
http://eprints.gla.ac.uk/7570/

Deposited on: 02 August 2012
Proteomic Profiling of \textit{Plasmodium} Sporozoite Maturation Identifies New Proteins Essential for Parasite Development and Infectivity

Edwin Lasonder\textsuperscript{1,2}\textsuperscript{*}, Chris J. Janse\textsuperscript{3,9}, Geert-Jan van Gemert\textsuperscript{4,9}, Gunnar R. Mair\textsuperscript{5,9}, Adriaan M. W. Vermunt\textsuperscript{1,9}, Bruno G. Douradinha\textsuperscript{4,9c}, Vera van Noort\textsuperscript{2,9}, Martijn A. Huynen\textsuperscript{2}, Adrian J. F. Luty\textsuperscript{4}, Hans Kroeze\textsuperscript{3}, Shahid M. Khan\textsuperscript{3}, Robert W. Sauerwein\textsuperscript{4}, Andrew P. Waters\textsuperscript{3,9e}, Matthias Mann\textsuperscript{5,9f}, Hendrik G. Stunnenberg\textsuperscript{1,9}

\textsuperscript{1}Department of Molecular Biology, NCMLS, Radboud University Nijmegen, Nijmegen, The Netherlands, \textsuperscript{2}Center for Molecular and Biomolecular Informatics, NCMLS, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, \textsuperscript{3}Leiden Malaria Research Group, Department of Parasitology, Centre for Infectious Diseases, Leiden University Medical Center, Leiden, The Netherlands, \textsuperscript{4}Department of Medical Microbiology, NCMLS, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands, \textsuperscript{5}Center for Experimental BioInformatics, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark

Abstract

\textit{Plasmodium falciparum} sporozoites that develop and mature inside an \textit{Anopheles} mosquito initiate a malaria infection in humans. Here we report the first proteomic comparison of different parasite stages from the mosquito—early and late oocysts containing midgut sporozoites, and the mature, infectious salivary gland sporozoites. Despite the morphological similarity between midgut and salivary gland sporozoites, their proteomes are markedly different, in agreement with their increase in hepatocyte infectivity. The different sporozoite proteomes contain a large number of stage specific proteins whose annotation suggest an involvement in sporozoite maturation, motility, infection of the human host and associated metabolic adjustments. Analyses of proteins identified in the \textit{P. falciparum} sporozoite proteomes by orthologous gene disruption in the rodent malaria parasite, \textit{P. berghei}, revealed three previously uncharacterized \textit{Plasmodium} proteins that appear to be essential for sporozoite development at distinct points of maturation in the mosquito. This study sheds light on the development and maturation of the malaria parasite in an \textit{Anopheles} mosquito and also identifies proteins that may be essential for sporozoite infectivity to humans.

Introduction

The life cycle of human malaria parasite \textit{Plasmodium falciparum} within the mosquito vector begins when gametocytes are taken up in an infected blood meal; after forming gametes and fertilisation, the resulting zygote differentiates into a motile ookinete that traverses the midgut epithelium and transforms within 36–48 hours into an oocyst (OOC) between the midgut epithelial cells and the basal lamina. The oocyst is an asexually replicating form of the parasite, which produces up to 2000–4000 sporozoites in about two weeks. Rupture of mature oocysts releases oocyst-derived sporozoites (ODS) into the hemocoele of the mosquito. The movement of the hemolymph brings the ODS in contact with the salivary glands, which they then invade. The sporozoites mature inside the salivary glands and then are stored ready for transmission to the mammalian host upon the next blood meal. A limited number of the salivary gland sporozoites (SGS) are injected during a mosquito bite and only a few of these complete the necessary migration from the skin to the liver to establish an infection inside hepatocytes. Clearly, the sporozoite has to complete a number of functions and metabolic readjustments both before and after injection into a mammalian host. The sporozoite has to be capable of actively exiting an oocyst, travelling through the hemolymph (the mosquito circulatory system), and
Author Summary

Human malaria is caused by *Plasmodium falciparum*, a unicellular protozoan parasite that is transmitted by *Anopheles* mosquitoes. An infectious mosquito injects saliva containing sporozoite forms of the parasite and these then migrate from the skin to the liver, where they establish an infection. Many intervention strategies are currently focused on preventing the establishment of infection by sporozoites. Clearly, an understanding of the biology of the sporozoite is essential for developing new intervention strategies. Sporozoites are produced within the oocyst, located on the outside wall of the mosquito midgut, and migrate after release from the oocysts to the salivary glands where they are stored as immature forms. Comparison of the proteomes of sporozoites derived from either the oocyst or from the salivary gland reveals remarkable differences in the protein content of these stages despite their similar morphology. The changes in protein content reflect the very specific preparations the sporozoites make in order to establish an infection of the liver. Analysis of the function of several previously uncharacterized, conserved proteins revealed proteins essential for sporozoite development at distinct points of their maturation.

In this study we have performed a detailed proteomic analysis that we previously applied to generate the glands which were obtained by hand-dissection of infected mosquitoes. Our analyses resulted in a proteome of oocysts (n=127), oocyst-derived sporozoites (n=450) and salivary gland sporozoites (n=477), which represent 728 individual *Plasmodium* proteins, of which 250 were exclusively detected in the oocyst/sporozoite stages when compared to the *P. falciparum* blood stage proteomes generated in a previous study [15]. The identification of proteins and their relative distributions within the different proteomes suggest specific metabolic adaptations and other biological functions of the maturing sporozoite. Moreover, we analyzed the function of eight sporozoite-specific proteins identified in our proteome analyses that were specifically annotated as hypothetical proteins, by targeted gene disruption of the orthologous genes of the rodent malaria parasite, *P. berghei*. We were able to demonstrate an essential and distinct role for three of these proteins in sporozoite development.

**Results**

**Mosquito stage proteome**

Protein samples derived from infected mosquito midguts and salivary glands were analyzed by nano-liquid chromatography tandem mass spectrometry (nLC-MS/MS) essentially as previously described [15]. The MS/MS spectra were searched against a combined database of all possible predicted tryptic peptides derived from all *P. falciparum*, human, and mosquito (*Anopheles gambiae*) proteins. The proteomic analysis of *P. falciparum* oocysts, oocyst-derived sporozoites, and salivary gland sporozoites resulted in a total of 4611 unique peptides mapping to 728 non-redundant *P. falciparum* proteins; they are distributed over the three stages with 127, 450 and 477, respectively and depicted as a Venn diagram in Figure 1A. Identified tryptic peptides and corresponding *Plasmodium* proteins of the mosquito stages are provided as supplementary material (Table S1). In our previous analysis of infected human red blood cells we identified 741 asexual blood stage parasite proteins from a mixture of schizonts and trophozoites and an additional 931 gametocyte and 645 gamete proteins [15]. Merging these datasets with the proteomes of the mosquito stages resulted in the identification of 250 *Plasmodium* proteins (Table S1) that are specifically detected in mosquito stages and 809 proteins that are expressed only in the blood stages (Figure 1B). However, it is important to note that due to the incomplete nature of all proteome datasets, absence of proteins from one dataset may also be due to the limits of detection and not the actual absence of expression. Parasite samples derived from infected mosquitoes were considerably contaminated with mosquito proteins with total parasite protein fractions of 35% for ODS, 31% for SGS and for OOC only 11% of the sequenced proteins were parasite in origin. Therefore this relatively high degree of contamination resulted in overall lower numbers of proteins compared to our previous *Plasmodium* infected blood stage proteome study. In particular, only 127 *P. falciparum* proteins in a pool of 987 mosquito proteins were identified for the oocyst sample that presumably represents the more abundantly expressed parasite proteins. Therefore, further analysis of the identified proteins and additional functional analyses are mainly focused on the proteins identified in the ODS and SGS. In total, we analyzed six different stages of *Plasmodium* (both from this study and our previous work) and have identified a total of 1543 *Plasmodium* proteins. The proportion of ‘stage specific’ proteins in the different life cycle stages ranged from 12% (gametes) to 28% and the stage specificity of proteins in the mosquito stages ranged between 15–24% (Figure 1C).

**Comparison with existing RNA/protein mosquito stage studies**

Genome-wide proteome and transcriptome studies have previously been reported for salivary gland sporozoites of *P. falciparum* [10,16], for oocysts and sporozoites of *P. berghei* [11] and
P. falciparum distribution of detected different life-cycle stages. (A) Venn diagram depicting the distribution of detected P. falciparum proteins over three different mosquito stage proteomes (oocysts, oocyst-derived sporozoites and salivary gland sporozoites). Numbers represent the number of proteins, that are either shared between 2 or 3 stages (overlapping areas) or that are detected in a single stage. (B) Comparison of the expression of P. falciparum proteins detected in the three mosquito stage proteomes to the blood stage proteomes described previously [15]. (C) The percentage of proteins exclusively detected in only one proteome out of 6 different life cycle stage proteomes, i.e. ASX - asexual blood stages; GCT – gametocytes; GAM – gametes; OOC – oocysts; ODS - oocyst-derived sporozoites; SGS - salivary gland sporozoites. 

