Quantitative proteomic analysis of simian primary hepatocytes reveals candidate molecular markers for permissiveness to relapsing malaria *Plasmodium cynomolgi.*

Laurent Dembele[1,2§], Devendra Kumar Gupta[1,3], Bamaprasad Dutta[4], Adeline C.Y. Chua[1,5], Siu Kwan Sze[4], Pablo Bifani[1,5,6,7§].

[1] Novartis Institute for Tropical Diseases, 10 Biopolis Road, #05-01 Chromos, Singapore, 138670

[2] Université des Sciences, des Techniques et des Technologies de Bamako (USTTB) ; MRTC – DEAP – Faculty of Pharmacy, Point G, P.O. Box: 1805, Bamako, Mali

[3] Novartis Institute for Tropical Diseases, 5300 Chiron way, Emeryville, California 94608, USA

[4] School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551

[5] Singapore Immunology Network (SIgN), A*STAR, 8A Biomedical Grove, Immunos Building, Singapore 138648, Singapore

[6] Department of Microbiology and Immunology, Yong Loo Lin School of Medicine, National University of Singapore, 119077 Singapore

[7] Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK

[§]Correspondence: laurent@icermali.org and micpb@nus.edu.sg

Running title: Biomarkers Relapsing *Plasmodium* liver infection

Abbreviations

Received: 13/01/2019; Revised: 07/07/2019; Accepted: 22/08/2019

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/pmic.201900021.

Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand (AFRIMS).

Asialoglycoprotein receptor (ASGPR)

Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC)

Bursting Schizont (Bur Schz)

Dithiothreitol (DTT)

Exo-Erythrocytic Forms (EEF)

False discovery rate (FDR)

Fresh monolayer day 1 (FMD1)

Hypnozoite (Hyp)

Institutional Animal Care and Use Committee (IACUC)

Liquid chromatograph coupled to tandem mass spectrometry (LC-MS/MS)

Old monolayer day 7 (OMD7)

Plasmodium cynomolgi (P. cynomolgi)

Sandwich gel day 7 (SGD7)

Schizont (Schz)

SingHealth Experimental Medicine Centre (SEMC)

Tandem mass tags (TMT)

United States Army Medical Directorate, Armed Forces Research Institute of Medical Sciences (USAMD-AFRIMS)

World Health Organization (WHO).

Abstract

A major obstacle impeding malaria research has been the lack of an *in vitro* system capable of supporting infection through the entire liver stage cycle of the parasite, including that of the dormant form known as hypnozoites. Primary hepatocytes lose their specialized functions in long-term *in vitro* culture. The malaria parasite *Plasmodium* initiates infection in hepatocyte. This corresponds to the first step of clinically silent infection and development of malaria parasite *Plasmodium* in the liver. Thus the liver stage is an ideal target for development of novel antimalarial interventions and vaccines. However, drug discovery against *Plasmodium* liver stage is severely hampered by the poor understanding of host-parasites interactions during the liver stage infection and development. In this study, we have performed tandem mass tags (TMT) labelling based quantitative proteomic analysis in simian primary hepatocytes cultured in three different systems of susceptibility to *Plasmodium* infection. Our results display potential candidate molecular markers including asialoglycoprotein receptor (ASGPR), apolipoproteins, squalene synthase and scavenger receptor B1 (SR-BI) that facilitate productive infection and full development in relapsing *Plasmodium* species. The identification of these candidate proteins required for constructive infection and development of hepatic malaria liver stages paves the way to explore them as therapeutic targets.

Statement of study significance

Current malaria control and elimination tools such as dug treatments are not targeting hypnozoites of *P. vivax* and *P. ovale*. This hypnozoite reservoir is a hurdle to achieve disease elimination goals. The objective of this current study is to identify and comprehensively characterize host molecular factors facilitating relapsing *Plasmodium* infection and development to enable malaria liver stages including hypnozoite active compounds rapid drug discovery. Our current data revealed candidate proteins required for constructive Plasmodium parasite infection and development to establish in the liver. This should now facilitate development of cell lines permissive to hypnozoite producing malaria parasites that would be used for high throughput screen of hypnozoite cidal compounds. Discovery of such malaria interventions would significantly support global malaria effective control and elimination. Disease control and elimination are important for national poverty reduction strategies in malaria-endemic countries because of the social and

economic impacts of malaria. Our study would be of interest to the clinical, drug discovery and basic research malaria communities as it could help understand the interplay between the host and the parasite. The study could also be useful for the discovery and development of novel vaccines and therapeutic candidates against relapsing malaria

Introduction:

Malaria was estimated to cause over 219 million infections resulting in approximately 435,000 deaths annually[1]. Global elimination of *P. vivax* and *P. ovale* is hampered due to several bottlenecks. Most notably, *P. vivax* and *P. ovale* produce a quiescent hepatic form "hypnozoites" that are responsible for malaria relapses[2] At present, there is no safe available treatment to tackle hypnozoite induced relapse. The issue of relapse is compound by lack of an adequate continuous in vitro cultivation system for the hepatic stage of the parasite, rendering developing a basic understand of the parasite's biology an arduous process. Currently, primaquine and tafenoquine are the only licensed drugs available in the market that could curtail the hypnozoite reservoir[3]. However, both drugs have serious toxicity liabilities and unsuitable for widespread use. This is a major hurdle to successful disease elimination[4]. As such, there is a growing need to develop new anti-hypnozoitocidal compounds to tackle this reservoir of the parasite and prevent malaria relapses. In the last decade, a number of new compounds have entered the antimalarial pipeline, most of which originated from high throughput screens on the asexual blood stage of *P. falciparum*, a malaria parasite that does not form hypnozoites[5]. This is because the development of liver stage screens entails a number of biological and logistical obstacles. The primary hepatocytes rapidly loses their liver specialized functions when grown under *in vitro* conditions on collagen coated plates[6] and hepatic cell lines lack many cellular functions of their *in vivo* counter parts that enable *P. vivax* to complete the liver-stage cycle. Furthermore, as hypnozoites do not multiply and infect new cell, the number of infected hepatocytes become diluted as cell lines continue to double. Physiological and functional features of primary hepatocytes can be retained to some extend when using matrigel, a protein extract from Engelbreth-Holm-Swarm mouse sarcoma cells[7]. Due to the difficulties in culturing hepatocytes

in vitro, only few studies have been undertaken to investigate the liver stage biology of relapsing parasites[8] [9] [10] and host hepatocyte-parasite interactions.

A number of host-cell receptors associated with *Plasmodium* liver-stage sporozoites infection of hepatocytes have been identified. The tetraspanin CD81[11] has been shown to be essential in human *P. falciparum* and rodent *P. yoelii* sporozoites infection. Likewise, scavenger receptor B1 (SR-BI) enhances hepatocytes permissiveness to *Plasmodium* infection and is shown to be critical for *Plasmodium* liver stage development[12] [13]. Interaction between heparan sulfate proteoglycans and *Plasmodium* circumsporozoite protein (CSP) triggers the host invasion signal in *Plasmodium* sporozoites and promote host cells entry[14]. Overexpression of host receptor liver-fatty acid binding protein (L-FABP) has been shown to promote liver stage malaria parasite growth[15]. Recently, the ephrin type-A receptor 2 (EphA2) was shown to be associated with the onset of productive invasion and development by the malaria parasite[16]. However, none of these receptors have been reported to be associated with the liver stages of *Plasmodium* relapsing species infection and development.

