or roles to the PHIST protein family and/or the three subgroups.

### **Comparison of PHIST and PRESAN Domains**

There has been confusion in the literature on the definition of PHIST and PRESAN domains which in fact are virtually identical. The confusion was generated when, independent to the sequence analysis published by Sargeant and colleagues (15), transcriptome analysis of *P. falciparum* parasites grown under elevated temperatures mimicking febrile conditions, revealed a number of up-regulated genes that coded for proteins containing a DnaJ domain (Pfam ID: PF00226) (18). Subsequent alignments identified an extended protein family with at least 67 members, all sharing a particular N-terminal domain with a DnaJ domain only present in some members. Because some of the DnaJ domain-containing proteins showed similarity to the *P. falciparum* ring-infected erythrocyte antigen (RESA), this N-terminal domain was termed PRESAN (*Plasmodium* RESA N-terminal) (Pfam ID: PF09687). Examination of the sequence alignments by Sargeant and colleagues (15), and Oakley and colleagues (18) reveals that the domain boundaries of PHIST and PRESAN are virtually identical. Secondary structure predictions by Oakley and colleagues (18) suggested



the presence of six  $\alpha$ -helices in PRESAN domains, however these overlapped with the predicted four  $\alpha$ -helices of the PHIST domain (15). Thus, the PRESAN domain can be regarded as highly similar to the PHIST domain with differences mainly depending on the prediction algorithms. An example for such close similarity is shown with the PHIST / PRESAN protein PF3D7\_0532400 (Figure 5B). The overall domain structure of all expressed PHIST proteins is shown in Figure 3. Crystallographic and NMR studies provided clear evidence that the PHIST / PRESAN domain forms a four helical bundle (19). Subsequently, Tarr and colleagues (20) defined an 'extended PRESAN' domain N-terminal to the original PHIST domain and which was thought to comprise a domain targeting membranes. This extended domain includes additional helix-forming sequences.

Confusingly, the current Interpro database uses the term 'PHIST domain' (IPR006526) to refer to a small protein fragment spanning approximately the N-terminal half of the PHIST / PRESAN domains as defined by the original authors (15, 18). In contrast, the Interpro (IPR019111) and Pfam (PF09687) entries for the PRESAN domain are identical to those original alignments. This discrepancy needs urgently correction. In this review, we will refer to this conserved domain type as PHIST.

### **PHISTs in Other *Plasmodium* Species**

PHIST proteins are exclusively found in the genus *Plasmodia* and the protein family has been significantly expanded in *P. falciparum* and in the Laveranian species *P. reichenowi*, in which there is a one to one representation of the PHIST genes (21). While little is known about PHIST proteins in *P. falciparum*, even less is known on PHIST proteins in other *Plasmodium* species. There are fewer PHIST proteins in other *Plasmodia* but their exact number is not yet clear and different publications

report varying counts. Initially, Sargeant et al. reported 39 PHIST proteins for *P. vivax* (15) but the complete analysis of the *P. vivax* genome (22) revealed the presence of a gene family (Pv\_fam\_e) which contained 44 *rad* genes and 21 *phist* genes, both groups showing structural similarities. However, PlasmoDB currently has only 18 genes annotated as *phist* (PlasmoDB) for *P. vivax*. Supplementary materials of two publications provide expression data for *P. vivax* PHIST proteins (23, 24). For *P. knowlesi* Sargeant and colleagues (15) initially reported 27 PHIST proteins but this has been re-annotated to 38 proteins by Pain et al. (25), and PlasmoDB currently lists 39 records. For the monkey malaria parasite *P. cynomolgi* 21 PHIST genes were identified, whilst the number of PHIST genes in rodent malaria is unclear and varies in different publications or databases. For *P. berghei*, 1 to 3 *phist* genes have been found and 1 to 2 in both *P. chabaudi* and *P. yoelii* (15, 16, 26). Moreira et al. recently characterized the two PHIST proteins in *P. berghei* and demonstrated potentially similar roles for these PHIST as has been attributed to *P. falciparum* PHISTs (26). So far, to our knowledge, no *phist* genes have been identified in the avian *Plasmodium* lineage. Whether other gene families with similar structure or function exist in those species, remains to be seen.

Except for CVC-81<sub>95</sub>, a PHIST protein of *P. vivax* that has been investigated in more detail, we will not further discuss PHIST proteins in other *Plasmodia*.

## **The PHIST Subfamilies**

### **PHISTa**

Proteins of the PHISTa subgroup are very short, and besides the PHIST domain consist of only a signal sequence and a PEXEL motif if present. PHISTa proteins are exclusively found in *P. falciparum* (15) and currently amount to 26 different proteins (Table 1a). In contrast to the other subgroups, PHISTa and PHISTa-like proteins

possess only two conserved tryptophan residues (see Figure 4). These proteins were previously already grouped together and described as a subtelomeric protein superfamily (27) although two of them, PF3D7\_1301500 and PF3D7\_1301300, are not located subtelomerically and are referred to as PHISTa-like proteins today (Table 1e).

Except for PF3D7\_0402000 and PF3D7\_1253300 PHISTa proteins have been reported to be transcriptionally silent in the reference strain 3D7 (15, 28). Initially, it was proposed that this transcriptional silencing might be caused by mutually exclusive expression (15). However, as shown in the heat map (Figure 5) several studies recently showed that other PHISTa proteins are up-regulated in pregnancy-associated malaria or in cerebral malaria (29-34). Additional studies reported members of the PHISTa group to be differentially expressed in parasites committed to become gametocytes (27, 35-37). Whether the transcriptional silencing of PHISTa proteins in 3D7 is an adaptation to *in vitro* growth remains to be tested but PHISTa proteins in natural infections seem to play an important role in pathogenesis.

### **PHISTb**

With 24 members, PHISTb proteins make the second largest subgroup (Table 1b). Members of this subgroup are slightly longer than PHISTa proteins with 300-600 residues. Characteristic for the PHISTb subgroup is a unique, long C-terminal amino acid stretch that follows the PHIST domain (15) and which might indicate a unique and different function of these proteins as compared to proteins of the other PHIST subgroups. It is conceivable that the PHIST domain provides a general binding motif whilst the C-terminus might provide a more specific interaction domain. Such dual binding capacity was recently shown by Oberli and colleagues (38) for the PHIST protein PF3D7\_0532400. C-terminal interaction motifs have indeed been identified in

several other PHISTb proteins (39). In this respect it is important to note that all PHISTb proteins characterized today were localized at and might interact with the host cell cytoskeleton (19, 20, 39, 40).

The non-PHIST protein MESA (PF3D7\_0500800) interacts via an N-terminal 19 amino acid motif with band 4.1R. This motif has been termed MESA erythrocyte cytoskeleton-binding motif (MEC) (39) and has been found in 14 other exported *P. falciparum* proteins of which 9 belong to the PHISTb subfamily (Table 1b). Each of these 9 PHIST proteins has a MEC motif C-terminal downstream of the PHIST domain and in two cases is followed by a DnaJ domain. Three of these PHIST proteins also have been found to bind to inside-out vesicles (IOVs) suggesting an interaction between the MEC motif and a cytoskeleton interaction partner. Co-precipitation with the MEC motif of the PHISTb-DnaJ protein PF3D7\_1038800 revealed band 4.1R as interaction partner (39). Although these PHIST proteins were not further functionally characterized, the presence of a MEC motif with its binding capacity to band 4.1R suggests their involvement in remodeling of the iRBC cytoskeleton and thus a possible contribution to pathology of malaria.

Seven PHISTb members (Table 1c) including the well-known ring infected erythrocyte surface antigen (RESA) also comprise a DnaJ domain, referred to as PHISTb-DnaJ. Proteins with a DnaJ domain belong to the Hsp40 family. The J domains can act as co-chaperones for proteins of the DnaK/Hsp70 families, and can associate with unfolded polypeptide chains to prevent their aggregation (18, 41). PHISTb and PHISTb-DnaJ differ only in one of the four conserved tryptophan positions (Figures 4 and 5A) indicating their close relatedness. Another characteristic that sets the PHISTb-DnaJ proteins apart from the PHISTb subgroup is the extended length ranging from approximately 800-1100 amino acids (15).

Several PHISTb proteins have been shown to localize at the host cell periphery and solubility assays with GFP-tagged PHISTb proteins suggested interaction with components of the host cytoskeleton (11). The authors also investigated sequence requirements for peripheral localization and showed that the PHIST domain or the region N-terminal of the PHIST domain alone is not sufficient for peripheral localization. The PHIST domain together with parts of the N-terminal region however conferred peripheral localization (11). This was shown for the PHISTb protein PF3D7\_0401800 and the PHISTb-DnaJ protein PF3D7\_0102200 and it remains to be confirmed whether this applies to all PHISTb proteins and what precise sequence or structural requirements are necessary. It also needs to be shown whether PHISTs of the other subfamilies have similar requirements for correct localization. Although localization does not predict function, these observations in general confirm that PHISTb proteins tend to associate with the iRBC cytoskeleton.

### **PHISTc**

Most information is available for the PHISTc subgroup which is entirely shared with *P. vivax* and *P. knowlesi* (15) indicating that the expansion of this subgroup occurred before the lineage diverged. The PHISTc subgroup is also the most diverse group in length varying from less than 200 to over 1200 amino acids. In most of the 18 PHISTc proteins (Table 1d) the PHIST domain is found very near the C-terminus of the protein (Figure 3), similar to the PHISTa subgroup. In contrast to PHISTb proteins which mostly are associated with the iRBC cytoskeleton, several PHISTc proteins have been found at structures such as Maurer's clefts (19, 42) and exosome-like vesicles (43), and are thought to be involved in protein trafficking (44, 45). There is recent evidence for PF3D7\_0936800 (PFI1780w) also to be localized at the host cell membrane (19).

319

## 320 **Other PHISTs**

321 PlasmoDB lists 14 additional proteins that are annotated as PHISTa-like or simply as  
322 PHIST proteins but originally were not included in the PHIST family (Table 1e). Some  
323 of the PHISTa-like proteins were included in the subtelomeric protein superfamily  
324 identified by Eksi et al. (27) of which most have now been included in the PHISTa  
325 subgroup. There seems to be a close relationship between the PHISTa and the  
326 PHISTa-like proteins indicated by sequence alignment and the pattern of conserved  
327 tryptophan residues (Figure 4 and 5A).

328

## 329 **PHIST Gene Expression**

330 Transcriptome analysis of the *P. falciparum* 3D7 strain showed that most *phist* genes  
331 were expressed at an early stage during the intraerythrocytic development cycle.  
332 Throughout the second half of the cycle almost all *phist* genes were switched off,  
333 while some were up-regulated again towards the very end of the cycle (Figure 6)  
334 (46). Genes with a similar expression pattern were described to be involved in  
335 parasite specific processes such as host cell invasion (47). In *Plasmodium*, protein  
336 presence is often delayed after transcription and the appearance of PHIST proteins  
337 was approximately delayed on average by 11 hours (48). Thus, most PHIST proteins  
338 are present and exported during the first half of the intraerythrocytic development  
339 cycle, again strongly suggesting an important role of PHIST proteins in host cell  
340 remodeling.

341 Some *phist* genes have been reported to be differentially expressed for example  
342 during gametocytogenesis (35), in pregnancy-associated malaria (29, 49) or in  
343 cerebral malaria (32). Variable expression generally has been also observed in field  
344 isolates (50).

Figure 6 presents expression data obtained only from asexual blood stage parasites but a proteomic study revealed a number of PHIST proteins to be enriched in early gametocytes, the sexual blood stage of *P. falciparum*. Of 26 putatively exported proteins enriched during early gametocyte stages, 9 belong to the PHIST protein family (35), indicated by a yellow box in Figure 6. Two of these, namely GEXP5/PF3D7\_0936600 and PfPTP2/PF3D7\_0731100 have been shown by microscopy to localize to gametocytes, the former having been shown to be exclusively expressed in gametocytes (51). A function during sexual development remains to be shown (43, 51). The presence of some PHIST proteins in gametocytes suggests a transcriptional up-regulation, and indeed, the expression of about 20 *phist* genes including GEXP5/PF3D7\_0936600 was shown to be under PfHP1 regulation (52). HP1 is a negative regulator of AP2-G, a transcription factor needed for sexual conversion which binds to a promoter motif common to several early gametocyte genes including the *phist* gene PF3D7\_1477300 (53, 54). This provides strong evidence that gene regulation by HP1 affects the expression of *phist* genes and their involvement in gametocytogenesis.

Almelli et al. (32) compared transcriptional differences between samples from patients with cerebral malaria and samples from asymptomatic malaria cases to the 3D7 reference strain. A number of *phist* genes from all subfamilies were either differentially up- or down-regulated (Figure 6, Table 1).

Mackinnon et al. (50) compared gene expression of the *P. falciparum* 3D7 strain with field isolates and also showed that a number of genes were differentially expressed. Among the 20 most regulated genes were also 7 *phist* genes (PF3D7\_0202100, PF3D7\_0424000, PF3D7\_0702100, PF3D7\_0832200.1, PF3D7\_0936600, PF3D7\_0936800, PF3D7\_1477700) suggesting that some PHIST proteins might be dispensable in culture but not for *in vivo* growth.

It also has been shown that 14 *phist* genes were variably expressed in *P. falciparum* 3D7 (Figure 6) (55) indicating an active role in adaptation to changing environments such as heat shock (febrile illness) or nutrient depletion.

All of these studies were performed on transcriptome or proteome level and reported general patterns and trends in gene expression or protein abundance. These studies repeatedly showed *phist* genes or proteins to be involved in different processes and support the notion that PHIST proteins play a central role in host cell remodeling. However, the functional and physical characterization of these proteins lags far behind. So far, we know that PHIST proteins are involved in cellular processes in the iRBC and in disease associated functions, but our understanding of their actual function and interactions is very limited, at least for the large majority of them.

Below we review and discuss individual PHIST proteins for which more detailed information is available, most of them belonging to the PHISTb or PHISTc subgroup. Further information on other PHIST proteins is summarized in Table 1.

#### **PF3D7\_0532400 (LyMP)**

The lysine-rich membrane-associated PHISTb protein, LyMP (PF3D7\_0532400 or PFE1605w) is a PEXEL containing PHISTb protein that is exported during the first half of the intraerythrocytic development cycle to the erythrocyte membrane where it can localize at parasite induced protrusions on the red blood cell membrane called knobs. Its transient localization at Maurer's clefts prior to its final destination correlates in space and time with that of PfEMP1 (19). The PHIST domain of LyMP (amino acids 122-335) interacts with the intracellular acidic terminal segment (ATS) of PfEMP1. Conditional down-regulation of LyMP reduced binding to CD36 by >60% indicating that the interaction between the PHIST domain of LyMP and the ATS domain of PfEMP1 is important for the cytoadhesive properties of iRBCs (19, 38, 40,



56). Similar conditional down-regulation of LyMP in iRBCs preselected for binding to different adhesion receptors displaying different PfEMP1 variants on the surface strongly differed in the reduction of cytoadherence (38) which led to the hypothesis that different PHISTs might be responsible to anchor different PfEMP1 variants to the cytoskeleton.

It was further shown that the C-terminal segment of LyMP (amino acids 319-528) was able to bind inside-out vesicles (IOVs), which are used to study protein interactions involving cytoskeletal proteins (40). We recently were able to show that this part of LyMP binds directly to the human membrane protein band 3 which is linked to the cytoskeleton via ankyrin (38). Together with the ATS interaction mediated by the LyMP PHIST domain (19), it is evident that LyMP can act as a linker between the virulence complex of *P. falciparum* PfEMP1 and the host cytoskeleton.

#### **PF3D7\_0936800**

PF3D7\_0936800 (PFI1780w) is a PHISTc protein that was also shown to interact with the ATS domain of PfEMP1 albeit much weaker than LyMP (56). A crystallographic structure of its PHIST domain (residues 85-247) has been obtained and is the first available structure for any PHIST protein. It confirms the predicted four alpha helical structure of the PHIST domain with a very short first alpha helix. The remaining three helices form a three helix bundle with weak structural similarity to spectrin (19, 56).

PF3D7\_0936800 has been classified as non-canonical PEXEL protein with the first position of its PEXEL motif rendered K to R, which recently has been shown to be correctly cleaved and N-acetylated, confirming that this PHIST is correctly exported from the parasite into the iRBC (12).

PF3D7\_0936800 localizes underneath the iRBC membrane but shows absence from knobs (19). It is speculated that it contains an interaction epitope for binding of the iRBC membrane/cytoskeleton components (12). Here it is also noteworthy that PF3D7\_0936800 has been found to be variantly expressed in 3D7 (55) and that no published PF3D7\_0936800 knockout parasite exists.

#### **PF3D7\_0731100 (PfPTP2)**

The PHISTc protein PF3D7\_0731100 (PfPTP2 or PfEMP1 trafficking protein 2) seems to play a role in PfEMP1 trafficking or surface display. Depletion of this protein leads to PfEMP1 accumulation in Maurer's clefts and its absence from knobs. Under flow conditions, PfPTP2-KO parasites do not adhere to CSA suggesting that the main role of PfPTP2 might be trafficking of PfEMP1 from Maurer's clefts to knobs (42, 44). Recently, PfPTP2 was also described to be located at exosome-like vesicles of approximately 70 nm that are involved in cell-cell communication and allow nucleic acid transfer between *Plasmodium* parasites. PfPTP2-deficient parasites showed reduced cell-cell communication and Regev-Rudzki et al. (43) suggested that PfPTP2 is essential for cell-cell communication between *P. falciparum* infected erythrocytes leading to an increase of gametocytogenesis *in vitro*. PfPTP2 was the only protein involved in knob formation or PfEMP1 display (besides SBP1 (57)) but is also found to be enriched in early gametocytes (35, 43). In contrast, transcriptional analysis shows PfPTP2 down-regulated in gametocytes, thus PfPTP2 abundance in early gametocytes might be acquired through exosome-like vesicles and this might in turn induce sexual conversion. Such hypothetical model would accommodate both the fact that PfPTP2 is up-regulated in asexual stages but also enriched in gametocytes.

Immuno-electron microscopy and Proteinase K assays revealed localization of PfPTP2 on the cytosolic face of Maurer's clefts, on budding vesicles, and on the surface of exosome-like vesicles (43, 44).

#### **PF3D7\_0402000**

The PHISTa protein PF3D7\_0402000 was identified in a yeast-2-hybrid assay to determine potential interaction partners of the human erythrocyte cytoskeleton protein band 4.1R. The protein interacted with the N- and alpha lobe of the FERM domain (Pfam ID: PF00373) of band 4.1R through helices two and three of the PHIST domain but the entire domain was required for maximal interaction (58). The host cytoskeleton protein band 4.1R binds to actin filaments and links these to membrane proteins such as band 3 and Glycophorin C, maintaining the biconcave shape (reviewed in (58, 59)). In this context, it is noteworthy that other *P. falciparum* proteins, e.g. MESA (PF3D7\_0500800) which is not a PHIST protein, bind to this host cell cytoskeleton interaction hub (60). In immunofluorescence imaging, subpopulations of both band 4.1R and PF3D7\_0402000 were shown to co-localize at the parasitophorous vacuole membrane (58). It remains unclear what role PF3D7\_0402000 plays in changes of membrane rigidity or host actin recruitment.

#### **PF3D7\_0936600 (GEXP5)**

GEXP5 (Gametocyte Exported Protein 5) was originally classified as PHISTb (15) but is now grouped amongst the PHISTc proteins (16) which is confirmed by the position of its conserved tryptophan residues (see Figure 4). It was the first PHIST protein detected in gametocytes. GEXP5-GFP fusion protein, when episomally expressed under the control of the endogenous promoter in asexual stages, was not detected by fluorescence imaging (51) but could be detected when expressed under the *pfcam*

promoter and confirmed the cytosolic localization (20). This indicated that endogenous GEXP5 is only expressed and exported in sexual stages and is found in the iRBC cytosol of stage I-IV gametocytes. Its uniform distribution throughout the iRBC and its presence in the soluble protein fraction suggest it to be a soluble protein (51). GEXP5 is already found 14 hours post-invasion in ring stage parasites committed to gametocytogenesis and is now considered to be the earliest gametocyte marker (35). This early appearance of GEXP5 in committed parasites might explain the presence of transcripts in the expression data sets in PlasmoDB. The expression and export of GEXP5 is independent of transcription factors normally associated with gametocytogenesis such as PfAP2-G, but GEXP5 is not able to drive gametocyte maturation alone in the absence of these transcription factors (51). It is therefore assumed that GEXP5 is not involved in processes driving gametocytogenesis but is used to remodel or prepare the host cell to accommodate sexual development.