Figure 1. Distribution of identified P. falciparum proteins over different life-cycle stages. (A) Venn diagram depicting the distribution of detected P. falciparum proteins over three different mosquito life-cycle stages (oocysts, oocyst-derived sporozoites and salivary gland sporozoites). Numbers represent the number of proteins, that are either shared between 2 or 3 stages (overlapping areas) or that are detected in a single stage. (B) Comparison of the expression of P. falciparum proteins detected in the three mosquito stage proteomes to the blood stage proteomes described previously [15]. (C) The percentage of proteins exclusively detected in only one proteome out of 6 different life cycle stage proteomes, i.e. ASX - asexual blood stages; GCT – gametocytes; GAM – gametes; OOC – oocysts; ODS - oocyst-derived sporozoites; SGS - salivary gland sporozoites.

doi:10.1371/journal.ppat.1000195.g001
Upregulated In Sporozoites (UIS) genes (Table S4)). It is however interesting to observe that between the two SSH studies only 2 out of 30 genes appear clearly up-regulated in both P. yoelii and P. berghei sporozoites.

Functional annotation of mosquito stage proteins

A global functional characterization of the ‘mosquito stage’ proteome was performed by an enrichment analysis of Gene Ontology (GO) annotations, for both the proteins that are shared between blood stages and mosquito stages (n = 478) and for the mosquito stage specific proteins (n = 250). The set of 478 genes commonly expressed in both mosquito and blood stages showed enrichment in GO annotations in all classes (i.e. Molecular Function, Cellular Component and Biological process (Figure S1)) and this enrichment is principally associated with housekeeping genes (Figure 2). The mosquito stage specific proteome did not reveal significant (p<0.01) enrichment in GO annotations nor did additional analyses for GO enrichment of the mosquito stage specific proteins using BINGO [34] and Ontologizer [35] (data not shown). In Figure 2B GO categories (Molecular Function) are shown for the mosquito stage specific proteome that contain more than 5 proteins. The lack of enrichment could be caused by the high proportion of genes annotated as hypothetical (300 out of 728) and consequently the relatively large number of proteins in the mosquito stage specific proteome (124 out of 250) without a GO annotation. Since our analysis did not reveal a significant GO enrichment for proteins known to be important in sporozoite function (e.g. motility and motor activity (Figure 2)) we analyzed our mosquito stage proteome for previously reported proteins, for which a function during sporozoite development is described and supported by strong experimental evidence (e.g. gene-knockout and/or antibody-inhibition studies). These proteins, in total 23, are listed in Table 1 and 15 out of 23 proteins are present in the mosquito stage proteome reported here. Based on a total number of 5410 genes in the genome of P. falciparum and 728 proteins in our mosquito stage proteome, these 15 proteins represent a 4.8 fold functional enrichment relative to the annotated genome and is highly significant (p<0.001 using Ontologizer). A good agreement exists between the function of the sporozoite proteins as shown in Table 1 and their expression pattern in the different mosquito stages. For example, proteins with multiple roles during sporozoite maturation (e.g. CS and TRAP) were identified in all stages (OOC, ODS and SGS) whereas proteins involved in hepatocyte traversal, such as SPECT1, SPECT2 (sporozoite microneme protein essential for cell traversal 1 and 2) and CeITOS (cell-traversal protein for ookinetes and sporozoites) were exclusively identified in mature SGS.

Sporozoites, like other motile stages (except male gametes) of Apicomplexan organisms, move on substrates by a mechanism known as gliding motility which is driven by an actomyosin motor complex [3,36,37]. Although there was no enrichment with high confidence (p<0.05) of the GO Molecular Function category ‘motor activity’ for mosquito stage specific proteins (Figure 2), several proteins known to be involved in the actomyosin motor complex are well represented and include TRAP, myosin A, MyoATail Domain Interacting Protein (MTIP), actin and F-1,6-BP aldehyde (3.6 fold enrichment with low confidence). Additionally, sporozoites encode a variety of surface molecules for both motility and invasion of host cells. For apicomplexan parasites members of the TRAP/MIC2 family have been shown to be important for host cell recognition and motility. The general architecture of this family is typified by one or more thrombospondin type I (TSP1) domains in their extracellular regions which may in addition also posses von Willebrand factor A (vWA) extracellular domains [38]. Our sporozoite proteome shows a 4.7 fold enrichment for proteins that contain one or multiple TSP1 domains (Table 2) compared to the P. falciparum proteome of 5410 proteins.

Two sporozoite proteomes – ODS versus SGS

Although the morphology of oocyst-derived and salivary gland sporozoites is identical at the level of light microscopy, ODS of P. berghei are significantly less infective to the mammalian host than SGS [39]. This marked difference in infectivity suggests significant developmental changes between these forms and was indicated by the analyses of gene transcription of different sporozoite stages by either SSH screens or SAGE analysis, which alludes to changes in protein expression in the sporozoite during the period of egress from the oocyst and the establishment of infection of the salivary glands [18,20,21]. In agreement with these observations, we found a large number of proteins expressed in SGS that were absent or relatively low expressed in ODS (Table S1). Several proteins involved in metabolic pathways show clear differences in distribution between ODS and SGS (Figure S2). For example, 8 out of 9 enzymes of the glycolytic pathway for ATP production were detected, all which were either more abundant or exclusive to SGS (SGS 8 proteins with 140 peptides; ODS 4 proteins and 48 peptides). A similar profile is observed for proteins involved in the production of NADPH via the pentose phosphate pathway with an up-regulation of these proteins in SGS (5 proteins and 26 peptides) compared to ODS (1 protein and 2 peptides). A third up-regulated metabolic pathway is the tricarboxylic acid (TCA) cycle (7 proteins and 85 peptides in SGS compared to 4 proteins and 26 peptides in ODS). Interestingly, several genes (4 out of 10) of the TCA cycle are most abundantly expressed in SGS, not only when compared to ODS but also in comparison with the blood stages, indicating an important role of the TCA cycle in mature sporozoites. Also the enzyme phosphoenolpyruvate carboxykinase (PF13_0234) is upregulated in the salivary gland sporozoites (3 peptides in ODS and 17 in SGS), which is again in agreement with the upregulation of enzymes involved in the TCA cycle and glycolysis [40]. It also appears that SGS prepare for enhanced protein synthesis: 9 of the 11 detected tRNA ligases are only detected in the SGS proteome and not in the ODS proteome (Table S1) as are ribosomal proteins, translation elongation factors and the TCF chaperonin complex proteins, which are either exclusively detected in SGS or are represented in the SGS proteome by substantially more peptides compared to the ODS proteome. As is shown in Table 1, proteins that are known to play a role in traversal and invasion of hepatocytes are highly enriched in SGS. On the other hand, the expression of MAEBL that is expressed along with CS and well before AMA-1 [41] and is known to function in attachment and invasion of the salivary gland [42] is more abundantly expressed in ODS. Therefore, it would appear that the proteomes of the sporozoite characterised by this study at different stages of development accurately reflect the functionality of either the ODS or SGS.

Consequently, based on the expression pattern and relative abundance of the peptides in the proteomes from OOC, ODS and SGS (see Materials and Methods section) the mosquito stage specific proteins can be regarded as belonging to one of 3 distinct groups (Table S1): Group I consists of 112 ODS proteins highly enriched for the ODS stage, putatively involved in sporozoite maturation inside the oocyst and in salivary gland invasion; similarly Group II which contains 74 proteins up-regulated in SGS potentially involved in infection of the mammalian host; and finally Group III that contains 59 proteins that are shared between the different mosquito stage proteomes and therefore may be
involved in sporozoite functions necessary both in the mosquito vector and the mammalian host (e.g. proteins involved in gliding motility and invasion such as CS [43,44] and TRAP [45,46] (Table 1)). These three groups formed the basis for selection of genes for further functional analysis of their encoded proteins through targeted disruption of the orthologous genes in the rodent malaria parasite, \textit{P. berghei}. The three groups were further refined for subsequent functional analysis using the following criteria (see also Materials and Methods section): i) high expression level as determined by the number of uniquely detected peptides per protein, ii) presence of gene sequences encoding putative transmembrane regions, signal peptides and/or GPI anchors, and iii) presence exclusively in the mosquito stage proteomes. This resulted in selection of genes as shown in Table 3. Further, in order to enrich for proteins that may define \textit{Plasmodium} specific functions, we preferentially selected not only genes that were annotated as hypothetical but also had no domains predicted by either the SMART or Pfam algorithms (i.e. with no indication of predicted function).

\textbf{Functional characterization of sporozoites-specific proteins}

In total eight genes identified in this study were selected (Table 4) for functional analysis by targeted gene disruption of their corresponding orthologs in \textit{P. berghei}, specifically, 3 ODS specific
genes (Group I), 2 SGS specific genes (Group II) and 3 from Group III (shared ODS/SGS). The sequences of the eight P. berghei gene orthologs (as well as their corresponding up and downstream sequences) were retrieved from the on-line *Plasmodium* genome databases, http://www.plasmodb.org and http://www.genedb.org/genedb/pberghei. However, for 4 of the 8 genes the *P. berghei* orthologs were fragmented and complete genes were manually assembled from a number of different *P. berghei* sequences by performing BLAST sequence searches of the full length *P. falciparum* genome and diagnostic PCR was performed using well established methods [48] and details of these analyses are shown in Figure S3.