To better understand host-parasite interactions, we have set out to analyse the host proteome using a tandem mass tags (TMT) based quantitative proteomic analysis of simian primary hepatocytes cultured under conditions permissive and inhibitory of *Plasmodium* liver infection and development. In the process, we identified the asialoglycoprotein receptor (ASGPR), apolipoproteins, and scavenger receptor B1 (SR-BI) as key component for constructive infection of the hepatocyte. Several of the up-regulated proteins observed in *Plasmodium* permissive cells culture conditions include previously identified specialized receptors in binding and entry of other pathogens. Additionally, a number of proteins involved in diverse biosynthetic pathways including cholesterol biosynthesis were up regulated in cells permissive to *Plasmodium* infection. Taken together, our data displayed that multiple proteins are required for a constructive *Plasmodium* parasite infection and development to establish. The definitive role of each individual protein for malaria parasite constructive infection and development remains to be investigated. However, identification of these important candidate proteins during malaria liver stage infection could help understand the interplay between the host hepatic cells and the

malaria parasite and could potentially be useful for the discovery and development of novel vaccines and therapeutic candidates against relapsing malaria.

Material and Experimental Procedures:

Ethical committees and animal welfare. Simian hepatocytes were isolated from segments of liver taken from healthy *Macaca fascicularis* monkeys provided by the institutions, Novartis Laboratory of Large Animal Services, New Jersey, U.S.A., (Novartis-LAS); and SingHealth, Singapore. The Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand (AFRIMS) provided *Macaca fascicularis* monkeys that were used for sporozoites production were from The Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand (AFRIMS). Prior to commencement of the study; Novartis Animal Welfare first evaluated and certified all institutions. The collection and use of materials were performed in accordance with the recommendations of the Association for the Assessment and Accreditation of Laboratory Animal Care (AAALAC) Standards and the Guide for the Care and Use of Laboratory Animals. Efforts were made to minimize discomfort of all animals. All experiments with animal were approved by the "Novartis Institutional Animal Care and Use Committees", prior to the study start. This project was reviewed for compliance with all established regulatory and accreditation standards at AFRIMS. AFRIMS' Institutional Animal care and Use Committee (IACUC) reviewed and approved that the project was in line with the humane animal care and use standards expressed in the Guide for the Care and Use of Laboratory Animals (National Research Council, 2011), the Ethical Principles and Guidelines for the Use of Animals for Scientific Purposes (National Research Council of Thailand, 2006) and Animal for Scientific Purposes Act 2015. Any noncompliance issues are reported to the AFRIMS Director as well at the Office of Laboratory Animal Welfare, National Institutes of Health, Washington D.C., USA. All animals used in this study are maintained in a facility that is fully accredited by the Association for the Assessment and Accreditation of Laboratory Animal Care International (AAALAC International). Environmental conditions were also maintained as recommended in the Guide for the Care and Use of Laboratory Animals. The United States Army Medical Directorate, Armed Forces Research Institute of Medical Sciences (USAMD-AFRIMS) IACUC and the Animal Use Review Division, U.S. Army Medical Research

and Materiel Command reviewed and approved the animals used in these studies under the Protocol Number PN13-06, in accordance with established principles under the Guide for the Care and Use of Laboratory Animals (ILAR, 2011). The USAMD-AFRIMS Animal Care and Use Program has been considered an excellent accredited program by Association for Assessment and Accreditation for Laboratory Animal Care (AAALAC) International. In SingHealth animal were housed in accordance with the recommendations of the Association for the Assessment and Accreditation of Laboratory Animal Care International (AAALAC) Standards and with the recommendations in the Guide for the Care and Use of Laboratory Animals. All animals discomfort was minimized and projects reviewed by SingHealth Experimental Medicine Centre (SEMC) prior to the Institutional Animal Care and Use Committee (IACUC). IACUC Ref No.: 2015/SHS/1024 SingHealth IACUC Secretariat @ SingHealth Office of Research.

Parasites:

P. cynomolgi (B strain) sporozoites were obtained from infected *Anopheles dirus*. Salivary glands were collected 14-30 days after a blood meal from an infected monkey. Infected salivary glands were removed and sporozoites were recovered as previously reported [4]. All *P. cynomolgi* infected *A. dirus* mosquitoes were obtained from USAMD-AFRIMS, Bangkok 10400, Thailand.

Primary hepatocyte isolation and culture conditions (**FMD1, SGD7 and OMD7**) **establishment:**

Simian primary hepatocytes were isolated from segments of liver collected from healthy *Macaca fascicularis* as previously described [6, 10]. Hepatocytes were cultured in complete media: William's Medium E (Life Technologies) supplemented with 10% Fetal Clone II serum (ref: SH30066.03), 5×10^{-5} M water soluble hydrocortisone (Sigma, H0396), 5 µg per ml insulin (Sigma), 2 mM L-glutamine and 200 U per ml penicillin, 200 µg per ml streptomycin (Life Technologies). Hepatocytes were seeded at a final cell density of 250,000 cells per cm^2 in 96 wells (Greiner bio One, 655956) and 48-well (Nest Scientific 748001) plates coated with collagen (SIGMA-MERCK # 11179179001). Hepatocytes were cultured in three conditions as depicted in the

schematic diagram in figure 1. All three samples were incubated in 5% CO_2 at 37°C. The three culture conditions are as follows: Condition 1: fresh monolayer day 1 (FMD1), corresponding to freshly isolated hepatocytes cultured in collagen coated plate for one day. Condition 2: sandwich gel day 7 (SGD7), consisting of FMD1 overlaid with matrigel (BD Biosciences) at (t = D 1), forming a culture sandwich collagen-matrigel and cultured for seven days. Condition 3: old monolayer day 7 (OMD7), consisting of FMD1 cultured for seven days in monolayer (collagen monolayer without Matrigel overlaid). Culture conditions were infected with *P. cynomolgi* sporozoites (figure 1 A) for parasite infectivity and developments assessment. Uninfected cultures (figure 1 B) were used to identify potential host proteins that aid in establishing *P. cynomolgi* infection. Culture condition FMD1 was infected with *P. cynomolgi* sporozoites at time D1 and ended at time D7; while SGD7 and OMD7 were infected with *P. cynomolgi* sporozoites at time D7 then ended at time D14 (figure 1A). Uninfected cultures were collected at time D1 for FMD1 and at time D7 for SGD7 and OMD7 for proteomic analysis (figure 1B).