#### **PF3D7\_0102200 (RESA)**

RESA (ring-infected erythrocyte surface antigen) contains both a PHIST and a DnaJ domain and is one of seven members of the PHISTb-DnaJ subgroup. Its name is a misnomer since RESA is an intracellular exported protein (20). RESA is expressed in the mature stage parasite (61) and is stored in the dense granules of newly formed merozoites (62). Within minutes after invasion of a merozoite into an erythrocyte the contents of the dense granules, including RESA, are discharged into the parasitophorous vacuole (63-65). From here, RESA is then exported into the iRBC and localizes at the iRBC cytoskeleton where it remains for approximately 24 hours (64, 65). At the cytoskeleton RESA is phosphorylated (66, 67) and interacts with spectrin (68). A fragment of 108 amino acids has been

identified to interact with repeat 16 of the  $\beta$ -chain of spectrin (here it is important to note that these residues 663-770 correspond only to a partial DnaJ X-domain and not to the DnaJ J-domain). As a result of this interaction, the spectrin tetramers are stabilized and rigidify the iRBC (69), however it needs to be proven whether this interaction and the accompanying rigidification would also occur with the full length domain or full length RESA protein.

There is controversy whether this interaction impairs invasion of merozoites since one study showed with erythrocytes pre-packed with a recombinant form of RESA that the RESA-spectrin interaction had no significant effect on invasion (70). The recombinant RESA used in this study contained residues 322-1073, lacking the entire PHIST domain (residues 174-294) (70) (see Figure 3). However, using a similar set up, Pei et al. showed that an even shorter fragment of RESA (residues 663-770) was able to reduce invasion efficiency (69). Importantly, this recombinant peptide lacked both the PHIST domain and the DnaJ domain and hence might have created an artificial inhibitory interaction.

A domain of 70 amino acids found in RESA shares 39% homology with the DnaJ heat shock protein of *E. coli* (41). Proteins with a DnaJ domain belong to the Hsp40 family and act as co-chaperones to the DnaK/Hsp70 chaperones which are involved in protein assembly and trafficking (18, 41). Proteins with a DnaJ domain have been described to be membrane-associated (41), a feature shared by RESA. Initially it was proposed that RESA destabilizes the RBC cytoskeleton (64) but now there is evidence that RESA does in fact the contrary, namely stabilizing the cytoskeleton with the effect that it rigidifies ring stage infected parasites (65, 69, 71). It has been shown that the RESA-spectrin interaction protects the iRBC from thermal damage (65, 69, 70), and it has been speculated that the DnaJ domain of RESA might prevent spectrin from unfolding under elevated temperatures or it might directly bind

to hydrophobic regions of partially unfolded spectrin molecules (70). This rigidification of ring stage infected erythrocytes, potentially impairing passage through the spleen (72) is rather surprising but this might be simply a side effect of the protective function of RESA against thermal damage to the host cell.

At the end of each 48 hour cycle, the iRBC ruptures and releases new merozoites. This event is accompanied by fever peaks that last for a few hours (73, 74). During these fever episodes, the parasites are in ring stages and this correlates with the time when RESA interacts with the cytoskeleton protecting it from thermal damage caused by fever. After approximately 24 hours, RESA dissociates from spectrin (61) and other proteins might be responsible for further rigidification of the cytoskeleton (71). This apparent functional contradiction might be an evolutionary trade-off between the need for survival during fever with costs of increased cell rigidity to a level that still allows passage through the spleen. The fact that less deformable ring-iRBCs are cleared by the spleen supports this idea of trade-off (72).

Whilst the function of the DnaJ domain in RESA has been described in much detail, the function of the PHIST domain still remains enigmatic.

#### **PVX\_093680 (CVC-81<sub>95</sub>)**

The only PHIST protein not from *P. falciparum* that has been further characterized is the *P. vivax* PHIST CVC-81<sub>95</sub> (PVX\_093680). Caveola-vesicle complexes (CVCs) are parasite induced indentations in the *P. vivax* infected erythrocyte cell membrane. The exact structure, protein composition, and function are not yet fully understood but CVC-81<sub>95</sub> is exported and a predominant protein in these CVCs and is found on tubular extension going inwards from the CVC (75, 76). Its orthologue in *P. cynomolgi* shows a similar localization (75). CVC-81<sub>95</sub> was found by Acharya et al. (77) in *P. vivax* infected peripheral blood and another study showed that over 80% of

the patient sera reacted positive to CVC81<sub>95</sub> (78). Based on these findings, CVC81<sub>95</sub> was suggested to be involved in immune evasion (78) as the PHIST protein family has been previously suggested to be involved in this process (24). Therefore, also in *P. vivax* PHIST proteins seem to be localized to parasite induced structures in the iRBC and might be essential for host-parasite interactions. However, more functional analyses are required to fully understand the role of this PHIST in the CVC of *P. vivax*.

### ***phist* Gene Knockout Study**

In a large gene knockout study in *P. falciparum* to functionally characterize exported proteins 17 *phist* genes were targeted for deletion (44). This included 1 *phista* gene, 7 *phistb* genes, 6 *phistb-dnaj* genes, and 3 *phistc* genes. Of these all but four (PF3D7\_0936800, PF3D7\_0401800, PF3D7\_0402100, PF3D7\_1149000) were successfully disrupted indicating that the majority was dispensable for *in vitro* growth. Except for PF3D7\_0731100 (PfPTP2), none of the *phist* knockout parasites from this study have been analyzed in greater detail. Some were reported to have an influence on altered knob morphology (PF3D7\_0424600-KO), reduced (PF3D7\_0102200/RESA-KO), or increased iRBC rigidity (PF3D7\_0220100-KO) (44). It would be urgently needed to have these parasite lines further phenotypically characterized to better understand the role of PHISTs in host cell remodeling.

### **Structure and Function**

It is highly conceivable that PHIST proteins might fulfill similar roles as chaperones since a number of chaperones contain also DnaJ domains. They may also be responsible for close association to the host cytoskeleton and cell membrane due to the intrinsic positive charge of many PHIST proteins. The few studies on PHISTs

have mostly looked at the PHIST domain only; however, Tarr et al. (20) suggested that the PHIST domain alone is not sufficient to correctly target the protein to its destination, at least for those PHIST proteins investigated. Mayer et al. (56) analyzed the binding affinities of various PHIST domains to a number of ATS domains but only the PHIST protein PF3D7\_0532400 has yet been studied in greater detail with respect to structure and function (19, 56). This is rather surprising for a protein family comprising a large proportion of the complete exportome of *P. falciparum*.

## CONCLUSION

The major three PHIST subfamilies not only differ in sequence but they also seem to differ in function. PHISTb proteins mostly interact with and localize to the iRBC cytoskeleton and seem to be involved in changes of cell properties such as rigidity. PHISTc proteins seem to be mostly involved in protein trafficking of other exported proteins, whilst the PHISTa subfamily remains enigmatic and there is not yet sufficient information available to draw a conclusion on the function.

The presence of PHIST proteins at various localizations in the host cell makes them highly interesting candidates as interaction partners for other exported proteins and therefore it would be important to understand what determines the destination of PHIST proteins. The identification of export requirements for PHISTs would increase our understanding of the interaction network of exported proteins.

It seems that some PHIST proteins have at least two binding domains, the PHIST domain and a second one towards the C-terminus, e.g. the MEC motif, the spectrin binding site, the band 3 interaction epitope, or a DnaJ-domain. This would make PHIST proteins ideal molecules for multifunctional interactions at the iRBC cytoskeleton or in protein export with the PHIST domain serving as a general adaptor whilst the C-terminal domain could function as a highly specific binding site.



However, only few additional binding motifs have been identified in some PHIST proteins and more PHIST proteins remain to be studied.

Our knowledge on PHIST proteins is still very limited and mainly based on studies of the 3D7 laboratory strain which does not suffice to understand the full extent of PHIST functions for *in vivo* parasite survival and malaria-associated disease and pathology. PHIST proteins should be studied in various disease presentations from mild to severe malaria as well as in pregnancy-associated malaria. Furthermore, the role of PHISTs in gametocytogenesis and also in mosquito stages remains completely enigmatic. Since many PHIST proteins seem to be central for host cell remodeling, it is essential to understand their function in this crucial process.

#### **Funding Information**

JW and HPB were supported by the Swiss National Science Foundation Grant Number 31003A\_149297/1. IV was supported by a Wellcome Trust RCD fellowship (088497/Z/09/Z). The funders had no role in determining the content of the paper or in the decision to submit this work for publication.

#### **Acknowledgments**

We would like to thank Beatrice Schibler and Alexander Oberli for carefully reading and commenting on the manuscript. We would like to acknowledge Christoph Schmid and Anita Lerch for their help with the alignments and phylogenetic tree construction. JW conducted the literature review. JW, IV, and HPB wrote the manuscript. The authors claim no conflicting interest.

## References

1. **Boddey JA, Cowman AF.** 2013. Plasmodium nesting: remaking the erythrocyte from the inside out. *Annu Rev Microbiol* **67**:243-269.
2. **Tuteja R.** 2007. Malaria - an overview. *FEBS J* **274**:4670-4679.
3. **WHO.** 2015. World Malaria Report 2015. WHO,
4. **Amino R, Giovannini D, Thiberge S, Gueirard P, Boisson B, Dubremetz JF, Prevost MC, Ishino T, Yuda M, Menard R.** 2008. Host cell traversal is important for progression of the malaria parasite through the dermis to the liver. *Cell Host Microbe* **3**:88-96.
5. **Sturm A, Amino R, van de Sand C, Regen T, Retzlaff S, Rennenberg A, Krueger A, Pollok JM, Menard R, Heussler VT.** 2006. Manipulation of host hepatocytes by the malaria parasite for delivery into liver sinusoids. *Science* **313**:1287-1290.
6. **Gardner JP, Pinches RA, Roberts DJ, Newbold CI.** 1996. Variant antigens and endothelial receptor adhesion in *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **93**:3503-3508.
7. **Turner L, Lavstsen T, Berger SS, Wang CW, Petersen JE, Avril M, Brazier AJ, Freeth J, Jespersen JS, Nielsen MA, Magistrado P, Lusingu J, Smith JD, Higgins MK, Theander TG.** 2013. Severe malaria is associated with parasite binding to endothelial protein C receptor. *Nature* **498**:502-505.
8. **Bannister L, Mitchell G.** 2003. The ins, outs and roundabouts of malaria. *Trends Parasitol* **19**:209-213.
9. **Marti M, Good RT, Rug M, Knuepfer E, Cowman AF.** 2004. Targeting malaria virulence and remodeling proteins to the host erythrocyte. *Science* **306**:1930-1933.

- 651 10. **Hiller NL, Bhattacharjee S, van Ooij C, Liolios K, Harrison T, Lopez-**  
652 **Estrano C, Haldar K.** 2004. A host-targeting signal in virulence proteins  
653 reveals a secretome in malarial infection. *Science* **306**:1934-1937.
- 654 11. **Boddey JA, Carvalho TG, Hodder AN, Sargeant TJ, Sleebs BE, Marapana**  
655 **D, Lopaticki S, Nebl T, Cowman AF.** 2013. Role of plasmepsin V in export of  
656 diverse protein families from the *Plasmodium falciparum* exportome. *Traffic*  
657 **14**:532-550.
- 658 12. **Schulze J, Kwiatkowski M, Borner J, Schluter H, Bruchhaus I, Burmester**  
659 **T, Spielmann T, Pick C.** 2015. The *Plasmodium falciparum* exportome  
660 contains non-canonical PEXEL/HT proteins. *Mol Microbiol* **97**:301-314.
- 661 13. **Spielmann T, Gilberger TW.** 2010. Protein export in malaria parasites: do  
662 multiple export motifs add up to multiple export pathways? *Trends Parasitol*  
663 **26**:6-10.
- 664 14. **Heiber A, Kruse F, Pick C, Gruring C, Flemming S, Oberli A, Schoeler H,**  
665 **Retzlaff S, Mesen-Ramirez P, Hiss JA, Kadekoppala M, Hecht L, Holder**  
666 **AA, Gilberger TW, Spielmann T.** 2013. Identification of new PNEPs indicates  
667 a substantial non-PEXEL exportome and underpins common features in  
668 *Plasmodium falciparum* protein export. *PLoS Pathog* **9**:e1003546.
- 669 15. **Sargeant TJ, Marti M, Caler E, Carlton JM, Simpson K, Speed TP,**  
670 **Cowman AF.** 2006. Lineage-specific expansion of proteins exported to  
671 erythrocytes in malaria parasites. *Genome Biol* **7**:R12.
- 672 16. **Frech C, Chen N.** 2013. Variant surface antigens of malaria parasites:  
673 functional and evolutionary insights from comparative gene family  
674 classification and analysis. *BMC Genomics* **14**:427.

- 675 17. **Shi J, Blundell TL, Mizuguchi K.** 2001. FUGUE: sequence-structure  
676 homology recognition using environment-specific substitution tables and  
677 structure-dependent gap penalties. *J Mol Biol* **310**:243-257.
- 678 18. **Oakley MS, Kumar S, Anantharaman V, Zheng H, Mahajan B, Haynes JD,**  
679 **Moch JK, Fairhurst R, McCutchan TF, Aravind L.** 2007. Molecular factors  
680 and biochemical pathways induced by febrile temperature in intraerythrocytic  
681 *Plasmodium falciparum* parasites. *Infect Immun* **75**:2012-2025.
- 682 19. **Oberli A, Slater LM, Cutts E, Brand F, Mundwiler-Pachlatko E, Rusch S,**  
683 **Masik MF, Erat MC, Beck HP, Vakonakis I.** 2014. A *Plasmodium falciparum*  
684 PHIST protein binds the virulence factor PfEMP1 and comigrates to knobs on  
685 the host cell surface. *FASEB J* **28**:4420-4433.
- 686 20. **Tarr SJ, Moon RW, Hardege I, Osborne AR.** 2014. A conserved domain  
687 targets exported PHISTb family proteins to the periphery of *Plasmodium*  
688 infected erythrocytes. *Mol Biochem Parasitol* **196**:29-40.
- 689 21. **Otto TD, Rayner JC, Bohme U, Pain A, Spottiswoode N, Sanders M, Quail**  
690 **M, Ollomo B, Renaud F, Thomas AW, Prugnolle F, Conway DJ, Newbold**  
691 **C, Berriman M.** 2014. Genome sequencing of chimpanzee malaria parasites  
692 reveals possible pathways of adaptation to human hosts. *Nat Commun*  
693 **5**:4754.
- 694 22. **Carlton JM, Adams JH, Silva JC, Bidwell SL, Lorenzi H, Caler E, Crabtree**  
695 **J, Angiuoli SV, Merino EF, Amedeo P, Cheng Q, Coulson RM, Crabb BS,**  
696 **Del Portillo HA, Essien K, Feldblyum TV, Fernandez-Becerra C, Gilson**  
697 **PR, Gueye AH, Guo X, Kang'a S, Kooij TW, Korsinczky M, Meyer EV,**  
698 **Nene V, Paulsen I, White O, Ralph SA, Ren Q, Sargeant TJ, Salzberg SL,**  
699 **Stoeckert CJ, Sullivan SA, Yamamoto MM, Hoffman SL, Wortman JR,**  
700 **Gardner MJ, Galinski MR, Barnwell JW, Fraser-Liggett CM.** 2008.

701 Comparative genomics of the neglected human malaria parasite *Plasmodium*  
702 *vivax*. *Nature* **455**:757-763.

703 23. **Westenberger SJ, McClean CM, Chattopadhyay R, Dharia NV, Carlton**  
704 **JM, Barnwell JW, Collins WE, Hoffman SL, Zhou Y, Vinetz JM, Winzeler**  
705 **EA**. 2010. A systems-based analysis of *Plasmodium vivax* lifecycle  
706 transcription from human to mosquito. *PLoS Negl Trop Dis* **4**:e653.

707 24. **Bozdech Z, Mok S, Hu G, Imwong M, Jaidee A, Russell B, Ginsburg H,**  
708 **Nosten F, Day NP, White NJ, Carlton JM, Preiser PR**. 2008. The  
709 transcriptome of *Plasmodium vivax* reveals divergence and diversity of  
710 transcriptional regulation in malaria parasites. *Proc Natl Acad Sci U S A*  
711 **105**:16290-16295.

712 25. **Pain A, Bohme U, Berry AE, Mungall K, Finn RD, Jackson AP, Mourier T,**  
713 **Mistry J, Pasini EM, Aslett MA, Balasubrammaniam S, Borgwardt K,**  
714 **Brooks K, Carret C, Carver TJ, Cherevach I, Chillingworth T, Clark TG,**  
715 **Galinski MR, Hall N, Harper D, Harris D, Hauser H, Ivens A, Janssen CS,**  
716 **Keane T, Larke N, Lapp S, Marti M, Moule S, Meyer IM, Ormond D, Peters**  
717 **N, Sanders M, Sanders S, Sargeant TJ, Simmonds M, Smith F, Squares**  
718 **R, Thurston S, Tivey AR, Walker D, White B, Zuiderwijk E, Churcher C,**  
719 **Quail MA, Cowman AF, Turner CM, Rajandream MA, Kocken CH, et al.**  
720 2008. The genome of the simian and human malaria parasite *Plasmodium*  
721 *knowlesi*. *Nature* **455**:799-803.

722 26. **Moreira CK, Naissant B, Coppi A, Bennett BL, Aime E, Franke-Fayard B,**  
723 **Janse CJ, Coppens I, Sinnis P, Templeton TJ**. 2016. The *Plasmodium*  
724 PHIST and RESA-Like Protein Families of Human and Rodent Malaria  
725 Parasites. *PLoS One* **11**:e0152510.

- 726 27. **Eksi S, Haile Y, Furuya T, Ma L, Su X, Williamson KC.** 2005. Identification  
727 of a subtelomeric gene family expressed during the asexual-sexual stage  
728 transition in *Plasmodium falciparum*. *Mol Biochem Parasitol* **143**:90-99.
- 729 28. **Scholz M, Fraunholz MJ.** 2008. A computational model of gene expression  
730 reveals early transcriptional events at the subtelomeric regions of the malaria  
731 parasite, *Plasmodium falciparum*. *Genome Biol* **9**:R88.
- 732 29. **Tuikue Ndam N, Bischoff E, Proux C, Lavstsen T, Salanti A, Guitard J,**  
733 **Nielsen MA, Coppee JY, Gaye A, Theander T, David PH, Deloron P.** 2008.  
734 *Plasmodium falciparum* transcriptome analysis reveals pregnancy malaria  
735 associated gene expression. *PLoS One* **3**:e1855.
- 736 30. **Daily JP, Le Roch KG, Sarr O, Ndiaye D, Lukens A, Zhou YY, Ndir O,**  
737 **Mboup S, Sultan A, Winzeler EA, Wirth DF.** 2005. In vivo transcriptome of  
738 *Plasmodium falciparum* reveals overexpression of transcripts that encode  
739 surface proteins. *J Infect Dis* **191**:1196-1203.
- 740 31. **Claessens A, Adams Y, Ghumra A, Lindergard G, Buchan CC, Andisi C,**  
741 **Bull PC, Mok S, Gupta AP, Wang CW, Turner L, Arman M, Raza A,**  
742 **Bozdech Z, Rowe JA.** 2012. A subset of group A-like var genes encodes the  
743 malaria parasite ligands for binding to human brain endothelial cells. *Proc Natl*  
744 *Acad Sci U S A* **109**:E1772-1781.
- 745 32. **Almelli T, Nuel G, Bischoff E, Aubouy A, Elati M, Wang CW, Dillies MA,**  
746 **Coppee JY, Ayissi GN, Basco LK, Rogier C, Ndam NT, Deloron P, Tahar**  
747 **R.** 2014. Differences in gene transcriptomic pattern of *Plasmodium falciparum*  
748 in children with cerebral malaria and asymptomatic carriers. *PLoS One*  
749 **9**:e114401.
- 750 33. **Mok BW, Ribacke U, Winter G, Yip BH, Tan CS, Fernandez V, Chen Q,**  
751 **Nilsson P, Wahlgren M.** 2007. Comparative transcriptomal analysis of

isogenic *Plasmodium falciparum* clones of distinct antigenic and adhesive phenotypes. *Mol Biochem Parasitol* **151**:184-192.