**Table 1.** Characterized proteins involved in sporozoite development and invasion of host cells.

<table>
<thead>
<tr>
<th>Accession nr</th>
<th>protein name (1)</th>
<th>protein involved in (2)</th>
<th>nr unique pept/protein in life cycle stages (3)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PFI14_0067</td>
<td>CCP3</td>
<td>sporozoite development</td>
<td>- 31 10 - - - -</td>
<td>Claudianos [72], Pradel [73]</td>
</tr>
<tr>
<td>PFI14_0532</td>
<td>CCP2</td>
<td>sporozoite development</td>
<td>- 39 20 - - - -</td>
<td>Pradel [73]</td>
</tr>
<tr>
<td>PFC0180c</td>
<td>IMC1</td>
<td>sporozoite development (cell shape)</td>
<td>- - - 0 9 25</td>
<td>Khatre [74]</td>
</tr>
<tr>
<td>PFC0210c</td>
<td>CS</td>
<td>sporozoite development, salivary gland and hepatocyte invasion</td>
<td>- - - 2 5 9</td>
<td>Menard [43], Wang [44]</td>
</tr>
<tr>
<td>PF80325c</td>
<td>cysteine protease</td>
<td>egress from oocyst</td>
<td>- - - - 16 23</td>
<td>Aly [56]</td>
</tr>
<tr>
<td>PF13_0233</td>
<td>myosin A</td>
<td>sporozoite gliding motility</td>
<td>- - - 2 39 49</td>
<td>Bergman [73]</td>
</tr>
<tr>
<td>PFL2225w</td>
<td>MTIP</td>
<td>sporozoite gliding motility</td>
<td>1 1 - 2 4 3</td>
<td>Bergman [73]</td>
</tr>
<tr>
<td>PF13_0201</td>
<td>SSP2/TRAP</td>
<td>sporozoite gliding motility, salivary gland and hepatocyte invasion</td>
<td>- - - 1 6 35</td>
<td>Rogers [45], Sultan [46]</td>
</tr>
<tr>
<td>PFI1055w</td>
<td>CRMP1</td>
<td>salivary gland invasion</td>
<td>- 1 - - - -</td>
<td>Thompson [76]</td>
</tr>
<tr>
<td>MAL7P1.92</td>
<td>CRMP2</td>
<td>salivary gland invasion</td>
<td>- - - - - -</td>
<td>-</td>
</tr>
<tr>
<td>PFI11_0486</td>
<td>MAEBL</td>
<td>salivary gland invasion</td>
<td>- - - - - 33 1</td>
<td>Kariu [42]</td>
</tr>
<tr>
<td>MAL13P1.212</td>
<td>SPECT1</td>
<td>cell traversal hepatocytes</td>
<td>1 - - - -</td>
<td>Ishino [77]</td>
</tr>
<tr>
<td>PF04030c</td>
<td>SPECT2</td>
<td>cell traversal hepatocytes</td>
<td>1 - - - -</td>
<td>Ishino [78]</td>
</tr>
<tr>
<td>PFEL800c</td>
<td>CellTOS</td>
<td>cell traversal hepatocytes</td>
<td>- - - - -</td>
<td>Kariu [79]</td>
</tr>
<tr>
<td>PFF1420w</td>
<td>PL</td>
<td>cell traversal hepatocytes</td>
<td>- - - - -</td>
<td>-</td>
</tr>
<tr>
<td>PFI11_0344</td>
<td>AMA-1</td>
<td>hepatocyte invasion</td>
<td>1 - - - - 2 19</td>
<td>Silvie [81]</td>
</tr>
<tr>
<td>PF80570w</td>
<td>SPATR</td>
<td>hepatocyte invasion</td>
<td>- - - - - 5</td>
<td>Chattopadhyay [82]</td>
</tr>
<tr>
<td>PF80095c</td>
<td>PIEMP3</td>
<td>hepatocyte invasion</td>
<td>3 - - - - -</td>
<td>Gruner [83]</td>
</tr>
<tr>
<td>PF07_0006</td>
<td>STARP</td>
<td>hepatocyte invasion</td>
<td>- - - - -</td>
<td>Pasquetto [84]</td>
</tr>
<tr>
<td>GI:1477963</td>
<td>SALSA</td>
<td>hepatocyte invasion</td>
<td>- - - - -</td>
<td>Puenteses [85]</td>
</tr>
<tr>
<td>PF02000w</td>
<td>TRSP</td>
<td>hepatocyte invasion</td>
<td>- - - - - 1 3</td>
<td>Lababeid [86]</td>
</tr>
<tr>
<td>PF02155c</td>
<td>pf52 protein/P36p</td>
<td>hepatocyte invasion - development</td>
<td>- - - - -</td>
<td>Ishino [29], van Dijk [30]</td>
</tr>
<tr>
<td>PF02105c</td>
<td>pbs36 homologue/P36</td>
<td>hepatocyte invasion - development</td>
<td>- - - - -</td>
<td>Ishino [29]</td>
</tr>
</tbody>
</table>

1CCP (members of the LCCL protein family), IMC (inner membrane complex), CS (circumsporozoite), MTIP (myosin A tail domain interacting protein), SSP2 (sporozoite surface protein), TRAP (Thrombospondin related anonymous protein), CRMP (cysteine repeat modular protein), MAEBL (membrane antigen erythrocyte binding like protein ), SPECT (sporozoite microneme protein essential for cell traversal ), CellTOS (cell-traversal protein for ookinetes and sporozoites), PL (phosphoplipase), AMA (apical membrane antigen), SPATR (secreted protein with altered thrombospondin domain), SALSA (sporozoite and liver stage antigen); TRSP (thrombospondin related protein).

1The role of these proteins has been determined by functional analysis in gene-knockout and/or antibody-inhibition studies.

1ASX - asexual blood stages; GCT – gametocytes; GAM – gametes; OOC – oocysts; ODS - oocyst-derived sporozoites; SGS - salivary gland sporozoites.

doi:10.1371/journal.ppat.1000195.s001

It was not possible to select mutant parasites for two genes, one belonging to Group I (orthologous to PFI14_0607) and the other belonging to Group III (orthologous to PFA0205w) in independent transfection experiments, suggesting that both these proteins may have an additional and essential role during blood stage development. For the remaining 6 genes mutants were generated in two independent transfection experiments per gene (Table 4) and correct disruption of the target genes was shown for all mutants (Figure S3). All 6 mutant lines showed normal asexual growth and also gametocyte and ookinete production that was comparable to wild type parasites (data not shown). As an initial phenotype screen of mosquito stage development, uncloned parental populations of the 6 mutant lines were allowed to infect mosquitoes. Oocyst numbers and salivary gland sporozoite numbers were determined at day 6 and 20 after infection, respectively, and infected mosquitoes were allowed to feed at day 20-22 on naïve mice. In 3 out of the 6 mutant lines (orthologous to APF11_0528, APF14_0074 and APFF1195c) parasite development inside the mosquito (oocyst number and salivary gland sporozoites number) was not significantly different from wild type
suggests the existence of significant redundancy in the function of S3). The lack of a clear effect of disruption of these 3 genes on the correct gene disruption genotype in blood stages of all 3 parasites after mosquito transmission of these 3 mutants revealed mutants. Genotype characterization by Field Inverse Gel Electrophoresis indicating ‘wild type’ infectivity of the sporozoites of these 3 parasitemias between 0.1 and 0.5 at day 4 after infection, parasites (Table 4). After infection of mice by bite of mosquitoes infected with any of these three mutant lines, all mice developed parasitemias between 0.1 and 0.5 at day 4 after infection, indicating ‘wild type’ infectivity of the sporozoites of these 3 mutants. Genotype characterization by Field Inverse Gel Electrophoresis (FIGE) analysis and diagnostic PCR of blood stage parasites after mosquito transmission of these 3 mutants revealed the correct gene disruption genotype in blood stages of all 3 mutants, demonstrating normal mosquito transmission of the mutant, rather than breakthrough of wild type parasites (Figure S3). The lack of a clear effect of disruption of these 3 genes on sporozoite production and infectivity to the mammalian host suggests the existence of significant redundancy in the function of these mosquito stage specific proteins.

The remaining 3 mutant lines (orthologous to ΔPFI14.0435, ΔPF00210552 and ΔMAL8P1.66) showed an aberrant development during mosquito development. The phenotypes of cloned lines of these mutants were therefore analyzed in more detail. Clones of all 3 gene-disrupted lines produced wild type numbers of oocysts ranging from 150–250 oocysts per mosquito on day 7/8 post infection. The development of parasites deficient in line 841cl1 was largely blocked at the oocyst stage. However, low numbers of sporozoites were formed that were able to invade the salivary gland (2750-6250 oocyst sporozoites per mosquito and 875–6600 SGS per salivary gland). Despite the low numbers of sporozoites that emerge from the oocyst, salivary gland invasion appears not to be affected since ODS and SGS numbers were comparable. In contrast to sporozoites of mutant 841cl1, salivary gland sporozoites of 843cl1 injected either intravenously (1×10³ sporozoites) or by mosquito bite were not infective for mice (2 experiments with 2 mice). Interestingly, 843cl1 sporozoites demonstrated the same or greater hepatocyte traversal rate than wild type sporozoites and they were also able to traverse and invade hepatocytes in vitro (Figure 3). This suggests that the lack of sporozoite infectivity to mice may be due to a defect in liver stage development after invasion of the hepatocyte.

**Table 2.** Expression of *Plasmodium* proteins containing thrombospondin type 1 (TSP1) and/or von Willebrand factor A (vWA) domains in different life cycle stage proteomes.