Infection and parasite quantification:

Three million (3,000,000) sporozoites /ml of *P. cynomolgi* were re-suspended in the above complete medium used for hepatocyte culture to infect cultures in 96 well plates (60,000 sporozoites in 50 µl/well). The infected culture plates were centrifuged for 10 min at 2000 rpm, allowing fast parasite sedimentation, followed by incubation incubated at 37°C with 5% CO_2. At 3 hours post infection, the medium was removed, hepatocytes were washed three times and fresh media was added back to the cultures and the infected hepatocytes incubated at 37°C, 5% CO_2. Matrigel (BD Biosciences) was added as per the manufacturer's recommendations. The medium was refreshed every 48 hours. At defined time points shown in figure 1, the cultures were fixed with 100% cold methanol for 3 min (figure 1) for imaging. Pre-erythrocytic (Hypnozoite: Hyp, Schizont: Schz and bursting Schizont: Bur Schz) parasites were detected by immunofluorescence. Following fixation, the cultures plates were specifically stained with a rabbit anti-*P. cynomolgi* HSP70 (Genscript 299349-3, 1:5,000) and a mouse anti-*P. cynomolgi* UIS4 polyclonal antibodies (1:200) for 1 hour incubation at 37°C with the antibodies diluted 1:5,000 and 1:200 in 1X PBS respectively for HSP70 and UIS4. Binding was visualized with (red) Alexa 594-

conjugated goat anti-mouse immunoglobulin for US4 (ref 11032 Invitrogen) and (green) Alexa 488-conjugated goat anti-rabbit immunoglobulin for HSP70 (ref 11034 Invitrogen). Parasite and cell nuclei were stained with 1 µg per ml of DAPI (Sigma). Parasites were enumerated under a fluorescence microscope with a 40X magnification and images were obtained from NIKON ECLIPSE TS100 microscope. GraphPad Prism software and ANOVA Statistical test were used to compare the infectivity of the different cell culture conditions and the parasite development. The populations compared were derived from three independent biological replicates. A p -value of 0.05 or less was considered to be statistically significant.

Sample collection and protein extraction:

Samples from SGD7, FMD1 and OMD7 cultures were collected at the defined time points shown in figure 1. The culture medium was removed and the cells were washed thrice with PBS. The cells were lysed in 100 µl lysis buffer (2% SDS, 100 mM Tris-HCl [pH 7.4], 1µM dithiothreitol (DTT), Protease inhibitor cocktail) by sonication at 40% amplitude for 1min. Insoluble cell debris were removed by centrifugation at 20000×g for 10 min at 4^{o}C. The protein samples were purified by acetone precipitate. The purified protein pellets were reconstituted in 250 µl urea buffer (8 M urea and protease inhibitor cocktail in 100 mM ammonium bicarbonate buffer (ABB), pH 8). Protein amount of each sample was measured by BCA based protein quantitation method.

In-solution protein digestion and TMT (Tandem mass tags) labelling:

One-hundred µg equivalent amount of protein from each sample was taken and used for subsequent proteomic experiment. The urea concentration of the samples was reduced to <1 M by diluting with 100 mM ABB and reduced with 10 mM Tris 2-carboxyethyl phosphine hydrochloride (TCEP) at 37^{o}C for 2 hours. Reduced samples were then alkylated with 20 mM Iodoacetamide (IAA) at room temperature for 45 minutes in a dark. Protein samples were subsequently digested with trypsin at a trypsin-protein ratio of 1:40 overnight at 37^{o}C. The reaction was quenched by adding 10% TFA (final concentration of 0.4%). Tryptic peptides were desalted by using Sep-Pak C18 cartridges (Waters, Milford, MA) and dried in a SpeedVac (Eppendrop,

USA). TMT labelling was performed using TMT 10-plex kits according to the manufacturer's protocol (Thermo scientific, USA). The samples were labelled accordingly with following tags as listed in the table 1. The TMT-labelled peptides were pooled together and desalted.

TMT labelled peptides were fractionated by reverse phase chromatography at high pH (~pH8). Briefly, labelled peptides were reconstituted in buffer A (0.02% NH4OH in water) and fractionated using X-Bridge C18 column (4.6 × 200 mm, 5 µm particle size, 130 Å pore size) (Waters, Milford, MA) on a HPLC unit (Prominence™, Shimadzu, Kyoto, Japan) at flow rate of 1ml/min. The HPLC gradient comprised 100-95% buffer A for 3 min; 5-35% buffer B (80% ACN and 0.02% NH_4OH) for 40 min; then 35-70% buffer B for 12min, followed by 70% buffer B for 5min. The chromatograms were recorded at 280 nm. A total of 60 fractions were collected within 65min time period. The collected fractions were combined into 15 fractions by using concatenated pooling strategy and concentrated to dryness using vacuum centrifugation.

LC-MS/MS and data analysis:

The LC-MS/MS analysis was performed using a Q Exactive mass spectrometer coupled with online nano-HPLC system (Thermo Scientific; USA). Data acquisition was performed using Xcalibur 2.2 software (Thermo Scientific; USA). Protein identification and quantification were performed using Proteome Discoverer 1.4.1.14 (Thermo Fisher, MA) together with the Sequest search engine (parameters included cysteine alkylation with MMTS as static modification; methionine oxidation, and asparagine and glutamine deamidation as dynamic modifications; full trypsin digestion and maximum 2 missed and/or non-specific cleavages as digestion parameter; 10ppm precursor mass and 0.02 Da fragment mass tolerance). The NCBI proteins of *Macaca fascicularis* proteins (downloaded on December, 2016, including 73,374 sequences and 58,200,443 residues) was used as a search database with a cut-off false discovery rate of <0.05% and <0.01% at the peptide and protein level accordingly. The resulting data set was automatically bias-corrected to eliminate potential variation due to unequal mixing when combining

samples. All LC-MS/MS analysis was performed in triplicate. All statistical calculations were performed using the biological triplet. Cluster analysis of the identified proteins was performed using the online bioinformatics tool Gene Pattern (http://genepattern.broadinstitute.org) followed by hierarchical clustering and Pearson correlation.

Statistical analysis

The quantitative proteomic results were first examined for quality using a Pearson correlation that showed a satisfactory linear correlation among all three biological replicates (n = 3) (figure 3a-b). The proteomic data distribution patterns and cut-off levels of statistical significance were determined using Student's t-Test and volcano plots (Figure 4a-c). Distribution of the Log2 TMT ratio was used to determine the cut-off for the differentially regulated hits (Figure 4a). The OMD7 samples were used as denominator for the relative quantification conditions given that they had the least infections. Proteins with ratios < 0.76 (Log2Ratio < -0.4) were classified as down-regulated hits, whereas proteins with ratios > 1.32 (Log2Ratio > 0.4) were consider as up-regulated hits (Figure 4a).

Results:

P. cynomolgi* liver stage productive infection and full development depend on host-cell culture conditions *in vitro.

We assessed the hepatocytes permissiveness to *P. cynomolgi* sporozoites infection through proteomic analysis of cells cultured under different conditions of permissiveness (Figure 1). These conditions included FMD1, OMD7 and SGD7 described above in method section and shown in figure 1.

Development of hepatic stages were analysed using antibodies raised against *Plasmodium* proteins HSP70 (heat shock protein 70) and UIS4 (Up-regulated in Infectious Sporozoite-4, a parasitophorous vacuole membrane (PVM) marker)[17]. At seven days post infection, *P. cynomolgi* (Pc) infectivity, measured with anti-PcHSP70 antibody (green), was similar in both FMD1 and SGD7 cultures (Figure 2a and 2b). In contrast, infection of the OMD7 cultured hepatocytes was significantly lowered as compared to that in FMD1 ($p = 0.0383$) and SGD7 ($p = 0.0006$) cultures (Figure 2a and 2b). The diameters of the developing parasites (schizonts) were noticeably larger in FMD1 ($p = 0.0024$) and SGD7 ($p = 0.0002$) cultures as compared to that of OMD7 (Figure 2c). All the parasite stages, including hypnozoites and schizonts were observed in all three conditions except for the presence of mature bursting schizonts in OMD7 (Figure 2d). The rupture of the mature bursting schizonts and release of the pre-erythrocytic merozoites could be clearly discerned using the anti-PcUIS4 antibody (red) displaying UIS4 rupture in SGD7 and FMD1 cultures (Figure 2d). Our results suggest that significant productive sporozoites infection of the hepatocytes were only possible in the fresh (FMD1) and matrigel overlaid (SGD7) hepatocyte culture conditions (Figure 2a-b). Having demonstrated that the culture conditions of the host cells significantly pre-determine their permissiveness to *Plasmodium* sporozoites infection and parasite development; we set out to identify the required host molecular markers involved.

