Extensive Transcriptional and Translational Regulation Occur During the Maturation of Malaria Parasite Sporozoites

Scott E. Lindner^{1,*,#}, Kristian E. Swearingen^{2,*}, Melanie J. Shears^{3,6}, Michael P. Walker¹, Erin N. Vrana¹, Kevin J. Hart¹, Allen M. Minns¹, Photini Sinnis³, Robert L. Moritz², and Stefan H.I. Kappe^{4,5,#}

Author Affiliations:

1. Department of Biochemistry and Molecular Biology, The Huck Center for Malaria Research, Pennsylvania State University, University Park, PA

2. Institute for Systems Biology, Seattle, WA

3. Department of Molecular Microbiology & Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

4. Center for Global Infectious Disease Research, Seattle Children's Research Institute, Seattle, WA

5. Department of Global Health, University of Washington, Seattle, WA

6. Current Address: Department of Laboratory Medicine, University of Washington, Seattle, WA

*. Equal Contributions

#. Co-Corresponding Authors. Scott E. Lindner (Scott.Lindner@psu.edu, ORCID: 0000-0003-1799-3726); Stefan Kappe (stefan.kappe@seattlechildrens.org, ORCID: 0000-0003-1540-1731)

Keywords: Plasmodium, sporozoite, translational repression, RNA-seq, proteomics

Additional ORCID's:

Kristian Swearingen ORCID: 0000-0002-6756-4471, Robert Moritz ORCID: 0000-0002-3216-9447,

Kevin Hart ORCID: 0000-0001-8250-7822

Abstract:

Plasmodium sporozoites are transmitted from an infected mosquito to mammals in which they infect the liver. The infectivity profile of sporozoites changes as they egress from oocysts on the mosquito midgut into the hemocoel, and then invade the salivary glands, where they maintain a poised and infectious state until transmission occurs. Upon transmission, the sporozoite must then navigate the host skin, vasculature, and liver. All of these feats require distinct repertoires of proteins and capabilities that are coordinated in an appropriate temporal manner. Here, we report the comprehensive and dynamic transcriptomes and proteomes of both oocyst sporozoite and salivary gland sporozoite stages in both rodent-infectious *Plasmodium yoelii* parasites and human-infectious *Plasmodium falciparum* parasites. These data robustly define mRNAs and proteins that are Upregulated in Oocyst Sporozoites (UOS) or Upregulated in Infectious Sporozoites (UIS), which include critical gene products for sporozoite functions, as well as many of unknown importance that are similarly regulated. Moreover, we found that *Plasmodium* uses two overlapping, extensive, and independent programs of translational repression across sporozoite maturation to temporally regulate specific genes necessary to successfully navigate the mosquito vector and mammalian host environments. Finally, gene-specific validation experiments of selected, translationally repressed transcripts in *P. yoelii* confirmed the interpretations of the global transcriptomic and proteomic datasets. Together, these data indicate that two waves of translational repression are implemented and relieved at different times in sporozoite maturation to promote its successful life cycle progression.

Introduction:

Malaria remains one of the great global health problems today, taking a large toll on people in the tropics and subtropics. This disease, caused by *Plasmodium* parasites, affects over 200 million people annually and kills over 400,000 (WHO World Malaria Report 2018). While a protein-based subunit vaccine (RTS,S) has recently been licensed and is being used for pilot implementation in three Sub-Saharan African countries, its protection has been limited and relatively short-lived in clinical trials ¹. Identifying and fully testing an effective and long-lasting malaria vaccine remains a chief goal that has yet to be achieved. Accomplishing this will require greater knowledge of the basic biology of the pre-erythrocytic sporozoite and liver stage parasites. Promising new whole-parasite vaccine candidates, based upon the sporozoite form of the parasite, are on the horizon and hopefully will meet these needs ².

Plasmodium parasites are transmitted between hosts by an infected female *Anopheles* mosquito (reviewed in ³). Following uptake of male and female gametocytes by the mosquito during a blood meal of an infected host, these parasites activate into male and female gametes in the midgut, fuse into a zygote, and develop into a motile ookinete, which burrows through the midgut wall and establishes an oocyst under the basal lamina. Within the oocyst, the parasite undergoes sporogony to produce up to five thousand oocyst sporozoites ⁴. These oocyst sporozoites are weakly infectious if injected directly into a naïve mammalian host ⁵, but become highly infectious following proteolytic rupture of the oocyst wall and transit through the mosquito hemocoel. Sporozoites further gain infectivity after invasion of the salivary glands ^{5,6}. Interestingly, the gain of infectivity for the mammalian host is concomitant with a loss of infectivity for the salivary glands. Within the glands, the sporozoites await transmission as salivary gland sporozoites, which occurs when the mosquito takes its next blood meal. Infection of a host begins with the injection of salivary gland sporozoites into the skin. Following this, sporozoites exit the bite site in the skin, locate and enter the vasculature, and passively travel to the liver, where

sporozoites can productively infect hepatocytes, which initiates the life cycle progression in the mammalian host ⁷. Moreover, because relatively few sporozoites are injected during a mosquito bite ⁸, this transmission event bottleneck has been the focus of intervention efforts using drugs, subunit vaccines, and attenuated whole parasite vaccines ².

Fundamental studies of sporozoite biology have informed efforts to inhibit and/or arrest the parasite during development. To date, genetically-attenuated parasite (GAP) vaccine candidates were generated by the deletion of genes whose transcripts are Up-regulated in Infective (Salivary Gland) Sporozoites (UIS). The UIS gene family was originally determined by a subtractive cDNA hybridization approach ⁹, where 30 expressed sequence tags (ESTs) were identified as being more abundant in salivary gland sporozoites compared to oocyst sporozoites. Of these, 23 map to currently annotated genes, with several of the original ESTs mapping to the top hit (UIS1). With the advent of microarray-based transcriptomics, a renewed effort to identify both UIS and Up-regulated in Oocyst Sporozoites (UOS) genes that may be similarly exploited to produce effective GAP vaccines identified 124 UIS and 47 UOS genes ¹⁰. Interestingly, only 7 of the original 23 UIS genes were confirmed in this expanded study. However, these UIS genes (UIS1, UIS2, UIS3, UIS4, UIS7, UIS16, UIS28) have proven to encode some of the most important proteins for the transmission and transformation of the sporozoite into a liver stage parasite, as well as for liver stage development ¹¹⁻¹⁴.

In addition to transcriptional controls placed upon key transcripts, the malaria parasite also imposes translational repression upon specific mRNAs in female gametocytes, and for at least a few mRNAs in salivary gland sporozoites (reviewed in ^{15,16}). This model allows for the proactive production of mRNAs and restriction of their translation before transmission, and yet enables a strategy of just-in-time production of these proteins after transmission when they are needed. However, this strategy (high transcription, low/no translation) is energetically costly, and model eukaryotes and human cells have evolved to avoid this gene regulatory combination (“Crick Space”) except in specific, beneficial situations

¹⁷. In light of this, it is notable that *Plasmodium* parasites have evolved to use translational repression for transmission to the mosquito, which has been clearly seen for mRNAs (*e.g.* *p28*) in female gametocytes ¹⁸. The mechanisms underlying this have been established most thoroughly in *P. berghei*, where DOZI (a DEAD-box RNA helicase orthologous to human DDX6) and CITH (an Lsm14 orthologue) bind, stabilize, and translationally repress specific mRNAs in female gametocytes ¹⁸⁻²⁰. Recently, the extent of translational repression in *P. falciparum* female gametocytes was assessed by mass spectrometry-based proteomics and RNA-seq ²¹. In this stage, the parasite expresses over 500 transcripts with no evidence for their encoded proteins, with over half of these maternal gene products being uncharacterized. Enriched in this set of translationally repressed mRNAs are those that encode for functions needed post-transmission and include gene families previously shown to be translationally repressed. These data support the model that large-scale preparations are made by female gametocytes, despite high energetic costs, to prepare and await transmission by storing and protecting specific mRNAs needed to establish the infection of the mosquito. However, the extent and precise mechanisms of how translational repression is imposed in sporozoites has not been established beyond targeted studies that suggest that the PUF2 RNA-binding protein may act upon *cis* elements found within the coding sequence of the *uis4* mRNA ¹³.

Many previous efforts that aimed to determine the global transcriptomes and proteomes of *Plasmodium* sporozoites were greatly restricted due to substantial contamination with material from the mosquito vector and its microbiome in these samples ²²⁻²⁴. In response to this, we have developed a scalable, discontinuous density gradient purification approach for sporozoites that greatly reduces contamination from the mosquito and its microbes ²⁵. The resulting fully infectious sporozoites have allowed extensive ChIP, transcriptomic (RNA-seq) and proteomic (nanoLC-MS/MS) analyses of sporozoites from both rodent-infectious (*P. yoelii*) and human-infectious (*P. falciparum*, *P. vivax*) parasites ²⁶⁻³². Taken together, these studies demonstrate that ‘omics’ level efforts on sporozoites are

now experimentally practical, and thus reopen long standing questions of mechanisms underlying critical sporozoite functions.

Here we have addressed one of the foremost of these questions: how and when does the sporozoite prepare molecularly for transmission from the mosquito vector to the mammalian host? Using RNA-seq-based transcriptomics and nanoLC-MS/MS-based proteomics, we have characterized both oocyst sporozoites and salivary gland sporozoites of both rodent-infectious (*P. yoelii*) and human-infectious (*P. falciparum*) species. Together, these data provide a comprehensive assessment of mRNA and protein abundances, provide evidence for extensive post-transcriptional regulation of the most abundant mRNAs, and demonstrate that two distinct and likely orthogonal translational repression programs are active during sporozoite maturation.

Results:

Dynamic Transcriptional Regulation during Plasmodium Sporozoite Development

Important insights into how sporozoites mature and become infectious were gained from studies of the Upregulated in Oocyst Sporozoites (UOS) and Upregulated in Infectious Sporozoites (UIS) transcripts in *Plasmodium*. Moreover, a number of these UIS genes turned out to be essential to hepatocyte infection and the early liver stage parasite. However, prior studies were limited by the methods and instrumentation available, thus resulting in an incomplete view of transcriptional regulation in the sporozoite. By leveraging RNA-seq and greatly improved sporozoite purification strategies, we could now achieve a more comprehensive transcriptome and differential expression analyses of sporozoites from rodent-infectious (*P. yoelii*) and human-infectious (*P. falciparum*) species. Additionally, as the sporozoite undergoes a transition from being weakly infectious to highly infectious for the mammalian host, which occurs while in transit through the hemocoel from the oocyst to the salivary glands and

within the salivary glands^{5,6}, we assessed both the oocyst sporozoite and salivary gland sporozoite transcriptomes (Figure 1, Table 1, Table S1).

First, using *P. yoelii* (17XNL non-lethal strain) rodent-infectious parasites we identified 4195 and 3887 RNAs with detectable and unambiguous sequence reads present in *P. yoelii* oocyst sporozoites and salivary gland sporozoites, respectively. Similarly, with *P. falciparum* (NF54 strain) human-infectious parasites, we identified 3535 and 3575 detectable and unambiguous RNAs in oocyst sporozoite and salivary gland sporozoite stages, respectively. Many well-characterized genes were among the most abundant transcripts in these two stages, including *apical membrane antigen 1 (ama1)*, *circumsporozoite protein (csp)*, *membrane-associated erythrocyte binding-like protein (maeb1)*, *perforin-like protein 1 (plp1/pp1p1/spect2)*, *thrombospondin-related anonymous protein (trap)*, *trap-like protein (t1p)*, *upregulated in infectious sporozoites 4 (uis4)*, and others (Table 1, Table S1)³³⁻³⁸. However, numerous mRNAs encode for proteins with un/under-characterized roles in the parasite. Notably, several of the most abundant mRNAs in oocyst and salivary gland sporozoites in both species encode for uncharacterized proteins, some of which (e.g. *py17x_0208200*, *py17x_0835500*, *py17x_1354300*) undergo the same extreme swings in transcript abundance between these stages as does *pyuis4* (>1000-fold). Finally, we found that a recently described sporozoite var gene (SpzPfEMP1) is robustly expressed in not only oocyst sporozoites as previously reported, but also in salivary gland sporozoites and thus may simplify the recently described model of how this interesting var gene is regulated (Table 1)³². Given the transcript abundance of the novel and uncharacterized genes of these lists, they warrant a prioritized assessment.