<table>
<thead>
<tr>
<th>Accession nr</th>
<th>Protein name</th>
<th>Domain</th>
<th>nr unique pept/prot in life cycle stages (1)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>PF13_0201</td>
<td>TRAP</td>
<td>TSP1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>PFC02105c</td>
<td>CS</td>
<td>TSP1</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>PFA0200w</td>
<td>TRSP</td>
<td>TSP1</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>MAL8P1.45</td>
<td>hypothetical protein</td>
<td>TSP1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PB0570w</td>
<td>SPAAT</td>
<td>TSP1</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>PFF0800w</td>
<td>TRAP-like protein</td>
<td>TSP1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>PFL0870w</td>
<td>PTRAMP</td>
<td>TSP1</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>PF10_0281</td>
<td>MTRAP</td>
<td>TSP1</td>
<td>-</td>
<td>Baker [88]</td>
</tr>
<tr>
<td>PF08_0136b</td>
<td>WARP</td>
<td>TSP1</td>
<td>-</td>
<td>Yuda [89]</td>
</tr>
<tr>
<td>PFC0640w</td>
<td>CTRP</td>
<td>TSP1</td>
<td>-</td>
<td>Baker [88]</td>
</tr>
<tr>
<td>PFL0875w</td>
<td>hypothetical protein</td>
<td>TSP1</td>
<td>-</td>
<td>1</td>
</tr>
</tbody>
</table>

(1 ASX - asexual blood stages; GCT – gametocytes; GAM – gametes; OOC – oocysts; ODS - oocyst-derived sporozoites; SGS - salivary gland sporozoites.

Discussion

The proteome analyses of the three mosquito stages of *Plasmodium falciparum*, oocysts, oocyst-derived sporozoites and salivary gland sporozoites, resulted in the identification of 728 proteins of which 250 are ‘mosquito stage specific’, having not been detected in our previous analysis of blood stage parasites [15]. Although the total number of proteins identified in the mosquito stages is lower compared to blood stages [15], which is in all likelihood due to sample purity and not reduced protein expression, we show a clear developmental progression of the
<table>
<thead>
<tr>
<th>Accession nr (1)</th>
<th>Protein Name</th>
<th>MW (kDa)</th>
<th>Group (2)</th>
<th>MOTIF PREDICTION (3)</th>
<th>ORTHOLOGY PREDICTION (4)</th>
<th>nr unique peptide/protein (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ortho MCL group</td>
<td>nr species</td>
</tr>
<tr>
<td></td>
<td></td>
<td>SP</td>
<td>GPI</td>
<td>nr TM (mean)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PEXEL</td>
<td>SMART</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Plasma</td>
<td>ApI</td>
<td>other</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PF11_0528</td>
<td>hypothetical protein</td>
<td>710</td>
<td>I</td>
<td>5.3</td>
<td>OG2_80072</td>
<td>6</td>
</tr>
<tr>
<td>PF11_0486</td>
<td>MAEBL</td>
<td>243</td>
<td>I</td>
<td>1.3</td>
<td>OG2_70757</td>
<td>6</td>
</tr>
<tr>
<td>MAL13P1.66</td>
<td>hypothetical protein</td>
<td>346</td>
<td>I</td>
<td>1.8</td>
<td>OG2_93592</td>
<td>6</td>
</tr>
<tr>
<td>PFF0645c</td>
<td>integral membrane protein</td>
<td>163</td>
<td>I</td>
<td>6.5</td>
<td>DUF1222</td>
<td>6</td>
</tr>
<tr>
<td>PF14_0404</td>
<td>hypothetical protein</td>
<td>408</td>
<td>I</td>
<td>0.8</td>
<td>OG2_71079</td>
<td>6</td>
</tr>
<tr>
<td>PF14_0722</td>
<td>CRMP4</td>
<td>699</td>
<td>I</td>
<td>9.0</td>
<td>3 * GCC2_GCC3</td>
<td>OG2_81006</td>
</tr>
<tr>
<td>PF14_0435</td>
<td>hypothetical protein</td>
<td>144</td>
<td>I</td>
<td>6.8</td>
<td>OG2_89567</td>
<td>5</td>
</tr>
<tr>
<td>PF13_0338</td>
<td>hypothetical protein (PF92)</td>
<td>93</td>
<td>I</td>
<td>1.0</td>
<td>s48-s45</td>
<td>OG2_117922</td>
</tr>
<tr>
<td>MAL8P1.126</td>
<td>serine protease</td>
<td>101</td>
<td>I</td>
<td>1.5</td>
<td>Trypsin</td>
<td>OG2_77036</td>
</tr>
<tr>
<td>PF14_0607</td>
<td>hypothetical protein</td>
<td>127</td>
<td>I</td>
<td>4.8</td>
<td>OG2_79434</td>
<td>6</td>
</tr>
<tr>
<td>MAL8P1.135</td>
<td>hypothetical protein</td>
<td>116</td>
<td>I</td>
<td>8.8</td>
<td>OG2_74819</td>
<td>6</td>
</tr>
<tr>
<td>PF14_0694</td>
<td>protein disulfide isomerase</td>
<td>66</td>
<td>I</td>
<td>0.8</td>
<td>Thioredoxin</td>
<td>OG2_78996</td>
</tr>
<tr>
<td>PF11_0229</td>
<td>hypothetical protein</td>
<td>57</td>
<td>I</td>
<td>1.0</td>
<td></td>
<td>OG2_95024</td>
</tr>
<tr>
<td>PFE0340c</td>
<td>hypothetical protein</td>
<td>87</td>
<td>I</td>
<td>5.0</td>
<td>Rhomboid</td>
<td>OG2_70960</td>
</tr>
<tr>
<td>MAL13P1.119</td>
<td>cAMP-specific 3',5'-cyclic</td>
<td>86</td>
<td>I</td>
<td>4.5</td>
<td>Hdc</td>
<td>OG2_93460</td>
</tr>
<tr>
<td>PFD0295c</td>
<td>hypothetical protein</td>
<td>85</td>
<td>I</td>
<td>1.0</td>
<td>CCP</td>
<td>OG2_89419</td>
</tr>
<tr>
<td>PF11_0064</td>
<td>hypothetical protein</td>
<td>85</td>
<td>I</td>
<td>1.5</td>
<td>OG2_92898</td>
<td>6</td>
</tr>
<tr>
<td>PF11_0141</td>
<td>UDP-galactose transporter</td>
<td>39</td>
<td>I</td>
<td>8.0</td>
<td>UAA; TPT; DUF6</td>
<td>OG2_72118</td>
</tr>
<tr>
<td>PF10_0223</td>
<td>hypothetical protein</td>
<td>49</td>
<td>I</td>
<td>4.3</td>
<td></td>
<td>OG2_92751</td>
</tr>
<tr>
<td>PF11_0344</td>
<td>AMA-1</td>
<td>72</td>
<td>II</td>
<td>1.5</td>
<td>AMA-1</td>
<td>OG2_81783</td>
</tr>
<tr>
<td>PFD0430c</td>
<td>SPECT2 (PLP1)</td>
<td>94</td>
<td>II</td>
<td>1.5</td>
<td>MACPF</td>
<td>OG2_81960</td>
</tr>
<tr>
<td>PF14_0074</td>
<td>hypothetical protein</td>
<td>47</td>
<td>II</td>
<td>0.5</td>
<td>OG2_84483</td>
<td>6</td>
</tr>
<tr>
<td>MAL13P1.212</td>
<td>SPECT1</td>
<td>28</td>
<td>II</td>
<td>1.0</td>
<td>OG2_105334</td>
<td>4</td>
</tr>
<tr>
<td>PFL1080c</td>
<td>CalTOS</td>
<td>20</td>
<td>II</td>
<td>0.5</td>
<td>OG2_92921</td>
<td>5</td>
</tr>
<tr>
<td>PF80570w</td>
<td>SPATR</td>
<td>29</td>
<td>II</td>
<td>0.8</td>
<td>TSP1</td>
<td>OG2_89620</td>
</tr>
<tr>
<td>PFF1195c</td>
<td>hypothetical protein</td>
<td>75</td>
<td>II</td>
<td>1.0</td>
<td></td>
<td>OG2_89698</td>
</tr>
<tr>
<td>PFE0395c</td>
<td>hypothetical protein (PF38)</td>
<td>41</td>
<td>II</td>
<td>1.3</td>
<td>s48-s45</td>
<td>OG2_93113</td>
</tr>
<tr>
<td>PFC0905c</td>
<td>hypothetical protein</td>
<td>360</td>
<td>III</td>
<td>1.0</td>
<td></td>
<td>OG2_86694</td>
</tr>
<tr>
<td>PF13_0233</td>
<td>myosin a</td>
<td>92</td>
<td>III</td>
<td>0.8</td>
<td>Myosin_head</td>
<td>OG2_70807</td>
</tr>
</tbody>
</table>
Table 3. Cont.