Quantitative proteomic analysis of simian primary hepatocytes revealed candidate molecular markers for permissiveness to *Plasmodium cynomolgi* infection.

Protein samples of the three cultures conditions from three monkey donors (Figure 1) were analysed simultaneously using tandem mass tag (TMT) quantitative proteomic analysis by liquid chromatograph coupled to tandem mass spectrometry (LC-MS/MS). The quantitative proteomic results were first examined for quality using a Pearson correlation that showed a satisfactory linear correlation among all the three biological replicates (figure 3a-b). The combined three biological replicates proteomic data distribution patterns and cut-off levels of statistical significance were determined using Student's t-Test and volcano plots (Figure 4a-c). A total of 5,763 proteins were identified with false discovery rate (FDR) <1%, including 4,588 proteins that incorporated more than one peptide. All the proteins and peptides identified are listed in the supplementary data set 1 and 2. Out of the 5,763 proteins identified, 1,627 were found to be non-annotated proteins. In FMD1 or OMD7 condition, a total of 5,163 proteins follow a similar trend with less than 20% deviation among the three biological replicates. Under SGD7 conditions 5,012 proteins were identified. Further analysis focused on 2,187 proteins that were significantly enriched (p-value < 0.05) (Figure 4a-c and supplementary data set 1). Distribution of the Log2 TMT ratio was used to determine the cut-off for the differentially regulated hits (Figure 4a). The OMD7 samples were used as denominator for the relative quantification conditions given that they had the least infections and would not highly express required proteins for successful infection. Proteins with ratios < 0.76 (Log2Ratio < -0.4) were classified as down-regulated hits, whereas proteins with ratios > 1.32 (Log2Ratio > 0.4) were considered as up-regulated hits (Figure 4a). These criteria resulted in a total of 477 down regulated and 574 up regulated hits in FMD1 condition while 89 down regulated and 50 up regulated in the SGD7 culture as compared to OMD7 (Figure 4d). Of the hits identified in FMD1 and SGD7, there was an overlap of 28 up regulated and 89 down-regulated proteins respectively (Figure 4d).

We focused on the 28 proteins that were commonly up regulated in FMD1 and SGD7 samples (Figure 5, Table 2 - 4). This cluster consists of proteins involved in general metabolism of hepatocyte, isoprenoid, cholesterol and fatty acid biosynthesis, signal transduction, protein recognition, and glycoprotein surface receptors (Figure 5, Table 2 -4). Notably, from the up-regulated proteins dataset, we identified the asialoglycoprotein receptor (ASGPR) in both FMD1 and SGD7, which

is known to interact with other human pathogens during infection. These pathogens included *Mycobacterium tuberculosis*, *Helicobacter pylori*, *Candida albicans*, *Leishmania spp, hepatitis C* and HIV-1[18]. In addition, apolipoproteins were identified in this screen that have previously been reported to interact with the malaria parasite *P. falciparum* PFE1590w/ETRAMP[19] and *P. berghei* parasitophorous vacuole membrane exported protein 1 (EXP-1) during liver-stage development[20]. Squalene synthase, an essential host hepatocyte enzyme required for sterol production[21] was also significantly up regulated in both FMD1 and SGD7. Previously, it has been shown that squalestatin, an squalene synthase inhibitor kills the blood stage *Plasmodium* parasites[22]. In future studies, it would be interesting to test its inhibitory effect on the dormant liver stage hypnozoite stage.

Given that a single protein may not be required for the infection but rather collection of proteins involved in a specific pathway(s) or in a mechanism to facilitate infection and development, we sought to identify key pathways that allow hepatocytes to be permissive to infection. A pathway enrichment analysis was performed on FMD1 and SGD7 (Table 2 and 3) (GO and KEGG pathways). Proteins in several pathways were uniquely up regulated in FMD1 samples but not in SGD7. These include hepatocyte developmental and metabolism related pathway, mitochondrial related pathways, peroxisome, and endoplasmic reticulum (Table 2). In contrast, there were only few pathways in which proteins were up regulated in SGD7 samples. These comprised pathways related to extracellular spaces of hepatocytes. The only pathway in which proteins were commonly up regulated in both FMD1 and SGD7 was the extracellular exosome secretion pathway.

Noteworthy is the observation that most of the documented host surface receptors required for *P. falciparum* sporozoites infection were not significantly altered in either FMD1 or SGD7 samples (Table 4) when compared to OMD7. Here, we only identified scavenger receptor class B member 1 (SR-B1) to be significantly up-regulated in simian FMD1 samples, while the liver-fatty acid binding protein L-FABP was significantly down-regulated in both FMD1 and SGD7 samples. Further studies are required to test the role of SR-BI in sporozoites infection for relapsing malaria species.

Discussion

Malaria parasites initial site of infection and development in the vertebrate host is the liver, however very little is known about the host molecular markers and pathways required for permissiveness and parasite development. Previously, four host receptors SR-B1[12], EphA2[16], CD81[23] and alpha-V containing integrin[24] were shown to be required for *P. falciparum* sporozoites infection to hepatocytes. However, none of these receptors are solely essential for *P. falciparum* sporozoites infection and development, suggesting that they interact in a mechanism that remains to be established. While, host receptors required for sporozoite infection and development of the liver stage cycle have been comprehensively characterized in *P. falciparum* and rodent malaria species. No receptor for the liver stage has been identified and reported for relapsing parasites; receptors required for relapsing parasites *P. vivax* and *P. ovale* remain elusive. The current study aims to identify candidate proteins potentially required in relapsing parasites infection and development. We report that the above known receptors for malaria entry in liver stages did not show significant changes in their expression level in all culture conditions. The exception was SR-BI, which was significantly up regulated in only FMD1 (*P. value* = 0.009) as compared to OMD7. Interestingly SR-BI expression level in FMD1 correlated with sporozoites infection rate and development in our dataset suggesting SR-BI role in *P. cynomolgi* sporozoites infection and full development. SR-BI is a glycosylated membrane protein present in several cell types including primary hepatocytes and was not up regulated in SGD7 hepatocytes. Like FMD1; significant infection rate and parasite full development was also found in SGD7 samples. In agreement with other observations[12] [13]; this result suggests that SR-BI might play a facilitator role in the bidirectional movement of cholesterol, which is required for malaria parasite infection and development[11]. The ASGPR receptor has been identified as a common host factor involved in entry for several other infections of bacterial, parasitic and viral origin[25]. Here, we are the first to identify ASGPR up regulation in conditions conducive to permissive infection (FMD1 and SGD7) by *P. cynomolgi* sporozoites infection. ASGPR is a heterodimeric C-type lectin receptors expressed on the surface of hepatocyte[25] . ASGPR and SR-BI receptors have been reported together to bind lipoproteins[26] [27]. Other studies have shown apolipoprotein as important ligand for the high-density lipoprotein and SR-BI[28]. Apolipoprotein also interacts with malaria parasite[19] [20]. Like ASGPR,

apolipoprotein expression level was markedly high also only in both FMD1 and SGD7 but not OMD7, suggesting a potential interaction with malaria relapsing species. ASGPR, apolipoprotein and SR-BI could be modulators of entry and development of pathogens such as malaria parasite and hepatitis C virus. We also identified obscurin like protein and squalene synthase that were significantly up regulated in both FMD1 and SGD7 samples. The squalene synthase is essential for cholesterol biosynthesis[22], which is also required for malaria parasite and hepatitis C virus infection and development [11] [22]. It has been reported that there is interaction between mosquito titin, an ortholog of human obscurin and dengue NS3 virus[29]. Collectively, our data suggest contribution of at least few proteins in a pathway or a mechanism for successful malaria parasite infection and development to establish.