Previous definitions of UOS or UIS mRNAs were assigned using lower thresholds of >2-fold increases in transcript abundance for any detectable transcript between oocyst and salivary gland sporozoites, which in part were dictated by the power of subtractive cDNA hybridization or microarray approaches available at the time^{9,10}. With greatly improved approaches, we have made the definitions of UIS and

160 UOS RNAs more stringent by assigning thresholds whereby transcripts must be both in the top decile of
161 abundance, and must be >6-fold more abundant in one stage compared to the other. Using these
162 parameters, we have defined 164 UOS mRNAs and 88 UIS mRNAs in *P. yoelii*, and 101 UOS mRNAs and
163 68 UIS mRNAs in *P. falciparum* (Table S2). Few of the UOS transcripts previously defined remain so using
164 these more stringent thresholds, but robustly include the previous top UOS hit: TREP/UOS3¹⁰.
165 Additional UOS transcripts include those that encode for proteins important for sporozoite functions in
166 the mosquito and the initial infection of the new host, and include those that encode for RNA metabolic
167 processes, protein translation, heat shock proteins (HSP20), the glideosome/inner membrane complex
168 (GAPM3, IMC1m), vesicular trafficking, and transporters. Similarly, a core set of the most abundant UIS
169 transcripts remain defined as such here: *uis1*, *uis2*, *uis3*, *uis4*, *uis7*, *uis8*, and *uis28*. Strikingly, *pyuis4*
170 transcript abundance increases 1500-fold and reaffirms the use of its promoter for highly enriched
171 expression of transgenes in salivary gland sporozoites^{28,39}. An additional 71 *P. yoelii* and 53 *P. falciparum*
172 transcripts that were not in the top decile of RNA abundance, but were in the seventh to ninth decile,
173 increase >10-fold in abundance in salivary gland sporozoites vs. oocyst sporozoites. These include *pfccr4*,
174 the *pfdbp10* and *pydbp10* RNA helicase, *pfsllarp/pfsap1* (40-fold), *pynek1* (38-fold), *pypplp3*, *pyplp5*,
175 *pfptex150*, *pyuis5* (22-fold) and *pyuis12* (40-fold) (Table S3). Beyond the previously defined UIS mRNAs,
176 several other transcripts are notable as their protein products are or may be important to the function
177 of salivary gland sporozoites: FAS-II pathway proteins, fatty acid modifiers, plasma membrane
178 transporters, adhesins/surface proteins (*p113*, *speld*, *tlp*), traversal-related proteins (*celtos*), heat shock
179 proteins, and ApiAP2 specific transcription factors (e.g. *py17x_0523100*, *pf3d7_0420300*)⁴⁰⁻⁴⁷. Together,
180 these transcripts encode for proteins that encompass many essential attributes necessary for sporozoite
181 development, transmission, and infectivity. However, it is striking that while a similar regulatory strategy
182 is employed by sporozoites of both parasite species, there is less overlap in the transcripts that are
183 regulated than might be expected, especially in oocyst sporozoites. These findings underscore the

strengths and importance of using comparative approaches with human- and rodent-infectious species to identify the important and conserved molecular components of infection.

Proteomic Comparisons of Plasmodium Sporozoites

While transcriptomics can provide an important window into gene expression, the inclusion of proteomics provides a much more comprehensive understanding of the parasite and its functions. To determine the presence and steady-state abundance of proteins found within *Plasmodium* sporozoites, the global proteomes of both *P. yoelii* and *P. falciparum* oocyst sporozoites were determined by nanoLC/MS/MS and were compared to our previously published salivary gland sporozoite global proteomes²⁹. This approach (steady-state protein abundance) was used, as it is compatible with currently feasible sporozoite production and purification capabilities, whereas ribosome profiling remains technically unfeasible with sporozoite samples due to the number of highly purified sporozoites that are required. Together, these four datasets now allow a more complete understanding of the oocyst and salivary gland sporozoite stages, and also allows for the definition of UOS Proteins and UIS Proteins for these two species of the malaria parasite.

Using approximately four million purified sporozoites per biological replicate, protein lysates were separated in a single lane of a gradient SDS-polyacrylamide gel, digested with trypsin, and the resulting tryptic peptides were extracted and subjected to nanoLC-MS/MS. Resulting mass spectra were assessed with the Trans-Proteomic Pipeline (TPP) to identify peptides and to infer identities of proteins. In sum, re-analysis of our previously acquired *P. falciparum* data identified 2037 salivary gland sporozoite proteins²⁹ and we now here also identify 1430 oocyst sporozoite proteins; similarly, from our previously acquired *P. yoelii* data, we identified 1773 salivary gland sporozoite proteins, and here identify 1760 oocyst sporozoite proteins (Figure 2, Table 2, see Table S1 for a complete list). In addition to detecting

well-characterized sporozoite proteins (CSP, CelTOS, TRAP, IMC/glideosome proteins, ALBA proteins, SIAP1) and abundant housekeeping proteins (histones, HSPs, GAPDH, translation-related proteins), dynamic changes in the abundance of specific proteins between these sporozoite stages were also identified. For instance, in both *P. yoelii* and *P. falciparum*, CelTOS, GEST and SPELD were not detected or were only weakly expressed in oocyst sporozoites, but were among the most abundant proteins in salivary gland sporozoites (Table 2, Table S1). Moreover, the presence/absence of cellular regulators such as specific ApiAP2s, histone modifiers, RNA-binding proteins and other proteins (Table S1) agree with previous reports describing how these types of regulation may be used by the sporozoite^{32,40,48}.

With proteomic data from both oocyst sporozoites and salivary gland sporozoites, we have expanded the UIS and UOS designations to proteins that are differentially abundant in one stage or the other. These designated UIS and UOS proteins in *P. yoelii* and *P. falciparum* were identified using the same stringent threshold as was applied to mRNA abundances (>6-fold more abundant in one stage compared to the other). Moreover, this was applied only to the top half of detected proteins of oocyst sporozoites (for UOS proteins) or salivary gland sporozoites (for UIS proteins), as differences in protein abundances quantified by spectral counting methods are most robust among higher-abundance proteins⁴⁹. Based upon this analysis, we identified 30 UOS proteins and 114 UIS proteins in *P. falciparum*, and 65 UOS and 65 UIS proteins in *P. yoelii*. UOS proteins detected in both species include UOS3/TREP, PCRMP2, and PCRMP4 (Figure 3, Table S4), which all have clearly been shown to be expressed in and important to oocyst sporozoites^{10,50,51}. Similarly, UIS proteins in both species include 6-Cys proteins (P38, P36, P52, B9), CLAMP, GEST, PLP1/PPLP1/SPECT2, PUF2, Sir2A, SPATR, SPELD, TRAMP, and UIS2, which have roles in salivary gland sporozoite infectivity, enable the sporozoite to navigate the host skin and liver, or traverse and productively infect hepatocytes. As with differential expression of RNA, species-specific differences in protein abundance changes were observed, with notable proteins being ApiAP2-SP (*P. falciparum*), and SPECT1, UIS3 and GAMER (*P. yoelii*). These data sets include many of the best-

characterized sporozoite proteins, which are also known to be critical to sporozoite maturation and transmission. However, in both species and in both stages, 28-49% of the proteins now defined as UOS and UIS proteins remain uncharacterized and are likely to be important to sporozoite functions in the mosquito vector and the mammalian host.

Comparison of UOS/UIS Designations Within and Across Species

When comparisons across species for UOS RNAs and proteins (Figure 3A) or UIS RNAs and proteins (Figure 3B) are made, several key features emerge. First, there are very few gene products that receive the same UOS designations across species (e.g. 14 UOS mRNAs, 6 UOS proteins), few that are both UOS mRNAs and proteins (4 in *P. falciparum*, 7 in *P. yoelii*), and only a single instance of a syntenic gene that encodes a UOS mRNA and protein in both species: TREP/UOS3. However, these proteins are known to be important to the oocyst sporozoite. For example, we identified TREP/UOS3 as well as *Plasmodium* Cysteine Repeat Modular Protein 2 and 4 (PCRMP2, PCRMP4) as UOS proteins in both *P. yoelii* and *P. falciparum*. TREP/UOS3 and PCRMP2 are important for sporozoite targeting to the salivary gland^{10,51}, whereas PCRMP4 is important for oocyst egress⁵⁰. Similarly, relatively few gene products receive the same UIS designations across species (Figure 3B), but those that do include several gene products known to be important to salivary gland sporozoites. For instance, 10 cross species UIS mRNAs and 25 UIS proteins were detected, and include gene products that enable the sporozoite to preserve its infectivity (PUF2), relieve translational repression (UIS2), traverse through host cells (PLP1, CelTOS) and more. Two of these gene products are both UIS mRNAs and UIS proteins in both species: CelTOS and SPELD. While less is known about SPELD, which is found on the sporozoite surface and is essential for liver stage development, CelTOS has been the focus of substantial study and is a promising target for therapeutic interventions^{41,52}. As above, we anticipate that the uncharacterized gene products

identified here will play similar roles to those that have already been studied, and deserve prioritization in future work. Together, these data indicate that similar gene regulatory strategies are used by *P. yoelii* and *P. falciparum* sporozoites, but in order to navigate and interact with specific host environments, they must regulate their most abundant gene products differently. Relaxation of the more stringent thresholds used to define UOS and UIS gene products by inclusion of additional expression deciles and/or requiring a lower fold change yields substantially more overlap in the regulated gene products (Table S1).

Evidence of Two Independent Translational Repression Programs in Plasmodium Sporozoites

Plasmodium parasites have adopted the use of translational repression in female gametocytes in a manner analogous to the maternal-to-zygotic transition of metazoans, with translation of stored and protected mRNAs occurring post transmission to the mosquito¹⁸. However, far less is known about whether a similar, energetically unfavorable regulatory strategy is used in sporozoites. Currently, few transcripts have been shown to be translationally repressed in sporozoites through reverse genetic studies. The best studied example is the *uis4* transcript, which has *cis* control elements located in the coding sequence itself to limit translation of the UIS4 protein prior to transmission¹³. Additionally, a translational repressor, PUF2, has been shown to be essential for the preservation of sporozoite infectivity during an extended residence in the salivary glands^{28,53-55}. Recently, transcriptomic and proteomic data from *P. vivax* sporozoites has indicated that translational repression occurs in this species as well⁵⁶.

In order to identify putatively translationally repressed transcripts in sporozoites, we analyzed our transcriptomic and proteomic data for evidence of highly abundant transcripts for which no protein could be detected. Existing data suggests that translational repression is imperfect, meaning that translationally repressed mRNAs may still produce a detectable amount of protein. Therefore, in these

comparisons we used the following highly stringent criteria to define a translationally repressed transcript: 1) transcripts must be in the top decile of mRNA abundance, 2) the corresponding protein must be either undetected or exhibit a disproportionately low abundance (e.g. bottom 50th percentile), and 3) must encode for a protein with detectable tryptic peptides (Table S5). Through comparison of the combined RNA-seq and proteomics datasets, we found that translational repression is extensively imposed upon many of the most abundant mRNAs of both oocyst sporozoite and salivary gland sporozoite stages of both species (Table S6). The extent of translational repression of transcripts in the top decile of abundance is comparable across both species and both sporozoite stages, with each species having transcripts with no evidence (~40-50% of mRNAs), or no or low amounts (~68-80% of mRNAs) of protein detected. Specifically, 115 of 164 UOS mRNAs and 71 of 88 UIS mRNAs are translationally repressed in *P. yoelii*, whereas 69 of 101 UOS mRNAs and 50 of 68 UIS mRNAs are translationally repressed in *P. falciparum*. Complete lists of the top 10% most abundant transcripts that are translationally repressed are provided (Table S1).

These datasets also reveal that *Plasmodium* has implemented two discrete and likely orthogonal translational repression programs during sporozoite maturation and transmission. One program imposes translational repression in oocyst sporozoites, which is relieved in salivary gland sporozoites to allow for the production of highly abundant proteins (“TR-oospz to UIS Proteins program”) (Table 3). A second program imposes and retains translational repression upon mRNAs throughout sporozoite maturation (“Pan-Sporozoite TR program”), which may allow for de-repression in the liver stage parasite as in the case of PyUIS4 (Table 4). However, formal demonstration of the full scale of a post-transmission release from the Pan-Sporozoite TR program awaits technical advances to enable total proteomics of early liver stage parasites.

Strikingly, in the case of both programs, well-characterized mRNAs are regulated to allow production of their encoded proteins when they are required for the parasite’s activities. For instance, the TR-oospz to

UIS Protein program (Table 3) controls production of PLP1/PPLP1/SPECT2, CelTOS, and TLP, which are critical or essential for the sporozoite to navigate the host skin, vasculature, and liver^{34,46,47,57-59}. Analyses of the complete TR-oospz to UIS Protein dataset reveal significant GO terms noting roles in the apical invasion complex, the parasite cell surface, movement in host environments, and interaction with and entry into host cells (Table S6). These data are in full agreement with original studies of PLP1/PPLP1/SPECT2 and CelTOS in *P. berghei*, which used IFA, western blotting, and immuno-EM to show that neither protein is present in oocyst sporozoites, but that both become abundant in salivary gland sporozoites^{34,47}. Work on other proteins provide supporting evidence for their expression in and importance to sporozoite functions in the salivary gland and early steps in the infection of the mammalian host⁶⁰⁻⁶².

Similarly, the Pan-Sporozoite Translational Repression program affects transcripts that encode for proteins that are known to be important/essential for subsequent stages of the parasite, with notable overlapping regulation of MTRAP, ApiAP2-G3, ApiAP2-I, an unnamed ApiAP2, MSP4, SEA1, and PAIP1 in both species and with similar timing (Table 4). Additionally, in *P. yoelii*, several of the historically defined UIS mRNAs (UIS1, UIS4, UIS7, UIS12, UIS28), ApiAP2-SP3, ApiAP2-L, and others are regulated by this program. In *P. falciparum*, autophagy-related proteins, liver stage associated protein 2, WARP, ApiAP2-O4 and other regulator proteins are affected. This indicates that the sporozoite is capable of immediate regulation of these mRNAs before any significant translation can occur, and is consistent with models that position cytosolic granules near the nuclear pore complex to receive exported mRNAs⁶³. Taken together, these data indicate that sporozoites have evolved two overlapping and independent translational repression programs to prepare and remain poised for their next required functions.