<table>
<thead>
<tr>
<th>Accession nr (1)</th>
<th>Protein Name</th>
<th>MW (kDa)</th>
<th>Group (2)</th>
<th>MOTIF PREDICTION (3)</th>
<th>ORTHOLOGY PREDICTION (4)</th>
<th>nr unique peptide/protein (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ortho MCL group</td>
<td>nr species</td>
</tr>
<tr>
<td>PFD0425w</td>
<td>hypothetical protein</td>
<td>113</td>
<td>III</td>
<td>1</td>
<td>1.8</td>
<td>OG2_86498</td>
</tr>
<tr>
<td>PF13_0201</td>
<td>TRAP</td>
<td>65</td>
<td>III</td>
<td>1</td>
<td>1.8</td>
<td>VWA; TSP1</td>
</tr>
<tr>
<td>PFC0210c</td>
<td>CS</td>
<td>43</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PFB0325c</td>
<td>cysteine protease</td>
<td>78</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PFC0180c</td>
<td>membrane skeletal protein</td>
<td>98</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MAL8P1.73</td>
<td>hypothetical protein</td>
<td>93</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PFA0205w</td>
<td>hypothetical protein</td>
<td>66</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF08.0008</td>
<td>hypothetical protein</td>
<td>85</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PFL1165w</td>
<td>hypothetical protein</td>
<td>96</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MAL3P1.254</td>
<td>hypothetical protein</td>
<td>31</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF14_0495</td>
<td>hypothetical protein</td>
<td>249</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PF14_0286</td>
<td>glutamate dehydrogenase</td>
<td>57</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MAL8P1.66</td>
<td>hypothetical protein</td>
<td>26</td>
<td>III</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

1Genes shown in bold have been further characterized by functional analysis through targeted gene disruption (see Table 4).
2Classification of proteins based on their protein expression patterns: group I (ODS-enriched), group II (SGS-enriched), and group III (ODS/SGS-enriched).
3Abbreviations used for motif predictions are: Signal Peptide (SP), Glycosylphosphatidylinositol anchored proteins (GPI), Trans Membrane regions (average of number of TM regions predicted by four algorithms) and N-terminal PEXEL (Plasmodium export element within first 200 amino acids obtained from http://www.plasmodb.org). Abbreviations for predicted SMART domains: Apical membrane antigen (AMA), Complement control proteins (CCP), Domain of Unknown Function (DUF), Gu/Leu/Phenylalanine/dehydrogenase dimerisation domain (ELFV_dehydro_N), Glutamate/Phenylalanine/Valine dehydrogenase (ELFV_dehydro), Domain containing 5 cysteine conserved residues (GCC2_GCC3), Metal dependent phosphohydrolases with conserved ‘HD’ motif (HDc), Membrane-attack complex/perforin domain family (MACPF), Phosphoenolpyruvate carboxykinase (PEPCK_ATP), Papain family cysteine protease (Pept_C1), Sexual stage antigen s48-45 domain (s48-45), Thrombospondin type 1 (TSP1), Triose-phosphate Transporter family (TPT), UAA transporter family (UAA), Von Willibrand factor type A domain (VWA).
4Orthology prediction from OrthoMCL database (http://www.orthomcl.org) shows OrthoMCL group, nr species Plasm (the number of Plasmodium species, additional nr species Api (number of additional Apicomplexan species without Plasmodium species), additional species other (number of species without Apicomplexan species (including Plasmodium)).
5OOC - oocysts; ODS - oocyst-derived sporozoites; SGS - salivary gland sporozoites. The mosquito fraction (mosq fr.) of a protein is calculated by the number of detected peptides in mosquito stages divided by the sum of the number of peptides in mosquito and blood stages.

doi:10.1371/journal.ppat.1000195.t003
Table 4. Genes selected for further functional analysis of their encoded proteins through targeted disruption of their orthologs in the *P. berghei* rodent model.

<table>
<thead>
<tr>
<th>P. fal. Ac. Nr.</th>
<th>Group (1)</th>
<th>P. berghei accession nr.</th>
<th>Exp nr (mutant) (2)</th>
<th>Successful disruption (3)</th>
<th>Oocyst no. mean (s.d.) (2-4 exp) (4)</th>
<th>spor. no. mean (s.d.) (2-3 exp) (5)</th>
<th>spor. infectivity to mice (mosquito bite) (6)</th>
<th>Phenotype (7)</th>
<th>Remarks (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control (WT)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>169 (39)</td>
<td>104000 (14500)</td>
<td>4 (n = 4)</td>
<td>wild type</td>
<td>-</td>
</tr>
<tr>
<td>PF14_0528</td>
<td>I</td>
<td>PB001073.00.0/PB000529.02.0/PB000863.00.0</td>
<td>800, 836</td>
<td>Yes (2 x)</td>
<td>160 (48)</td>
<td>64800 (12500)</td>
<td>2 (n = 2)</td>
<td>Wild type</td>
<td>Redundant function?</td>
</tr>
<tr>
<td>PF14_0435</td>
<td>I</td>
<td>PB101363.00.0/PB000829.02.0/PB105739.00.0</td>
<td>802, 838</td>
<td>Yes (2 x)</td>
<td>181 (60)</td>
<td>0</td>
<td>0 (n = 4)</td>
<td>Complete block of sporozoite development inside oocyst</td>
<td></td>
</tr>
<tr>
<td>PF14_0607</td>
<td>I</td>
<td>PB000015.03.0</td>
<td>803, 829, 839</td>
<td>No (3 x)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Essential for blood stage?</td>
<td></td>
</tr>
<tr>
<td>PF14_0074</td>
<td>II</td>
<td>PB000372.00.0</td>
<td>807, 842</td>
<td>Yes (2 x)</td>
<td>153 (27)</td>
<td>80200 (16000)</td>
<td>2 (n = 2)</td>
<td>Wild type</td>
<td>Redundant function?</td>
</tr>
<tr>
<td>PFF1195c</td>
<td>II</td>
<td>PB107027.00.0/PB107193.00.0/PB001101.03.0</td>
<td>801, 837</td>
<td>Yes (2 x)</td>
<td>165 (47)</td>
<td>97300 (12200)</td>
<td>2 (n = 2)</td>
<td>Wild type</td>
<td>Redundant function?</td>
</tr>
<tr>
<td>PFD0425w</td>
<td>III</td>
<td>PB000251.01.0</td>
<td>806, 841</td>
<td>Yes (2 x)</td>
<td>154 (59)</td>
<td>375 (125)</td>
<td>0 (n = 4)</td>
<td>Arrest of sporozoite egress from oocyst Oocyst sporozoites are infective to mice</td>
<td></td>
</tr>
<tr>
<td>PF0205w</td>
<td>III</td>
<td>PB301531.01.0</td>
<td>809, 830, 844</td>
<td>No (3 x)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Essential for blood stage?</td>
<td></td>
</tr>
<tr>
<td>MAL8P1.66</td>
<td>III</td>
<td>PB402680.00.0</td>
<td>808, 843</td>
<td>Yes (2 x)</td>
<td>188 (51)</td>
<td>2625 (415)</td>
<td>0 (n = 4)</td>
<td>(Partial) block of sporozoite development inside oocyst Few sporozoites formed; not infective to mice</td>
<td></td>
</tr>
</tbody>
</table>

1 Classification of proteins based on their protein expression patterns. Group I ODS-enriched, group II SGS-enriched, and group III ODS/SGS-enriched.
2 The experiment numbers (= mutant number) of 2–3 independent transfection experiments.
3 Successful disruption was determined by Southern analysis of separated chromosomes and diagnostic PCR of selected, pyrimethamine resistant parasites as shown in Figure S3 and Table S5.
4 Mean number of oocysts per mosquito and standard deviation (s.d.).
5 Mean number of salivary gland sporozoites per mosquito and standard deviation (s.d.).
6 Inoculation of sporozoites was tested by infecting mice by mosquito bite. The number of mice that showed blood stage infection comparable to wild type infections is shown.
7,8 See Figure 3 for details of the characterization of the 3 mutants with a phenotype different from wild type.

OOC – oocysts; ODS – oocyst-derived sporozoites; SGS – salivary gland sporozoites.

doi:10.1371/journal.ppat.1000195.s004
parasite through the mosquito that is reflected in changes of its protein repertoire.

Analysis of the ‘stage specificity’ of proteins in six different life cycle (mammalian and mosquito) stage proteomes demonstrated that expression of proteins restricted to a single stage ranges from 12 to 28% with the highest percentage of ‘stage specificity’ in the gametocyte and reaching 24% in ODS. The 478 proteins common to blood and mosquito stages are significantly enriched in house keeping proteins involved in metabolic processes. The absence of specific enrichment of GO annotations in the 250 proteins of the mosquito stage specific proteome can most likely be ascribed to the fact that a relatively small number of these proteins posses a GO designation. Many of the mosquito stage specific proteins are still annotated as hypothetical and probably have functions that are specific for sporozoites and/or Plasmodium. This concept is supported by the observation that 15 of the 23 Plasmodium proteins known to have a sporozoite specific function are present in the 250 mosquito stage proteins identified in this study, a 4–5 fold enrichment. Moreover, their stage specific expression in our different proteomes also confirms that in general the timing of protein expression coincides with observation of function as inferred from gene deletion studies. For example, proteins involved in traversal and invasion of the hepatocyte (e.g. SPECT1/2, CeITOS, AMA-1, STARP, TRSP, P36p and P36) (Table 1) are either exclusively or much more highly expressed in SGS than ODS. Such changes in protein composition and abundance demonstrate that sporozoites go through dynamic changes and may exist as clearly defined developmental stages – currently ODS and SGS – that express stage specific proteins. These clear differences seem unexpected in the light of the morphological similarity of the two stages but on the other hand are in good agreement with the significant rise in mammalian host infectivity observed during the maturation and migration of sporozoites from oocysts to salivary glands [20,39]. These changes are not only restricted to proteins directly involved in these processes, but extend also to enzymes implicated in metabolic housekeeping processes such as glycolysis, production of NADPH and the TCA cycle that might be expected to coincide with subcellular reorganisation at the level of the organelles. Mature, salivary gland sporozoites might be considered to be in the resting phase (G0) of the cell cycle and are able to persist and remain infectious within the salivary glands of the mosquito for the remainder of its life. Therefore, the abundance and storage of these proteins may suggest that the salivary gland sporozoite contains stockpiles of proteins which are deployed only upon activation in the vertebrate host and growth (G1) and multiplication (S, M phases) inside the hepatocyte. Alternatively, some of these proteins could specifically be required by the parasite in the salivary glands of the mosquito host and therefore do not depend on activation in the vertebrate host.