A significant release of extracellular exosomes was found commonly in both FMD1 and SGD7 as compared to OMD7. Exosomes are formed from invagination of endosomal membranes forming vesicles that display cell surface lipids and proteins on their exterior face[30] . Released extracellular exosomes are shown to interact with target pathogens or cells through three major mechanisms: receptor-mediated binding, membrane fusion and nonspecific entry through the endocytic pathway[30] . Production of these vesicles during infection has been correlated with malaria severity in humans and animal[31]. Earlier studies have shown that the host exosomes to interact with other pathogens[30, 32]. The overall results of this study suggest that the proteins commonly up regulated in FMD1 and SGD7 hepatic cells that respectively displayed significant parasite count ($p = 0.0383$ and $p = 0.0006$) and development ($p = 0.0024$ and $p = 0.0002$) when compared to OMD7 are potentially required for *P. cynomolgi* sporozoites productive infection and full development (figure 2 and figure 4d). Identifying conditions for human hepatocyte cultures that enable constructive malaria infection brings the field one step closer to establishing an *in vitro* system that will permit for further studies in hypnozoite biology and drug development. However, further thorough investigations are required to determine the potential role of these proteins individually or collectively.

Availability of data and material

The mass spectrometry proteomics raw data files along with Proteome Discoverer search data file (including protein summary) have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository[33]. ProteomeXchange provides globally coordinated proteomics data submission and dissemination with the dataset identifier PXD013310. The Reviewer account details is Username: eviewer54597@ebi.ac.uk and Password: W6fipGrM. The proteins and peptides information from Proteome Discoverer search are available in the supplementary data set 1 and 2.

Acknowledgments

The Novartis Institute Tropical Diseases (NITD) financed the project. Many thanks to the European and Developing Countries Clinical Trials Partnership (EDCTP) for Dr Laurent Dembele fellowship # TMA2017CDF-1892—HypnoBio that enables him to achieve this manuscript. We are grateful to Andy Yip and Jessica Ong for their assistance with primary hepatocytes isolation and Dr Thierry Diagana head of NITD for his kind improvement of the manuscript. The authors declare that they have no conflict of interest.

Authors contributions:

L.D., D.G., A.C.Y.C. co-designed and performed the experiments. B.D., S.K.S. extracted proteins and performed mass spectrometry analysis. L.D., D.G., B.D., S.K.S. analysed the data. P.B., L.D., co-designed the project and wrote the manuscript. All authors critically revised the manuscript.

References

[1] WHO, WORLD HEALTH ORGANIZATION 2018.

[2] T. N. C. Wells, J. N. Burrows, J. K. Baird, Trends in Parasitology 2010, 26, 145.

[3] J. N. Burrows, S. Duparc, W. E. Gutteridge, R. Hooft van Huijsduijnen, W. Kaszubska, F. Macintyre, S. Mazzuri, J. J. Möhrle, T. N. C. Wells, Malaria Journal 2017, 16, 26.

[4] L. Dembélé, J.-F. Franetich, A. Lorthiois, A. Gego, A.-M. Zeeman, C. H. M. Kocken, R. Le Grand, N. Dereuddre-Bosquet, G.-J. van Gemert, R. Sauerwein, J.-C. Vaillant, L. Hannoun, M. J. Fuchter, T. T. Diagana, N. A. Malmquist, A. Scherf, G. Snounou, D. Mazier, Nature Medicine 2014, 20, 307.

[5] J. N. Burrows, R. H. van Huijsduijnen, J. J. Möhrle, C. Oeuvray, T. N. C. Wells, Malaria journal 2013, 12, 187.

[6] J. Keemink, M. Oorts, P. Annaert, in *Protocols in In Vitro Hepatocyte Research*, (Eds: M. Vinken, V. Rogiers), Springer New York, New York, NY 2015, 175.

[7] D. Deharde, C. Schneider, T. Hiller, N. Fischer, V. Kegel, M. Lübberstedt, N. Freyer, J. G. Hengstler, T. B. Andersson, D. Seehofer, J. Pratschke, K. Zeilinger, G. Damm, Archives of Toxicology 2016, 90, 2497.

[8] A. Voorberg-van der Wel, G. Roma, D. K. Gupta, S. Schuierer, F. Nigsch, W. Carbone, A.-M. Zeeman, B. H. Lee, S. O. Hofman, B. W. Faber, J. Knehr, E. Pasini, B. Kinzel, P. Bifani, G. M. C. Bonamy, T. Bouwmeester, C. H. M. Kocken, T. T. Diagana, eLife 2017, 6, e29605.

[9] N. L. Bertschi, A. Voorberg-van der Wel, A.-M. Zeeman, S. Schuierer, F. Nigsch, W. Carbone, J. Knehr, D. K. Gupta, S. O. Hofman, N. van der Werff, I. Nieuwenhuis, E. Klooster, B. W. Faber, E. L. Flannery, S. A. Mikolajczak, V. Chuenchob, B. Shrestha, M. Beibel, T. Bouwmeester, N. Kangwanrangsan, J. Sattabongkot, T. T. Diagana, C. H. M. Kocken, G. Roma, eLife 2018, 7, e41081.

[10] D. K. Gupta, L. Dembele, A. Voorberg-van der Wel, G. Roma, A. Yip, V. Chuenchob, N. Kangwanrangsan, T. Ishino, A. M. Vaughan, S. H. I. Kappe, E. L. Flannery, J. Sattabongkot, S. A. Mikolajczak, P. Bifani, C. H. M. Kocken, T. T. Diagana, eLife 2019, 8, e43362.

[11] O. Silvie, S. Charrin, M. Billard, J.-F. Franetich, K. L. Clark, G.-J. van Gemert, R. W. Sauerwein, F. Dautry, C. Boucheix, D. Mazier, E. Rubinstein, Journal of Cell Science 2006, 119, 1992.

[12] S. Yalaoui, T. Huby, J.-F. Franetich, A. Gego, A. Rametti, M. Moreau, X. Collet, A. Siau, G.-J. van Gemert, R. W. Sauerwein, A. J. F. Luty, J.-C. Vaillant, L. Hannoun, J. Chapman, D. Mazier, P. Froissard, Cell Host & Microbe 2008, 4, 283.

[13] C. D. Rodrigues, M. Hannus, M. Prudêncio, C. Martin, L. A. Gonçalves, S. Portugal, S. Epiphanio, A. Akinc, P. Hadwiger, K. Jahn-Hofmann, I. Röhl, G.-J. van

Gemert, J.-F. Franetich, A. J. F. Luty, R. Sauerwein, D. Mazier, V. Koteliansky, H.-P. Vornlocher, C. J. Echeverri, M. M. Mota, Cell Host & Microbe 2008, 4, 271.
[14] U. Frevert, P. Sinnis, C. Cerami, W. Shreffler, B. Takacs, V. Nussenzweig, The Journal of experimental medicine 1993, 177, 1287.