Validation of Translational Repression of Select P. yoelii Transcripts in Sporozoites

To further validate the regulation of select mRNAs by the Pan-Sporozoite TR program, we have used a gold standard, gene-by-gene assessment of wild-type and transgenic *P. yoelii* salivary gland sporozoites by fluorescence microscopy. For this, we have selected genes that exhibit RNA abundances in either the top decile (UIS4 and PY17X_1354300) or the third decile (UIS12), but that by mass spectrometry have protein abundances in the lower half. Previous characterizations of UIS4 in both *P. yoelii* and *P. berghei* have yielded conflicting data on the presence/abundance of this protein using either fluorescence microscopy or mass spectrometry-based proteomics. To address this, we have generated rabbit polyclonal antisera against recombinant PyUIS4 to monitor protein abundance in wild-type sporozoites. By IFA, nearly all salivary gland sporozoites isolated 14 days post-infectious blood meal have no evidence of the UIS4 protein, with few sporozoites having PyUIS4 protein levels that were detectable above background (Figure 4A). This is consistent with a previous report that showed UIS4 protein levels increase over the residence time of *P. berghei* sporozoites in the salivary gland⁶⁴. Together with our current findings, these data are consistent with a robust but incomplete translational repression of UIS4, which becomes increasingly leaky over time in sporozoites, even in the earliest isolatable salivary gland sporozoites. We hypothesize that this might be attributed to our incomplete understanding of how to minimally perturb sporozoites upon extraction from the mosquito.

We have further investigated whether these classifications of translational repression apply to uncharacterized gene products with mRNAs in the top decile of abundance. To this end, we chose PY17X_1354300, which is one of the most abundant mRNAs in *P. yoelii* salivary gland sporozoites (99.5th percentile) but was among the least abundant proteins detected (Table S1). Thus, we created PY17X_1354300::GFP transgenic salivary gland sporozoites, and in agreement with the proteomic data, did not detect the presence of PY17X_1354300::GFP protein using anti-GFP antibodies (Figure 4B). Finally, while we have restricted our definition of translationally repressed transcripts to include only the

most abundant mRNAs, it is also likely that less abundant mRNAs are similarly regulated as well. To address this, we have assessed PyUIS12 protein expression in salivary gland sporozoites, as it has high mRNA expression (80th percentile) but was barely detected by mass spectrometry. Using live fluorescence microscopy with PyWT-GFP or PyUIS12::GFP sporozoites, we clearly observed GFP expression in control *P. yoelii* WT-GFP sporozoites, but did not detect UIS12::GFP protein when transcribed from its native locus (Figure 4C). In agreement with translational repression being relieved post-transmission, IFA micrographs clearly show UIS12::GFP expression in the cytosol of 24 hour old liver stage parasites (Figure 4D). Taken together, these data indicate that *Plasmodium* parasites can impose translational repression upon sporozoite transcripts, and can do so beyond what our conservative definition applied to only the top decile of RNA abundance encompasses. However, it is notable that because the consistency and completeness of this regulation varies across individual sporozoites, both global and single cell approaches (such as those applied here) are informative and required to understand this regulatory process.

Discussion

Plasmodium sporozoites offer an intriguing model of parasite infection biology with distinct infectivity profiles in the mosquito vector site of development (oocysts) and site of sequestration for transmission to the mammalian host (salivary glands). Here we report a comprehensive and comparative assessment of the transcriptomes and proteomes of both *P. yoelii* and *P. falciparum* sporozoites. We have captured these gene expression profiles for immature sporozoites from the mosquito midgut (oocyst sporozoites) and mature, infectious sporozoites from the mosquito salivary glands. From these extensive data sets, several important features of transcriptome and proteome regulation can be deciphered that are likely controlling the distinct sporozoite phenotypes in the mosquito vector and mammalian host.

First, these datasets provide a robust classification of transcript regulation across sporozoite maturation at both the mRNA and protein levels. Previous work identified mRNAs upregulated in oocyst sporozoites or salivary gland sporozoites, but relied upon less comprehensive instrumentation and low stringency thresholds. The use of current RNA-seq methodologies and improved genome annotation employed here provides a far more extensive and robust classification of UOS and UIS transcripts, and now does so for both rodent-infectious (*P. yoelii*) and human-infectious (*P. falciparum*) sporozoites (Figure 1, Table 1, Tables S1-3). Moreover, we have also assessed these parasites for large-scale changes in protein abundance through mass spectrometry-based proteomics (Figure 2, Tables 2, S1, and S4). Together, these data strongly align with the expression levels and timing reported for individually studied mRNAs and proteins, and will provide the foundation for a systems analysis of the regulatory networks that govern sporozoite infection biology.

Second, we uncovered evidence that extensive translational repression occurs in both *P. falciparum* and *P. yoelii*, in both oocyst sporozoites and salivary gland sporozoites. In analyzing our data, we first applied rigorous thresholds to interrogate the most abundant transcripts and proteins with the goal of identifying putative targets with the highest possible confidence. We felt that this was prudent, as detection of mRNAs by RNA-seq (which includes sample amplification approaches) is more sensitive than detection of proteins by mass spectrometry (which cannot benefit from sample amplification). Among the top decile of mRNAs by abundance, the encoded proteins for nearly half were not detected at all by mass spectrometry, and the encoded proteins for another quarter were detected at a disproportionately low abundance (Table S6). It is notable that relaxation of these thresholds reveals that translational repression also occurs with less abundant mRNAs, which we also observed through microscopy in the validation of PyUIS12 expression (Figure 4).

Moreover, we find that two translational repression programs appear to be functioning in sporozoites, with some transcripts being translationally repressed in oocyst sporozoites but highly translated in

salivary gland sporozoites (“TR-oospz to UIS Protein program”) while others remain translationally repressed throughout sporozoite maturation (“pan-sporozoite TR program”) (Figure 5). In the case of those proteins that have been characterized for their roles in sporozoite maturation and functions in the mosquito and host, clear patterns arise. The TR-oospz to UIS Protein program would provide for the rapid production of proteins in salivary gland sporozoites, and would be well suited for proteins that are needed immediately after transmission for host cell traversal in the skin, vasculature and liver, and/or for productive infection of hepatocytes. In agreement with this, we find several proteins with known roles in cell traversal (PLP1/PPLP1/SPECT2, CelTOS, GAMER, TLP; Table 3). The second Pan-sporozoite TR program, particularly those UIS transcripts that are expressed only in salivary gland sporozoites but that are translationally repressed, would regulate the establishment of a new intrahepatocytic liver stage of infection by allowing for the rapid translation of these mRNAs after hepatocyte invasion. In the absence of robust global proteomic analysis of the early liver stage parasite, which will be exceedingly difficult to achieve, this dataset constitutes a best possible platform from which to assess the liver stage proteome in a candidate-based approach.

Together, our findings indicate that a multi-tiered, temporal translational repression mechanism is at work in *Plasmodium* sporozoites. This regulatory system aligns with the several windows of functionality that are required for sporozoites’ journey as they egress from oocysts on the mosquito midgut, migrate through the hemocoel, invade the salivary glands, remain there poised for transmission and when transmitted, migrate in the mammalian tissue, traverse cells, and ultimately infect hepatocytes. As the post-transcriptional control of specific mRNAs is energetically unfavorable as compared to *de novo* transcription, we hypothesize that reducing the time between receipt of an external/environmental stimulus and the availability of a protein is critical to the parasite. The use of a TR-oospz to UIS Protein program is straightforward, as it would require a short duration of translational repression until invasion of the salivary glands occurs. However, it is less clear why an energetically unfavorable Pan-Sporozoite

program would be activated in oocyst sporozoites, instead of simply transcribing these mRNAs in salivary gland sporozoites. One scenario that could explain the use of both programs is one where the sporozoite invades the salivary gland and then is immediately transmitted, as this would allow immediate responses to both events. While population-level approaches (like those used here) are practical and informative, single-cell approaches should be coupled with them to uncover meaningful differences in the variance of gene expression across individual sporozoites. Enabling technology for single-cell RNA sequencing is currently feasible, and single-cell proteomics is on the horizon.

Finally, new questions emerge from these data. For instance, what are the *trans* factors and *cis* elements responsible for these two likely orthogonal translational repression systems? While several RNA-binding proteins (RBPs) have been implicated in the preparation of salivary gland sporozoites for transmission, specific RBPs have not been associated with specific transcripts in the sporozoite. Moreover, as the TR-oospz to UIS Protein program initiates in the oocyst sporozoite, these experiments must also be pursued in that stage as well. Additionally, at what point are mRNAs governed by the Pan-Sporozoite TR program released for translation? The prevailing model based upon a few gene-specific examples suggests that it should be relieved after hepatocyte infection in early liver stage. However, total proteomic analyses of this stage are not available and will be technically challenging to conduct due to the relatively large proportion of host-to-parasite material. Also, could another program be active in those species that produce latent liver stage forms, called hypnozoite stages? As the commitment to active or latent liver stage forms likely occurs early, having a translational repression program available in sporozoites could allow for this, and could also contribute to determining the ratio of active to latent liver stage parasites that form. Lastly, perhaps the most appealing questions of all revolve around the uncharacterized and under-characterized gene products identified here. These may provide new clues to unappreciated parasite functions, or produce proteins so very different from host proteins that they can be therapeutically targeted.

Materials and methods:

Production and purification of Plasmodium yoelii and Plasmodium falciparum Sporozoites

Wild-type *Plasmodium yoelii* (17XNL strain) sporozoites were produced as previously described ²⁸.

Briefly, 6-to-8 week old Swiss Webster mice were infected by intraperitoneal (IP) injection of cryopreserved infected blood and were monitored until the peak day of male gametocyte exflagellation.

Mice were then anesthetized with an IP injection of ketamine/xylazine and exposed to 150-200

Anopheles stephensi mosquitoes for 15 minutes with periodic movements on the cage to promote

consistency in the transmission of parasites to the mosquito population. Oocyst sporozoites were

collected by microdissection and grinding of mosquito midguts on day 10 post-blood meal, whereas

salivary gland sporozoites were similarly collected from salivary glands on day 14 post-blood meal.

All animal care adheres to the Association for Assessment and Accreditation of Laboratory Animal Care

(AAALAC) guidelines, and all experiments conformed to approved IACUC protocols at the Center for

Infectious Disease Research (formerly Seattle Biomedical Research Institute, Protocol ID #SK-02 to

Stefan Kappe) or at Pennsylvania State University (Protocol ID #42678 to Scott Lindner). To this end, all

work with vertebrate animals was conducted in strict accordance with the recommendations in the

Guide for Care and Use of Laboratory Animals of the National Institutes of Health with approved Office

for Laboratory Animal Welfare (OLAW) assurance.

Wild-type *Plasmodium falciparum* (NF54 strain) sporozoites were produced as previously described ³¹ by

Seattle Children's (formerly the Center for Infectious Disease Research, Seattle Biomedical Research

Institute) and Johns Hopkins University. *P. yoelii* and *P. falciparum* sporozoites were purified by DEAE

sepharose and/or two sequential Accudenz gradients as previously described ^{25,29,65}.

Reverse Genetic Modification of P. yoelii Parasites

Plasmodium yoelii (17XNL strain) was genetically modified using conventional, double homologous recombination approaches with the pDEF plasmid vector as previously described²⁸. Oligonucleotides used for the creation of targeting sequences are listed in Table S7. The 3' end of *py17X_1354300* or *pyuis12* (PY17X_0507300) was modified by the addition of the GFPmut2 coding sequence prior to the stop codon. Transgenic parasites were identified by genotyping PCR, with independent transgenic clones being isolated by limiting dilution cloning. Clonal parasites were transmitted to *Anopheles stephensi* mosquitoes to produce salivary gland sporozoites as described above.

Live Fluorescence and Indirect Immunofluorescence Assays

Wild-type and transgenic *P. yoelii* sporozoites (PY17X_1354300::GFP, PyUIS12::GFP) were subjected to live fluorescence assays and/or an indirect immunofluorescence assay (IFA) (described in detail in⁶⁶) to characterize the extent of translational repression of these candidates. For live fluorescence microscopy of PyUIS12::GFP, freshly produced salivary gland sporozoites placed on glass slides in VectaShield, overlaid with a cover glass slip, and visualized by fluorescent microscopy using a Zeiss Axioscope A1 with 8-bit AxioCam ICc1 camera and Zen imagine software from the manufacturer. Alternatively, fresh salivary gland sporozoites were fixed in 10% v/v formalin for 10 min, and then air dried to a well on a glass slide defined by a hydrophobic coating. Sporozoites were treated for IFA as previously described, using either rabbit polyclonal anti-PyUIS4 (antigen consisting of AA80-224), produced by Pocono Rabbit Farm and Laboratory, Canadensis, PA) or rabbit polyclonal anti-GFPmut2 as primary antibodies and anti-rabbit IgG antibodies conjugated to Alexa Fluor 488 as a secondary antibody.