Protein and gene expression studies of SGS have previously been performed in *P. falciparum* [10,16] as well as for the rodent parasites *P. berghei* [11,20,21] and *P. yoelii* [18,19]. The relatively low overlap between the proteins detected in the various proteomes of sporozoites can in part be ascribed to the difficulties in collecting material of sufficient purity and quantity. This limitation results in the frequent sequencing of peptides derived from mosquito proteomes which reduces the total number of identified parasite proteins. However, both the degree of overlap between the proteomes and the degree of certainty in protein calling can be improved if more strict selection criteria are used for protein calling [31]. When we compared only proteins that were identified by at least 2 or more fully-tryptic peptides in all datasets (i.e. ours, Flores [10] *P. falciparum* SGS and Hall [11] *P. berghei* SGS) we found a greater than 50% overlap in proteins. Moreover, in the Hall *P. berghei* SGS and OOC proteomes it is observed that more than 80% of these proteins have a direct ortholog in *P. falciparum*. Further, when we again only compare ‘fully tryptic proteomes’ we find 75% of the *P. berghei* SGS proteins are also
expressed in the SGS of *P. falciparum* indicating that sporozoites of different *Plasmodium* species employ similar processes of maturation and invasion. Despite the relatively low overlap in total numbers of proteins detected in the different proteomes, there is good correlation of protein abundance between our SGS proteome and the previously reported SGS proteome of *P. falciparum* [10] based on peptide counting methods.

Interestingly, nearly all the expressed genes of *P. yoelii* sporozoites detected by EST analyses [18] are also present in our proteome. Similarly, in a recent microarray analysis of *P. yoelii*, where 3500 expressed genes were measured in the ODS/SGS stages we find that all of our 601 (i.e. 601 of the 728 *P. falciparum* genes that have a *P. yoelii* ortholog) mosquito stage specific *P. falciparum* proteins are also detected as mRNA [22].

We found a lower percentage of shared proteins between our proteome and the transcripts detected in sporozoites of *P. berghei* [11,20,21]. The variation in overlap between the various proteome and transcriptome studies is certainly influenced by the varying and often small number of identified genes/proteins and indicates that a comprehensive expression profile of the salivary gland sporozoite has still to be realized. Comparison of mRNA species detected in *P. falciparum* SGS [22] with our proteome showed that for a large percentage of genes, mRNA production coincides with the presence of its protein (463 mRNA species for 477 proteins; 97%). The simultaneous presence of transcripts and protein expression has also been observed during blood stage development, supporting the ‘just in time’ model [9]. However, more than 2100 genes demonstrate an up-regulation of transcription in sporozoites [22], many of which were not detected as a protein in the various proteome studies. Moreover, a low correlation exists between the abundance levels of our SGS proteins (i.e. by emPAI) and the mRNA abundance of previously reported large-scale SGS transcriptome studies (i.e. r = 0.31–0.33; [16,22]). This is in line with the observations made by Le Roch et al (2004) where transcript levels are not always well correlated with same stage protein expression, suggesting a delay between mRNA and protein accumulation [49]. It is interesting to speculate whether these differences in expression between RNA and protein could be in part explained by translation repression as is observed in gametocytes that contain pools of translationally repressed transcripts that are only translated following zygote formation [11,50,51]. However, as discussed above, the proteome of sporozoites may not be comprehensive enough to draw conclusions about the relationship between specific mRNA and protein expression patterns.

The sporozoite proteomes, despite not being exhaustive, provide for the first time information on parasite protein expression both at the mosquito midgut and salivary gland stages. This has allowed for the identification of hitherto uncharacterized proteins which in turn has informed the selection of genes for targeted orthologous gene disruption studies in the rodent malaria parasite, *P. berghei*. Mutant *P. berghei* parasites lacking mosquito stage specific proteins have proven to be an efficient way to obtain an understanding into the function of such proteins [32]. We were able to generate 6 mutants for 8 hypothetical proteins that were selected from our proteomes for further functional analysis in *P. berghei* of which 3 showed distinct phenotypes, demonstrating an important and essential role of these proteins in sporozoite development and maturation. The knock-out parasite lines of 3 genes that do not exhibit a clear phenotypic difference from wild type parasites indicate either a redundancy in function for the proteins encoded by these genes or else phenotypes that are presently too subtle for us to detect with our current methodologies. However, functional redundancy is a well-established phenomenon for a number of *Plasmodium* proteins that are expressed in the blood and sexual stages of the parasite [53–55].

The *P. falciparum* protein PF14_0435 is highly and exclusively expressed in sporozoites obtained from the oocyst stage and the phenotype of the orthologous gene knock-out mutant in *P. berghei*, 802cl1, is an abnormal development of the oocyst and the complete absence of sporozoite production. This example demonstrates not only the validity of the orthologous gene studies in *P. berghei* but also the informative power of this combination of proteome-reverse genetic approach in the characterization of proteins at discreet stages of the parasite life-cycle. Furthermore, the number of oocysts produced by the 802cl1 mutant is not different from wild type levels and a defect appears to occur prior to sporozoite development indicating that the role of PF14_0435 is upstream of sporozoite production. The phenotype of a second *P. berghei* mutant, line 841c11, which lacks the orthologue of PF0425w closely resembles the egress defects observed with the cysteine protease ECP1 (or SER2A in *P. falciparum*) and CS mutants that are mutated in their thrombospondin repeat; where sporozoites are unable to exit from midgut oocysts [44,56]. Although ECP1 mutant sporozoites are not infectious, it has been suggested that ECP1 may be involved in the cleavage of CS and thereby release of sporozoites from the oocyst [4]. Interestingly, while oocyst-derived sporozoites that lack ECP1 or express mutated CS are not infective to mice, the mechanically liberated oocyst-derived sporozoites of mutants lacking PF0425w are able to establish an infection in mice by i.v. inoculation, and this implies that PF0425w - in contrast to ECP1 - has no additional function during infection of the mammalian host. Its role appears to be restricted and directly involved in sporozoite release from the oocyst and has a more immediate/causeutive function in the release of sporozoites from oocysts. In contrast, the protein MAL8P1.66 appears to have multiple roles during sporozoite development within the oocyst and infectivity to the mammalian host. Mutants lacking this protein (i.e. line 843) are affected in the production of sporozoites within oocysts. However, the low numbers of sporozoites formed are able to invade salivary glands and hepatocytes in vitro but are unable to infect mice, suggesting an additional role during further development inside the hepatocyte. Interestingly and in line with the expectation, the expression of MAL8P1.66 has recently been identified in liver stage of *Plasmodium* [57]. The exact role during development of the liver stages awaits further analysis.

This study sheds light not only on the development and maturation of the malaria parasite in an *Anopheles* mosquito but also identifies proteins that are uniquely synthesized as the sporozoite becomes increasingly infectious to humans. Infection initiated by injection of *P. falciparum* sporozoites into humans represents the culmination of many precise, sequential and critical developmental steps of the malaria parasite through the mosquito. Moreover, transmission is a bottle neck in the life cycle of *Plasmodium* and the full maturation of sporozoites is essential in the survival of the parasite. The changes in the different sporozoite proteomes documented here emphasise that each event from oocyst development to egress and invasion of salivary gland and injection is tightly regulated. Intervention studies are now being conducted that aim to exploit the tightly regulated pathways that the parasite has evolved to ensure transmission. This has been recently demonstrated with the use of genetically attenuated sporozoites that have rapidly become an important focus in the development of new vaccines. The disruptions of individual genes that encode sporozoite proteins sufficiently weakens the parasite such that development in the liver is blocked, enabling the mammalian host to generate a strong protective immunity against
subsequent infection. Clearly, the targeted disruption of genes encoding proteins identified in this study, which are involved in essential mature sporozoite functions, namely hepatocyte traversal, invasion and intracellular survival may also accelerate the identification of new protective attenuated parasite lines. Understanding the sporozoite and all its various developmental steps during the establishment of an infection continues to represent a promising approach in the hunt for new weapons in the fight against malaria.

Materials and Methods

Collection of *P. falciparum* oocysts and sporozoites

*Anopheles stephensi* mosquitoes (Sind-Kasur strain, 3–5 days old) [58] were infected with *P. falciparum* gametocytes (NF54) [59] by membrane feeding. Unfed and partially fed mosquitoes were removed and fully fed mosquitoes were kept at 26±1°C at 80% humidity. After one day, a 5% glucose solution soaked in cotton wool was offered to the mosquitoes and mosquitoes were allowed to take an extra (uninfected) blood meal at day 8–10 after infection [60]. Oocysts and oocyst-derivated sporozoites were collected from midguts at 7–8 and 13–14 days after infection, respectively. Approximately 100–200 mosquito midguts were hand-dissected and homogenized in a home made glass tissue grinder in 200 µl of PBS pH 7.2 at 4°C. Salivary gland sporozoites were collected from salivary glands 18–22 days after infection. Approximately 70 salivary glands were hand-dissected and treated in a similar way as the oocyst samples. For the parasite preparations (OOC, ODS and SGS), four, three and two batches respectively were generated and processed further for analysis by nLC-MS/MS.