[15] S. A. Mikolajczak, V. Jacobs-Lorena, D. C. MacKellar, N. Camargo, S. H. I. Kappe, International Journal for Parasitology 2007, 37, 483.

[16] A. Kaushansky, A. N. Douglass, N. Arang, V. Vigdorovich, N. Dambrauskas, H. S. Kain, L. S. Austin, D. N. Sather, S. H. I. Kappe, Science (New York, N.Y.) 2015, 350, 1089.

[17] Sebastian A. Mikolajczak, Ashley M. Vaughan, N. Kangwanrangsan, W. Roobsoong, M. Fishbaugher, N. Yimamnuaychok, N. Rezakhani, V. Lakshmanan, N. Singh, A. Kaushansky, N. Camargo, M. Baldwin, Scott E. Lindner, John H. Adams, J. Sattabongkot, Stefan H. I. Kappe, Cell Host & Microbe 2015, 17, 526.

[18] T. B. H. Geijtenbeek, S. I. Gringhuis, Nature Reviews Immunology 2009, 9, 465.

[19] M. Vignali, A. McKinlay, D. J. LaCount, R. Chettier, R. Bell, S. Sahasrabudhe, R. E. Hughes, S. Fields, Malaria journal 2008, 7, 211.

[20] C. Sá E Cunha, B. Nyboer, K. Heiss, M. Sanches-Vaz, D. Fontinha, E. Wiedtke, D. Grimm, J. M. Przyborski, M. M. Mota, M. Prudêncio, A.-K. Mueller, Proceedings of the National Academy of Sciences of the United States of America 2017, 114, E1138.

[21] L. H. Cohen, M. Griffioen, C. W. T. van Roermund, R. J. A. Wanders, Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism 1992, 1126, 114.

[22] K. Saito, Y. Shirasago, T. Suzuki, H. Aizaki, K. Hanada, T. Wakita, M. Nishijima, M. Fukasawa, Journal of virology 2014, 89, 2220.

[23] O. Silvie, E. Rubinstein, J.-F. Franetich, M. Prenant, E. Belnoue, L. Rénia, L. Hannoun, W. Eling, S. Levy, C. Boucheix, D. Mazier, Nature Medicine 2002, 9, 93.

[24] K. Dundas, M. J. Shears, Y. Sun, C. S. Hopp, C. Crosnier, T. Metcalf, G. Girling, P. Sinnis, O. Billker, G. J. Wright, Proceedings of the National Academy of Sciences 2018, 115, 4477.

[25] D. Roggenbuck, M. G. Mytilinaiou, S. V. Lapin, D. Reinhold, K. Conrad, Auto-immunity highlights 2012, 3, 119.

[26] M. Sharma, A. Von Zychlinski-Kleffmann, C. M. Porteous, G. T. Jones, M. J. A. Williams, S. P. A. McCormick, Journal of lipid research 2015, 56, 1318.

[27] X.-P. Yang, M. J. Amar, B. Vaisman, A. V. Bocharov, T. G. Vishnyakova, L. A. Freeman, R. J. Kurlander, A. P. Patterson, L. C. Becker, A. T. Remaley, Journal of lipid research 2013, 54, 2450.

[28] M. C. de Beer, D. M. Durbin, L. Cai, A. Jonas, F. C. de Beer, D. R. van der Westhuyzen, Journal of Lipid Research 2001, 42, 309.

[29] D. Mairiang, H. Zhang, A. Sodja, T. Murali, P. Suriyaphol, P. Malasit, T. Limjindaporn, R. L. Finley, Jr., PloS one 2013, 8, e53535.

[30] A. J. Szempruch, L. Dennison, R. Kieft, J. M. Harrington, S. L. Hajduk, Nature Reviews Microbiology 2016, 14, 669.

[31] F. El-Assaad, J. Wheway, N. H. Hunt, G. E. R. Grau, V. Combes, PLOS Pathogens 2014, 10, e1003839.

[32] J. S. Schorey, Y. Cheng, P. P. Singh, V. L. Smith, EMBO reports 2015, 16, 24.

[33] J. A. Vizcaíno, E. W. Deutsch, R. Wang, A. Csordas, F. Reisinger, D. Ríos, J. A. Dianes, Z. Sun, T. Farrah, N. Bandeira, P.-A. Binz, I. Xenarios, M. Eisenacher, G. Mayer, L. Gatto, A. Campos, R. J. Chalkley, H.-J. Kraus, J. P. Albar, S. Martinez-Bartolomé, R. Apweiler, G. S. Omenn, L. Martens, A. R. Jones, H. Hermjakob, Nat Biotechnol 2014, 32, 223.

Figure 1.