Comparative RNA-seq of Oocyst and Salivary Gland Sporozoites

For all oocyst sporozoite and salivary gland sporozoite replicates, RNA was prepared using the Qiagen RNeasy kit with two sequential DNaseI on-column digests, and was quality controlled by analysis on a BioAnalyzer. Barcoded libraries were created using the Illumina TruSeq Stranded mRNA Library Prep Kit, according to the manufacturer's protocol. Sequencing was conducted on an Illumina HiSeq 2500 using 100 nt single read length on three biological replicates per sample type. The resulting data was mapped to the respective reference genomes (*P. yoelii* 17XNL strain, plasmodb.org v30; *P. falciparum* 3D7 strain, plasmodb.org v30) using Tophat2 in a local Galaxy instance (version 2.1.0). Count files were generated using htseq-count (version 0.6.1) with a minimum alignment quality value set at 30 and a union mode setting. These count files compare the aligned BAM files to a reference GFF file (plasmodb.org v30 for both Py17xNL and Pf3D7) to evaluate the number of reads mapping to each feature, or gene. The count files are combined and compared across conditions using DEseq2 (version 2.11.38), which outputs complete transcript abundance comparisons and performs best among current differential expression tools for three biological replicates. Normalization values for these data are determined by DEseq2 across compared datasets for oocyst sporozoites and salivary gland sporozoites. Statistical metrics utilized by DEseq2 have been previously described⁶⁷. The average number of counts across biological replicates and their standard error of the mean were calculated to allow ranking of transcripts detected over background. Gene ontology terms (components, functions, and processes) were retrieved from PlasmoDB.org (v30). RNA-seq data reported here is available through the GEO depository (Accession #GSE113582).

Mass Spectrometry-Based Proteomics of Oocyst and Salivary Gland Sporozoites

SDS-PAGE fractionation

Purified sporozoites were subjected to SDS-PAGE pre-fractionation and in-gel tryptic digestion essentially as described in ^{29,30}. Briefly, samples were electrophoresed through a 4-20 % w/v SDS-polyacrylamide gel (Pierce Precise Tris-HEPES). Gels were stained with Imperial Stain (Thermo Fisher Scientific), de-stained in Milli-Q Water (Millipore), and cut into equal-sized fractions (26 fractions pooled into 13 LC-MS samples for the *P. yoelii* gel and 24 fractions analyzed as 24 LC-MS samples for the *P. falciparum* gel). Gel pieces were then de-stained with 50 mM ammonium bicarbonate (ABC) in 50 % acetonitrile (ACN) and dehydrated with ACN. Disulfide bonds were reduced with 10 mM DTT and cysteines were alkylated with 50 mM iodoacetamide in 100 mM ABC. Gel pieces were washed with ABC in 50 % ACN, dehydrated with ACN, and rehydrated with 6.25 ng/μL sequencing grade trypsin (Promega). After incubating overnight at 37 °C, the supernatant was recovered and peptides were extracted by incubating the gel pieces with 2 % v/v ACN/1 % v/v formic acid, then ACN. The extractions were combined with the digest supernatant, evaporated to dryness in a centrifugal vacuum concentrator, and reconstituted in liquid chromatography (LC) loading buffer consisting of 2 % v/v ACN/0.2 % v/v trifluoroacetic acid (TFA).

Liquid chromatography-mass spectrometry

LC and MS parameters were essentially as described previously ^{29,31}. Briefly, LC was performed using an Agilent 1100 nano pump with electronically controlled split flow at 300 nL/min (*P. falciparum* sample) or an Eksigent nanoLC at 500 nL/min (*P. yoelii* sample). Peptides were separated on a column with an integrated fritted tip (360 μm outer diameter (O.D.), 75 μm inner diameter (I.D.), 15 μm I.D. tip; New

Objective) packed in-house with a 20 cm bed of C18 (Dr. Maisch ReproSil-Pur C18-AQ, 120 Å, 3 µm). Prior to each run, sample was loaded onto a trap column consisting of a fritted capillary (360 µm O.D., 150 µm I.D.) packed with a 1 cm bed of the same stationary phase and washed with loading buffer. The trap was then placed in-line with the separation column for the separation gradient. The LC mobile phases consisted of buffer A (0.1 % v/v formic acid in water) and buffer B (0.1 % v/v formic acid in ACN). The separation gradient was 5 % B to 35 % B over 60 min (*P. falciparum* sample) or 90 min (*P. yoelii* sample). Tandem MS (MS/MS) was performed with a Thermo Fisher Scientific LTQ Velos Pro-Orbitrap Elite (*P. falciparum*) or LTQ Velos-Orbitrap (*P. yoelii*). Data-dependent acquisition was employed to select the top 20 precursors for collision-induced dissociation (CID) and analysis in the ion trap. Dynamic exclusion and precursor charge state selection were employed. Three nanoLC-MS technical replicates were performed for each fraction.

Peak list generation

The raw MS data from our previously reported analysis of salivary gland sporozoites²⁹ were re-analyzed using the same databases and parameters described here. Mass spectrometer output files were converted to mzML format using msConvert version 3.0.6002⁶⁸ and searched with Comet version 2015.02 rev.0⁶⁹. The precursor mass tolerance was ±20 ppm, and fragment ions bins were set to a tolerance of 1.0005 *m/z* and a monoisotopic mass offset of 0.4 *m/z*. Semi-trypsin peptides and up to 2 missed cleavages were allowed. The search parameters included a static modification of +57.021464 Da at Cys for formation of S-carboxamidomethyl-Cys by iodoacetamide and potential modifications of +15.994915 Da at Met for oxidation, -17.026549 Da at peptide N-terminal Gln for deamidation from formation of pyroGlu, -18.010565 Da at peptide N-terminal Glu for loss of water from formation of pyroGlu, -17.026549 Da at peptide N-terminal Cys for deamidation from formation of cyclized N-

terminal S-carboxamidomethyl-Cys, and +42.010565 for acetylation at the N-terminus of the protein, either at N-terminal Met or the N-terminal residue after cleavage of N-terminal Met. The spectra were searched against a database comprising either *P. falciparum* 3D7⁷⁰ or *P. yoelii yoelii* 17X⁷¹ (PlasmoDB v.30, www.plasmodb.org⁷²) appended with *Anopheles stephensi* Indian Astel2.3⁷³ (VectorBase, www.vectorbase.org⁷⁴), and a modified version of the common Repository of Adventitious Proteins (v.2012.01.01, The Global Proteome Machine, www.thegpm.org/cRAP) with the Sigma Universal Standard Proteins removed and the LC calibration standard peptide [Glu-1] fibrinopeptide B appended. Decoy proteins with the residues between tryptic residues randomly shuffled were generated using a tool included in the TPP and interleaved among the real entries. The MS/MS data were analyzed using the Trans-Proteomic Pipeline (TPP)⁷⁵ version 5.0.0 Typhoon. Peptide spectrum matches (PSMs) were assigned scores in PeptideProphet, peptide-level scores were assigned in iProphet⁷⁶, and protein identifications were inferred with ProteinProphet⁷⁷. In the case that multiple proteins were inferred at equal confidence by a set of peptides, the inference was counted as a single identification and all relevant protein ID's were listed. Only proteins with ProteinProphet probabilities corresponding to a false discovery rate (FDR) less than 1.0 % (as determined from the ProteinProphet mixture models) were reported.

Protein quantification

Relative protein abundance within and between samples was estimated using label-free proteomics methods based on spectral counting. Briefly, the spectral counts for a protein were taken as the total number of high-quality PSMs (identified at an iProphet probability corresponding to an FDR less than 1.0 %) that identified the protein. Spectral counts were quantified using the StPeter program in the TPP⁷⁸. The distributed spectral counts model was used to divide PSMs from degenerate peptides (peptides

whose sequences were found in multiple proteins in the database) among proteins containing that peptide in a weighted fashion⁷⁹. Relative protein abundance within samples was ranked using the normalized spectral abundance factor (NSAF)^{49,80}. Relative protein abundance ratios based on spectral counts were normalized and p-values were assigned as previously described³⁰. The raw and fully analyzed data files for these mass spectrometry-based proteomic experiments have been deposited in PRIDE (Accession # PXD009726, PXD009727, PXD009728, PXD009729).

Prediction of tryptic peptides

The CONSeQuence algorithm was used to identify proteins with detectable fully tryptic peptides with no missed cleavages⁸¹. A threshold of a Rank score ≥ 0.5 (derived from the combined predictors) was applied, a cutoff that had a sensitivity >70% with a false positive rate <50% when tested on datasets other than the training data as reported by the developers (21813416). Application of this algorithm with this threshold to our published *P. falciparum* salivary gland sporozoite proteome only misidentified 2.9% of all proteins as having no detectable peptides²⁹.

Acknowledgements

We appreciate and acknowledge ongoing scientific discussions with Istvan Albert, Aswathy Sebastian, Manuel Llinás and his research group, as well as advice given on sequencing services provided by the Penn State Genomics Core Facility (University Park, PA). The authors would like to thank the Johns Hopkins Malaria Research Institute Insectary and Parasitology core facilities, particularly, Christopher Kizito for expert rearing of mosquitoes, and Drs. Abhai Tripathi and Godfree Mlambo for production of the *P. falciparum*-infected mosquitos. We are grateful to Bloomberg Philanthropies for support of these core facilities.

Research reported in this publication was supported by Penn State start-up funds (SEL), the National Institutes of Health National Institute of Allergy and Infectious Disease (<http://www.niaid.nih.gov/>) under award numbers 1K22AI101039 (SEL), 1R01AI123341 (SEL), K25AI119229 (KES), R01AI132359 (PS), R21AI133369 (SHIK), and R01AI134956 (SHIK), by the National Institutes of Health National Institute of General Medical Sciences (www.nigms.nih.gov) under award number R01GM087221 (RLM), by the National Institutes of Health National Center for Research Resources under award number S10RR027584 (RLM), by the National Science Foundation (www.nsf.gov) under award number 0923536 (RLM, KES), and by the Provost's Postdoctoral Diversity Fellowship from Johns Hopkins University (MS). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the National Science Foundation or the Bill and Melinda Gates Foundation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conducted experiments: SEL, KES, MS, MPW, ENV, KJH, AMM; Analyzed Data: SEL, KES, MPW, PS, SHIK;

Wrote manuscript: SEL, KES, PS, SHIK; Funded the work: SEL, KES, PS, RLM, SHIK

Competing interests

The authors declare no competing interests.

Figure Legends and Table Legends

Figure 1: Comparative transcriptomics of oocyst and salivary gland sporozoites. RNA from purified (A) *P. falciparum* or (B) *P. yoelii* sporozoites isolated from oocysts or the salivary gland was assessed by RNA-seq, and transcript abundances compared by DEseq2. Transcripts were binned based upon fold change and adjusted p-value, with mRNAs shaded in lighter (+/- 1 to 2.5 log2 fold change) or darker shades (>2.5 log2 fold change).

Figure 2: Comparative total proteomics of oocyst sporozoites and salivary gland sporozoites. Total protein from purified (A) *P. falciparum* or (B) *P. yoelii* sporozoites isolated from oocysts or the salivary gland was separated by a gradient SDS-PAGE, digested with trypsin, extracted, and subjected to analysis. Proteins with statistically significant differences in abundance are noted with red dots, and expression percentile thresholds are provided by dashed lines. Proteins that were only detected in oocyst sporozoites or salivary gland sporozoites are boxed in light blue and purple dashed lines, respectively.

Figure 3: Comparisons of UOS and UIS Gene Products. RNAs or proteins in the top decile of abundance that were also at least 6-fold more abundant in either oocyst sporozoites or salivary gland sporozoites were denoted at UOS or UIS gene products, respectively. Comparisons across molecule types and species for (A) UOS and (B) UIS identify many gene products that are critical/essential to sporozoite development and/or transmission.

Figure 4: Validation of translationally repressed transcripts in *P. yoelii*. The presence and relative abundance of proteins encoded by translationally repressed mRNAs was assessed by fluorescence microscopy. (A) PyUIS4 expression is largely absent in the population of wild-type sporozoites taken shortly after salivary gland invasion. Very few sporozoites have detectable PyUIS4 protein, which is weak and diffuse. (B) An uncharacterized protein, PY17X_1354300, was modified to encode a C-terminal GFP tag, and protein expression was monitored by IFA. In agreement with the total proteomic data, no protein was detectable. (C) Transgenic parasites expressing unfused GFP (WT-GFP control) or GFP fused to UIS12 were assessed by live fluorescence for evidence of protein expression. No PyUIS12::GFP could be detected above background in sporozoites. (D) However, robust expression of PyUIS12::GFP is detected by IFA post-transmission in 24 hour liver stage parasites.

Figure 5: A model for two independent translational repression programs in *Plasmodium* sporozoites. One regulatory program acts upon mRNAs that are highly abundant in oocyst sporozoites to impose translational repression, which is relieved in salivary gland sporozoites and enables protein translation to occur. Transcripts encoding for proteins important to host cell traversal and infection are regulated in this manner. A second regulatory program translationally represses mRNAs throughout sporozoite maturation, which is hypothesized to be relieved following transmission and infection of a hepatocyte to promote liver stage development.

Table 1: The top 25 most abundant mRNAs in oocyst sporozoites and salivary gland sporozoites of *P.*

falciparum and *P. yoelii*. Transcripts are defined as Upregulated in Oocyst Sporozoites (UOS) or

Upregulated in Infectious Sporozoites (UIS) mRNAs as well when their abundance is in the top decile,

and their fold change is >6-fold in one sporozoite stage compared to the other. A complete listing of

mRNA abundances and fold changes between sporozoite stages is provided in Table S1.