34. **Bertin GI, Sabbagh A, Argy N, Salnot V, Ezinmegnon S, Agbota G, Ladipo Y, Alao JM, Sagbo G, Guillonneau F, Deloron P.** 2016. Proteomic analysis of *Plasmodium falciparum* parasites from patients with cerebral and uncomplicated malaria. *Sci Rep* **6**:26773.

35. **Silvestrini F, Lasonder E, Olivieri A, Camarda G, van Schaijk B, Sanchez M, Younis Younis S, Sauerwein R, Alano P.** 2010. Protein export marks the early phase of gametocytogenesis of the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics* **9**:1437-1448.

36. **Joice R, Narasimhan V, Montgomery J, Sidhu AB, Oh K, Meyer E, Pierre-Louis W, Seydel K, Milner D, Williamson K, Wiegand R, Ndiaye D, Daily J, Wirth D, Taylor T, Huttenhower C, Marti M.** 2013. Inferring developmental stage composition from gene expression in human malaria. *PLoS Comput Biol* **9**:e1003392.

37. **Joice R, Nilsson SK, Montgomery J, Dankwa S, Egan E, Morahan B, Seydel KB, Bertuccini L, Alano P, Williamson KC, Duraisingh MT, Taylor TE, Milner DA, Marti M.** 2014. *Plasmodium falciparum* transmission stages accumulate in the human bone marrow. *Sci Transl Med* **6**:244re245.

38. **Oberli A, Zurbrugg L, Rusch S, Brand F, Butler ME, Day JL, Cutts EE, Lavstsen T, Vakonakis I, Beck HP.** 2016. *Plasmodium falciparum* PHIST Proteins Contribute to Cytoadherence and Anchor PfEMP1 to the Host Cell Cytoskeleton. *Cell Microbiol* doi:10.1111/cmi.12583.

39. **Kilili GK, LaCount DJ.** 2011. An erythrocyte cytoskeleton-binding motif in exported *Plasmodium falciparum* proteins. *Eukaryot Cell* **10**:1439-1447.

- 777 40. **Proellocks NI, Herrmann S, Buckingham DW, Hanssen E, Hodges EK,**  
778 **Elsworth B, Morahan BJ, Coppel RL, Cooke BM.** 2014. A lysine-rich  
779 membrane-associated PHISTb protein involved in alteration of the  
780 cytoadhesive properties of Plasmodium falciparum-infected red blood cells.  
781 FASEB J **28**:3103-3113.
- 782 41. **Bork P, Sander C, Valencia A, Bukau B.** 1992. A Module of the DnaJ Heat-  
783 Shock Proteins Found in Malaria Parasites. Trends Biochem Sci **17**:129-129.
- 784 42. **Maier AG, Cooke BM, Cowman AF, Tilley L.** 2009. Malaria parasite proteins  
785 that remodel the host erythrocyte. Nat Rev Microbiol **7**:341-354.
- 786 43. **Regev-Rudzki N, Wilson DW, Carvalho TG, Sisquella X, Coleman BM,**  
787 **Rug M, Bursac D, Angrisano F, Gee M, Hill AF, Baum J, Cowman AF.**  
788 2013. Cell-cell communication between malaria-infected red blood cells via  
789 exosome-like vesicles. Cell **153**:1120-1133.
- 790 44. **Maier AG, Rug M, O'Neill MT, Brown M, Chakravorty S, Szeszak T,**  
791 **Chesson J, Wu Y, Hughes K, Coppel RL, Newbold C, Beeson JG, Craig**  
792 **A, Crabb BS, Cowman AF.** 2008. Exported proteins required for virulence  
793 and rigidity of Plasmodium falciparum-infected human erythrocytes. Cell  
794 **134**:48-61.
- 795 45. **de Koning-Ward TF, Gilson PR, Boddey JA, Rug M, Smith BJ, Papenfuss**  
796 **AT, Sanders PR, Lundie RJ, Maier AG, Cowman AF, Crabb BS.** 2009. A  
797 newly discovered protein export machine in malaria parasites. Nature  
798 **459**:945-949.
- 799 46. **Llinas M, Bozdech Z, Wong ED, Adai AT, DeRisi JL.** 2006. Comparative  
800 whole genome transcriptome analysis of three Plasmodium falciparum strains.  
801 Nucleic Acids Res **34**:1166-1173.



- 802 47. **Bozdech Z, Llinás M, Pulliam BL, Wong ED, Zhu J, DeRisi JL.** 2003. The  
803 transcriptome of the intraerythrocytic developmental cycle of *Plasmodium*  
804 *falciparum*. *PLoS biology* **1**:e5.
- 805 48. **Foth BJ, Zhang N, Chael BK, Sze SK, Preiser PR, Bozdech Z.** 2011.  
806 Quantitative time-course profiling of parasite and host cell proteins in the  
807 human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics*  
808 **10**:M110 006411.
- 809 49. **Francis SE, Malkov VA, Oleinikov AV, Rossnagle E, Wendler JP,**  
810 **Mutabingwa TK, Fried M, Duffy PE.** 2007. Six genes are preferentially  
811 transcribed by the circulating and sequestered forms of *Plasmodium*  
812 *falciparum* parasites that infect pregnant women. *Infection and Immunity*  
813 **75**:4838-4850.
- 814 50. **Mackinnon MJ, Li J, Mok S, Kortok MM, Marsh K, Preiser PR, Bozdech Z.**  
815 2009. Comparative transcriptional and genomic analysis of *Plasmodium*  
816 *falciparum* field isolates. *PLoS Pathog* **5**:e1000644.
- 817 51. **Tiburcio M, Dixon MW, Looker O, Younis SY, Tilley L, Alano P.** 2015.  
818 Specific expression and export of the *Plasmodium falciparum* Gametocyte  
819 EXported Protein-5 marks the gametocyte ring stage. *Malar J* **14**:334.
- 820 52. **Flueck C, Bartfai R, Volz J, Niederwieser I, Salcedo-Amaya AM, Alako BT,**  
821 **Ehlgen F, Ralph SA, Cowman AF, Bozdech Z, Stunnenberg HG, Voss TS.**  
822 2009. *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci  
823 linked to phenotypic variation of exported virulence factors. *PLoS Pathog*  
824 **5**:e1000569.
- 825 53. **Brancucci NM, Bertschi NL, Zhu L, Niederwieser I, Chin WH, Wampfler R,**  
826 **Freymond C, Rottmann M, Felger I, Bozdech Z, Voss TS.** 2014.

- 827 Heterochromatin protein 1 secures survival and transmission of malaria  
828 parasites. *Cell Host Microbe* **16**:165-176.
- 829 54. **Goldowitz IS.** 2015. *Plasmodium's Crossroads: Deciphering the Molecular*  
830 *Pathway That Leads to Malaria Transmission.* Doctoral dissertation, Harvard  
831 University, Boston, USA.
- 832 55. **Rovira-Graells N, Gupta AP, Planet E, Crowley VM, Mok S, Ribas de**  
833 **Pouplana L, Preiser PR, Bozdech Z, Cortes A.** 2012. Transcriptional  
834 variation in the malaria parasite *Plasmodium falciparum*. *Genome Res* **22**:925-  
835 938.
- 836 56. **Mayer C, Slater L, Erat MC, Konrat R, Vakonakis I.** 2012. Structural  
837 analysis of the *Plasmodium falciparum* erythrocyte membrane protein 1  
838 (PfEMP1) intracellular domain reveals a conserved interaction epitope. *J Biol*  
839 *Chem* **287**:7182-7189.
- 840 57. **Dixon MW, Thompson J, Gardiner DL, Trenholme KR.** 2008. Sex in  
841 *Plasmodium*: a sign of commitment. *Trends in parasitology* **24**:168-175.
- 842 58. **Parish LA, Mai DW, Jones ML, Kitson EL, Rayner JC.** 2013. A member of  
843 the *Plasmodium falciparum* PHIST family binds to the erythrocyte cytoskeleton  
844 component band 4.1. *Malar J* **12**:160.
- 845 59. **Fowler VM.** 2013. The human erythrocyte plasma membrane: a Rosetta  
846 Stone for decoding membrane-cytoskeleton structure. *Curr Top Membr* **72**:39-  
847 88.
- 848 60. **Lustigman S, Anders RF, Brown GV, Coppel RL.** 1990. The mature-  
849 parasite-infected erythrocyte surface antigen (MESA) of *Plasmodium*  
850 *falciparum* associates with the erythrocyte membrane skeletal protein, band  
851 4.1. *Mol Biochem Parasitol* **38**:261-270.

- 852 61. **Brown GV, Culvenor JG, Crewther PE, Bianco AE, Coppel RL, Saint RB,**  
853 **Stahl HD, Kemp DJ, Anders RF.** 1985. Localization of the ring-infected  
854 erythrocyte surface antigen (RESA) of *Plasmodium falciparum* in merozoites  
855 and ring-infected erythrocytes. *J Exp Med* **162**:774-779.
- 856 62. **Aikawa M, Torii M, Sjolander A, Berzins K, Perlmann P, Miller LH.** 1990.  
857 Pf155/Resa Antigen Is Localized in Dense Granules of *Plasmodium-*  
858 *Falciparum* Merozoites. *Exp Parasitol* **71**:326-329.
- 859 63. **Riglar DT, Richard D, Wilson DW, Boyle MJ, Dekiwadia C, Turnbull L,**  
860 **Angrisano F, Marapana DS, Rogers KL, Whitchurch CB, Beeson JG,**  
861 **Cowman AF, Ralph SA, Baum J.** 2011. Super-resolution dissection of  
862 coordinated events during malaria parasite invasion of the human erythrocyte.  
863 *Cell Host Microbe* **9**:9-20.
- 864 64. **Coppel RL, Lustigman S, Murray L, Anders RF.** 1988. Mesa Is a  
865 *Plasmodium-Falciparum* Phosphoprotein Associated with the Erythrocyte-  
866 Membrane Skeleton. *Mol Biochem Parasitol* **31**:223-232.
- 867 65. **Silva MD, Cooke BM, Guillotte M, Buckingham DW, Sauzet JP, Le Scanf**  
868 **C, Contamin H, David P, Mercereau-Puijalon O, Bonnefoy S.** 2005. A role  
869 for the *Plasmodium falciparum* RESA protein in resistance against heat shock  
870 demonstrated using gene disruption. *Mol Microbiol* **56**:990-1003.
- 871 66. **Anders R, Murray L, Thomas L, Davern K, Brown G, Kemp D.** Structure  
872 and function of candidate vaccine antigens in *Plasmodium falciparum*, p 103-  
873 114. *In* (ed),
- 874 67. **Foley M, Murray LJ, Anders RF.** 1990. The Ring-Infected Erythrocyte  
875 Surface-Antigen Protein of *Plasmodium-Falciparum* Is Phosphorylated Upon  
876 Association with the Host-Cell Membrane. *Mol Biochem Parasitol* **38**:69-75.

- 877 68. **Foley M, Tilley L, Sawyer WH, Anders RF.** 1991. The ring-infected  
878 erythrocyte surface antigen of *Plasmodium falciparum* associates with spectrin  
879 in the erythrocyte membrane. *Mol Biochem Parasitol* **46**:137-147.
- 880 69. **Pei XH, Guo XH, Coppel R, Bhattacharjee S, Haldar K, Gratzer W,**  
881 **Mohandas N, An XL.** 2007. The ring-infected erythrocyte surface antigen  
882 (RESA) of *Plasmodium falciparum* stabilizes spectrin tetramers and  
883 suppresses further invasion. *Blood* **110**:1036-1042.
- 884 70. **Da Silva E, Foley M, Dluzewski AR, Murray LJ, Anders RF, Tilley L.** 1994.  
885 The *Plasmodium falciparum* protein RESA interacts with the erythrocyte  
886 cytoskeleton and modifies erythrocyte thermal stability. *Mol Biochem Parasitol*  
887 **66**:59-69.
- 888 71. **Mills JP, Diez-Silva M, Quinn DJ, Dao M, Lang MJ, Tan KS, Lim CT, Milon**  
889 **G, David PH, Mercereau-Puijalon O, Bonnefoy S, Suresh S.** 2007. Effect of  
890 plasmodial RESA protein on deformability of human red blood cells harboring  
891 *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **104**:9213-9217.
- 892 72. **Safeukui I, Correias JM, Brousse V, Hirt D, Deplaine G, Mule S, Lesurtel**  
893 **M, Goasguen N, Sauvanet A, Couvelard A, Kerneis S, Khun H, Vigan-**  
894 **Womas I, Ottone C, Molina TJ, Treluyer JM, Mercereau-Puijalon O, Milon**  
895 **G, David PH, Buffet PA.** 2008. Retention of *Plasmodium falciparum* ring-  
896 infected erythrocytes in the slow, open microcirculation of the human spleen.  
897 *Blood* **112**:2520-2528.
- 898 73. **Crutcher JM, Hoffman SL.** 1996. Malaria. *In* Baron S (ed), *Medical*  
899 *Microbiology*, 4th ed. University of Texas Medical Branch at Galveston,  
900 Galveston (TX).
- 901 74. **Oakley MS, Gerald N, McCutchan TF, Aravind L, Kumar S.** 2011. Clinical  
902 and molecular aspects of malaria fever. *Trends Parasitol* **27**:442-449.

- 903 75. **Akinyi S, Hanssen E, Meyer EVS, Jiang J, Korir CC, Singh B, Lapp S,**  
904 **Barnwell JW, Tilley L, Galinski MR.** 2012. A 95 kDa protein of *Plasmodium*  
905 *vivax* and *P. cynomolgi* visualized by three-dimensional tomography in the  
906 caveola-vesicle complexes (Schüffner's dots) of infected erythrocytes is a  
907 member of the PHIST family. *Mol Microbiol* **84**:816-831.
- 908 76. **Anderson DC, Lapp SA, Akinyi S, Meyer EV, Barnwell JW, Korir-Morrison**  
909 **C, Galinski MR.** 2015. *Plasmodium vivax* trophozoite-stage proteomes. *J*  
910 *Proteomics* **115**:157-176.
- 911 77. **Acharya P, Pallavi R, Chandran S, Chakravarti H, Middha S, Acharya J,**  
912 **Kochar S, Kochar D, Subudhi A, Boopathi AP, Garg S, Das A, Tatu U.**  
913 2009. A glimpse into the clinical proteome of human malaria parasites  
914 *Plasmodium falciparum* and *Plasmodium vivax*. *Proteomics Clin Appl* **3**:1314-  
915 1325.
- 916 78. **Lu F, Li J, Wang B, Cheng Y, Kong DH, Cui L, Ha KS, Sattabongkot J,**  
917 **Tsuboi T, Han ET.** 2014. Profiling the humoral immune responses to  
918 *Plasmodium vivax* infection and identification of candidate immunogenic  
919 rhoptry-associated membrane antigen (RAMA). *J Proteomics* **102**:66-82.
- 920 79. **Nyalwidhe J, Lingelbach K.** 2006. Proteases and chaperones are the most  
921 abundant proteins in the parasitophorous vacuole of *Plasmodium falciparum*-  
922 infected erythrocytes. *Proteomics* **6**:1563-1573.
- 923 80. **Jiang H, Patel JJ, Yi M, Mu J, Ding J, Stephens R, Cooper RA, Ferdig MT,**  
924 **Su XZ.** 2008. Genome-wide compensatory changes accompany drug-  
925 selected mutations in the *Plasmodium falciparum* crt gene. *PLoS One*  
926 **3**:e2484.

- 927 81. **Mwai L, Diriyé A, Masseno V, Muriithi S, Feltwell T, Musyoki J, Lemieux**  
928 **J, Feller A, Mair GR, Marsh K, Newbold C, Nzila A, Carret CK.** 2012.  
929 Genome wide adaptations of *Plasmodium falciparum* in response to  
930 lumefantrine selective drug pressure. *PLoS One* **7**:e31623.
- 931 82. **Mbengue A, Berry L, Braun-Breton C.** 2014. Establishment of *Plasmodium*  
932 *falciparum* Extracellular Compartments in its Host Erythrocyte, p 133-159,  
933 Heat Shock Proteins of Malaria. Springer.
- 934 83. **Crowley VM, Rovira-Graells N, Ribas de Pouplana L, Cortes A.** 2011.  
935 Heterochromatin formation in bistable chromatin domains controls the  
936 epigenetic repression of clonally variant *Plasmodium falciparum* genes linked  
937 to erythrocyte invasion. *Mol Microbiol* **80**:391-406.
- 938 84. **Bertin GI, Sabbagh A, Guillonneau F, Jafari-Guemouri S, Ezinmegnon S,**  
939 **Federici C, Hounkpatin B, Fievet N, Deloron P.** 2013. Differential Protein  
940 Expression Profiles Between *Plasmodium falciparum* Parasites Isolated From  
941 Subjects Presenting With Pregnancy-Associated Malaria and Uncomplicated  
942 Malaria in Benin. *J Infect Dis* **208**:1987-1997.
- 943 85. **Buchholz K, Burke TA, Williamson KC, Wiegand RC, Wirth DF, Marti M.**  
944 2011. A high-throughput screen targeting malaria transmission stages opens  
945 new avenues for drug development. *J Infect Dis* **203**:1445-1453.
- 946 86. **Goel S, Muthusamy A, Miao J, Cui L, Salanti A, Winzeler EA, Gowda DC.**  
947 2014. Targeted disruption of a ring-infected erythrocyte surface antigen  
948 (RESA)-like export protein gene in *Plasmodium falciparum* confers stable  
949 chondroitin 4-sulfate cytoadherence capacity. *J Biol Chem* **289**:34408-34421.
- 950 87. **Natalang O, Bischoff E, Deplaine G, Proux C, Dillies MA, Sismeiro O,**  
951 **Guigon G, Bonnefoy S, Patarapotikul J, Mercereau-Puijalon O, Coppee**  
952 **JY, David PH.** 2008. Dynamic RNA profiling in *Plasmodium falciparum*

953       synchronized blood stages exposed to lethal doses of artesunate. BMC  
954       Genomics **9**:388.

955   88.   **Cai H, Hong C, Gu J, Lilburn TG, Kuang R, Wang Y.** 2012. Module-based  
956       subnetwork alignments reveal novel transcriptional regulators in malaria  
957       parasite *Plasmodium falciparum*. BMC Syst Biol **6 Suppl 3**:S5.

958   89.   **Vincensini L, Richert S, Blisnick T, Van Dorsselaer A, Leize-Wagner E,**  
959       **Rabilloud T, Braun Breton C.** 2005. Proteomic analysis identifies novel  
960       proteins of the Maurer's clefts, a secretory compartment delivering  
961       *Plasmodium falciparum* proteins to the surface of its host cell. Mol Cell  
962       Proteomics **4**:582-593.

963   90.   **Lanzer M, Wickert H, Krohne G, Vincensini L, Braun Breton C.** 2006.  
964       Maurer's clefts: a novel multi-functional organelle in the cytoplasm of  
965       *Plasmodium falciparum*-infected erythrocytes. Int J Parasitol **36**:23-36.

966   91.   **Ralph SA, Bischoff E, Mattei D, Sismeiro O, Dillies MA, Guigon G,**  
967       **Coppee JY, David PH, Scherf A.** 2005. Transcriptome analysis of antigenic  
968       variation in *Plasmodium falciparum*--var silencing is not dependent on  
969       antisense RNA. Genome Biology **6**:R93.