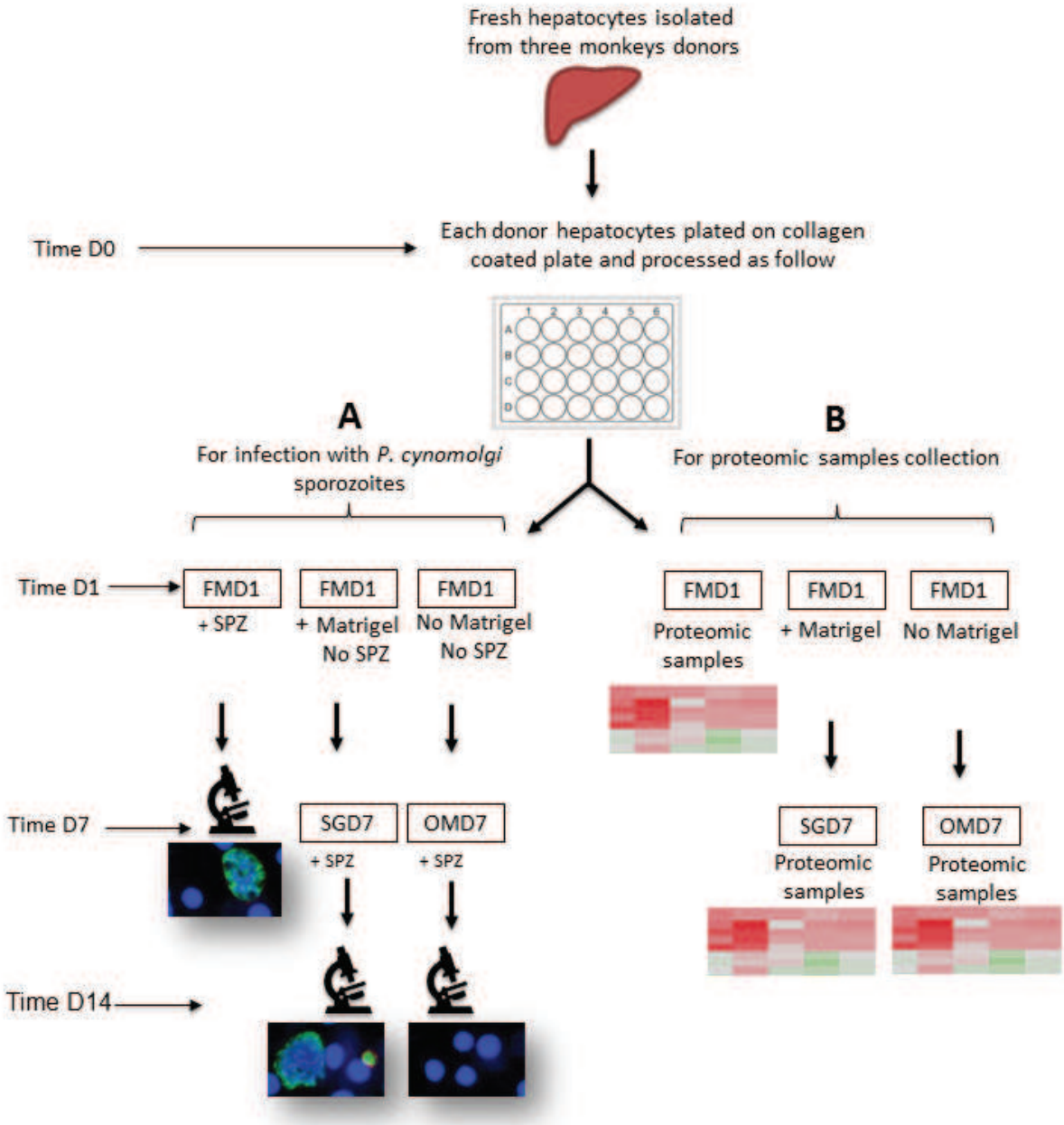

Figure 1. Study procedure schematic diagram. A) Cultures infected with *P. cynomolgi* sporozoite and used for parasite infectivity and developments assessment

in FMD1, OMD7 and SGD7 conditions. B). Cultures not infected, used for proteomic study to assess host cells candidate markers required for parasite infectivity and developments in FMD1 and SGD7 when compared OMD7. FMD1 consists of fresh monolayer day 1 cultures; OMD7 is old monolayer day 7 cultures and SGD7 is sandwich gel overlaid cultures day 7. In A) images, green correspond to *P. cynomolgi* anti-heat shock protein 70 (Pc-anti HSP70), red is parasite UIS4 and bleu is DAPI corresponding to parasite and hepatocytes nuclei.

Figure 2.

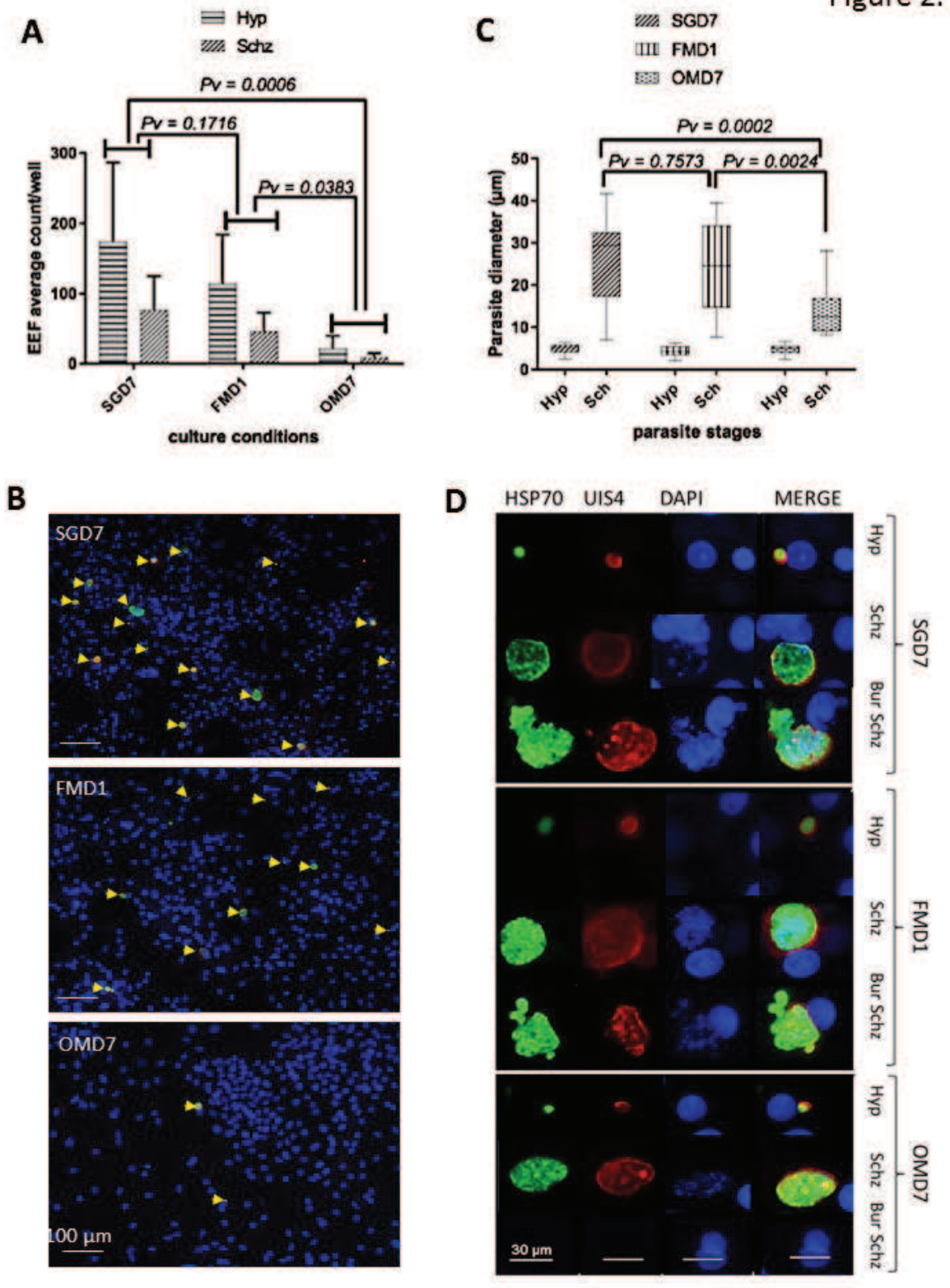

Figure 2. *P. cynomolgi* liver stage productive infection and full development depend on host cell culture conditions. A) Average count of exo-erythrocytic forms (EEF) of *P. cynomolgi* in SGD7, FMD1 and OMD7 at day seven post infection; B) 10X

immune-fluorescence images of EEF of *P. cynomolgi* in SGD7, FMD1 and OMD7. C) diameter of EEF of *P. cynomolgi* diameter in SGD7, FMD1 and OMD7 and D) 40X immune-fluorescence images of EEF of *P. cynomolgi* in SGD7, FMD1 and OMD7. *P. cynomolgi* anti-heat shock protein 70 (Pc-anti HSP70) labeled parasites in Green. *P. cynomolgi* anti UIS4 (PC-antiUIS4) labeled parasites in red. DAPI labeled nuclei of hepatic cell and parasites is in blue. The graph bar in figure 2A and 2C are derived from three independent biological assays.

Figure 3.