Table 2: The top 30 most abundant proteins in oocyst sporozoites and salivary gland sporozoites of *P.*

falciparum and *P. yoelii*. Proteins are defined as Upregulated in Oocyst Sporozoites (UOS) or

Upregulated in Infectious Sporozoites (UIS) proteins as well when their abundance is in the top half of

the proteome and their fold change is >6-fold in one sporozoite stage compared to the other. A

complete listing of protein abundances using label-free quantification and fold changes between

sporozoite stages is provided in Table S1.

Table 3: Gene products under the translational repression in oocyst sporozoite to UIS protein regulatory

program (“TR oospz to UIS protein program”) in *P. yoelii* and *P. falciparum*. Published gene deletion

phenotypes related to sporozoite traversal and infectivity of hosts is noted with references when

available. Four gene products (top four rows) are similarly regulated across species.

Table 4: Gene products under the Pan-Sporozoite translational repression program in *P. yoelii* and *P.*

falciparum. Gene ID, common gene name, description, and duration of regulation (OO+SG = oocyst

sporozoite and salivary gland sporozoite, SG-only = salivary gland sporozoite only). Several gene

products (top seven rows) are similarly regulated across species.

674 **Supporting Table Legends**

675 **Table S1:** Complete transcriptomic and proteomic datasets and their analyses.

676 **Table S2:** UOS and UIS mRNAs from *P. falciparum* and *P. yoelii*.

677 **Table S3:** Transcripts in the top seventh to ninth deciles that increase in abundance 10-fold or greater in
678 salivary gland sporozoites as compared to oocyst sporozoites.

679 **Table S4:** UOS and UIS proteins from *P. falciparum* and *P. yoelii*.

680 **Table S5:** Tryptic peptide analyses of putatively translationally repressed transcripts using the
681 CONSeQuence algorithm.

682 **Table S6:** Extended data and GO terms on translationally repressed transcripts.

683 **Table S7:** Oligonucleotides used in this study.

684 **References**

- 685 1 Olotu, A. *et al.* Seven-Year Efficacy of RTS,S/AS01 Malaria Vaccine among Young African
686 Children. *N Engl J Med* **374**, 2519-2529, doi:10.1056/NEJMoa1515257 (2016).
- 687 2 Itsara, L. S. *et al.* The Development of Whole Sporozoite Vaccines for Plasmodium falciparum
688 Malaria. *Front Immunol* **9**, 2748, doi:10.3389/fimmu.2018.02748 (2018).
- 689 3 Shaw, W. R. & Catteruccia, F. Vector biology meets disease control: using basic research to fight
690 vector-borne diseases. *Nat Microbiol* **4**, 20-34, doi:10.1038/s41564-018-0214-7 (2019).
- 691 4 Rosenberg, R. Malaria: some considerations regarding parasite productivity. *Trends in*
692 *parasitology* **24**, 487-491, doi:10.1016/j.pt.2008.07.009 (2008).
- 693 5 Vanderberg, J. P. Development of infectivity by the Plasmodium berghei sporozoite. *The Journal*
694 *of parasitology* **61**, 43-50 (1975).
- 695 6 Sato, Y., Montagna, G. N. & Matuschewski, K. Plasmodium berghei sporozoites acquire virulence
696 and immunogenicity during mosquito hemocoel transit. *Infection and immunity* **82**, 1164-1172,
697 doi:10.1128/IAI.00758-13 (2014).
- 698 7 Hopp, C. S. & Sinnis, P. The innate and adaptive response to mosquito saliva and Plasmodium
699 sporozoites in the skin. *Ann N Y Acad Sci* **1342**, 37-43, doi:10.1111/nyas.12661 (2015).
- 700 8 Medica, D. L. & Sinnis, P. Quantitative dynamics of Plasmodium yoelii sporozoite transmission by
701 infected anopheline mosquitoes. *Infection and immunity* **73**, 4363-4369,
702 doi:10.1128/IAI.73.7.4363-4369.2005 (2005).
- 703 9 Matuschewski, K. *et al.* Infectivity-associated changes in the transcriptional repertoire of the
704 malaria parasite sporozoite stage. *J Biol Chem* **277**, 41948-41953, doi:10.1074/jbc.M207315200
705 (2002).
- 706 10 Mikolajczak, S. A. *et al.* Distinct malaria parasite sporozoites reveal transcriptional changes that
707 cause differential tissue infection competence in the mosquito vector and mammalian host. *Mol*
708 *Cell Biol* **28**, 6196-6207, doi:10.1128/MCB.00553-08 (2008).
- 709 11 Real, E. *et al.* Plasmodium UIS3 sequesters host LC3 to avoid elimination by autophagy in
710 hepatocytes. *Nat Microbiol* **3**, 17-25, doi:10.1038/s41564-017-0054-x (2018).
- 711 12 Zhang, M., Mishra, S., Sakthivel, R., Fontoura, B. M. & Nussenzweig, V. UIS2: A Unique
712 Phosphatase Required for the Development of Plasmodium Liver Stages. *PLoS pathogens* **12**,
713 e1005370, doi:10.1371/journal.ppat.1005370 (2016).
- 714 13 Silvie, O., Briquet, S., Muller, K., Manzoni, G. & Matuschewski, K. Post-transcriptional silencing of
715 UIS4 in Plasmodium berghei sporozoites is important for host switch. *Mol Microbiol* **91**, 1200-
716 1213, doi:10.1111/mmi.12528 (2014).
- 717 14 Zhang, M. *et al.* The Plasmodium eukaryotic initiation factor-2alpha kinase IK2 controls the
718 latency of sporozoites in the mosquito salivary glands. *The Journal of experimental medicine*
719 **207**, 1465-1474, doi:10.1084/jem.20091975 (2010).
- 720 15 Holmes, M. J., Augusto, L. D. S., Zhang, M., Wek, R. C. & Sullivan, W. J., Jr. Translational Control
721 in the Latency of Apicomplexan Parasites. *Trends in parasitology* **33**, 947-960,
722 doi:10.1016/j.pt.2017.08.006 (2017).
- 723 16 Cui, L., Lindner, S. & Miao, J. Translational regulation during stage transitions in malaria
724 parasites. *Ann N Y Acad Sci* **1342**, 1-9, doi:10.1111/nyas.12573 (2015).
- 725 17 Hausser, J., Mayo, A., Keren, L. & Alon, U. Central dogma rates and the trade-off between
726 precision and economy in gene expression. *Nature communications* **10**, 68, doi:10.1038/s41467-
727 018-07391-8 (2019).
- 728 18 Mair, G. R. *et al.* Regulation of sexual development of Plasmodium by translational repression.
729 *Science* **313**, 667-669, doi:10.1126/science.1125129 (2006).

730 19 Guerreiro, A. *et al.* Genome-wide RIP-Chip analysis of translational repressor-bound mRNAs in
731 the Plasmodium gametocyte. *Genome biology* **15**, 493, doi:10.1186/s13059-014-0493-0 (2014).
732 20 Mair, G. R. *et al.* Universal features of post-transcriptional gene regulation are critical for
733 Plasmodium zygote development. *PLoS pathogens* **6**, e1000767,
734 doi:10.1371/journal.ppat.1000767 (2010).
735 21 Lasonder, E. *et al.* Integrated transcriptomic and proteomic analyses of P. falciparum
736 gametocytes: molecular insight into sex-specific processes and translational repression. *Nucleic*
737 *acids research* **44**, 6087-6101, doi:10.1093/nar/gkw536 (2016).
738 22 Lasonder, E. *et al.* Proteomic profiling of Plasmodium sporozoite maturation identifies new
739 proteins essential for parasite development and infectivity. *PLoS pathogens* **4**, e1000195,
740 doi:10.1371/journal.ppat.1000195 (2008).
741 23 Khan, S. M. *et al.* Proteome analysis of separated male and female gametocytes reveals novel
742 sex-specific Plasmodium biology. *Cell* **121**, 675-687, doi:10.1016/j.cell.2005.03.027 (2005).
743 24 Lasonder, E. *et al.* Analysis of the Plasmodium falciparum proteome by high-accuracy mass
744 spectrometry. *Nature* **419**, 537-542, doi:10.1038/nature01111 (2002).
745 25 Kennedy, M. *et al.* A rapid and scalable density gradient purification method for Plasmodium
746 sporozoites. *Malaria journal* **11**, 421, doi:10.1186/1475-2875-11-421 (2012).
747 26 El-Manzalawy, Y., Munoz, E. E., Lindner, S. E. & Honavar, V. Plasmosep: Predicting surface-
748 exposed proteins on the malaria parasite using semisupervised self-training and expert-
749 annotated data. *Proteomics* **16**, 2967-2976, doi:10.1002/pmic.201600249 (2016).
750 27 Jex, A. *et al.* Transcriptome and histone epigenome of Plasmodium vivax salivary-
751 gland sporozoites point to tight regulatory control and potential mechanisms for liver-stage
752 differentiation. 145250, doi:10.1101/145250 %J bioRxiv (2018).
753 28 Lindner, S. E. *et al.* Perturbations of Plasmodium Puf2 expression and RNA-seq of Puf2-deficient
754 sporozoites reveal a critical role in maintaining RNA homeostasis and parasite transmissibility.
755 *Cellular microbiology* **15**, 1266-1283, doi:10.1111/cmi.12116 (2013).
756 29 Lindner, S. E. *et al.* Total and putative surface proteomics of malaria parasite salivary gland
757 sporozoites. *Mol Cell Proteomics* **12**, 1127-1143, doi:10.1074/mcp.M112.024505 (2013).
758 30 Swearingen, K. E. *et al.* Proteogenomic analysis of the total and surface-exposed proteomes of
759 Plasmodium vivax salivary gland sporozoites. *PLoS Negl Trop Dis* **11**, e0005791,
760 doi:10.1371/journal.pntd.0005791 (2017).
761 31 Swearingen, K. E. *et al.* Interrogating the Plasmodium Sporozoite Surface: Identification of
762 Surface-Exposed Proteins and Demonstration of Glycosylation on CSP and TRAP by Mass
763 Spectrometry-Based Proteomics. *PLoS Pathog* **12**, e1005606, doi:10.1371/journal.ppat.1005606
764 (2016).
765 32 Zanghi, G. *et al.* A Specific PfEMP1 Is Expressed in P. falciparum Sporozoites and Plays a Role in
766 Hepatocyte Infection. *Cell reports* **22**, 2951-2963, doi:10.1016/j.celrep.2018.02.075 (2018).
767 33 Mueller, A. K. *et al.* Plasmodium liver stage developmental arrest by depletion of a protein at
768 the parasite-host interface. *Proc Natl Acad Sci U S A* **102**, 3022-3027,
769 doi:10.1073/pnas.0408442102 (2005).
770 34 Ishino, T., Chinzei, Y. & Yuda, M. A Plasmodium sporozoite protein with a membrane attack
771 complex domain is required for breaching the liver sinusoidal cell layer prior to hepatocyte
772 infection. *Cell Microbiol* **7**, 199-208, doi:10.1111/j.1462-5822.2004.00447.x (2005).
773 35 Silvie, O. *et al.* A role for apical membrane antigen 1 during invasion of hepatocytes by
774 Plasmodium falciparum sporozoites. *J Biol Chem* **279**, 9490-9496, doi:10.1074/jbc.M311331200
775 (2004).
776 36 Kariu, T., Yuda, M., Yano, K. & Chinzei, Y. MAEBL is essential for malarial sporozoite infection of
777 the mosquito salivary gland. *J Exp Med* **195**, 1317-1323 (2002).

778 37 Sultan, A. A. *et al.* TRAP is necessary for gliding motility and infectivity of plasmodium
779 sporozoites. *Cell* **90**, 511-522 (1997).

780 38 Yoshida, N., Potocnjak, P., Nussenzweig, V. & Nussenzweig, R. S. Biosynthesis of Pb44, the
781 protective antigen of sporozoites of Plasmodium berghei. *The Journal of experimental medicine*
782 **154**, 1225-1236 (1981).

783 39 Hopp, C. S. *et al.* Longitudinal analysis of Plasmodium sporozoite motility in the dermis reveals
784 component of blood vessel recognition. *Elife* **4**, doi:10.7554/eLife.07789 (2015).

785 40 Zhang, C. *et al.* Systematic CRISPR-Cas9-Mediated Modifications of Plasmodium yoelii ApiAP2
786 Genes Reveal Functional Insights into Parasite Development. *MBio* **8**, doi:10.1128/mBio.01986-
787 17 (2017).

788 41 Al-Nihmi, F. M. *et al.* A Novel and Conserved Plasmodium Sporozoite Membrane Protein SPELD
789 is Required for Maturation of Exo-erythrocytic Forms. *Sci Rep* **7**, 40407, doi:10.1038/srep40407
790 (2017).

791 42 Offeddu, V., Rauch, M., Silvie, O. & Matuschewski, K. The Plasmodium protein P113 supports
792 efficient sporozoite to liver stage conversion in vivo. *Mol Biochem Parasitol* **193**, 101-109,
793 doi:10.1016/j.molbiopara.2014.03.002 (2014).

794 43 Lindner, S. E. *et al.* Enzymes involved in plastid-targeted phosphatidic acid synthesis are
795 essential for Plasmodium yoelii liver-stage development. *Mol Microbiol* **91**, 679-693,
796 doi:10.1111/mmi.12485 (2014).