Sample preparation for Mass Spectrometry analysis

In order to estimate the number of sporozoites in the samples described above the total number of oocyst and salivary gland sporozoites per mosquito was determined as follows: midguts and salivary glands were dissected from 10 mosquitoes at day 13 and day 22 after feeding respectively. The midguts/salivary glands were homogenized in a home made glass capsule in 1000 µl of PBS pH 7.2 and sporozoites were counted in a Bürker-Türk counting chamber using phase-contrast microscopy (1–1.6×10^6 sporozoites obtained from salivary glands of one mosquito, and 0.5–5×10^5 sporozoites per mosquito midgut). Parasites samples from mosquito midguts and salivary glands (approx. 1–4×10^10 ODS and SGS sporozoites, and 1–2×10^5 oocysts from 63–200 mosquito midguts) were divided into a soluble and insoluble fraction by a freeze–thawing procedure similar to the oocyst samples. For the parasite preparations (OOC, ODS and SGS), four, three and two batches respectively were generated and processed further for analysis by nLC-MS/MS.

Identified peptide count analysis to determine protein abundance index values

To determine the protein abundance in our samples, mass spectrometric data was analyzed using an identified peptide per protein count analysis to compute the exponentially modified Protein Abundance Index (emPAI) values [16,22]. EmPAI values for all proteins in Table S1 were calculated as 10^PAI–1 (PAI = nobserved peptides/nobservable peptides). The number of ‘observable’ peptides per protein was calculated from the output of the program Protein Coverage Summarizer [http://nrcr.pnl.gov/software/]. The collected peptide list of this study (malaria peptides identified in the mosquito stages) is available in Table S1.
3.8×10⁷ sporozoites) and 2 SGS batches (1.3–2.5×10⁷ sporozoites).

Normalization between different batches was performed according to the median and 20 percent trimmed mean method [61]. Normalization methods and approaches for merging emPAI data were evaluated on performance in correlation studies with mRNA data of P.falciparum salivary gland sporozoites [16,22] (Table S3). Mean protein emPAI values of merged and median normalized data were calculated per stage and have been included in Table S1. This approach was also applied to our proteomic data set of blood stages [15] to calculate normalized emPAI values.

Correlation between protein expression data from different studies

Values for the level (abundance) of protein expression from different datasets were obtained for all individual proteins by calculated emPAI values. EmPAI values and mRNA levels of microarray analyses were log2 transformed before regression analysis to obtain normal distributions. Pearson correlation between datasets was performed using R [http://www.r-project.org/].

Gene ontology annotation

Gene Ontology SLIM terms were assigned using “Generic GO (http://go.princeton.edu/cgi-bin/GOTermMapper). A GO enrichment analysis for ‘Biological Process’, ‘Cellular Component’ and ‘Molecular Function’ using default GO association files was performed with “GO Term Finder” [http://go.princeton.edu/cgi-bin/GOTermFinder] where statistical significance (p-value) is calculated based on hypergeometric distribution with Bonferroni multiple testing correction and false discovery rate calculation as described [62]. To perform a GO enrichment analysis with adjusted GO association files, Ontologizer [http://www.charite.de/ch/medgen/ontologizer/] was used where statistical significance (p-value) is calculated as in ‘GO Term Finder’ (see above and [63]).

Selection criteria for ODS, SGS and ODS/SGS proteins

Proteins with more than 90 percent of the peptides detected in the mosquito stages (mosquito fraction≥0.9) were divided into three groups (OOC, ODS and ODS/SGS) based on their expression patterns. The mosquito fraction equals \( \frac{n_{\text{mosq}}}{n_{\text{mosq}}+n_{\text{blood}}} \) where \( n \) is the number of unique peptides per protein at the mosquito and blood stages, respectively. The mosquito enriched proteins were further subdivided into 112 ODS-specific proteins expressed in the ODS stage and not SGS (Group I); 74 SGS-specific proteins not expressed in ODS (Group II); and finally 59 Group III proteins that are shared between several mosquito life cycle stages. A further refinement of these groups was based on the following criteria. Only proteins with more than two peptides detected in the mosquito stages \( n_{\text{mosq}} \geq 3 \) were considered. In addition, only proteins were selected that contained Signal peptide (SP), Transmembrane (TM) or Glycosylphosphatidylinositol (GPI) domains and combinations of these motifs. Sequence-based prediction data for these domains was retrieved from PlasmoDB [http://www.plasmoDB.org] for SP and TM domain predictions based on TMHMM, TMAP, TMHMM2 and TOPRED2 algorithms; from http://gpi.unibe.ch/ for GPI predictions by a Kohonen Self Organizing Map; and from http://smart.embl-heidelberg.de/ for SMART protein domain searches. The number of TM domains is the average of four values obtained from the different TM prediction algorithms. Different criterions were set for combinations of predicted motifs. For less abundant proteins without predicted signal peptide (SP = 0, and 3≤n_{\text{mosq}} ≤15), only proteins with at least 4 predicted TM regions were included (average TM>4). For abundant proteins without signal peptide (SP = 0, n_{\text{mosq}} > 15) proteins with at least 0.5 predicted TM regions (average TM≥0.5) were included. For proteins with predicted signal peptide (SP = 1), all proteins with at least 0.5 predicted TM regions (average TM≥0.5) were included. Finally, all proteins with a predicted GPI anchor (GPI = 1) were selected independent of the presence of predicted signal peptide or TM regions.

Generation and characterization of gene knockout P. berghei parasite mutants

Eight P. falciparum proteins were selected for functional analysis by targeted gene disruption of their corresponding orthologs in P. berghei. The sequences of the eight P. berghei gene orthologs (as well their corresponding up and downstream sequences) were retrieved from the online PlasmoDB genome databases, http://www.plasmodb.org and http://www.genedb.genedb.org/genedb/pberghei. For P. berghei genes with incomplete sequence information in the database (4 out of 8), the complete genes were manually assembled from a number of different P. berghei sequences by performing BLAST sequence searches of the full length P. falciparum genes against the P. berghei genome and closing gaps by PCR and DNA sequencing (see for details Figure S3 and Table S3). Standard plasmid vectors were designed for targeted gene disruption by double cross-over homologous recombination [48]. To replace the protein coding sequences of the target genes with the dhfr/t gyrA pyrimethamine resistance marker from Toxoplasma gondii, we cloned the 5′ and 3′ flanking regions of the gene of interest up- and downstream of the selection cassette of pBluescript II KS− (Stratagene). After isolation of the 3′ fragment by AluI digestion, we cloned it into pBluescript II KS− to generate the vector pBluescript II KS−/G418 (the latter containing the P. berghei orthologous gene PB000029.02.0 - amplified from P. berghei genomic DNA using primer-pairs 2666-2653 and 2654-2655, respectively, and ligated into plasmid pBluescript II KS− yielding targeting plasmid pL1175). To replace the protein coding sequences of the target genes with the dhfr/t gyrA pyrimethamine resistance marker from Toxoplasma gondii, we cloned the 5′ and 3′ flanking regions of the gene of interest up- and downstream of the selection cassette of pBluescript II KS− (Stratagene). After isolation of the 3′ fragment by AluI digestion, we cloned it into pBluescript II KS− to generate the vector pBluescript II KS−/G418 (the latter containing the P. berghei orthologous gene PB000029.02.0 - amplified from P. berghei genomic DNA using primer-pairs 2666-2653 and 2654-2655, respectively, and ligated into plasmid pBluescript II KS− yielding targeting plasmid pL1175). To replace the protein coding sequences of the target genes with the dhfr/t gyrA pyrimethamine resistance marker from Toxoplasma gondii, we cloned the 5′ and 3′ flanking regions of the gene of interest up- and downstream of the selection cassette of pBluescript II KS− (Stratagene). After isolation of the 3′ fragment by AluI digestion, we cloned it into pBluescript II KS− to generate the vector pBluescript II KS−/G418 (the latter containing the P. berghei orthologous gene PB000029.02.0 - amplified from P. berghei genomic DNA using primer-pairs 2666-2653 and 2654-2655, respectively, and ligated into plasmid pBluescript II KS− yielding targeting plasmid pL1175).
mice (n = 2–4) to 40–50 mosquitoes, at day 20 after the infectious blood meal. Infection was monitored by analysis of blood stage infection in Giemsa stained films of tail blood at day 4 till day 8 after infection. Infectivity was recorded as ‘wild type’ if mice developed a parasitemia of 0.1–0.5% at day 4 after infection. Infectivity of sporozoites to mice of 2 mutant lines was also determined by intravenous injection of sporozoites that were mechanically liberated by a glass grinder from either midgut oocysts (1–2×10^6 oocyst sporozoites collected at day 20 from mutant line 841 and wild type line 507c1l) or collected from salivary glands (10^6 salivary gland sporozoites at day 27 for mutant line 843 and wild type line 507c1l). For obtaining oocysts and salivary gland sporozoites, mosquito midguts or salivary glands were dissected in a drop of RPMI culture medium and the transferred by a custom made needle into a glass grinder after which sporozoites were released by gently grinding. Blood stage infection in mice injected with sporozoites in 200 μl RPMI buffer was monitored as described for infection of mice via mosquito interrupted feeding.