970   92.   **Vignali M, Armour CD, Chen J, Morrison R, Castle JC, Biery MC, Bouzek**  
971       **H, Moon W, Babak T, Fried M, Raymond CK, Duffy PE.** 2011. NSR-seq  
972       transcriptional profiling enables identification of a gene signature of  
973       *Plasmodium falciparum* parasites infecting children. J Clin Invest **121**:1119-  
974       1129.

975   93.   **Le Roch KG, Johnson JR, Florens L, Zhou Y, Santrosyan A, Grainger M,**  
976       **Yan SF, Williamson KC, Holder AA, Carucci DJ, Yates JR, 3rd, Winzeler**  
977       **EA.** 2004. Global analysis of transcript and protein levels across the  
978       *Plasmodium falciparum* life cycle. Genome Res **14**:2308-2318.

- 979 94. **Wang D.** 2008. Discrepancy between mRNA and protein abundance: insight  
980 from information retrieval process in computers. *Comput Biol Chem* **32**:462-  
981 468.
- 982 95. **Nair S, Nkhoma S, Nosten F, Mayxay M, French N, Whitworth J,**  
983 **Anderson T.** 2010. Genetic changes during laboratory propagation: copy  
984 number At the reticulocyte-binding protein 1 locus of *Plasmodium falciparum*.  
985 *Mol Biochem Parasitol* **172**:145-148.
- 986 96. **Sam-Yellowe TY.** 2009. The role of the Maurer's clefts in protein transport in  
987 *Plasmodium falciparum*. *Trends Parasitol* **25**:277-284.
- 988 97. **Talevich E, Tobin AB, Kannan N, Doerig C.** 2012. An evolutionary  
989 perspective on the kinome of malaria parasites. *Philos Trans R Soc Lond B*  
990 *Biol Sci* **367**:2607-2618.
- 991 98. **Mu J, Awadalla P, Duan J, McGee KM, Keebler J, Seydel K, McVean GA,**  
992 **Su XZ.** 2007. Genome-wide variation and identification of vaccine targets in  
993 the *Plasmodium falciparum* genome. *Nature Genetics* **39**:126-130.
- 994 99. **LaCount DJ, Vignali M, Chettier R, Phansalkar A, Bell R, Hesselberth JR,**  
995 **Schoenfeld LW, Ota I, Sahasrabudhe S, Kurschner C, Fields S, Hughes**  
996 **RE.** 2005. A protein interaction network of the malaria parasite *Plasmodium*  
997 *falciparum*. *Nature* **438**:103-107.
- 998 100. **Spillman NJ, Beck JR, Goldberg DE.** 2015. Protein export into malaria  
999 parasite-infected erythrocytes: mechanisms and functional consequences.  
1000 *Annu Rev Biochem* **84**:813-841.
- 1001 101. **Pallavi R, Acharya P, Chandran S, Daily JP, Tatu U.** 2010. Chaperone  
1002 expression profiles correlate with distinct physiological states of *Plasmodium*  
1003 *falciparum* in malaria patients. *Malar J* **9**:236.



- 1004 102. **Pavithra SR, Kumar R, Tatu U.** 2007. Systems analysis of chaperone  
1005 networks in the malarial parasite *Plasmodium falciparum*. *PLoS Comput Biol*  
1006 **3**:1701-1715.
- 1007 103. **Oyelade J, Ewejobi I, Brors B, Eils R, Adebisi E.** 2011. Computational  
1008 identification of signalling pathways in *Plasmodium falciparum*. *Infect Genet*  
1009 *Evol* **11**:755-764.
- 1010 104. **Razak MR, Abdullah NR, Chomel R, Muhamad R, Ismail Z.** 2014. Effect of  
1011 choline kinase inhibitor hexadecyltrimethylammonium bromide on *Plasmodium*  
1012 *falciparum* gene expression. *Southeast Asian J Trop Med Public Health*  
1013 **45**:259-266.
- 1014 105. **Botha M, Pesce ER, Blatch GL.** 2007. The Hsp40 proteins of *Plasmodium*  
1015 *falciparum* and other apicomplexa: regulating chaperone power in the parasite  
1016 and the host. *Int J Biochem Cell Biol* **39**:1781-1803.
- 1017 106. **Florens L, Liu X, Wang Y, Yang S, Schwartz O, Peglar M, Carucci DJ,**  
1018 **Yates JR, 3rd, Wu Y.** 2004. Proteomics approach reveals novel proteins on  
1019 the surface of malaria-infected erythrocytes. *Mol Biochem Parasitol* **135**:1-11.
- 1020 107. **Spielmann T, Hawthorne PL, Dixon MW, Hannemann M, Klotz K, Kemp**  
1021 **DJ, Klonis N, Tilley L, Trenholme KR, Gardiner DL.** 2006. A cluster of ring  
1022 stage-specific genes linked to a locus implicated in cytoadherence in  
1023 *Plasmodium falciparum* codes for PEXEL-negative and PEXEL-positive  
1024 proteins exported into the host cell. *Mol Biol Cell* **17**:3613-3624.
- 1025 108. **Siau A, Toure FS, Ouwe-Missi-Oukem-Boyer O, Ciceron L, Mahmoudi N,**  
1026 **Vaquero C, Froissard P, Bisvigou U, Bisser S, Coppee JY, Bischoff E,**  
1027 **David PH, Mazier D.** 2007. Whole-transcriptome analysis of *Plasmodium*  
1028 *falciparum* field isolates: identification of new pathogenicity factors. *J Infect Dis*  
1029 **196**:1603-1612.

- 1030 109. **Penn WD.** 2010. Yeast two-hybrid analysis of host-parasite protein  
1031 interactions in *Falciparum malaria*. Doctoral Dissertation, Purdue University,  
1032 West Lafayette, IN, USA.
- 1033 110. **Fried M, Hixson KK, Anderson L, Ogata Y, Mutabingwa TK, Duffy PE.**  
1034 2007. The distinct proteome of placental malaria parasites. *Mol Biochem*  
1035 *Parasitol* **155**:57-65.
- 1036 111. **Hviid L.** 2011. The case for PfEMP1-based vaccines to protect pregnant  
1037 women against *Plasmodium falciparum* malaria. *Expert Rev Vaccines*  
1038 **10**:1405-1414.
- 1039 112. **Nunes MC, Sterkers Y, Gamain B, Scherf A.** 2008. Investigation of host  
1040 factors possibly enhancing the emergence of the chondroitin sulfate A-binding  
1041 phenotype in *Plasmodium falciparum*. *Microbes Infect* **10**:928-932.
- 1042 113. **Carret CK, Horrocks P, Konfortov E, Winzeler E, Qureshi M, Newbold C,**  
1043 **Ivens A.** 2005. Microarray-based comparative genomic analyses of the  
1044 human malaria parasite *Plasmodium falciparum* using Affymetrix arrays. *Mol*  
1045 *Biochem Parasitol* **144**:177-186.
- 1046 114. **Prajapati SK, Singh OP.** 2013. Remodeling of human red cells infected with  
1047 *Plasmodium falciparum* and the impact of PHIST proteins. *Blood Cells Mol Dis*  
1048 **51**:195-202.
- 1049 115. **Rodriguez LE, Curtidor H, Urquiza M, Cifuentes G, Reyes C, Patarroyo**  
1050 **ME.** 2008. Intimate molecular interactions of *P. falciparum* merozoite proteins  
1051 involved in invasion of red blood cells and their implications for vaccine  
1052 design. *Chem Rev* **108**:3656-3705.
- 1053 116. **Gamboa D, Ho MF, Bendezu J, Torres K, Chiodini PL, Barnwell JW,**  
1054 **Incardona S, Perkins M, Bell D, McCarthy J, Cheng Q.** 2010. A large

1055 proportion of *P. falciparum* isolates in the Amazon region of Peru lack pfhrp2  
 1056 and pfhrp3: implications for malaria rapid diagnostic tests. PLoS One **5**:e8091.

1057 117. **Abdallah JF, Okoth SA, Fontecha GA, Torres RE, Banegas EI, Matute ML,**  
 1058 **Bucheli ST, Goldman IF, de Oliveira AM, Barnwell JW, Udhayakumar V.**  
 1059 2015. Prevalence of pfhrp2 and pfhrp3 gene deletions in Puerto Lempira,  
 1060 Honduras. Malar J **14**:19.

1061 118. **Gelhaus C, Fritsch J, Krause E, Leippe M.** 2005. Fractionation and  
 1062 identification of proteins by 2-DE and MS: towards a proteomic analysis of  
 1063 *Plasmodium falciparum*. Proteomics **5**:4213-4222.

1064 119. **Singh K, Agarwal A, Khan SI, Walker LA, Tekwani BL.** 2007. Growth, drug  
 1065 susceptibility, and gene expression profiling of *Plasmodium falciparum*  
 1066 cultured in medium supplemented with human serum or lipid-rich bovine  
 1067 serum albumin [corrected]. J Biomol Screen **12**:1109-1114.

1068 120. **Ribacke U, Mok BW, Wirta V, Normark J, Lundeberg J, Kironde F,**  
 1069 **Egwang TG, Nilsson P, Wahlgren M.** 2007. Genome wide gene  
 1070 amplifications and deletions in *Plasmodium falciparum*. Mol Biochem Parasitol  
 1071 **155**:33-44.

1072 121. **Angus BJ, Chotivanich K, Udomsangpetch R, White NJ.** 1997. In vivo  
 1073 removal of malaria parasites from red blood cells without their destruction in  
 1074 acute falciparum malaria. Blood **90**:2037-2040.

1075 122. **Dixon MW, Kenny S, McMillan PJ, Hanssen E, Trenholme KR, Gardiner**  
 1076 **DL, Tilley L.** 2011. Genetic ablation of a Maurer's cleft protein prevents  
 1077 assembly of the *Plasmodium falciparum* virulence complex. Mol Microbiol  
 1078 **81**:982-993.

1079 123. **Rug M, Maier AG.** 2011. The heat shock protein 40 family of the malaria  
 1080 parasite *Plasmodium falciparum*. IUBMB Life **63**:1081-1086.

- 1081 124. **Le Roch KG, Johnson JR, Ahiboh H, Chung DW, Prudhomme J, Plouffe**  
 1082 **D, Henson K, Zhou Y, Witola W, Yates JR, Mamoun CB, Winzeler EA, Vial**  
 1083 **H.** 2008. A systematic approach to understand the mechanism of action of the  
 1084 bisthiazolium compound T4 on the human malaria parasite, *Plasmodium*  
 1085 *falciparum*. *BMC Genomics* **9**:513.
- 1086 125. **Acharya P, Kumar R, Tatu U.** 2007. Chaperoning a cellular upheaval in  
 1087 malaria: heat shock proteins in *Plasmodium falciparum*. *Mol Biochem Parasitol*  
 1088 **153**:85-94.
- 1089 126. **Badaut C, Guyonnet L, Milet J, Renard E, Durand R, Viwami F, Sagbo G,**  
 1090 **Layla F, Deloron P, Bonnefoy S, Migot-Nabias F.** 2015. Immunoglobulin  
 1091 response to *Plasmodium falciparum* RESA proteins in uncomplicated and  
 1092 severe malaria. *Malar J* **14**:278.
- 1093 127. **Caro F, Ah Yong V, Betegon M, DeRisi JL.** 2014. Genome-wide regulatory  
 1094 dynamics of translation in the *Plasmodium falciparum* asexual blood stages.  
 1095 *Elife* **3**:e04106.
- 1096 128. **Durand R, Migot-Nabias F, Andriantsoanirina V, Seringe E, Viwami F,**  
 1097 **Sagbo G, Lalya F, Deloron P, Mercereau-Puijalon O, Bonnefoy S.** 2012.  
 1098 Possible association of the *Plasmodium falciparum* T1526C *resa2* gene  
 1099 mutation with severe malaria. *Malar J* **11**:128.
- 1100 129. **Pelle KG, Oh K, Buchholz K, Narasimhan V, Joice R, Milner DA,**  
 1101 **Brancucci NM, Ma S, Voss TS, Ketman K, Seydel KB, Taylor TE,**  
 1102 **Barteneva NS, Huttenhower C, Marti M.** 2015. Transcriptional profiling  
 1103 defines dynamics of parasite tissue sequestration during malaria infection.  
 1104 *Genome Med* **7**:19.
- 1105 130. **Siau A, Silvie O, Franetich JF, Yalaoui S, Marinach C, Hannoun L, van**  
 1106 **Gemert GJ, Luty AJ, Bischoff E, David PH, Snounou G, Vaquero C,**

1107 **Froissard P, Mazier D.** 2008. Temperature shift and host cell contact up-  
 1108 regulate sporozoite expression of *Plasmodium falciparum* genes involved in  
 1109 hepatocyte infection. *PLoS Pathog* **4**:e1000121.

1110 131. **Speake C, Duffy PE.** 2009. Antigens for pre-erythrocytic malaria vaccines:  
 1111 building on success. *Parasite Immunol* **31**:539-546.

1112 132. **Duffy PE, Fried M.** 2011. Pregnancy malaria: cryptic disease, apparent  
 1113 solution. *Mem Inst Oswaldo Cruz* **106 Suppl 1**:64-69.

1114 133. **Davies DH, Duffy P, Bodmer JL, Felgner PL, Doolan DL.** 2015. Large  
 1115 screen approaches to identify novel malaria vaccine candidates. *Vaccine*  
 1116 **33**:7496-7505.

1117 134. **Radfar A, Diez A, Bautista JM.** 2008. Chloroquine mediates specific  
 1118 proteome oxidative damage across the erythrocytic cycle of resistant  
 1119 *Plasmodium falciparum*. *Free Radic Biol Med* **44**:2034-2042.

1120 135. **Fatumo S, Adebisi M, Adebisi E.** 2013. In silico models for drug resistance,  
 1121 p 39-65, In *Silico Models for Drug Discovery*. Springer.

1122 136. **Boddey JA, Hodder AN, Gunther S, Gilson PR, Patsiouras H, Kapp EA,**  
 1123 **Pearce JA, de Koning-Ward TF, Simpson RJ, Crabb BS, Cowman AF.**  
 1124 2010. An aspartyl protease directs malaria effector proteins to the host cell.  
 1125 *Nature* **463**:627-631.

1126 137. **Sam-Yellowe TY, Florens L, Wang T, Raine JD, Carucci DJ, Sinden R,**  
 1127 **Yates JR, 3rd.** 2004. Proteome analysis of rhoptry-enriched fractions isolated  
 1128 from *Plasmodium* merozoites. *J Proteome Res* **3**:995-1001.

1129 138. **Baum E, Badu K, Molina DM, Liang X, Felgner PL, Yan G.** 2013. Protein  
 1130 microarray analysis of antibody responses to *Plasmodium falciparum* in  
 1131 western Kenyan highland sites with differing transmission levels. *PLoS One*  
 1132 **8**:e82246.

- 1133 139. **Andrews KT, Gupta AP, Tran TN, Fairlie DP, Gobert GN, Bozdech Z.** 2012.  
1134 Comparative gene expression profiling of *P. falciparum* malaria parasites  
1135 exposed to three different histone deacetylase inhibitors. *PLoS One* **7**:e31847.
- 1136 140. **Le Roch KG, Zhou Y, Blair PL, Grainger M, Moch JK, Haynes JD, De La**  
1137 **Vega P, Holder AA, Batalov S, Carucci DJ, Winzeler EA.** 2003. Discovery  
1138 of gene function by expression profiling of the malaria parasite life cycle.  
1139 *Science* **301**:1503-1508.
- 1140 141. **Chookajorn T, Hartl DL.** 2006. Position-specific polymorphism of  
1141 *Plasmodium falciparum* Stuttering motif in a PHISTc PFI1780w. *Exp Parasitol*  
1142 **114**:126-128.
- 1143 142. **Otto TD, Wilinski D, Assefa S, Keane TM, Sarry LR, Böhme U, Lemieux J,**  
1144 **Barrell B, Pain A, Berriman M.** 2010. New insights into the blood-stage  
1145 transcriptome of *Plasmodium falciparum* using RNA-Seq. *Mol Microbiol* **76**:12-  
1146 24.
- 1147 143. **Carver T, Harris SR, Otto TD, Berriman M, Parkhill J, McQuillan JA.** 2013.  
1148 BamView: visualizing and interpretation of next-generation sequencing read  
1149 alignments. *Brief Bioinform* **14**:203-212.
- 1150 144. **Issar N, Roux E, Mattei D, Scherf A.** 2008. Identification of a novel post-  
1151 translational modification in *Plasmodium falciparum*: protein sumoylation in  
1152 different cellular compartments. *Cell Microbiol* **10**:1999-2011.
- 1153 145. **Josling GA, Llinas M.** 2015. Sexual development in *Plasmodium* parasites:  
1154 knowing when it's time to commit. *Nat Rev Microbiol* **13**:573-587.
- 1155 146. **Haase S, Herrmann S, Gruring C, Heiber A, Jansen PW, Langer C, Treeck**  
1156 **M, Cabrera A, Bruns C, Struck NS, Kono M, Engelberg K, Ruch U,**  
1157 **Stunnenberg HG, Gilberger TW, Spielmann T.** 2009. Sequence

1158 requirements for the export of the *Plasmodium falciparum* Maurer's clefts  
 1159 protein REX2. *Mol Microbiol* **71**:1003-1017.

1160 147. **Volz J, Carvalho TG, Ralph SA, Gilson P, Thompson J, Tonkin CJ,**  
 1161 **Langer C, Crabb BS, Cowman AF.** 2010. Potential epigenetic regulatory  
 1162 proteins localise to distinct nuclear sub-compartments in *Plasmodium*  
 1163 *falciparum*. *Int J Parasitol* **40**:109-121.

1164 148. **Lee AH, Symington LS, Fidock DA.** 2014. DNA Repair Mechanisms and  
 1165 Their Biological Roles in the Malaria Parasite *Plasmodium falciparum*.  
 1166 *Microbiol Mol Biol R* **78**:469-486.

1167 149. **Taylor DW, Parra M, Chapman GB, Stearns ME, Renner J, Aikawa M, Uni**  
 1168 **S, Aley SB, Panton LJ, Howard RJ.** 1987. Localization of *Plasmodium*  
 1169 *falciparum* histidine-rich protein 1 in the erythrocyte skeleton under knobs. *Mol*  
 1170 *Biochem Parasitol* **25**:165-174.

1171 150. **Watermeyer JM, Hale VL, Hackett F, Clare DK, Cutts EE, Vakonakis I,**  
 1172 **Fleck RA, Blackman MJ, Saibil HR.** 2016. A spiral scaffold underlies  
 1173 cytoadherent knobs in *Plasmodium falciparum*-infected erythrocytes. *Blood*  
 1174 **127**:343-351.

1175 151. **Mundwiler-Pachlatko E, Beck HP.** 2013. Maurer's clefts, the enigma of  
 1176 *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* **110**:19987-19994.

1177 152. **Cyrklaff M, Sanchez CP, Frischknecht F, Lanzer M.** 2012. Host actin  
 1178 remodeling and protection from malaria by hemoglobinopathies. *Trends*  
 1179 *Parasitol* **28**:479-485.

1180 153. **Cyrklaff M, Sanchez CP, Kilian N, Bisseye C, Simporé J, Frischknecht F,**  
 1181 **Lanzer M.** 2011. Hemoglobins S and C interfere with actin remodeling in  
 1182 *Plasmodium falciparum*-infected erythrocytes. *Science* **334**:1283-1286.