797 44 van Schaijk, B. C. *et al.* Type II fatty acid biosynthesis is essential for Plasmodium falciparum
798 sporozoite development in the midgut of Anopheles mosquitoes. *Eukaryot Cell* **13**, 550-559,
799 doi:10.1128/EC.00264-13 (2014).

800 45 Vaughan, A. M. *et al.* Type II fatty acid synthesis is essential only for malaria parasite late liver
801 stage development. *Cell Microbiol* **11**, 506-520, doi:10.1111/j.1462-5822.2008.01270.x (2009).

802 46 Moreira, C. K. *et al.* The Plasmodium TRAP/MIC2 family member, TRAP-Like Protein (TLP), is
803 involved in tissue traversal by sporozoites. *Cell Microbiol* **10**, 1505-1516, doi:10.1111/j.1462-
804 5822.2008.01143.x (2008).

805 47 Kariu, T., Ishino, T., Yano, K., Chinzei, Y. & Yuda, M. CelTOS, a novel malarial protein that
806 mediates transmission to mosquito and vertebrate hosts. *Mol Microbiol* **59**, 1369-1379,
807 doi:10.1111/j.1365-2958.2005.05024.x (2006).

808 48 Munoz, E. E. *et al.* ALBA4 modulates its stage-specific interactions and specific mRNA fates
809 during Plasmodium yoelii growth and transmission. *Mol Microbiol* **106**, 266-284,
810 doi:10.1111/mmi.13762 (2017).

811 49 Gokce, E., Shuford, C. M., Franck, W. L., Dean, R. A. & Muddiman, D. C. Evaluation of
812 normalization methods on GeLC-MS/MS label-free spectral counting data to correct for variation
813 during proteomic workflows. *J Am Soc Mass Spectrom* **22**, 2199-2208, doi:10.1007/s13361-011-
814 0237-2 (2011).

815 50 Douradinha, B. *et al.* Plasmodium Cysteine Repeat Modular Proteins 3 and 4 are essential for
816 malaria parasite transmission from the mosquito to the host. *Malar J* **10**, 71, doi:10.1186/1475-
817 2875-10-71 (2011).

818 51 Thompson, J. *et al.* Plasmodium cysteine repeat modular proteins 1-4: complex proteins with
819 roles throughout the malaria parasite life cycle. *Cell Microbiol* **9**, 1466-1480, doi:10.1111/j.1462-
820 5822.2006.00885.x (2007).

821 52 Alves, E. *et al.* Evaluation of Plasmodium vivax Cell-Traversal Protein for Ookinetes and
822 Sporozoites as a Preerythrocytic P. vivax Vaccine. *Clin Vaccine Immunol* **24**,
823 doi:10.1128/CVI.00501-16 (2017).

824 53 Miao, J. *et al.* Puf mediates translation repression of transmission-blocking vaccine candidates in
825 malaria parasites. *PLoS Pathog* **9**, e1003268, doi:10.1371/journal.ppat.1003268 (2013).

826 54 Muller, K., Matuschewski, K. & Silvie, O. The Puf-family RNA-binding protein Puf2 controls
827 sporozoite conversion to liver stages in the malaria parasite. *PLoS One* **6**, e19860,
828 doi:10.1371/journal.pone.0019860 (2011).

829 55 Gomes-Santos, C. S. *et al.* Transition of Plasmodium sporozoites into liver stage-like forms is
830 regulated by the RNA binding protein Pumilio. *PLoS Pathog* **7**, e1002046,
831 doi:10.1371/journal.ppat.1002046 (2011).

832 56 Consortium, V. S. Transcriptome and histone epigenome of Plasmodium vivax salivary-gland
833 sporozoites point to tight regulatory control and mechanisms for liver-stage differentiation in
834 relapsing malaria. *International journal for parasitology*, doi:10.1016/j.ijpara.2019.02.007
835 (2019).

836 57 Steel, R. W. J. *et al.* Plasmodium yoelii S4/CeTOS is important for sporozoite gliding motility and
837 cell traversal. *Cell Microbiol* **20**, doi:10.1111/cmi.12817 (2018).

838 58 Jimah, J. R. *et al.* Malaria parasite CeTOS targets the inner leaflet of cell membranes for pore-
839 dependent disruption. *Elife* **5**, doi:10.7554/eLife.20621 (2016).

840 59 Risco-Castillo, V. *et al.* Malaria Sporozoites Traverse Host Cells within Transient Vacuoles. *Cell*
841 *Host Microbe* **18**, 593-603, doi:10.1016/j.chom.2015.10.006 (2015).

842 60 Yang, A. S. P. *et al.* Cell Traversal Activity Is Important for Plasmodium falciparum Liver Infection
843 in Humanized Mice. *Cell Rep* **18**, 3105-3116, doi:10.1016/j.celrep.2017.03.017 (2017).

844 61 Akinosoglou, K. A. *et al.* Characterization of Plasmodium developmental transcriptomes in
845 Anopheles gambiae midgut reveals novel regulators of malaria transmission. *Cell Microbiol* **17**,
846 254-268, doi:10.1111/cmi.12363 (2015).

847 62 Matthews, K. *et al.* The Plasmodium translocon of exported proteins (PTEx) component
848 thioredoxin-2 is important for maintaining normal blood-stage growth. *Mol Microbiol* **89**, 1167-
849 1186, doi:10.1111/mmi.12334 (2013).

850 63 Sheth, U., Pitt, J., Dennis, S. & Priess, J. R. Perinuclear P granules are the principal sites of mRNA
851 export in adult C. elegans germ cells. *Development* **137**, 1305-1314, doi:10.1242/dev.044255
852 (2010).

853 64 Silva, P. A. *et al.* Translational Control of UIS4 Protein of the Host-Parasite Interface Is Mediated
854 by the RNA Binding Protein Puf2 in Plasmodium berghei Sporozoites. *PLoS One* **11**, e0147940,
855 doi:10.1371/journal.pone.0147940 (2016).

856 65 Mack, S. R., Vanderberg, J. P. & Nawrot, R. Column separation of Plasmodium berghei
857 sporozoites. *J Parasitol* **64**, 166-168 (1978).

858 66 Pei, Y. *et al.* Plasmodium yoelii inhibitor of cysteine proteases is exported to exomembrane
859 structures and interacts with yoelipain-2 during asexual blood-stage development. *Cell Microbiol*
860 **15**, 1508-1526, doi:10.1111/cmi.12124 (2013).

861 67 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for
862 RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).

863 68 Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. ProteoWizard: open source software
864 for rapid proteomics tools development. *Bioinformatics* **24**, 2534-2536,
865 doi:10.1093/bioinformatics/btn323 (2008).

866 69 Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: an open-source MS/MS sequence database
867 search tool. *Proteomics* **13**, 22-24, doi:10.1002/pmic.201200439 (2013).

868 70 Gardner, M. J. *et al.* Genome sequence of the human malaria parasite Plasmodium falciparum.
869 *Nature* **419**, 498-511, doi:10.1038/nature01097 (2002).

870 71 Otto, T. D. *et al.* A comprehensive evaluation of rodent malaria parasite genomes and gene
871 expression. *BMC Biol* **12**, 86, doi:10.1186/s12915-014-0086-0 (2014).

872 72 Aurrecochea, C. *et al.* PlasmoDB: a functional genomic database for malaria parasites. *Nucleic*
873 *acids research* **37**, D539-543, doi:10.1093/nar/gkn814 (2009).

874 73 Jiang, X. *et al.* Genome analysis of a major urban malaria vector mosquito, *Anopheles stephensi*.
875 *Genome Biol* **15**, 459, doi:10.1186/s13059-014-0459-2 (2014).
876 74 Giraldo-Calderon, G. I. *et al.* VectorBase: an updated bioinformatics resource for invertebrate
877 vectors and other organisms related with human diseases. *Nucleic Acids Res* **43**, D707-713,
878 doi:10.1093/nar/gku1117 (2015).
879 75 Deutsch, E. W. *et al.* Trans-Proteomic Pipeline, a standardized data processing pipeline for large-
880 scale reproducible proteomics informatics. *Proteomics Clin Appl* **9**, 745-754,
881 doi:10.1002/prca.201400164 (2015).
882 76 Shteynberg, D., Nesvizhskii, A. I., Moritz, R. L. & Deutsch, E. W. Combining results of multiple
883 search engines in proteomics. *Mol Cell Proteomics* **12**, 2383-2393,
884 doi:10.1074/mcp.R113.027797 (2013).
885 77 Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins
886 by tandem mass spectrometry. *Anal Chem* **75**, 4646-4658 (2003).
887 78 Hoopmann, M. R., Winget, J. M., Mendoza, L. & Moritz, R. L. StPeter: Seamless Label-Free
888 Quantification with the Trans-Proteomic Pipeline. *J Proteome Res* **17**, 1314-1320,
889 doi:10.1021/acs.jproteome.7b00786 (2018).
890 79 Zhang, Y., Wen, Z., Washburn, M. P. & Florens, L. Refinements to label free proteome
891 quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem* **82**, 2272-2281,
892 doi:10.1021/ac9023999 (2010).
893 80 Zybilov, B. *et al.* Statistical analysis of membrane proteome expression changes in
894 *Saccharomyces cerevisiae*. *Journal of proteome research* **5**, 2339-2347, doi:10.1021/pr060161n
895 (2006).
896 81 Eysers, C. E. *et al.* CONSeQuence: prediction of reference peptides for absolute quantitative
897 proteomics using consensus machine learning approaches. *Mol Cell Proteomics* **10**, M110
898 003384, doi:10.1074/mcp.M110.003384 (2011).

899

Figure 1: Lindner and Swearingen et al.

A.

B.

Figure 1: Comparative transcriptomics of oocyst and salivary gland sporozoites. RNA from purified (A) *P. falciparum* or (B) *P. yoelii* sporozoites isolated from oocysts or the salivary gland was assessed by RNA-seq, and transcript abundances compared by DEseq2. Transcripts were binned based upon fold change and adjusted p-value, with mRNAs shaded in lighter (+/- 1 to 2.5 log₂ fold change) or darker shades (>2.5 log₂ fold change).

Figure 2: Comparative total proteomics of oocyst sporozoites and salivary gland sporozoites. Total protein from purified (A) *P. falciparum* or (B) *P. yoelii* sporozoites isolated from oocysts or the salivary gland was separated by a gradient SDS-PAGE, digested with trypsin, extracted, and subjected to analysis. Proteins with statistically significant differences in abundance are noted with red dots, and expression percentile thresholds are provided by dashed lines. Proteins that were only detected in oocyst sporozoites or salivary gland sporozoites are boxed in light blue and purple dashed lines, respectively.

Figure 3: Lindner and Swearingen et al.

Figure 3: Comparisons of UOS and UIS Gene Products. RNAs or proteins in the top decile of abundance that were also at least 6-fold more abundant in either oocyst sporozoites or salivary gland sporozoites were denoted at UOS or UIS gene products, respectively. Comparisons across molecule types and species for (A) UOS and (B) UIS identify many gene products that are critical/essential to sporozoite development and/or transmission.

Figure 4: Lindner and Swearingen *et al.*

Figure 4: Validation of translationally repressed transcripts in *P. yoelii*. The presence and relative abundance of proteins encoded by translationally repressed mRNAs was assessed by fluorescence microscopy. (A) PyUIS4 expression is largely absent in the population of wild-type sporozoites taken shortly after salivary gland invasion. Very few sporozoites have detectable PyUIS4 protein, which is weak and diffuse. (B) An uncharacterized protein, PY17X_1354300, was modified to encode a C-terminal GFP tag, and protein expression was monitored by IFA. In agreement with the total proteomic data, no protein was detectable. (C) Transgenic parasites expressing unfused GFP (WT-GFP control) or GFP fused to UIS12 were assessed by live fluorescence for evidence of protein expression. No PyUIS12::GFP could be detected above background in sporozoites. (D) However, robust expression of PyUIS12::GFP is detected by IFA post-transmission in 24 hour liver stage parasites.

Figure 5 Lindner and Swearingen *et al.*

Figure 5: A model for two independent translational repression programs in *Plasmodium* sporozoites.

One regulatory program acts upon mRNAs that are highly abundant in oocyst sporozoites to impose translational repression, which is relieved in salivary gland sporozoites and enables protein translation to occur. Transcripts encoding for proteins important to host cell traversal and infection are regulated in this manner. A second regulatory program translationally represses mRNAs throughout sporozoite maturation, which is hypothesized to be relieved following transmission and infection of a hepatocyte to promote liver stage development.

Table 1: The top 25 most abundant mRNAs in oocyst sporozoites and salivary gland sporozoites of *P. falciparum* and *P. yoelii*.