In vitro hepatocyte traversal and invasion experiments were performed as described elsewhere [67,68] by adding purified sporozoites (5×10^5) to confluent monolayers of HepG2 cells in DMEM medium (note: medium had 10% FCS and 1% PenStrep). Mutant sporozoites were obtained as described above from either oocysts (day 20) or from salivary glands (day 27). Quantification of cell traversal and invasion was accomplished by using a cell-impermeable fluorescent marker molecule, rhodamine-dextran at 1 mg/ml that will visualize parasitized wounded cells specifically but not uninfected HepG2 cells. Sporozoites were incubated with HepG2 cells in the presence of fluorescent dextran for 2 hr, followed by washing the cells to remove the marker and incubation for an additional 24 hours to determine the development of exoerythrocytic forms (EEFs) of the parasite. Hepatocyte invasion was determined by counting the percentage of sporozoites inside dextran-negative cells because parasites do not develop successfully in wounded dextran-positive cells [68]. After fixation of the HepG2 cells, infection was quantified by staining EEFs with monoclonal antibody 2E6 against HSP70 [69] and compared to infection of wild type sporozoites. Hepatocyte cell traversal was determined by counting the percentage of dextran-positive cells 2 hours after adding sporozoites to HepG2 cells, and compared to wild type sporozoite cell traversal. In this procedure, monoclonal antibody 3D11 against CS was used.

Accession Numbers

All datasets will become available through the official Web site of the Plasmidome genome project, PlasmoDB [http://www.plasmodb.org [70,71]]. In the text and tables most genes and gene products are accompanied with their PlasmoDB Accession Number.

The PlasmoDB accession numbers for other genes and gene products discussed in this paper are for P. falciparum: CS (PFC0210c), TRAP (PF13_0201), UIS3 (PF13_0012), P36 (PF0201c), P36p (PF0215c), myosin A (PF13_0233), MTIP (PF12225w), actin (PF12215w) and F-1,6-BP aldolase (PF14_0425, AMA-1 (PF11_0344), TRSP (PFA0200w), RESA3 (PF0232c), SPECT1 (MAL13P1.212), SPECT2 (PFD0430c), CcTOS (PFL0800c), STARP (PF07_0006) and for P. berghei: UIS4 (PB100551.00.0), ECPI (PB000649.01.0).

The sequences of the eight P. berghei gene orthologs (as well their corresponding up and downstream sequences) have been analysed in gene-disruption studies were retrieved from the PlasmoDB database [http://www.plasmodb.org] and from the GeneDB database [http://www.genedb.genedb.org/]. For 4 out of 8 P. berghei genes with incomplete sequence information in the database, the complete genes were manually assembled from a number of different P. berghei sequences by performing BLAST sequence searches, PCR and DNA sequencing (see for details Figure S3 and Table S5). Primer sequences used in contig gap closure and location of primers relating to contigs and reads of the revised P. berghei gene models have been submitted to GenBank and are provided in Figure S3 and Table S5.

Supporting Information

Figure S1 Gene Ontology (GO) annotation for proteins from proteomes from two mosquito stages of P. falciparum, oocyst-derived sporozoites and salivary gland sporozoites. Found at: doi:10.1371/journal.ppat.1000195.s001 (0.05 MB DOC)

Figure S2 Pathway profiling with the number of unique peptides/protein detected in 5 different life-cycle stages (data obtained from this study and from Lasonder et al. [15]). Found at: doi:10.1371/journal.ppat.1000195.s002 (0.09 MB DOC)

Table S1 Peptides and proteins of P. falciparum identified in proteomes of oocysts, oocyst-derived sporozoites and salivary gland sporozoites. Page ‘Peptides’: peptides identified by nLC-MS/MS. Information provided in the table: 1) life cycle stage, 2) peptide sequence, 3) MCR (mass of charge of parent ion), 4) charge of parent ion, 5) measured mass of peptide, 6) calibrated mass of peptide after internal mass calibration, 7) peptide rank in Mascot searches, 8) peptide score in Mascot searches, 9) Mascot peptide delta score (which equals the Mascot score difference between a rank1 and a rank2 peptide), 10) accession number of protein identification from PlasmoDB version 5.3, 11) protein name, 12 reannotation PlasmoDB 2008/07/15 (genes with the most recently modified annotation by PlasmoDB), 13) protein mass (molecular weight), 14) protein pl (iso-electric point), 15) nr peptide/sample (the number of unique peptides per protein), 16) Residue Start (residue nr in the protein sequence of the N-terminal amino acid), 17) Residue End (residue nr in the protein sequence of the C-terminal amino acid), 18) sequence coverage (percentage of the protein covered by the identified peptides), and 19) protein emPAI value. Page ‘Proteins’: The corresponding proteins identified by the sequenced peptides, listing: 1) accession number (PlasmoDB version 5.3), 2) protein name, 3 reannotation PlasmoDB 2008/07/15 (genes with the most recently modified annotation by PlasmoDB), 4) protein mass (molecular weight), 5) protein pl (iso-electric point), 6) number of unique identified peptides, 7) protein emPAI value, and 8) mosquito fraction (which is calculated by the number of detected peptides in mosquito stages divided by the sum of the number of peptides in mosquito and blood stages). Page ‘OOC’: proteins detected in OOC; column headings the same as for Page ‘Proteins’. Page ‘ODS’: proteins detected in ODS; column headings the same as for Page ‘Proteins’. Page ‘SGS’: proteins detected in SGS; column headings the same as for Page ‘Proteins’. Page ‘Mosquito stage specific’: proteins exclusively detected in mosquito stages (mosquito fraction = 1); column headings the same as for Page ‘Proteins’. Page ‘Shared with RBG stages’: proteins detected in mosquito stages and blood stages (0<mosquito fraction<1); column headings the same as for Page ‘Proteins’. Page ‘OOC-enriched’: proteins that are ‘highly
enriched’ in OOC (mosquito fraction) >0.9); column headings the same as for Page ‘Proteins’. Page ‘ODS-enriched’: proteins that are ‘highly enriched’ in ODS (mosquito fraction) >0.9; column headings the same as for Page ‘Proteins’. Page ‘SIS genes:’ proteins identified in SIS (Sporozoite expressed gene) as identified by either subtractive hybridization (SSH) or cDNA quantification methods (SAGE). S-genes: 25 sporozoite (S) genes identified in a P. yoelii SSH screen [18]. UIS-genes: 30 UIS genes (Upregulated In Sporozoites) identified in a P. berghei SSH screen [20]. SIS genes: 123 SIS genes (Sporozoite expressed gene Identified by SAGE (SIS) genes) identified in a P. berghei SAGE analysis [21].

Table S2

Comparison of proteins identified in the mosquito stage proteomes of this study with the proteomes of salivary gland sporozoites (SGS) of P. falciparum as reported by Florens and colleagues [10] and the proteomes of oocysts (OOC) and SGS of P. berghei as reported by Hall and coworkers [11]. The tables contain information about proteins shared by our analysis and others, where our expression data is presented by the number of unique tryptic peptides per protein. The other data sets are presented in a similar way, but a distinction is made in expression data for detection either by all peptides (tryptic, half (non) tryptic), or by tryptic peptides.

Found at: doi:10.1371/journal.ppat.1000195.s006 (3.44 MB ZIP)

Table S3

Correlation of protein abundance (emPAI approach) identified in the SGS stage proteome of this study with mRNA levels of the SGS transcripts of P. falciparum as reported by Le Roch and coworkers [16] and Zhou and coworkers [22]. Tables show Pearson correlation coefficients (r), probabilities (p) and number of shared genes/proteins (n).

Found at: doi:10.1371/journal.ppat.1000195.s005 (1.58 MB ZIP)

Table S4

Comparison of proteins identified in the mosquito stage proteomes with genes transcribed in sporozoites in P. berghei and P. yoelii as identified by either subtractive hybridization (SSH) or cDNA quantification methods (SAGE). S-genes: 25 sporozoite (S) genes identified in a P. yoelii SSH screen [18]. UIS-genes: 30 UIS genes (Upregulated In Sporozoites) identified in a P. berghei SSH screen [20]. SIS genes: 123 SIS genes (Sporozoite expressed gene Identified by SAGE (SIS) genes) identified in a P. berghei SAGE analysis [21].

Found at: doi:10.1371/journal.ppat.1000195.s007 (0.06 MB XLS)

Table S5

Primer sequences used in this study. Primers used in KO targeting plasmid construction. Primers used to check for plasmid integration in mutant (KO) parasites. Primers used in contig gap closure and wild type (WT) PCR analysis

Found at: doi:10.1371/journal.ppat.1000195.s008 (0.03 MB XLS)

Acknowledgments

We would like to thank Marga van de Vegte-Bolmer, Jolanda Klaassen, Laura Pelsers-Posthumus, Astrid Pouwelsen, Jacqueline Kuhnen, and Jai Ramesar for technical assistance with the preparation of parasite samples, mosquito infection and transfection experiments. We would also like to thank the Wellcome Trust Sanger Institute, Hinxton, Cambridge, U.K. for making publically available sequence data for Plasmodium berghei.

Author Contributions

Conceived and designed the experiments: EL CJJ MAH AJFL SMK APW. Contributed reagents/materials/analysis tools: RWS APW HGS. Performed the experiments: EL CJJ GJvG GRM AMWV. Analyzed the data: EL CJJ MAH AJFL SMK APW. Wrote the paper: EL CJJ MAH APW SMK APW HGS. The authors declare that they have no competing interests.

References

56. Aly AS, Matuschewski K (2005) A malarial cysteine protease is necessary for
47. Janse CJ, Franke-Fayard B, Waters AP (2006) Selection by flow-sorting of
30. van Dijk MR, Douradinha B, Franke-Fayard B, Heussler V, van Dooren MW,