- 1183 154. **Spycher C, Rug M, Klonis N, Ferguson DJ, Cowman AF, Beck HP, Tilley**  
 1184 **L.** 2006. Genesis of and trafficking to the Maurer's clefts of *Plasmodium*  
 1185 *falciparum*-infected erythrocytes. *Molecular and cellular biology* **26**:4074-4085.
- 1186 155. **Pachlatko E, Rusch S, Muller A, Hemphill A, Tilley L, Hanssen E, Beck**  
 1187 **HP.** 2010. MAHRP2, an exported protein of *Plasmodium falciparum*, is an  
 1188 essential component of Maurer's cleft tethers. *Mol Microbiol* **77**:1136-1152.
- 1189 156. **Hawthorne PL, Trenholme KR, Skinner-Adams TS, Spielmann T, Fischer**  
 1190 **K, Dixon MWA, Ortega MR, Anderson KL, Kemp DJ, Gardiner DL.** 2004. A  
 1191 novel *Plasmodium falciparum* ring stage protein, REX, is located in Maurer's  
 1192 clefts. *Mol Biochem Parasitol* **136**:181-189.
- 1193 157. **Dietz O, Rusch S, Brand F, Mundwiler-Pachlatko E, Gaida A, Voss T,**  
 1194 **Beck HP.** 2014. Characterization of the small exported *Plasmodium*  
 1195 *falciparum* membrane protein SEMP1. *PLoS One* **9**:e103272.
- 1196 158. **Blisnick T, Eugenia M, Betoulle M, Barale JC, Uzureau P, Berry L,**  
 1197 **Desroses S, Fujioka H, Mattei D, Breton CB.** 2000. Pfsbp 1, a Maurer's cleft  
 1198 *Plasmodium falciparum* protein, is associated with the erythrocyte skeleton.  
 1199 *Mol Biochem Parasitol* **111**:107-121.
- 1200 159. **Rug M, Cyrklaff M, Mikkonen A, Lemgruber L, Kuelzer S, Sanchez CP,**  
 1201 **Thompson J, Hanssen E, O'Neill M, Langer C, Lanzer M, Frischknecht F,**  
 1202 **Maier AG, Cowman AF.** 2014. Export of virulence proteins by malaria-  
 1203 infected erythrocytes involves remodeling of host actin cytoskeleton. *Blood*  
 1204 **124**:3459-3468.
- 1205 160. **Kats LM, Proellocks NI, Buckingham DW, Blanc L, Hale J, Guo X, Pei X,**  
 1206 **Herrmann S, Hanssen EG, Coppel RL, Mohandas N, An X, Cooke BM.**  
 1207 2015. Interactions between *Plasmodium falciparum* skeleton-binding protein 1



1208 and the membrane skeleton of malaria-infected red blood cells. *Biochim*  
1209 *Biophys Acta* **1848**:1619-1628.

1210 161. **Kulzer S, Rug M, Brinkmann K, Cannon P, Cowman A, Lingelbach K,**  
1211 **Blatch GL, Maier AG, Przyborski JM.** 2010. Parasite-encoded Hsp40  
1212 proteins define novel mobile structures in the cytosol of the *P. falciparum*-  
1213 infected erythrocyte. *Cell Microbiol* **12**:1398-1420.

1214 162. **Külzer S, Charnaud S, Dagan T, Riedel J, Mandal P, Pesce ER, Blatch GL,**  
1215 **Crabb BS, Gilson PR, Przyborski JM.** 2012. *Plasmodium falciparum* -  
1216 encoded exported hsp70/hsp40 chaperone/co-chaperone complexes within  
1217 the host erythrocyte. *Cell Microbiol* **14**:1784-1795.

1218 163. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy  
1219 and high throughput. *Nucleic Acids Res* **32**:1792-1797.

1220 164. **Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, Park YM,**  
1221 **Buso N, Lopez R.** 2015. The EMBL-EBI bioinformatics web and  
1222 programmatic tools framework. *Nucleic Acids Res* **43**:W580-584.

1223 165. **Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R,**  
1224 **McAnulla C, McMenamin C, Nuka G, Pesseat S, Sangrador-Vegas A,**  
1225 **Scheremetjew M, Rato C, Yong SY, Bateman A, Punta M, Attwood TK,**  
1226 **Sigrist CJ, Redaschi N, Rivoire C, Xenarios I, Kahn D, Guyot D, Bork P,**  
1227 **Letunic I, Gough J, Oates M, Haft D, Huang H, Natale DA, Wu CH, Orengo**  
1228 **C, Sillitoe I, Mi H, Thomas PD, Finn RD.** 2015. The InterPro protein families  
1229 database: the classification resource after 15 years. *Nucleic Acids Res*  
1230 **43**:D213-221.

1231 166. **Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R,**  
1232 **McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG.** 2011.

1233 Fast, scalable generation of high-quality protein multiple sequence alignments  
 1234 using Clustal Omega. *Mol Syst Biol* **7**:539.

1235 167. **Hall TA**. BioEdit: a user-friendly biological sequence alignment editor and  
 1236 analysis program for Windows 95/98/NT, p 95-98. *In* (ed),

1237 168. **Cole C, Barber JD, Barton GJ**. 2008. The Jpred 3 secondary structure  
 1238 prediction server. *Nucleic Acids Res* **36**:W197-201.

1239 169. **Cuff JA, Barton GJ**. 2000. Application of multiple sequence alignment profiles  
 1240 to improve protein secondary structure prediction. *Proteins* **40**:502-511.

1241 170. **Saeed AI, Bhagabati NK, Braisted JC, Liang W, Sharov V, Howe EA, Li J,**  
 1242 **Thiagarajan M, White JA, Quackenbush J**. 2006. TM4 microarray software  
 1243 suite. *Methods Enzymol* **411**:134-193.

1244 **Table 1 - Summary and information on PHIST proteins**

1245 **A: PHISTa genes**

1246

Gene ID <sup>1</sup>	Other Gene IDs or name <sup>2</sup>	PEXEL <sup>3</sup>	Gene status <sup>4</sup>	MW <sup>5</sup>	Length <sup>5</sup>	KO <sup>6</sup>	References <sup>7</sup>
PF3D7_0102000	PFA0100c, MAL1P1.11		PG	30.4	252		
PF3D7_0115100	PFA0735w, MAL1P4.22	x		28.6	293		only 5 aa difference to PF3D7_0800600 (PlasmoDB)
PF3D7_0402000	PFD0090c, MAL4P1.18	x		41.7	357	x	interaction between band 4.1R (58), not silenced in the 3D7 strain (28) over-expressed in samples from patients with cerebral malaria (34)
PF3D7_0424900	PFD1185w, MAL4P1.232	x		25.3	216		
PF3D7_0425300	PFD1210w, MAL4P1.237		PG	27.5	232		exported despite lack of any export signal (79)
PF3D7_0425400	PFD1215w, MAL4P1.238	x		19.5	164		
PF3D7_0601700	PFF0085w, MAL6P1.21	x	PG	26.3	289		
PF3D7_0800600	MAL8P1.163	x		28.3	237		adjacent to a PfEMP1 variant on gene locus. Weak binding to the ATS domain of PfEMP1 as compared to other PHISTs tested (19) only 5 aa difference to PF3D7_0115100 (PlasmoDB)
PF3D7_1001100.1	PF10_0014, PF10_0015, acyl-CoA binding protein			36.5	313	x	grouped as PHISTa (15) deleted in parasites with PfCRT mutations, annotated as putative lipid transporter (80), up-regulated under Lumefantrine pressure, deletion of the gene found under Chloroquine pressure (81), unusual PHIST member with N-terminal Acyl-CoA binding protein domain (16)
PF3D7_1001100.2	PF10_0014, PF10_0015, acyl-CoA binding protein			31.0	267	x	same as PF3D7_1001100.1

<b>PF3D7_1001300</b>	PF10_0017	x		22.8	190	diffuse localization in the RBC cytosol (20), differentially expressed in parasites under Lumefantrine pressure (81)
<b>PF3D7_1100600</b>	PF11_0012		PG	33.3	281	
<b>PF3D7_1149700</b>	PF11_0514			10.5	85	once mentioned in relation to PHISTa structure (58)
<b>PF3D7_1253100</b>	PFL2555w, MAL12P1.506	x		26.5	221	HSP40 chaperone with a DNA-J domain (82), used as negative control in interaction study (58)
<b>PF3D7_1253300</b>	PFL2565w, MAL12P1.508	x	PG	18.3	157	not silenced in 3D7 (28)
<b>PF3D7_1253800</b>	PFL2590w, MAL12P1.513	x		26.2	219	
<b>PF3D7_1253900</b>	PFL2595w, 2600w, MAL12P1.514	x	PG	24.3	275	once mentioned in relation to PHISTa structure (58)
<b>PF3D7_1301100</b>	MAL13P1.11, MAL13P1.11a	x	PG	24.6	277	
<b>PF3D7_1301500</b>	MAL13P1.59			37.5	308	downstream of a clag gene and its expression is affected by H3K9me3 acetylation (83), not located in telomeric region (27)
<b>PF3D7_1372000</b>	MAL13P1.470	x		41.0	349	String database mining suggested interaction with PFI1785w, PFB0115w, and PFL0050c (84)
<b>PF3D7_1400900</b>	PF14_0009	x	PG	24.6	277	
<b>PF3D7_1477700</b>	PF14_0748, Pfg14-748			34.5	291	highest transcript concentration in bone marrow when comparing different organs (37), marker for early and mid-gametocytes, its promoter can drive gametocyte-specific gene expression (36, 37), expressed in sexually committed schizonts with more than 100 fold up-regulation (27). Used as marker for sexual commitment (85)
<b>PF3D7_1478000</b>	PF14_0752, GEXP17	x		25.5	215	over-expressed in field-samples compared to 3D7. Suggested surface protein (30), differentially expressed in different 3D7 clones (33)
<b>PF3D7_1478500</b>	PF14_0757	x	PG	25.1	264	
<b>PF3D7_1479200</b>	PF14_0763	x		26.2	219	
<b>PF3D7_1479300</b>	PF14_0764 PF14_0765	x	PG	24.6	277	

1247  
1248

## B: PHISTb genes

Gene ID <sup>1</sup>	Other Gene IDs or name <sup>2</sup>	PEXEL <sup>3</sup>	Gene status <sup>4</sup>	MW <sup>5</sup>	Length <sup>5</sup>	KO <sup>6</sup>	References <sup>7</sup>
PF3D7_0201600	PFB0080c, PF02_0016	x		48.0	394	x	affected by large chromosome break which causes loss of cytoadherence and permanent expression of <i>var2csa</i> . Regulates <i>var2csa</i> and <i>var</i> gene expression (86), displayed on iRBC surface (86), cytosolic localization with weak accumulation at the erythrocyte periphery (20), PHISTb domain-containing RESA-like protein 1 (PlasmoDB) over-expressed in samples from patients with cerebral malaria (34)
PF3D7_0401800	PFD0080c, MAL4P1.16, PFD80	x		54.8	512		up-regulated within 90 minutes after artesunate treatment (87), based on data mining it is suggested to interact with PFD0985w (88), putative Maurer's cleft protein (89, 90), differentially expressed in Pf-FRC parasites selected for CSA or CD36 binding (91), among the most significantly up-regulated genes in children with <i>P. falciparum</i> malaria (92), discrepancy observed between RNA and protein level (93, 94)
PF3D7_0402100	PFD0095c, MAL4P1.19	x		61.7	522		<i>var</i> gene <i>chr4var7</i> showed recombination with PFD0100c (95) essential for in vitro growth (44), apparently linked to invasion ligand RH1-abundance in FCR3 (80, 95)
PF3D7_0424600	PFD1170c, MAL4P1.229	x		26.2	221	x	required for correct KAHRP transport and knob formation. Deletion has similar effect as in KAHRP KO parasites but does not affect PfEMP1 transport (44), protein shows peripheral localization (20), no interaction with the ATS of PfEMP1 (19), involved in knob formation and cytoadherence (96), present in peripheral blood of malaria patients (77)
PF3D7_0424800	PFD1180w, MAL4P1.231	x		31.1	266		
PF3D7_0532300	PFE1600w, MAL5P1.314	x		49.9	419		exported and tyrosine-phosphorylated in the RBC cytoplasm (97) shown on immunoblot with candidate vaccine antigens (98), shown in interaction cluster with PFE1605w (99) (see Figure S7)

<b>PF3D7_0532400</b>	PFE1605w, MAL5P1.315, LyMP - lysine-rich membrane- associated PHISTb protein	x		50.6	439		showed interaction with ATS of PfEMP1 variants, localized to the knobs (19), localization at membrane cytoskeleton in between knobs (40), recently reviewed with figure indicating both knob and membrane localization (100), referred to as Hsp40 protein. Interaction with MSP1 (101), Hsp40 protein with J-domain (102), Yeast-two hybrid interaction with SBP1 and PFE1600w (102) (see Figure S7), also computationally predicted to be a nuclear pore protein and to be part of a signaling pathway in the FIKK protein family (103)
<b>PF3D7_0601500</b>	PFF0075c, MAL6P1.19	x		50.0	417		has two MEC motifs. Only 1 aa difference to PF3D7_0631100 (39)
<b>PF3D7_0631100</b>	PFF1510w	x		49.9	416		has two MEC motifs. Only 1 aa difference to PF3D7_0601500 (39)
<b>PF3D7_0702100</b>	MAL7P1.7	x	PG	69.1	586		part of an interaction network with SBP in the center (99) (see Figure S7), HDTAB treatment changed its expression (104), identified a RESA-like protein (82), identified as RESA-like protein but not a HSP40 (105)
<b>PF3D7_0731300</b>	MAL7P1.174, Pfg174	x		31.6	263	x	putative Maurer's cleft protein (89, 90), suggested location of a surface protein (106), soluble protein (107)
<b>PF3D7_0831000</b>	MAL8P1.2, GEXP09			51.0	426		listed as Hsp40 chaperone with a J-domain (102)
<b>PF3D7_0902700</b>	PFI0130c	x	PG	44.0	448		silencing of this gene resulted in inhibition of apoptosis without affecting parasite growth (108), potentially links an unknown surface protein to the iRBC cytoskeleton (109)
<b>PF3D7_0936900</b>	PFI1785w	x	PG	32.9	363		particularly abundant protein in samples from pregnant women, but not in samples from children (49, 110, 111), affected by deletions on Chromosome 9 (112, 113), String database mining suggested interaction with var2csa and MAL13P1.470-1 (84), one of the earliest up-regulated genes in the asexual cycle (28), often mentioned together with PFD1140w

<b>PF3D7_0937000</b>	PF11790w		43.0	357	x	yeast-two hybrid interaction with band 4.1R (99, 114), potential involvement in host cytoskeleton remodeling (39), located in a region prone to deletion in <i>P. falciparum</i> strains IT and FCR3 (113)
<b>PF3D7_1102500</b>	PF11_0037, GEXP2	x	64.6	547	x	Immuno-precipitation using this protein pulled down PTEX components HSP101, PTEX150 and EXP2 (45), differentially expressed in HP1 depleted parasites (53), involved in cytoadherence (109)
<b>PF3D7_1201000</b>	PFL0050c, MAL12P1.10	x	71.7	605	x	putative Maurer's cleft protein (89, 90), String database mining suggests interaction with var2csa and MAL13P1.470-1 (84)
<b>PF3D7_1252700</b>	PFL2535w, MAL12P1.502	x	42.1	363		RESA-like protein (115)
<b>PF3D7_1252800</b>	PFL2540w, MAL12P1.503	x	66.6	559		
<b>PF3D7_1372100</b>	MAL13P1.475, GEXP04	x	67.0	558	x	frequently deleted in field samples, affected by same deletion as HRP2 and HRP3 (116, 117)
<b>PF3D7_1401600</b>	PF14_0018	x	45.7	391	x	Knock out resulted in less rigid but viable parasites(39, 44), putative TM domain or GPI anchor (118)
<b>PF3D7_1476200</b>	PF14_0731 or PF14_0730	x	49.3	410		up-regulated in vitro when cultured with human serum as compared to Albumax (119)
<b>PF3D7_1476300</b>	PF14_0732	x	61.5	514	x	frequently deleted in HB3 and other strains (120)
<b>PF3D7_1477500</b>	PF14_0746	x	49.9	411		

1249

1250

1251

1252

1253

1254

1255

1256

1257  
1258

### C: PHISTb genes with DnaJ-domain

Gene ID <sup>1</sup>	Other Gene IDs or name <sup>2</sup>	PEXEL <sup>3</sup>	Gene Status <sup>4</sup>	MW <sup>5</sup>	Length <sup>5</sup>	KO <sup>6</sup>	References <sup>7</sup>
PF3D7_0102200	PFA0110w, MAL1P1.13, Pf155, RESA - Ring-infected erythrocyte surface antigen			126.5	1085	x	RESA-KO iRBCs have reduced cell rigidity (44, 65), RESA stiffens the iRBC and protects it from cell damage at febrile temperatures (71), binds to spectrin (68), expressed in ring stages of early gametocytes (51), part of an interaction network with SBP in the center (99), found in peripheral blood of malaria patients (77) RESA-positive, parasite-free RBCs observed (121) Subject of research in vaccine development (search term: <i>Plasmodium</i> Pf155 vaccine)
PF3D7_0201700	PFB0085c, PF02_0017	x		99.7	846	x	part of an up-regulated gene cluster in C4S binding parasites but suggested to be non-essential for PfEMP1 expression (86), HSP40 chaperone with a DnaJ domain, RESA-like protein (82), chromosomal deletion also affecting KAHRP and resulting in the absence of knobs and reduced cytoadherence (15, 122)
PF3D7_0220100	PFB0920w, PF02_0188	x		106.8	909	x	identified as type III Hsp40 (105, 123), deletion leads to increased rigidity and increased CS2 binding (44), DnaJ protein, putative (PlasmoDB)
PF3D7_1038800	PF10_0378	x		96.8	822	x	co-precipitated full length band 4.1R (39), identified as type III Hsp40 (123), RESA-like protein with PHIST and DnaJ domains (PlasmoDB)
PF3D7_1149200	PF11_0509, RESA 3 - Ring- infected erythrocyte surface antigen 3	x		117.2	1003		part of an interaction network with SBP in the center but also present in a different interaction network with RESA, SBP1 and other PHISTs (99), identified as type IV Hsp40 (105), absent in parasites treated with T4 which causes arrest of the cell cycle (124), less abundant in proteomic data than RESA (65), described as essential gene in Maier et al. (44), expression peak in late ring, early trophozoite stage (125), high sequence homology with RESA (126)



<b>PF3D7_1149500</b>	PF11_0512, RESA 2 - Ring- infected erythrocyte surface antigen 2	x	PG	89.0	839	x	although annotated as pseudogene, it has been found to be poorly translated with very low protein levels. Function not known, protein might be degraded (127), identified as RESA2 (65), mutation T1526C frequently found in severe malaria, up-regulated in vivo (128), identified as type IV Hsp40 (105), not expressed in lab strains according to Maier et al. (42) (2009) but highly up-regulated when comparing two isogenic 3D7 strains (33), and peak expression in trophozoite stage (125), over-expressed in in vivo samples (30), transcript found abundant in samples with cerebral malaria from Malawi (129) (see Figure S2)
	<b>PF3D7_1201100</b>	PFL0055c, MAL12P1.11	x	96.2	806		identified as type III Hsp40 (105, 123), RESA-like protein with PHIST and DnaJ domains (PlasmoDB)