<i>P. yoelii</i> Gene ID	Common Name	Product Description	oo-spz Avg	oo-spz Rank	sg-spz Avg	sg-spz Rank	UIS or UOS
PY17X_0405400	CSP	circumsporozoite (CS) protein	94087.7	1	201220	1	
PY17X_1354800	TRAP	thrombospondin-related anonymous protein	8560.84	2	36972	2	
PY17X_0502200	UIS4	early transcribed membrane protein	2.84664	2489	4797.6	3	UIS
PY17X_0902700	MAEBL	membrane associated erythrocyte binding-like protein	3639.67	3	1490.6	4	
PY17X_0504800		conserved Plasmodium protein, unknown function	11.0612	1125	1005.2	5	UIS
PY17X_0826700	PGK	phosphoglycerate kinase, putative	721.68	8	874.52	6	
PY17X_1218100		conserved Plasmodium protein, unknown function	61.4346	224	872.5	7	UIS
PY17X_0624100		RNA-binding protein, putative	1770.31	4	800.09	8	
PY17X_1117800		AP-2 complex subunit alpha, putative	116.718	108	790.11	9	UIS
PY17X_1338000		conserved Plasmodium protein, unknown function	31.2679	466	687.1	10	UIS
PY17X_1125300	SMS1	sphingomyelin synthase 1, putative	250	43	633.65	11	
PY17X_1008900	ALP5b	actin-like protein, putative	70.0941	196	589.39	12	UIS
PY17X_1336100		conserved Plasmodium protein, unknown function	226.207	51	588.45	13	
PY17X_0712100	HSP70	heat shock protein 70	67.4928	204	527.01	14	UIS
PY17X_0837000	CSE1	importin alpha re-exporter, putative	47.2627	286	516.69	15	UIS
PY17X_1119000		trafficking protein particle complex subunit 6A, putative	241.562	47	510.4	16	
PY17X_1007700	PLP1	sporozoite micronemal protein essential for cell traversal	639.961	11	485.79	17	
PY17X_0303200		3'-5' exonuclease, putative	1752.68	5	477.65	18	UIS
PY17X_0302700		conserved Plasmodium protein, unknown function	232.97	48	433.45	19	UIS
PY17X_0511900		conserved Plasmodium protein, unknown function	243.898	46	427.68	20	UIS
PY17X_1354300		conserved Plasmodium protein, unknown function	0.36065	3866	412.42	21	UIS
PY17X_0826600		para-aminobenzoic acid synthetase, putative	450.447	16	382.8	22	
PY17X_0505100		tRNA methyltransferase, putative	7.88187	1408	382.46	23	UIS
PY17X_0943700	RTCB	tRNA-splicing ligase RtcB, putative	188.95	67	372.63	24	
PY17X_0101400	UIS7	conserved rodent malaria protein, unknown function	11.0352	1133	357.46	25	UIS
PY17X_0902800		ATP synthase subunit delta, mitochondrial, putative	720.216	9	288.89	28	
PY17X_0712000		conserved Plasmodium protein, unknown function	442.748	17	279.69	31	
PY17X_0823600		protein disulfide isomerase, putative	494.356	15	276.28	33	
PY17X_1239500		ribosomal large subunit pseudouridylate synthase, putative	543.896	14	143.94	67	
PY17X_0207300		lipid/sterol:H ⁺ symporter, putative	713.603	10	105.71	119	UOS
PY17X_0611400		general transcription factor 3C polypeptide 5, putative	367.196	23	104.02	122	
PY17X_0611200		conserved Plasmodium protein, unknown function	387.309	20	97.417	128	
PY17X_0710400		conserved Plasmodium protein, unknown function	385.895	21	96.915	130	
PY17X_0314800		conserved Plasmodium protein, unknown function	612.291	12	90.188	135	UOS
PY17X_0307200		conserved Plasmodium protein, unknown function	353.599	24	89.85	136	
PY17X_1024600		conserved Plasmodium protein, unknown function	1065.97	6	67.497	178	UOS
PY17X_0307300		transporter, putative	565.378	13	54.733	229	UOS
PY17X_0405000		EH domain-containing protein, putative	411.48	19	54.706	230	UOS
PY17X_0839000	PKAc	cAMP-dependent protein kinase catalytic subunit	377.858	22	24.574	496	UOS
PY17X_0822500		50S ribosomal protein L3, apicoplast, putative	736.78	7	6.0115	1377	UOS
PY17X_0835500		conserved Plasmodium protein, unknown function	350.148	25	0.1199	3886	UOS
PY17X_0208200		conserved Plasmodium protein, unknown function	429.611	18	0.1199	3887	UOS

<i>P. falciparum</i> Gene ID	Common Name	Product Description	oo-spz Avg	oo-spz Rank	sg-spz Avg	sg-spz Rank	UIS or UOS
PF3D7_0304600	CSP	CSP, circumsporozoite protein	7608.3	1	9068.1	1	
PF3D7_1235200	VP2	V-type K ⁺ -independent H ⁺ -translocating inorganic pyrophosphatase	27.171	322	6221.4	2	UIS
PF3D7_1008700		tubulin beta chain	867.56	5	1646.9	3	
PF3D7_1026600		conserved Plasmodium protein, unknown function	675.62	10	1563.1	4	
PF3D7_1475800		conserved Plasmodium protein, unknown function	211.16	37	1104.1	5	UIS
PF3D7_1335900	TRAP	TRAP, thrombospondin-related anonymous protein	174.85	47	892.92	6	
PF3D7_0516500	MFS1	major facilitator superfamily domain-containing protein, putative	9.1314	788	686.03	7	UIS
PF3D7_1033100	AdoMetDC/ODC	S-adenosylmethionine decarboxylase/ornithine decarboxylase	332.67	21	628.38	8	
PF3D7_0919400	PDi9	protein disulfide isomerase 9	408	14	600.56	9	
PF3D7_1347200	NT1	nucleoside transporter 1	15.252	529	574.57	10	UIS
PF3D7_1023900	CHD1	CHD1, chromodomain-helicase-DNA-binding protein 1 homolog, putative	283.94	25	554.2	11	
PF3D7_1133400	AMA1	AMA1, apical membrane antigen 1	219.01	34	552.23	12	
PF3D7_0616500	TLP	TLP, TRAP-like protein	43.702	204	529.32	13	UIS
PF3D7_1147800	MAEBL	MAEBL, membrane associated erythrocyte binding-like protein	1495.1	2	491.99	14	
PF3D7_0511400		conserved Plasmodium protein, unknown function	99.759	80	407.63	15	
PF3D7_1365600		DNA topoisomerase VI, b subunit, putative	134.4	61	386.11	16	
PF3D7_0529200		sugar transporter, putative	24.067	371	374.12	17	UIS
PF3D7_0407500		mitochondrial carrier protein, putative	70.868	118	368.17	18	UIS
PF3D7_0102800		conserved Plasmodium protein, unknown function	44.672	202	368.16	19	UIS
PF3D7_1020800	DLAT	dihydrolipoamide acyltransferase component E2	70.932	116	364.85	20	UIS
PF3D7_1239600	ThzK	hydroxyethylthiazole kinase	1110.7	3	326.52	21	
PF3D7_0820600		conserved Plasmodium membrane protein, unknown function	369.95	18	306.12	22	
PF3D7_1346100	SEC61	SEC61, protein transport protein, subunit alpha	89.982	87	303.25	23	
PF3D7_1407700		conserved Plasmodium protein, unknown function	198.72	40	302.06	24	
PF3D7_0925700	HDAC1	HDAC1, histone deacetylase 1	315.68	22	300.86	25	
PF3D7_1361100	SEC24A	Sec24A, protein transport protein	75.14	8	259.65	32	
PF3D7_1015900	ENO	enolase	381.57	16	210.54	44	
PF3D7_0404900	P41	p41, 6-cysteine protein	429.74	13	207.83	45	
PF3D7_0315700		conserved Plasmodium membrane protein, unknown function	799.67	7	142.56	74	UOS
PF3D7_0515300	PI3K	PI3K, phosphatidylinositol 3-kinase	396.54	15	120.46	84	
PF3D7_0404800		conserved Plasmodium protein, unknown function	971.37	4	110.29	94	UOS
PF3D7_0809100	SpzPFEMP1	SpzPFEMP1	707.24	9	100.08	104	UOS
PF3D7_0829800		unspecified product	850.48	6	89.228	123	UOS
PF3D7_1442600	TREP	TREP/UOS3, TRAP-like protein	618.51	11	41.651	252	UOS
PF3D7_0721200		conserved Plasmodium protein, unknown function	354.69	20	38.564	270	UOS
PF3D7_0525700		conserved Plasmodium protein, unknown function	289.09	24	37.255	279	UOS
PF3D7_1030100	PRP22	PRP22, pre-mRNA-splicing factor ATP-dependent RNA helicase, putative	376.92	17	4.5642	1119	UOS
PF3D7_1338300	EF-1g	elongation factor 1-gamma, putative	359.2	19	3.6924	1254	UOS
PF3D7_0709000	CRT	CRT, chloroquine resistance transporter	305.02	23	3.4415	1290	UOS
PF3D7_0918200		50S ribosomal protein L3, apicoplast, putative	446.65	12	1.5301	1846	UOS

Table 2: The top 30 most abundant proteins in oocyst sporozoites and salivary gland sporozoites of *P. falciparum* and *P. yoelii*.

<i>P. yoelii</i> Gene ID	Common Name	Protein Description	Abundance			UIS or UOS	<i>P. falciparum</i> Gene ID	Common Name	Protein Description	Abundance			UIS or UOS
			oo-spz SAF	Rank (of 1760)	sg-spz SAF					oo-spz SAF	Rank (of 1430)	sg-spz SAF	
PY17X_1330200	GAPDH	glyceraldehyde-3-phosphate dehydrogenase	1.85	5	4.53	1	PF3D7_1105000	H4	histone H4	4.20	3	5.55	1
PY17X_1134800													
PY17X_1134900	EF-1alpha	elongation factor 1-alpha, putative	2.76	2	3.73	2	PF3D7_1008700	N/A	tubulin beta chain	4.93	2	5.15	2
PY17X_1424900	N/A	conserved Plasmodium protein, unknown function	1.10	16	3.16	3	PF3D7_0503400	ADF1	actin-depolymerizing factor 1	1.53	21	4.16	3
PY17X_0944400	N/A	histone H4, putative	2.73	3	2.81	4	PF3D7_1335900	TRAP	thrombospondin-related anonymous protein	0.89	49	3.64	4
PY17X_1210100	N/A	tubulin beta chain, putative	3.73	1	2.62	5	PF3D7_1105100	H2B	histone H2B	3.83	4	3.56	5
PY17X_1461900	ACT1	actin 1, putative	1.38	9	2.41	6	PF3D7_0304600	CSP	circumsporozoite (CS) protein	9.48	1	3.45	6
							PF3D7_1357000						
PY17X_1316500	GEST	gamete egress and sporozoite traversal protein, putative	N/D	N/A	2.30	7	PF3D7_1357100	N/A	elongation factor 1-alpha; elongation factor 1-alpha	2.49	10	3.25	7
PY17X_1312400	ALDO2	fructose-bisphosphate aldolase 2	0.55	69	2.18	8	PF3D7_0816500	HSP20	small heat shock protein HSP20, putative	2.94	7	2.95	8
PY17X_0823500	TrxL1	thioredoxin-like protein 1, putative	1.30	13	2.17	9	PF3D7_1137800	SPELD	sporozoite surface protein essential for liver stage development	0.07	632	2.81	9
PY17X_1434600	CellTOS	cell traversal protein for ookinets and sporozoites	0.01	1461	2.14	10	PF3D7_1449000	GEST	gamete egress and sporozoite traversal protein, putative	N/D	N/A	2.69	10
PY17X_0714500	HSP20	small heat shock protein HSP20, putative	2.43	4	1.94	11	PF3D7_1246200	ACT1	actin 1	2.98	6	2.65	11
PY17X_1344800	LDH	L-lactate dehydrogenase, putative	0.80	40	1.79	12	PF3D7_1462800	GAPDH	glyceraldehyde-3-phosphate dehydrogenase	2.73	8	2.64	12
PY17X_0712800	14-3-3i	14-3-3 protein, putative	1.32	12	1.77	13	PF3D7_1407700	N/A	conserved Plasmodium protein, unknown function	0.78	63	2.49	13
PY17X_1036900	N/A	conserved Plasmodium protein, unknown function	0.46	101	1.75	14	PF3D7_0818200	14-3-3i	14-3-3 protein	2.28	12	2.41	14
PY17X_0944300	H2B	histone H2B, putative	1.17	15	1.69	15	PF3D7_0903700	N/A	alpha tubulin 1	2.30	11	2.29	15
PY17X_1143100	N/A	ubiquitin-60S ribosomal protein L40, putative	0.88	30	1.69	16	PF3D7_1438900	Trx-Px1	thioredoxin peroxidase 1	1.36	28	2.18	16
PY17X_1220800	H2A.Z	histone H2A.Z, putative	0.95	26	1.63	17	PF3D7_1342600	MyoA	myosin A	1.58	18	2.10	17
PY17X_1354800	TRAP	thrombospondin-related anonymous protein	0.32	154	1.62	18	PF3D7_0919300	TrxL1	thioredoxin-like protein 1, putative	1.45	24	2.10	18
PY17X_1213100	N/A	conserved Plasmodium protein, unknown function	0.32	155	1.61	19	PF3D7_1133400	AMA1	apical membrane antigen 1	0.57	108	2.04	19
PY17X_1205100	IMC1c	inner membrane complex protein 1c, putative	0.97	24	1.60	20	PF3D7_1365900	N/A	ubiquitin-60S ribosomal protein L40	3.54	5	2.02	20
PY17X_1306600	TPx1	thioredoxin peroxidase, putative	1.03	18	1.58	21	PF3D7_1324900	LDH	L-lactate dehydrogenase	1.16	31	1.89	21
PY17X_0713100	HMG2	high mobility group protein 82	0.57	66	1.54	22	PF3D7_0408600	SIAP1	sporozoite invasion-associated protein 1	1.76	15	1.77	22
PY17X_0930100	PGM1	phosphoglycerate mutase, putative	0.70	49	1.44	23	PF3D7_1017500	ELC	essential light chain ELC	0.84	57	1.75	23
PY17X_1422300	H2B.Z	histone H2B variant, putative	0.53	74	1.44	24	PF3D7_0814600	N/A	conserved Plasmodium protein, unknown function	0.92	43	1.60	24
PY17X_0810800	RPS11	40S ribosomal protein S11, putative	0.83	35	1.43	25	PF3D7_1444800	FBPA	fructose-bisphosphate aldolase	0.60	96	1.55	25
PY17X_1114100	PNP	purine nucleoside phosphorylase, putative	0.96	25	1.43	26	PF3D7_0918000	GAP50	glideosome-associated protein 50	1.08	34	1.48	26
PY17X_1425300	ALBA1	DNA/RNA-binding protein Alba 1, putative	0.94	27	1.40	27	PF3D7_1120100	PGM1	phosphoglycerate mutase, putative	0.77	66	1.47	27
PY17X_0912300	SPELD	sporozoite surface protein essential for liver stage development	N/D	N/A	1.36	28	PF3D7_1457200	TRX1	thioredoxin 1	0.77	65	1.45	28
PY17X_1453700	N/A	p25-alpha family protein, putative	0.49	90	1.32	29	PF3D7_1323700	GAPM1	glideosome associated protein with multiple membrane spans 1	1.53	20	1.44	29
PY17X_0826700	PGK	phosphoglycerate kinase, putative	0.99	21	1.31	30	PF3D7_1460600	ISP3	inner membrane complex sub-compartment protein 3	0.78	62	1.41	30
PY17X_1361400	MyoA	myosin A	1.52	7	1.31	31	PF3D7_1222300	GRP94	endoplasmic, putative	2.50	9	1.34	31
PY17X_1411000	N/A	RNA-binding protein, putative	1.57	6	1.25	34	PF3D7_0818900	HSP70	heat shock protein 70	1.70	16	1.33	34
PY17X_0712100	HSP70	heat shock protein 70, putative	1.50	8	1.24	35	PF3D7_1406800	GAPM3	glideosome associated protein with multiple membrane spans 3	1.55	19	1.19	43
PY17X_1007600	SIAP1	sporozoite invasion-associated protein 1	0.99	20	1.13	40	PF3D7_0819600	N/A	conserved Plasmodium protein, unknown function	1.88	14	1.19	44
PY17X_1207600	ALBA3	DNA/RNA-binding protein Alba 3, putative	0.98	22	1.07	43	PF3D7_0917900	HSP70-2	heat shock protein 70	1.60	17	1.18	46
PY17X_0822200	HSP70-2	heat shock protein 70, putative	1.22	14	1.05	44	PF3D7_1222700	GAP45	glideosome-associated protein 45	1.45	25	1.16	49
PY17X_0404800	IMC1a	inner membrane complex protein 1a, putative	1.38	10	1.04	46	PF3D7_0304000	IMC1a	inner membrane complex protein 1a, putative	1.40	27	1.16	50
PY17X_0404900	IMC1e	inner membrane complex protein 1e, putative	0.90	29	0.78	77	PF3D7_0423500	GAPM2	glideosome associated protein with multiple membrane spans 2	1.45	23	1.15	51
PY17X_0420500	N/A	alpha tubulin 1	1.33	11	0.77	80	PF3D7_0714000	H2B.Z	histone H2B variant	1.41	26	1.07	57
PY17X_0808800	HSP90	heat shock protein 90, putative	0.98	23	0.70	93	PF3D7_1350600	N/A	conserved Plasmodium protein, unknown function	2.01	13	0.94	68
PY17X_0833700	N/A	G2 protein, putative	0.91	28	0.67	100	PF3D7_0708400	HSP90	heat shock protein 90	1.22	29	0.67	106
PY17X_1369100	N/A	conserved Plasmodium protein, unknown function	1.05	17	0.49	140	PF3D7_1131500	N/A	conserved Plasmodium protein, unknown function	1.18	30	0.66	112
PY17X_0718900	N/A	small GTP-binding protein sar1, putative	1.02	19	0.38	187	PF3D7_0931300	N/A	conserved Plasmodium protein, unknown function	1.50	22	0.34	243