1259

1260

1261

1262

1263

1264

1265

1266

1267

1268

1269

1270

1271

1272

1273

1274

1275

1276

1277  
1278**Table 1 d: PHISTc genes**

Gene ID <sup>1</sup>	Other Gene IDs or name <sup>2</sup>	PEXEL <sup>3</sup>	Gene status <sup>4</sup>	MW <sup>5</sup>	Length <sup>5</sup>	KO <sup>6</sup>	References <sup>7</sup>
PF3D7_0202100	PFB0105c, LSAP2 - liver stage associated protein 2	x		25.4	216		not recognized by pooled immune sera (98), continuously up-regulated, with predicted TM, granular pattern in IFA of older parasites and expressed in liver stages (130), located at periphery in liver stage parasites, similar to CSP (131), polymorphic gene with SNPs (80), differentially expressed in parasites under lumefantrine pressure (81)
PF3D7_0219700	PFB0900c, PF02_0184, GEXP20			35.3	295		
PF3D7_0219800	PFB0905c, PF02_0185	x		31.9	263		
PF3D7_0424000	PFD1140w, MAL4P1.223	x		35.2	296	x	antigenic properties, antibody recognition several times reports, potential vaccine candidate (29, 49, 111, 132, 133), stronger expression in culture when supplemented with albumin instead of human serum (119), String database mining suggests interaction with var2csa and PFB0115w (84), often referred to together with PFI1785w
PF3D7_0532200	PFE1595c, MAL5P1.313			27.3	226		involved in response to chloroquine treatment (134, 135)
PF3D7_0731100	MAL7P1.172, GEXP11 PfPTP2 - PfEMP1- trafficking protein 2	x		92.5	799	x	involved in cell-cell communication and plasmid transfer, located on budding vesicles from Maurer's clefts (43), Knock out with very reduced levels of surface PfEMP1 which was trapped in Maurer's clefts. No cytoadherence to CSA (44), there is indirect evidence for PfPTP2 to be located inside the lumen of Maurer's clefts (44), co-precipitated with Plasmeprin V (136), RESA-like protein found amongst soluble proteins from PV lumen and iRBC cytosol (82)
PF3D7_0801000	PF08_0137	x		147.1	1127		found amongst soluble proteins from PV lumen and iRBC cytosol (82), co-precipitated PTEX components such as HSP101 and PTEX150 (45), associated with J-dots (J. Przyborski, pers. Com.), <i>P. yoelli</i> orthologue is annotated as a putative dentin phosphoryn protein. Possibly involved in mineral nucleation or functions as an extracellular matrix protein (137),

					contains a relaxed PEXEL motif (1), elicits immune response (138)
<b>PF3D7_0830600</b>	MAL8P1.4	x	45.1	380	located in Maurer's clefts, no interaction with the ATS of PfEMP1 (19), expression is down-regulated in HB3, 3D7, and Dd2 (28), up-regulated under treatment with histone deacetylase inhibitors (TSA, SAHA) (139)
<b>PF3D7_0936600</b>	PF11770w, GEXP5	x	25.2	212	originally classified as PHISTb (15), reclassified as PHISTc (16), sole 'PHISTb' that showed only cytosolic localization (20), up-regulated in gametocytes (107, 140), earliest known post-invasion gametocyte marker, expressed 14 hours post invasion and independent of major gametocyte marker. Cannot promote gametocyte maturation alone when other factors are not present (51)
<b>PF3D7_0936800</b>	PF11780w	x	38.6	324	protein of unknown function, several TM domains predicted (15), interacts with the ATS of PfEMP1 (56), located at the iRBC periphery with PfEMP1, but not co-transported with PfEMP1 (19), stuttering motif identified not found in any other PHIST protein (141), non-canonical PEXEL motif that is correctly cleaved (12), suggested presence in figure displaying exported proteins (100)
<b>PF3D7_1001700</b>	PF10_0021	x	27.6	226	significantly down-regulated after Chloroquine treatment (80)
<b>PF3D7_1001800</b>	PF10_0022	x	23.8	204	corrected gene model with adjusted exon and intron sizes (142, 143)
<b>PF3D7_1016500</b>	PF10_0161	x	81.3	675	wrongly annotated as PF3D7_1016600 in (15), evidence for sumoylation (144)
<b>PF3D7_1016600</b>	PF10_0161 or PF10_0161a	x	27.8	226	had been annotated together with PF3D7_1016500 (15), evidence for sumoylation (144), IFA shows bright punctuate pattern in the iRBC (20)
<b>PF3D7_1016700</b>	PF10_0162		98.8	830	
<b>PF3D7_1016800</b>	PF10_0163	x	30.0	241	
<b>PF3D7_1148700</b>	PF11_0503, GEXP12	x	38.1	323	
<b>PF3D7_1200900</b>	PFL0045c, MAL12P1.9	x	37.6	310	non-canonical PEXEL. Localizes to small dotted structures in iRBC (12)

1279

1280

1281  
1282

## E: PHISTa-like and PHIST proteins

Gene ID <sup>1</sup>	Other Gene IDs or name <sup>2</sup>	PEXEL <sup>3</sup>	Gene status <sup>4</sup>	MW <sup>5</sup>	Length <sup>5</sup>	KO <sup>6</sup>	References <sup>7</sup>
PF3D7_0831300	MAL8P1.205, GEXP13	(x)		93.0	770		classified as PHIST protein (16) although no PHIST domain identified in Interpro. detected in early gametocytes (35) has a PEXEL motif (RKLSE), although this is not indicated in any publication yet
PF3D7_0831500				39.1	335		
PF3D7_0831750			PG	34.0	288		
PF3D7_0831900	MAL7P1.230		PG	33.5	287	x	deletion in field samples from South America (116), no deletion in samples from Central America (117)
PF3D7_0832200.1	MAL7P1.225			33.5	286		one of the earliest up-regulated genes in the asexual cycle (28), one splice form of MAL7P1.225 (GeneDB)
PF3D7_0832200.2	MAL7P1.225			34.3	293		same as PF3D7_0832200.1
PF3D7_0832300	MAL7P1.224			30.8	259		up-regulated during the commitment phase for the sexual developmental cycle (53, 145)
PF3D7_0832700	MAL7P1.220		PG	30.8	260		
PF3D7_1000700	PF10_0007, PF10_0008	x	PG	26.1	288		both old Gene IDs have different domains (27), PHISTa protein with N-terminal PHIST domain (15)
PF3D7_1201200	PFL0060w, MAL12P1.12	x		16.0	134		down-regulated in several strains (28), PHISTa-like protein (15)
PF3D7_1301300	MAL13P1.58		PG	28.9	247		not located in telomeric region (27)
PF3D7_1372300				24.2	206		
PF3D7_1477300	PF14_0744, Pfg14-744	x		22.8	196		strongly up-regulated at gametocytogenesis (27), expression peak in stage I gametocytes, present also in stage II (85, 145, 146), grouped as PHIST or PHISTa-like (1, 15, 16), tested as a gametocyte marker (129), expressed as early as in committed schizont stage, regulated by AP2-G, and HP1 is reported to have a silencing effect (54)
PF3D7_1477400	PF14_0745			36.2	303		expressed in parasites committed for the sexual cycle (27), found in nuclear proteomic fraction (147), classified as PHIST protein (16)

1283

1284 <sup>1</sup> current Gene ID from PlasmoDB  
1285 <sup>2</sup> other/previous names  
1286 <sup>3</sup> presence of a PEXEL motif indicated by 'X' as identified by (15) or (1)  
1287 <sup>4</sup> annotated as pseudogene (PG) in PlasmoDB.  
1288 <sup>5</sup> molecular weight (MW; in kDa) and length (in amino acids). For proteins with a PEXEL motif molecular weight and length are calculated for the PEXEL-cleaved  
1289 form of the protein (MW for PEXEL-cleaved proteins was calculated using the ExPASy web tool ([http://web.expasy.org/compute\\_pi/](http://web.expasy.org/compute_pi/))).  
1290 <sup>6</sup> existing viable knockout (KO) parasites or natural gene deletions are indicated by 'X' (44) or references as indicated)  
1291 <sup>7</sup> Information or references referring to the corresponding gene or protein.  
1292  
1293  
1294

## Figure legends

### Figure 1 – *The Plasmodium* life cycle

Life Cycle of *P. falciparum*: Upon the bite of a *Plasmodium falciparum*-infected female *Anopheles* mosquito, sporozoites are injected into the dermal tissue of the human host (right). Sporozoites quickly enter the bloodstream and are transported to the liver where they invade liver cells and develop into liver schizonts. Through merosomes thousands of merozoites are released into the blood stream where they invade erythrocytes, starting the asexual replication cycle. Each cycle, a few parasites cease replicating and commit to the sexual cycle and develop into gametocytes. Mature gametocytes are taken up by a mosquito during a blood meal, rapidly develop into male and female gametes which fuse together in the gut of the mosquito (left), form an ookinete that penetrates the gut wall, and undergo sexual replication in the oocyst, producing thousands of sporozoites. At the end, sporozoites migrate to the salivary gland ready to be transmitted to a human host during the next blood meal. Figure obtained with permission from Lee et al. (148).

### Figure 2 - PHIST proteins in the remodeled iRBC

A. The remodeled iRBC. During the asexual blood stage of *P. falciparum*, the human erythrocyte is subject to extensive remodeling. All erythrocyte proteins are colored in gray (reviewed in (59)). Parasite structures, compartments, and organelles are labeled in bold characters. PHIST proteins are labeled in bold red letters, and all parasite proteins are represented by colored shapes. References on PHIST proteins as indicated in their respective sections. Knobs are parasite-derived protrusion in the host cell membrane with KAHRP being a prominent protein of these structures (149, 150). Maurer's clefts (reviewed in (151)) are parasite-derived membranous structure

in the iRBC cytoplasm involved in protein trafficking and are connected with knobs via actin filaments (152, 153). MAHRP1 is a Maurer's clefts resident protein (154) that potentially interacts with MAHRP2, the tether protein anchoring Maurer's clefts to the iRBC membrane (155). REX1 (156), REX2 (107), SEMP1 (157), SBP1 (158), and PfPTP1 (159) are other exported parasite proteins localizing to Maurer's clefts, with the latter two being located in a high molecular weight complex (159) and SBP1 interacting with erythrocyte cytoskeleton proteins spectrin and band 4.1R (160). J-dots are mobile, dot-like structures in the iRBC (161, 162). PfPTP2 is associated with exosomes which are parasite-derived vesicles that are involved in cell-cell communication between iRBCs (43).

B. Cytoskeleton of the uninfected red blood cell. Reviewed in Fowler (59) and adapted from Maier et al. (42) and Spillman et al. (100).

### **Figure 3 – Phylogenetic tree and protein domain prediction of PHIST proteins**

The amino acid sequences of all 89 PHIST proteins were obtained from PlasmoDB. Of 19 PHIST proteins annotated as pseudogenes, 16 contained one or more premature stop codons in the amino acid sequence and were removed from the list. All those of the remaining 73 PHIST proteins with a PEXEL motif were PEXEL-cleaved *in silico* using the PEXEL motifs as provided by Sargeant et al. (15), Boddey et al. (11), or Schulze et al. (12). The amino acid sequences were aligned with MUSCLE (163). The alignment is represented in a phylogenetic tree using the phylogenetic tree tool built into MUSCLE on the EMBL-EBI website (164). The branch lengths are drawn in cladogram style and do not represent actual phylogenetic distances. The colored bars next to the gene IDs represent the different PHIST subgroups as indicated.

The structure of the amino acid sequences was then analyzed with InterPro (<https://www.ebi.ac.uk/interpro/>) (165). A schematic representation of the results for each PHIST is shown next to its respective gene ID. The different domains are highlighted as follows: PHIST domain (brown), DnaJ domain as defined by Pfam (PF00226) (light blue), DnaJ-containing protein with X-domain as defined Interpro (IPR026894) or Pfam (PF14308) (dark blue), acyl-CoA binding protein domain as defined by Pfam (PF00887) (green), coil domains as identified in Interpro (yellow), signal peptides as defined by SignalP or transmembrane domains (grey), and the MEC motif as defined by Kilili and LaCount (39) (red).

#### **Figure 4 – Sequence alignment of PHIST proteins**

The processed amino acid sequences as described for Figure 3 were sorted into the PHIST subgroups and individual alignments for each subgroup were obtained with CLUSTAL (166). The output was analyzed using BioEdit (167). Conserved amino acids at a threshold of 83% were highlighted. This figure shows the core of the alignment using the position of the first conserved tryptophan to align the different alignment blocks of the subgroups. Tryptophan residues conserved above or below the 83% threshold are marked as indicated.

#### **Figure 5 – PHIST domain modeling**

A. Visual representation of the position of the conserved tryptophan residues of the PHIST subgroups, treating PHISTb and PHISTb-DnaJ as well as PHISTa and PHISTa-like/PHIST as separate subgroups to show the slight variations between them. A consensus sequence was generated for each subgroup alignment from Figure 4. The consensus sequence was then split into blocks of ten residues represented by gray boxes (boxes containing no tryptophans were shortened). The



positions of the tryptophan residues conserved within a subgroup are marked by yellow bars.

B. To exemplify the differences in structure prediction Jpred4 (168, 169) output is shown for the PHIST protein PF3D7\_0532400. The upper sequence is the consensus sequence used for PHIST identification by Sargeant et al. (15). JNetPRED displays the consensus prediction: helices are marked as red tubes, and sheets as dark green arrows. JNetCONF provides confidence estimates for the prediction. High values mean high confidence. JNetHMM profile based prediction: the six predicted helices are marked as red tubes, and sheets as dark green arrows. JNETPSSM based prediction: the four helices are marked as red tubes, and sheets as dark green arrows.

#### **Figure 6 - Heatmap**

A. Heatmap of *phist* gene expression of the *P. falciparum* 3D7 strain over the course of 53 hours: Microarray data on transcript abundance was obtained from (46) (accessed through PlasmoDB, Release 28, March 31, 2016). The heatmap was constructed using TMeV4.9 (170). *phist* genes were clustered into their respective subgroups and were then ordered by gene ID. Genes annotated as '*phista*-like' or as '*phist*' were grouped together as '*phist* others'. A yellow box to the right of the heatmap indicates PHIST proteins present in the early gametocyte proteome (35), differentially expressed *phist* genes in pregnancy-associated malaria (PAM) (29, 49, 50, 84, 92, 112), and in cerebral malaria (32, 34), or if variant expression has been reported (55). Green to red colors represent fold changes (log fold changes from -6 to 6).

1396 B. For a number of *phist* genes, no expression data was available in the data set and  
1397 those genes were grouped first by PHIST subgroup and then by gene ID. Additional  
1398 information for selected PHIST proteins is provided in Table 1.

1399 The colored bars next to the gene IDs indicate the PHIST subgroup.

1400

1401

## **Authors' biographies**

**Jan Warncke** obtained his B.Sc. in biology with an emphasis on microbiology from Brigham Young University - Idaho, Rexburg, ID. He then moved on to the Swiss Tropical and Public Health Institute and the University of Basel, Switzerland where he received his M.Sc. in Infection Biology. Currently, he is enrolled in a Ph.D. program at the Swiss Tropical and Public Health Institute, working in the Molecular Parasitology and Epidemiology unit under the supervision of Professor Hans-Peter Beck. His interest is in exported proteins of *P. falciparum* and their role in host cell remodeling during the asexual blood stage of malaria.

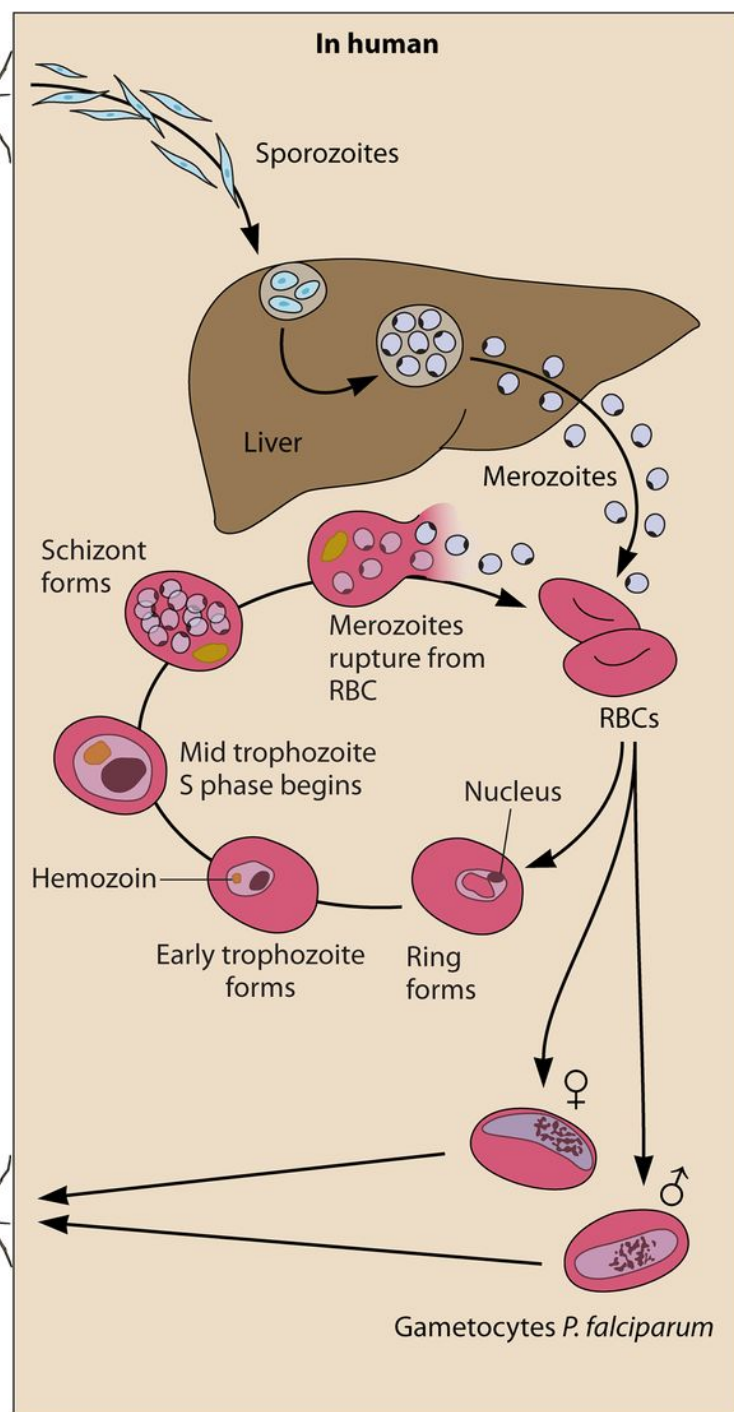
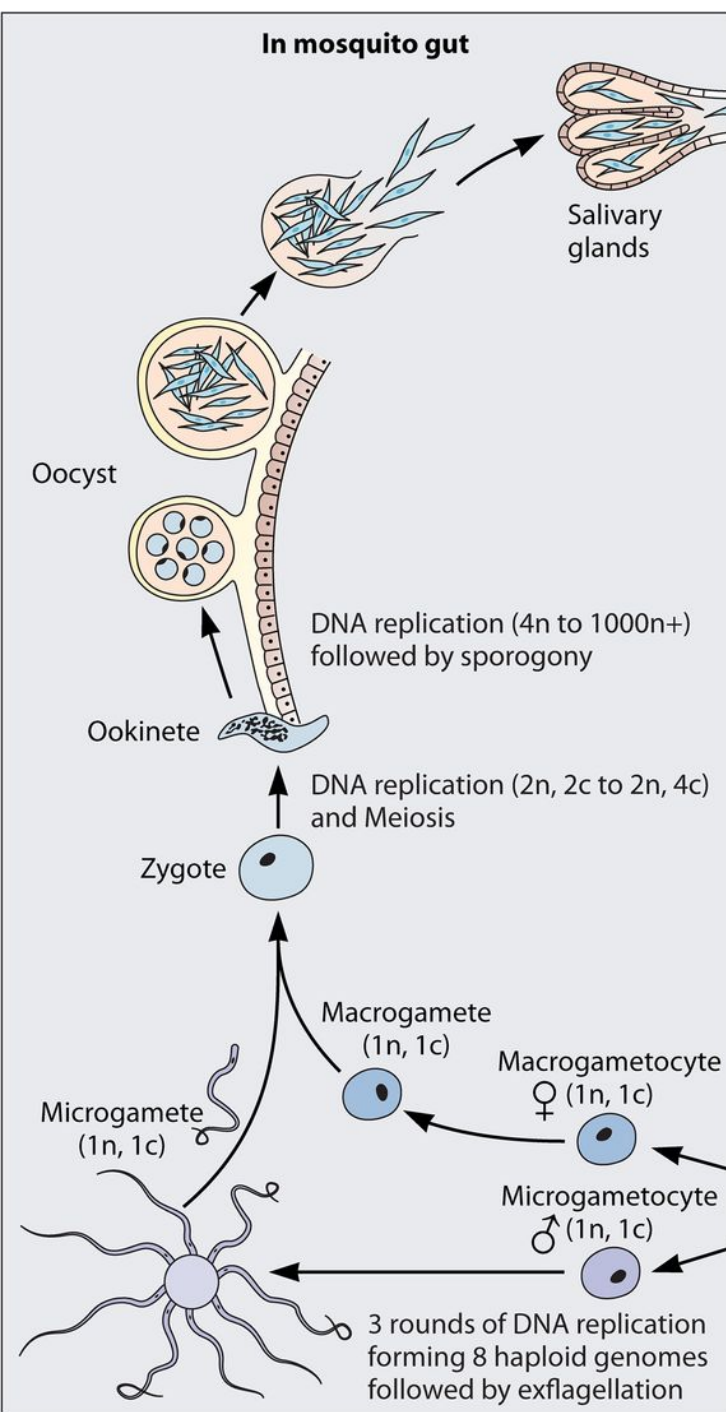
**John (Ioannis) Vakonakis** has a B.Sc. in Biology from the University of Crete. He obtained his Ph.D. in Biochemistry from Texas A&M University, where he pioneered the structural analysis of circadian clock proteins from cyanobacteria. His postdoctoral work at the University of Oxford focused on the mechanisms behind cell adhesion and the formation of extracellular matrix in animals. He was trained in X-ray crystallography during a second postdoc at the Swiss Light Source, prior to starting his own lab in Oxford Biochemistry as Wellcome Trust Career Development Fellow. John is now Associate Professor in Structural Biology and Biophysics at the University of Oxford. Over the last six years his research aims to understand how large molecular machines form in cells, such as the cytoadherent assemblies created upon *P. falciparum*-infection of human erythrocytes.

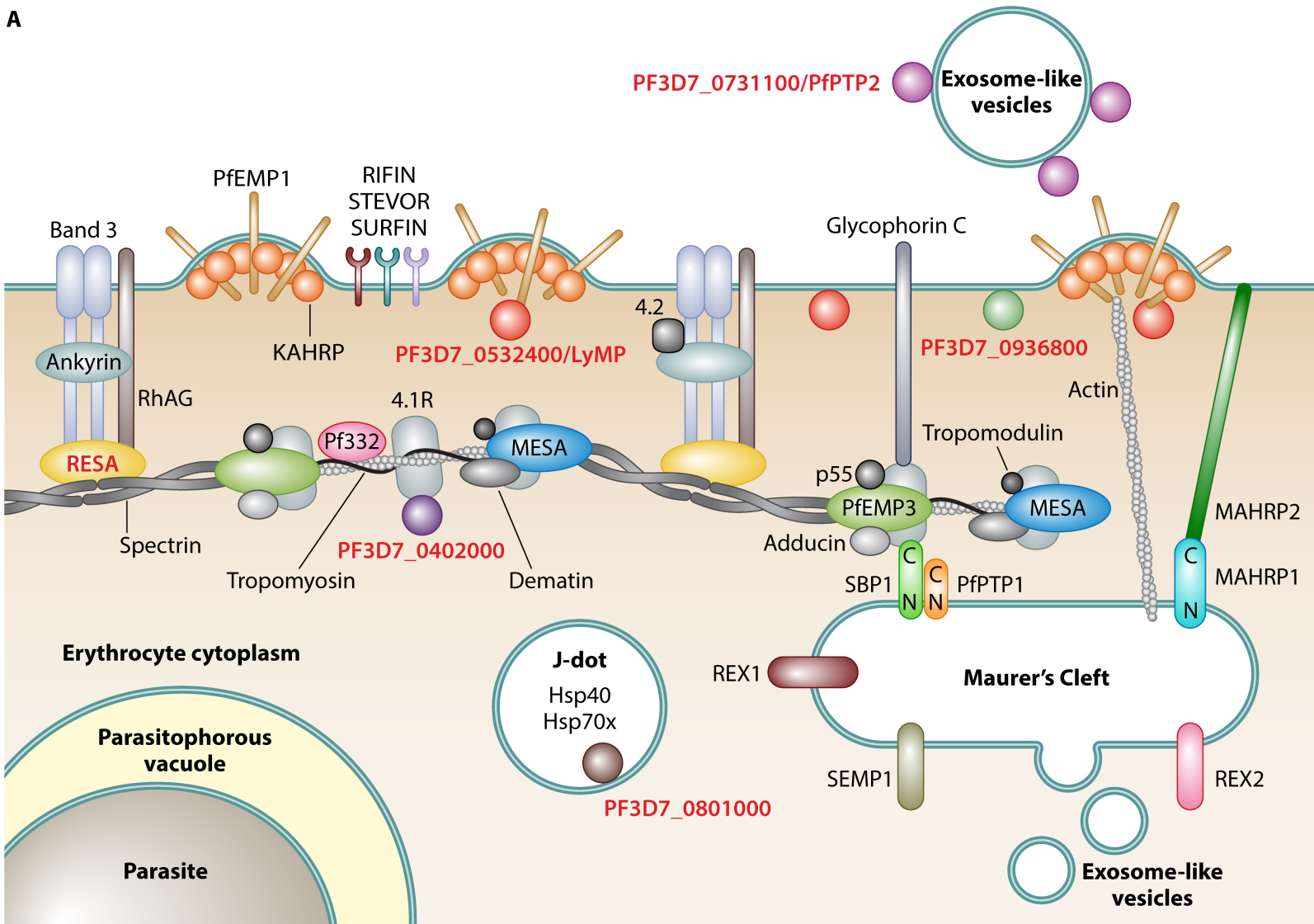
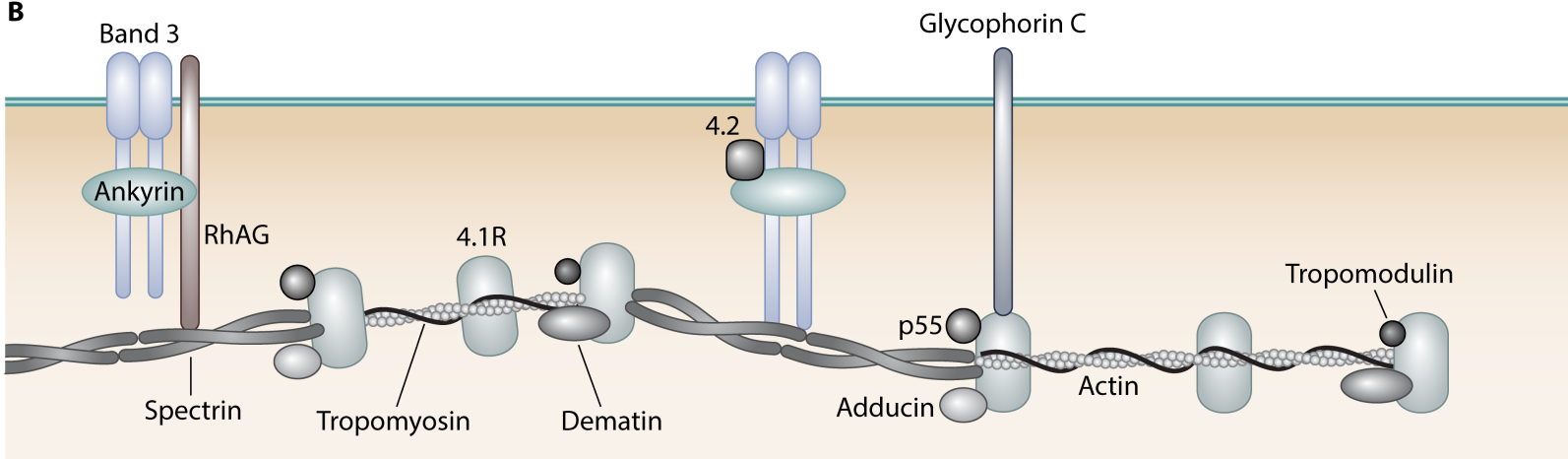
**Hans-Peter Beck** obtained his MSc and PhD in cell biology at the University of Tübingen. After his Post-Doc position at the Wellcome Centre of Molecular Parasitology in Edinburgh/Glasgow where he worked on the bovine parasite *Theileria annulata* he moved as visiting scientist to the Walter and Eliza Hall Institute in

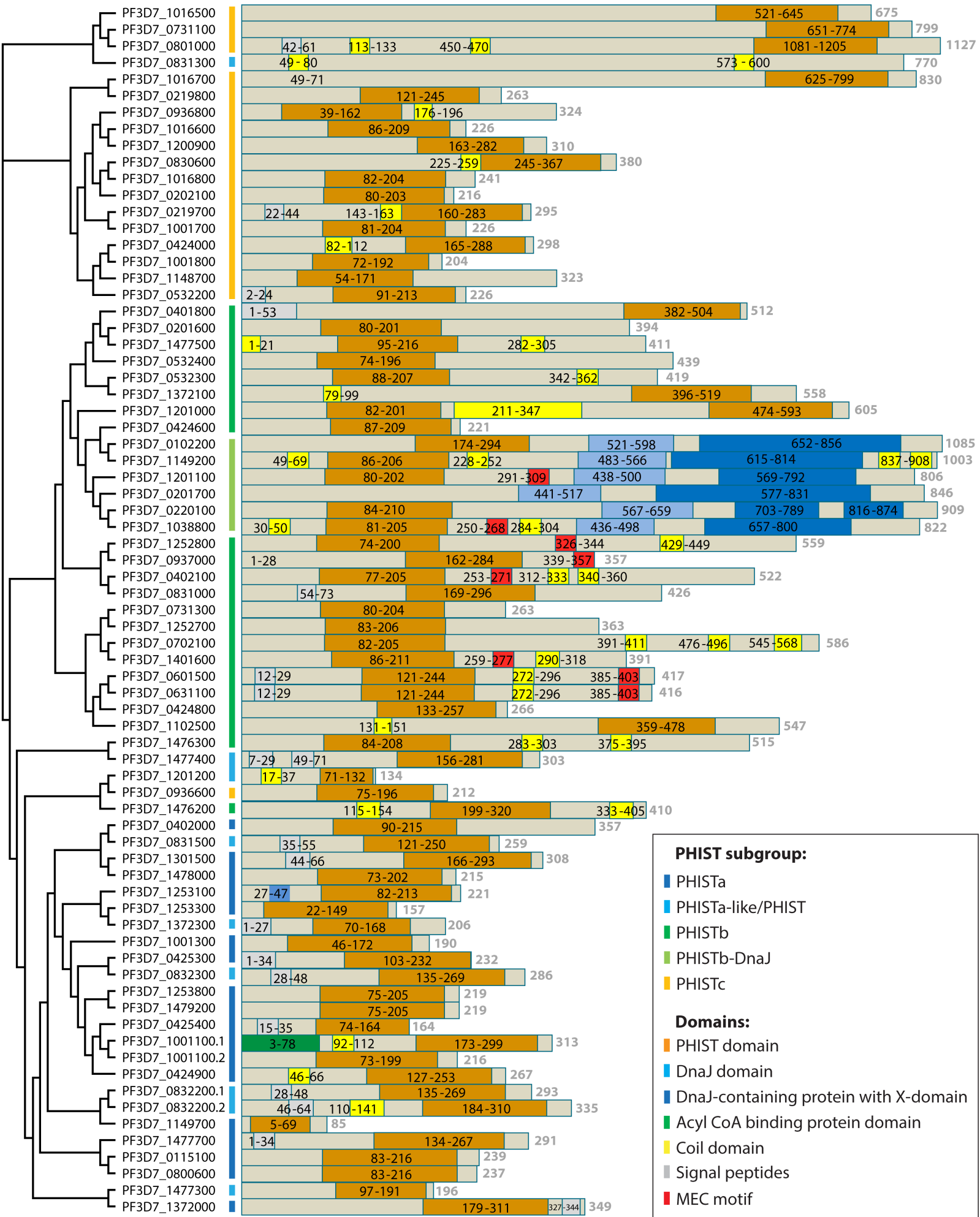
1428 Melbourne before he took up the position as group leader in Molecular Parasitology  
1429 at the Papua New Guinean Institute of Medical Research. During this period he  
1430 pioneered molecular epidemiology for *P. falciparum* infections and established a  
1431 research group working on the cell biology of the parasites. After five years in Papua  
1432 New Guinea, Hans-Peter moved to the University of Witten-Herdecke, Germany, as  
1433 group leader and subsequently moved to the Swiss Tropical and Public Health  
1434 Institute where he holds the position as unit head of a research group and holds a  
1435 Professorship for Molecular Parasitology at the University of Basel. His research  
1436 focus is on molecular-epidemiology and cell biology of the human malaria parasite *P.*  
1437 *falciparum*. His research group in particular focuses on the interaction network of  
1438 exported proteins of the parasite.

1439

1440



**A****B**





PHISTa

PF3D7\_1477700 VVYVYNSRNREINARFGRNDDTPRNDISVNHVRRNRDGMVINSLLYVNVFNRYNKGKGLDVKEILEELKIDKESLRLFKTSSLKESSCDFKRYNDYITNNEKKFELKDLNSYKFADDTKKKIYH  
PF3D7\_0115100 DINYNLTNNNRKRLALETNKTTPRNDINRSHVVGNRGEIDEIDKIAYTEYVYNKHNKHNDINEMKTFNINNSKGNDFKIECVVKKACEDSVFSQRFYKARKKKKKELKKLKSYYADAKKKEYE  
PF3D7\_0800600 DINYNLTNNNRKRLALETNKTTPRNDINRSHVVGNRGEIDEIDKIAYTEYVYNKHNKHNDINEMKTFNINNSKGNDFKIECVVKKACEDSVFSQRFYKARKKKKKELKKLKSYYADAKKKEYE  
PF3D7\_1301500 NLNYNLTTFKFEIPDFQNSKIPSDDEDENAKHHTSAKEGLGRKKYFVLGDMDEYEE--QKQKCFCKRKEISEDFMDEINETSSEKSTYKFKYKARKKRTERKNFPAFCKKINLYLH  
PF3D7\_1253100 VHNNNFASKQFPEFYDYTYVVFPERKKILLAWQITGYNQGLDLKKEYNFLPNRVYRNKQNSE--GRSFYKIIISREKIDECVINCREDAIKKVELTKGYIDANTKKKARNFESQEQKANNFCN  
PF3D7\_1372000 NLNYNLTTRQNSYISDQKGNINPNEDVNMVQITGCKSGGHHISNNKNEYKSCLSYKAILPHYGGFPREKNPEEYKCKLYDFEVKLNASVEYNSKTYDANKKDFRKEFEFSCLEFSNKKQLYK  
PF3D7\_1253300 DSKFESITEQITREYEFDLVQVFPRTYLLNESQNGCRQGTGKLLKNRGIAPKLTYRHIG--NPNYVGKSPPKITWGQCSYDCNMNSTLETEQTNRYNINKKAPIDEIKSFIRSCIDEFDKHTDLVY  
PF3D7\_1001300 TILINDSRQFPEFYDYNSIEIPPKRVINLNYQS-LNIAKDMRELKKEGYVDEYLNKNPN--YDCDNFINDNNSINDCLDNDIVETSYEEMIEREHNANQDITNLVNEFYKONFYFDEKKKLF  
PF3D7\_1149700 -----MPCIRNTKRYINYSKMSHDIGVITSSITYMEHALNFYIKYGTSSRKKKRYVFKYYDTNKKDLFN  
PF3D7\_0402000 SCRYNDNKKFIECKRNINSDEIPPRNDIEKTNHAKTANSGETSRKKRKEYEQRYGRCYEERPN-----RFGSYEQVLISQPHFENRKHVENDTVTFYEDLDKDPTECKNNTSFGQNLIDFLN  
PF3D7\_1478000 DINYNDISQITEQFDVNSKVPVPPKENIHNHHTIGNKKGLDLIKDLKYFPKNSISDRFELL--EFRMTEDCYTKMFDMSDFGI-ATMNETEHTRSFYGERQPSCLRNYSFSEYQRYNELYE  
PF3D7\_0425300 NKCYNDMSKITEKIREVNSIECPSKEYIMNHSHTGAKEGLDNLKEKELIQEYLDNDIYPIF--NKYGANTFLYDITWKGILFNLCGTASEEVKTKSLSLNDKHTCLLNLYLFYFQIKKELHE  
PF3D7\_1253800 SINNNDSKNTEKREYVNSIECPSNEDRKNTNTITAKGGFLLQKEKELIQEYLDNDIYPIF--RC-IGKFGYGSIDENIFRFYRTANEVEQYNSDNYRNDNDEHTELKLFYYSFCHFKTKKELQE  
PF3D7\_1479200 SINNNDSKNTEKREYVNSIECPSNEDRKNTNTITAKGGFLLQKEKELIQEYLDNDIYPIF--RC-IGKFGYGSIDENIFRFYRTANEVEQYNSDNYRNDNDEHTELKLFYYSFCHFKTKKELQE  
PF3D7\_0425400 HINYNDSKNTEKREYKSFKECPKKNRNHHTGAKEGFYNLKDKGSIQKYLDNDIYVGV--HEDDDKIYLYERMKNYISRFYETINEEK  
PF3D7\_0429400 NINYNDSKNTEKREYNSFECPPNDDRNHNSAKAKEGFDNLKEKGIQNYLDNDIQLKN-G--KLLYKDTNDRIVLKFCTIAEEAETRKEYSNDKHTNLVKLYSFENFTKEELHE  
PF3D7\_1001100.1 NINYNDSKNTEKREYNSIECPNEDRNHSHAGANEGFDNLQKEKTLIQKYLDNDIDLGN-G----KVQYNDIKGIVLSFCTIGTEVVEVTNNFLGNREHTCLLKFNSFDFKTKVKLHK  
PF3D7\_1001100.2 HINYNDSKNTEKREYNSIECPNEDRNHSHAGANEGFDNLQKEKTLIQKYLDNDIDLGN-G----KVQYNDIKGIVLSFCTIGTEVVEVTNNFLGNREHTCLLKFNSFDFKTKVKLHK

PHISTa-like/PHIST

PF3D7\_1477300 NIKLEIEISPLSKLISHISRNKDVFEYMMNLQAFVSNDLV-DVNLNEYIQYVKNQVQKIE-----LTSNGNSYKKNCNVELAQAFLFEHIEYKTKYKVIITRHM-----  
PF3D7\_1201200 DGKYDTLEQWNEKQNKKLDLDIHSIDRLILQAQIGQSERIKTYSVYTRIRLFKSLKEYNLKQKDKIEF-----ESINIIILRNFSHEVSEKYPPLLMYKNRKIKLEIKPFNFIIVQYIWAFFELDRIKRDI  
PF3D7\_1477400 TKKCTDLASFQKDIYILKSEIPKCDLGAQYAIYIARNS-KQHEKKEYMCHPENYKPEVLCDFK-----  
PF3D7\_0831300 NIDYNKFIKDLKEDRNDLPHGSSYPLKDLGKWEALSIKGLFELTRDFKSMYTYLDKYEYHNRSTY-EYPIFWISYIGSWGRCILNLGMPASEQKFSRNHYLKNKTANLEDDINFIYSGMMYFEKLKRL  
PF3D7\_0831500 YLYNNDMTQLTKLAFYLLSLKCKPFGRRNANWHAHALGVSIDLEDEKIDINVIQYKKKY-----ESRRKQSYR-ISKSGKNSYKLEDFERMEQRTYSSNKKHLHKSSELDDIKNFHTFIDDLKELINYM  
PF3D7\_0832300 NKPNVNDMSKNTKREYVNSIEECFPEEDINWHAHALRAGKGF--VNVNVEKILQYLDNDIYNTCCCKELVY-----RSWEKNCEFFDYDELVELYKSYNGFELNGKHTLDDILEYISFLQFILKRL  
PF3D7\_0832200.1 NINYNDSKQLTLKSHSLDDLEKSTTEDYNNVQVLGIAKDGFDGLTESYVVEEYHLEYEYERYHYGGRRFVSMKNVRETWYKSMFIEGALSTSDVKNTLDYYSFKSGASIDEMKNFYEFIKYDTLKNL  
PF3D7\_0832200.2 NINYNDSKQLTLKSHSLDDLEKSTTEDYNNVQVLGIAKDGFDGLTESYVVEEYHLEYEYERYHYGGRRFVSMKNVRETWYKSMFIEGALSTSDVKNTLDYYSFKSGASIDEMKNFYEFIKYDTLKNL  
PF3D7\_1372300 DSKFESVTEQLTRREYEFDLVQVFPRTYLLNHNHNGICRQGTDLKNRGIAPKLT-----YRHTGNTNQYKGSPPKITWGQCSYCNMNSTLETEQTNRYNINKKAPIDEIKSFIRSCIDEFDKHTDL