Table 3: Gene products under the translational repression in oocyst sporozoite to UIS protein regulatory program ("TR oo-spz to UIS protein program") in *P. yoelii* and *P. falciparum*.

<i>P. yoelii</i> Gene ID	Common Name	Product Description	Gene Deletion Phenotypes/Roles in Sporozoites	Refs	<i>P. falciparum</i> Gene ID	Common Name	Product Description	Gene Deletion Phenotypes/Roles in Sporozoites	Refs
PY17X_1007700	PLP1/SPECT2	perforin-like protein 1	Host Traversal Defect	[34], [59]	PF3D7_0408700	PLP1	perforin-like protein 1	Traversal Defect in Host	[34], [59]
PY17X_1436000	TRAMP	thrombospondin-related apical membrane protein			PF3D7_1218000	TRAMP	thrombospondin-related apical membrane protein		
PY17X_1112100	N/A	conserved Plasmodium protein			PF3D7_0511400	N/A	conserved Plasmodium protein		
PY17X_0829300	RRP8	ribosomal RNA-processing protein 8			PF3D7_0925200	RRP8	ribosomal RNA-processing protein 8		
PY17X_1434600	CeTOS	cell traversal protein for ookinetes and sporozoites	Host Traversal Defect	[47], [57], [58]	PF3D7_0616500	TLP	TRAP-like protein	Host Traversal Defect	[46]
PY17X_1228600	GAMER	gamete release protein	Reduced Infectivity of Host via Mosquito Bite	[61]	PF3D7_1350900	ApiAP2-O4	AP2 domain transcription factor	No sporozoites form	[40]
PY17X_0214800	NMD3	60S ribosomal export protein NMD3			PF3D7_1019000	eIF2A	eukaryotic translation initiation factor subunit	Regulated is critical to translational repression	[12]
PY17X_1359400	DBR1	RNA lariat debranching enzyme			PF3D7_1116700	DPA1	dipeptidyl aminopeptidase 1		
PY17X_0515300	CLAMP	claudin-like apicomplexan microneme protein			PF3D7_1116800	HSP101	heat shock protein 101		
PY17X_0712600	FCF1	rRNA-processing protein FCF1			PF3D7_1122800	CDPK6	calcium-dependent protein kinase 6		
PY17X_1360900	IRP	aconitate hydratase			PF3D7_0801900	LSD2	lysine-specific histone demethylase		
PY17X_0814400	SKI2W	helicase SKI2W			PF3D7_1235300	NOT4	CCR4-NOT transcription complex subunit 4		
PY17X_0835200	EF-Tu	elongation factor Tu			PF3D7_1308900	DCP2	mRNA-decapping enzyme 2		
PY17X_1442000	CCHL	cytochrome c heme lyase			PF3D7_1328800	SIR2A	transcriptional regulatory protein sir2a		
PY17X_0924600	N/A	kelch domain-containing protein			PF3D7_1456000	ApiAP2	AP2 domain transcription factor, unnamed		
PY17X_0509400	N/A	conserved Plasmodium protein			PF3D7_1414500	ABCK2	atypical protein kinase, ABC-1 family		
PY17X_1012000	N/A	conserved Plasmodium membrane protein			PF3D7_1430100	PTPA	serine/threonine protein phosphatase 2A activator		
PY17X_1025700	N/A	conserved Plasmodium membrane protein			PF3D7_0709300	CG2	Cg2 protein		
PY17X_1125100	N/A	conserved Plasmodium protein			PF3D7_0108300	ARP	conserved Plasmodium protein		
PY17X_1244500	N/A	conserved Plasmodium protein			PF3D7_0704600	UT	E3 ubiquitin-protein ligase		
PY17X_1431400	N/A	conserved Plasmodium protein			PF3D7_0806300	Ferlin	ferlin		
					PF3D7_1110400	N/A	RNA-binding protein		
					PF3D7_1147700	N/A	ATP synthase subunit delta, mitochondrial		
					PF3D7_1008100	N/A	zinc finger protein		
					PF3D7_1018600	N/A	tRNA wybutosine-synthesizing protein		
					PF3D7_0405700	N/A	lysine decarboxylase		
					PF3D7_0416000	N/A	RNA-binding protein		
					PF3D7_0519800	N/A	EELM2 domain-containing protein		
					PF3D7_0207100	N/A	conserved Plasmodium protein		
					PF3D7_0212500	N/A	conserved Plasmodium protein		
					PF3D7_0317300	N/A	conserved Plasmodium protein		
					PF3D7_0609700	N/A	conserved Plasmodium protein		
					PF3D7_0723300	N/A	conserved Plasmodium protein		
					PF3D7_0814700	N/A	conserved Plasmodium protein		
					PF3D7_1023800	N/A	conserved Plasmodium protein		
					PF3D7_1026600	N/A	conserved Plasmodium protein		
					PF3D7_1202600	N/A	conserved protein		
					PF3D7_1207400	N/A	conserved Plasmodium protein		
					PF3D7_1407600	N/A	conserved Plasmodium protein		

Table 4: Gene products under the Pan-Sporozoite translational repression program in *P. yoelii* and *P. falciparum*.

<i>P. yoelii</i> Gene ID	Gene Name	Description	OO+SG or SG-only	<i>P. falciparum</i> Gene ID	Gene Name	Description	OO+SG or SG-only
PY17X_0513900	MTRAP	merozoite TRAP-like protein	OO+SG	PF3D7_1028700	MTRAP	merozoite TRAP-like protein	OO+SG
PY17X_1417400	ApiAP2-G3	AP2 domain transcription factor AP2-G3	OO+SG	PF3D7_1317200	ApiAP2-G3	AP2 domain transcription factor AP2-G3	OO+SG
PY17X_0305000	MSP4/5	merozoite surface protein 4/5	SG-only	PF3D7_0207000	MSP4	merozoite surface protein 4	SG-only
PY17X_0507100	SEA1	schizont egress antigen-1	SG-only	PF3D7_1021800	SEA1	schizont egress antigen-1	SG-only
PY17X_0523100	Unnamed ApiAP2	AP2 domain transcription factor	SG-only	PF3D7_0420300	Unnamed ApiAP2	AP2 domain transcription factor	SG-only
PY17X_0942100	PAIP1	polyadenylate-binding protein-interacting protein 1	SG-only	PF3D7_1107300	PAIP1	polyadenylate-binding protein-interacting protein 1	OO+SG
PY17X_1209100	ApiAP2-I	AP2 domain transcription factor AP2-I	SG-only	PF3D7_1007700	ApiAP2-I	AP2 domain transcription factor AP2-I	OO+SG

<i>P. yoelii</i> Gene ID	Gene Name	Description	OO+SG or SG-only	<i>P. falciparum</i> Gene ID	Gene Name	Description	OO+SG or SG-only
PY17X_0215800	ApiAP2-L	AP2 domain transcription factor AP2-L	OO+SG	PF3D7_0216700	ATG11	autophagy-related protein 11	OO+SG
PY17X_0305400	SERA4	serine repeat antigen 4	OO+SG	PF3D7_0404900	P41	6-cysteine protein	OO+SG
PY17X_1243900	NEK2	NIMA related kinase 2	OO+SG	PF3D7_0507500	SUB1	subtilisin-like protease 1	OO+SG
PY17X_0101400	UIS7	conserved rodent malaria protein, unknown function	SG-only	PF3D7_0824400	NT2	nucleoside transporter 2	OO+SG
PY17X_0207200	UIS1	eukaryotic translation initiation factor 2-alpha kinase 2	SG-only	PF3D7_1116800	HSP101	heat shock protein 101	OO+SG
PY17X_0502200	UIS4	early transcribed membrane protein	SG-only	PF3D7_1122800	CDPK6	calcium-dependent protein kinase 6	OO+SG
PY17X_0507300	UIS12	RNA-binding protein, putative	SG-only	PF3D7_1350900	ApiAP2-O4	AP2 domain transcription factor AP2-O4	OO+SG
PY17X_1019000	UIS28	lipase	SG-only	PF3D7_0103200	NT4	nucleoside transporter 4	SG-only
PY17X_1123200	ApiAP2-SP3	AP2 domain transcription factor AP2-SP3	SG-only	PF3D7_0202100	LSAP2	liver stage associated protein 2	SG-only
PY17X_1354300	Unnamed Protein	conserved Plasmodium protein, unknown function	SG-only	PF3D7_0801300	WARP	von Willebrand factor A domain-related protein	SG-only
PY17X_1413100	eIF4G	eukaryotic translation initiation factor 4 gamma	SG-only	PF3D7_0811300	CAF1	CCR4-associated factor 1	SG-only
PY17X_1457400	UIS8	conserved Plasmodium protein, unknown function	SG-only	PF3D7_0829600	ETRAMP8	early transcribed membrane protein 8	SG-only
				PF3D7_1240100	ETRAMP12	early transcribed membrane protein 12	SG-only
				PF3D7_1417300	ATG4	cysteine protease ATG4	SG-only