PHISTb

PF3D7\_1476200 NLSVQISLEF--QAL--NSLDSVLSAKEHNNVNRGIRFLLEDTVSDFTVYHDLLETYNL-TAKETEHWKNCISDARHSQSYAETQCNVDHNNIRKK--NLTQEAKNYITQSEKSKLRDEHDKGETY  
PF3D7\_0401800 AHGKYNNERVENKIDIE--DGTDTQYRKYSLSNENE--EEDIDIRKLCYCASKNKSLSHR-TPEESEQAWKYVQRLIKLYYNDFDKHFKEVLLNALS-EGFDVEKMLAANLLWKLSDKQESKVM  
CSLLSDSNITKEINNDDQIT-DHINNRPKYSVWDMKNEKLYNLTIKYKHHQNKIEKYKNINQYIAEMQWKECRKVIILARIKERYKINNI--CKINKP--TKTSSTFKKINKNLTWMLTNYIQTYSSES  
PF3D7\_1252700 NTFIDNVNNEEDDEESFLN-GGNYDREKINRWQCMRNELYSVSVNKKYKVVYKARKKYNV-DNSYAEKNWNECKRNIIVGRVEYQNYINKVLDLNGD--SINNDIKILRTCVTKKLRTEIKALKKKS  
PF3D7\_1401600 YSELASMTIEEDDEKIME-YTNNNREKYSVWQCMRNERNVYVIMKNNYFFELKKKYNV-DNSFSEKQWNICKRVKKRNLRYERSINTTLNKNLS--LLRKDYRKLKANLKHLSIKTERKACHNN  
PF3D7\_0831000 NYPFVKNKKEKMKQKLLSINEDNKKEHKKITDFTDKLNSLDELKELSLDFEEKSYQNI-KDEYVNNWNECLDLIETSGMSMNTNLNAYKKNYKT-KVLIDNIFFCGILVWMLFSSSEISCKDI  
PF3D7\_0937000 FPIVLMDPKLEINRKYTDY--FFKNKKRYRTINSEMKF--LNLFLFKHEIEFYKNMADYNI-SNTYVGNFVCKYMLVMDILNSEYHNLVLYNFKNS--LKKTKYKMLNACVYALYLRKHQSCEDI  
PF3D7\_1476300 IKGTYNNKILKYNKKEIN--DEILDKKEIKYNNKSNNEFYMYITKVKMLMHRKIKYKE-DNIFTNDLWKKDCSNYLSEKKNVHLLKEMDDLVHG--HKLKDFKIWLGTLIYQLLNNKKNKEIK  
SGSMVLNEKIOIEKELKG-DNNNIRESKYVNDLRNRRRYENLRDSYKFYNDKRSVHKV-PKSFAEERWQCECNILIEGQNNVTMDNIYDIDNNDNNNNINERNIMDANFAAKNNKSCCEGL  
PF3D7\_0402100 NSDKRYNQEGEGRERDAIK-YGCSDRMLYTNWCMKKNEDLYLLKEBNHFYEMRQSKFV-PLDISNRLWCSCNDIIFINGINNVPYIDGIEIFNNT--ETNIRKRIKIDATLSWKLKEDLNSCKEI  
PF3D7\_0631100 NSDKRYNQEGEGRERDAIK-YGCSDRMLYTNWCMKKNEDLYLLKEBNHFYEMRQSKFV-PLDISNRLWCSCNDIIFINGINNVPYIDGIEIFNNT--ETNIRKRIKIDATLSWKLKEDLNSCKEI  
PF3D7\_0731300 KHRYSNENKKNAYEYKAK-SKKNRYEKYQYRTVQNEENILSVKRYEVEFNFYKYSV-HYKSEEEIMNELRTITFDLRNVTFNMMDLEIRKIDN-DVLNKGHALDGNVWNLNKKETAFAKEI  
PF3D7\_1102500 ENGLYIDDLKLENNLSN--SKNDVNTLRDNNWNRNKEISVITRYNLSLYNKNKYVDV-FVHSTNELNNAYGKFLHNNYIKSYFNGINENIKTS--EGNIKSRILIAACILWKKKNITLQSLHEI  
PF3D7\_0424800 KCTLLTNKMLQSEDLIG--NESEETNKLDDWKKMKNESLTYLLKRRYQGHKLRNRSRL-SHDKLNNLINCENVLVKYNNYMDKINTNIEKASTVT--PHNIFKRLKACLTWKSISKNAHAYTEL  
PF3D7\_1372100 ETNYSVQKTFSNEDSSK--SYNMCVKCSDNWNMKNRGTFNHIITDYRYFRTIKKKYV-SGNFAFNWETVNSNRLRYKNMAEYILNTIENINNN--RLNIGKMLMNNFLWKKTKKDYDINKNI  
PF3D7\_0532400 YMGVSQDITTYIEKELGL-DYENIIDSRYKLNMMKNGKFLGLKNRYKQFEFFPINVT-INNNKKRFDCEYRERIDLGADIIENGLNKLKSCIK--NVLDRKNQKVTANFLKRLSLDQECQNV  
PF3D7\_1252800 HLIPFYNEKNSNNNDSDG-RDNLIRLKYNTLNMQCMQYVYGRHSNRYKYNKNQYNI-KIESNDKALKECNEICLSAHLQRLHNNMNEFKK--GFVSISQNTIANHGVVLTITNKKKYDIT  
PF3D7\_0424600 YNPFVYKNAFVLYYDNGY-CKGDTIKAEIIEATNMMRRYATNENKRYLYDDIRKRII-RPRTYTNKLNSCNQIMSESEKKIESHLSSEMNKFFQ--KSLFLDDNRFTACIGVLTINAKHQCDRI  
PF3D7\_1201000 FGLVYVYKINIEADADNDE-LSETFKDDAYNFWKMKNESYFQNFVDYMYLEDTKKNLIL-DNRYEDKLNECRHIVELPAQKLPLGLTRIEDNTN--EILYKKEFTILNSVATWALTNYALVECKNIF  
PF3D7\_0532300 DIRNGKHEMININIEFKCE-LDDN-RKKDFKETEPEKKNYQIYDSTYVLETRKSNKVV-PKSFAEERWQCECNILIEGQNNVTMDNIYDIDNNDNNNNINERNIMDANFAAKNNKSCCEGL  
PF3D7\_0201600 GNSEYNDILENNKVLNN-ETNS-ELRHITNNMKSEDFNSWYMYFNNNFYLRKKYKT-PFNYAKPTCNQCEFFALSCKKYIENSFNKVNKFKNN-VYLDVNEFRVIMACILLKRTLATKEEGMLY  
PF3D7\_1477500 LKGIYINKYPPENRNILNK-ETRN-RRKNEHYLLCNHRLNENFLHEYFSELYDTRNSHKS-PMYAKETFOOLNEFVVLGKYVVENTCSNAKKKLKNN-DDLNICQENLILACILWRTLTKKEGRKY

PHISTb-DnaJ

PF3D7\_0102200 EII---VQVLFYNEVDNGSGGDLKKQTLNIDINKRKYDSKEQITYSYVQYDPKLYE-SKTQCCLDQCEGNERNSCKNRYRQYINLALRVTLVNLAKALSNQGYTRKNS  
PF3D7\_1149200 EII---VQVLFYNEVDNGSGGDLKKQTLNIDINKRKYDFTKTEQITYSYVQYDPKLYE-SKGQCLLDQCGDNERNITQKKNRYRYNIEEERTVNLAKALSNQGYTRKNS  
PF3D7\_1201100 EYT---YNNKDIIRKHEMVKRQHRLEDSDSEVLNNESSYNTYNEKLYYDLYQKLLKRRND-DIINNYSKFEIEKKELEYKINNVVYKYAYNIEVNSCKAKALSNQGYTRKNTIY  
PF3D7\_0201700 EYF---NNKQDILNNDEENRELDITKRLNINKNEKNFASKCFNQNYPNNHNPKEQYE-KGCNLCKEIEIEKYKILNSVYENYDQVQALYKIERCKAKALSNQGYINLN  
PF3D7\_0220100 VYNNNNNNKKDDILITHHNGNGEDLKTILMLFENEKEYYLKGRIYNNFNNKYTIDYPRKKNNYNDYKSKLDEANMKMYEYKKNNEERKATVLCIGKALYNNIKKSHS  
PF3D7\_1038800 EYK---YNDIDILNEEDNGNGEDLSENNNGVNNNEKSILSLTHRHYYNIVYKPNY-HNAKKECYKRMGEIDKRNKMKCEFCQKLYEGLSVTACILAKALSNQGYTRKNSV

PHISTc

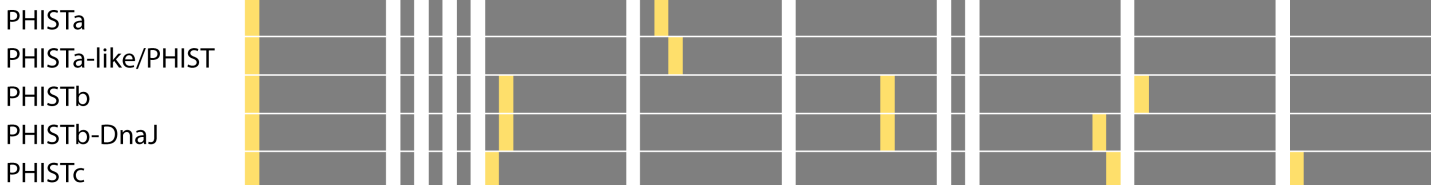
PF3D7\_1148700 GYFNNLNFNLEYTHLQTS-DPKTSIQCTINFDNENNAKNHYIEAENVSLILAHQNDPEKHRKSTCKIKHDAQCDFIKEHPNHSNDVFIRKI---QSIDFESYKECIIHSNNFTEKKLQYFFKY  
PF3D7\_0532200 -----RAAKEYIPIRINICTFDRYTSKNVYDITYKHIDYKAFRLRVDEWFTITFAKENKSENYSKYWECCNIIQSNLTKNDDRGNDKFLSQPDPFPTQETIDNRKILNDEFMFKKNDMLLIK  
PF3D7\_0731100 -----HEITEDLTMMRSDNIPSINLHWQREVEHHEAMQYRMKYIGHLNTYDLEDFAKDLVAYSRYATKLEKKNFYKNVYSFIKRE-A-VTKHALECFSDCVKSHQYMLYNKKKQCDRIA  
PF3D7\_0219800 -----KNLIDKIDIMQRHHDITKDKFSTWINRYNRYRVYVDMNNEWSYVKEESDKNNFSDIAFNKIKWLYPPISEFREKNNNYNDFSIFNKE-I-CDPDIYNNFNTTRKNNEIICVMRYKWITIP  
PF3D7\_1001800 -----E-DMS-SINKQKENVPSKDIRNWFKHCNEINPHSIIYSRKIHDKYFSDTNKNEELGKKVNNKCCATTEELLKLSIQNSQHNLMQGE-I-VTLQEDDFKFSVNGYKNTKKTETKICIKKK  
PF3D7\_0424000 -----NLISDCHIDMNNNDIASFEFHNNWNECKSEYFELYIYDRKLYEELIDDIDIEEEELLETVCVFGKIQVRASGYNDLNNLLSDE-Y-VSKKEFDDEKERNRLSIRSQLKDRCKEKI  
PF3D7\_0936600 -----KQWEDKIEILNSLDIPNKFFULTIMQGRSEILMHNLYGRLLYKELINKYNPMESCTHILNCYDITNYHLIYHENNINLDTLTK-E-LTRTDFDFKDTKNKNKLRNELRETVRKDI  
PF3D7\_0936800 -----INKCDMKKYTAHIMNNSSENFINKNNIISYHESDEEKKKEENFKFIQSIVETYKPDYKMRKFAFAHFGQYALKCKELL-EAFLSLNG-KLCERKKEVEYVKRDEIFRKSMEFVWKEKA  
PF3D7\_0830600 -----NEDDEEFHEKMEKNNNGTLGTCRNWKCFIENKFTYNNQMKWNFCEILAFEYKQSTLANEKKIYTYNEELIQHNLNISDENTLVFNK-I-YDALDFEFKHKRESKQFRESEEPVKNNT  
PF3D7\_1016800 -----EVLTEFHNKRNIIRGGIFLYEYGNHYNQRERYTTRQEPKWVCEDLSVKYNPHECTMKKEKKVRYKKEELLKKK-EDERDQYFVEDN-I-YTRWEYAFIDTKRESKVFRRHMKERKWDML  
PF3D7\_0202100 -----KVTEFLERMIEFGKLERKLYLWHYSHSLDIFNKKSSWSICGKLAHEKPFKIKMKKWKCCGHTDELLIKHDDYNSIYNYINNE-S-SSRECFIFNNMIKHSYTFFMETFIKCRIS  
PF3D7\_0801000 -----KLTEFLMRRRDYTVSVKILFNNHLAHEKIKYTKQEYMYYSQVLEKTYLPDAFRKKYWRVHYMTTEVVVRRTDNLIDHQFLRKG-S-CEKREYFNSKRRKGADLTETMKNIWMER  
PF3D7\_1200900 RYYKSNSLSHISNDNVLYFKKNYKGMTMEYVYVYHEIEIKLYNCKEYLYCEKLANKYPPDDYKIKIKKLANNDHGFIYLNKKGFIYKIKRKKKNG-----ESRNTLNLILKDTKKFREMMNNMWSN  
PF3D7\_1001700 -----VLBNQIKKIDSGEVVYVCHKEFDLHDNEKKEFNKEDHEYSKLSQYDPENIKKKKILKASYEETLKNNEFFFNHVFANIG-E-RRRSLDAVGYGKKKEVRYTVDETWEYVY  
PF3D7\_1016500 -----TLDEEFYKMDTGVLNCKKLYLNFNSFEKLYLLGEDFINCEWYSSVKNPVNVKQYMIKIAFYCTKDYLDSITFYNYYSFINRT-VYIDKQDLKNEIKTKNNRLEKMEYKKKYS  
PF3D7\_0219700 -----RLSDCHINDKNDSEVSKKLYLNFNSFNNOCKYYNIDNWTFCESLCSYNEPKKILFNLKAYNIDICTHLLNYSMKQYELFDKN-E-CSRNNYQFDILGESYNNLTKKMKNNWNTI  
PF3D7\_1016600 -----DLSEFSENNLDVNNKNEIYNNYNSKEFHYLNKNSCFENLAYNDPDIERNELKAYYITDILLINKNFHNYVHLISYG-R-CDKNFFEVIDIHINSNRIRRSMMNLWTKKN  
PF3D7\_1016700 -----NYTEKINDISKNEFVSLKDFIENHAKRIEKKELCEQYNEYIERLKISNNPTYIKRKKCAKVNPHNTDYLDIERILQNYTMLYRG-S-MSKSEQLNEQRDNLNQRRQLNVWKKDIE

\* conserved W

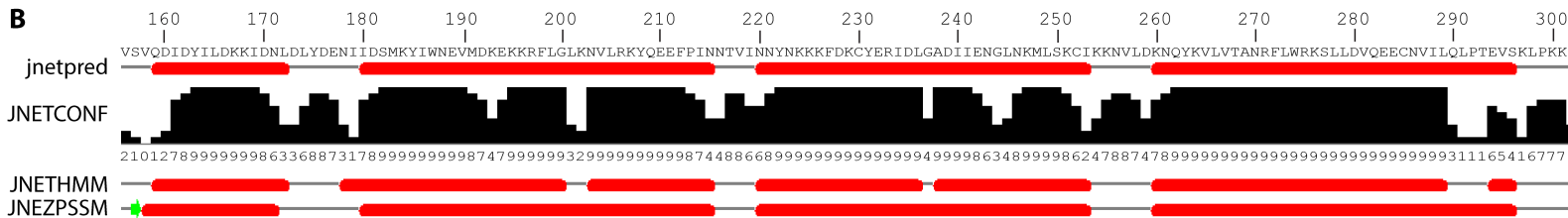
(\*) less conserved W



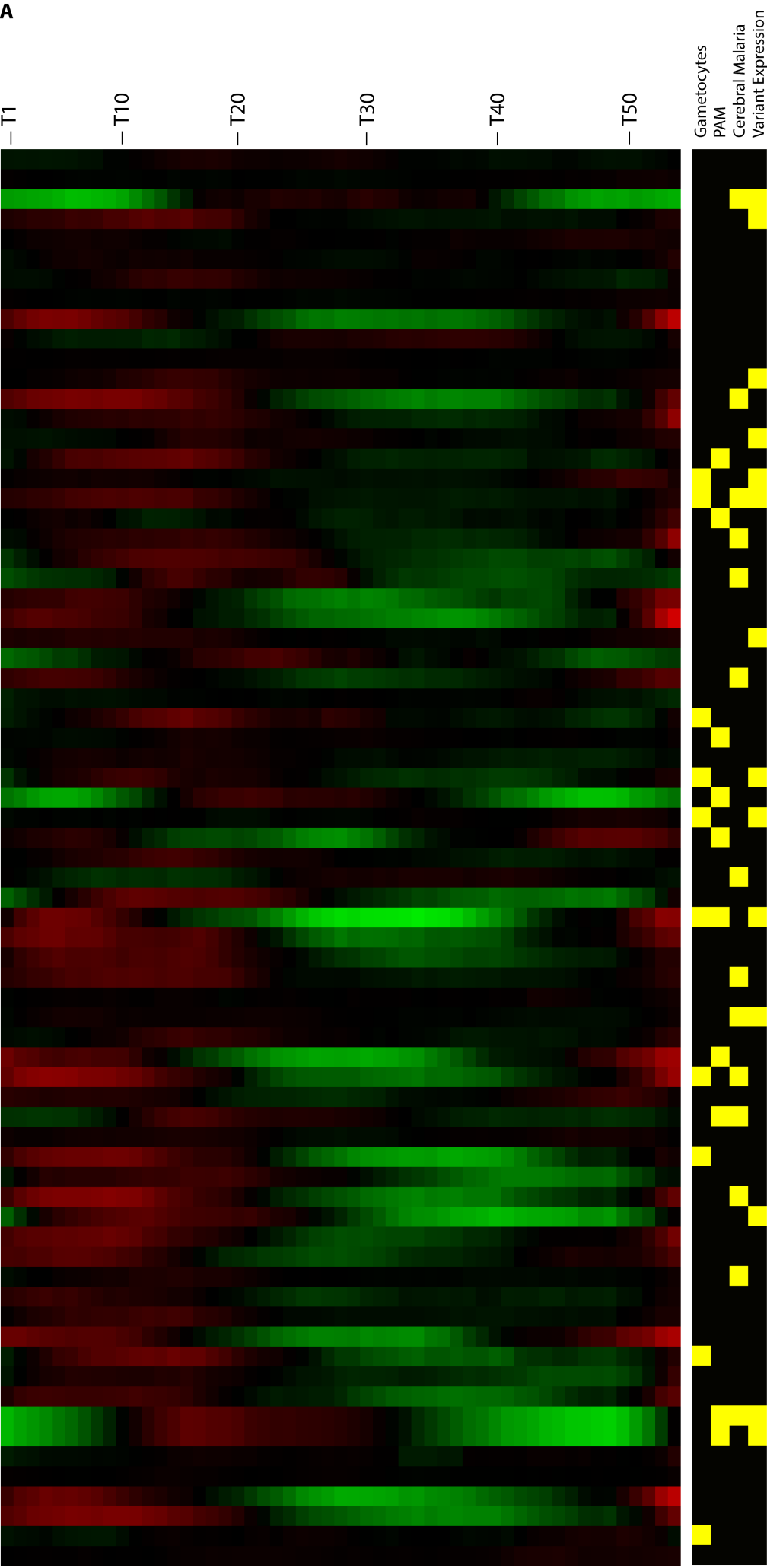
**A**



**B**



A



B