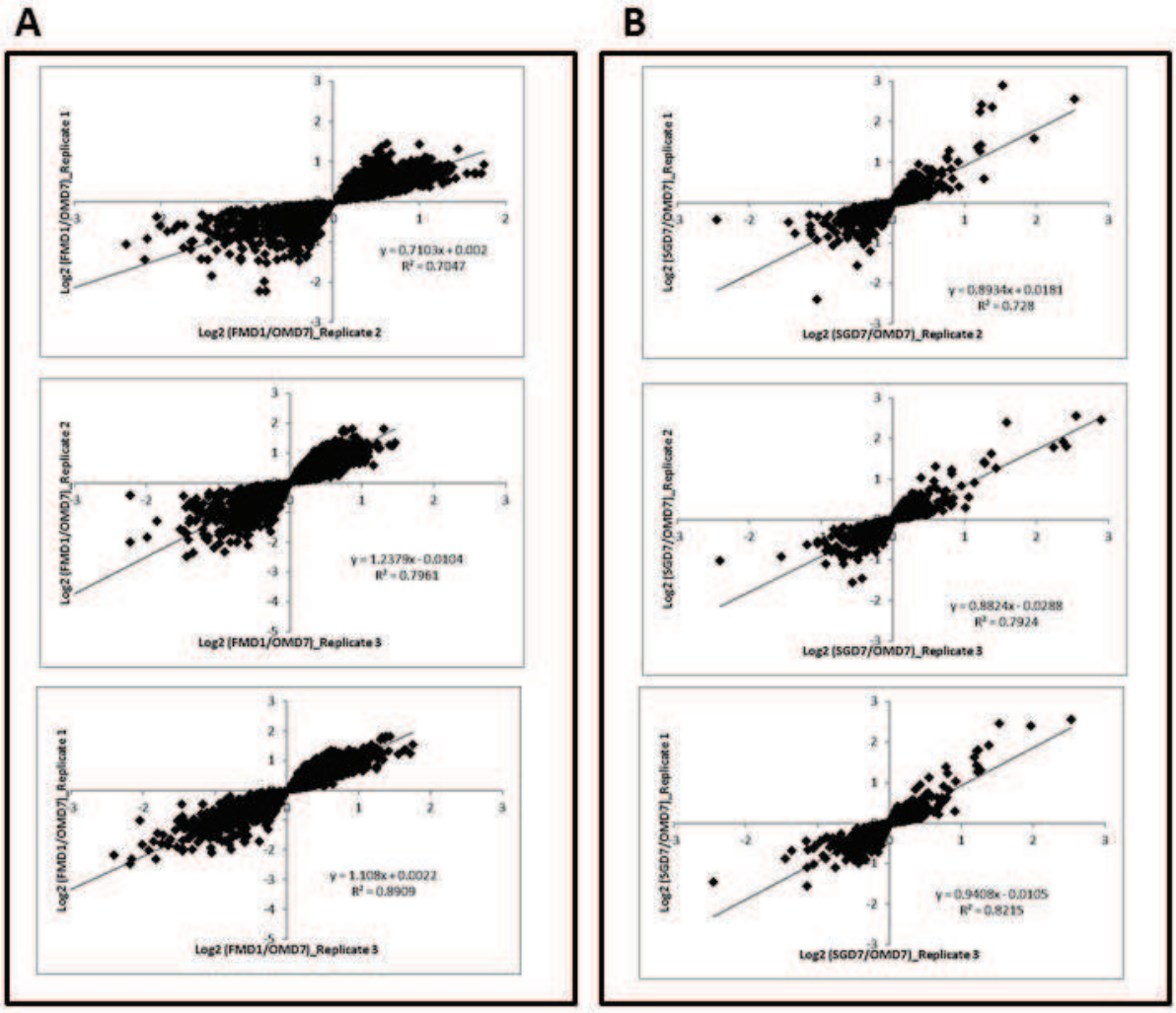

Figure 3. Scatter plot of differentially expressed proteins between biological replicates for SGD7 and FMD1 samples. The x-axis and y-axis represents the log2

fold change data. Scatter plot showing the tight positive correlation among 3 biological replicates for A) FMD1 samples and B) SGD7 samples.

Figure 4.

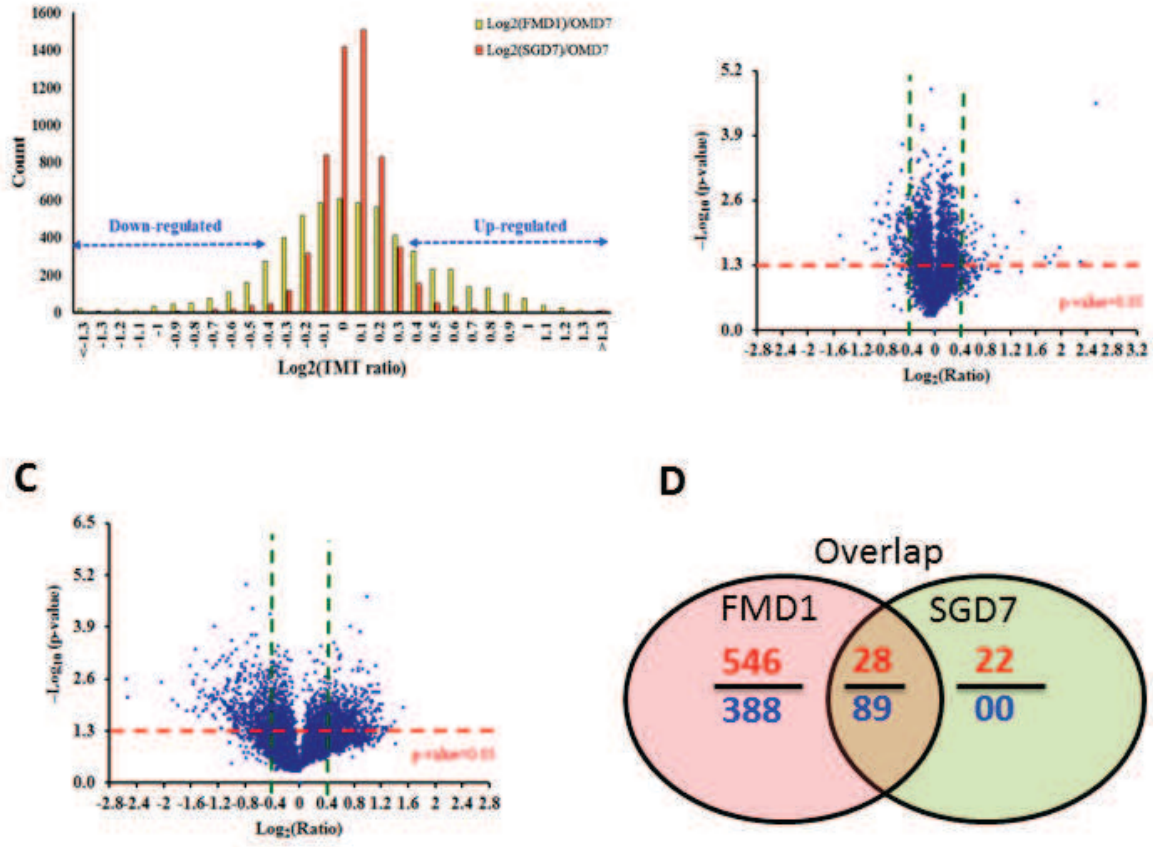

Figure 4. A) Frequency distribution plot of Log2TMT ratio. B-C) Volcano plot highlighting significantly differentially regulated proteins for SGD7 and FMD1 samples. Every data point in the volcano plot serves as fold change difference

between FMD1 and OMD7 B) and between SGD7 and OMD7 C) plotted against their respective p-values. The x-axis represents the log2 fold change whereas the y-axis is the -log10 p-values. D) Venn diagram showing the proteins that were differentially expressed in FMD1 and SGD7 samples. Up-regulated hits (red) and down-regulated hits (blue). There are 117 proteins that are commonly regulated in both FMD1 and SGD7 samples. This is from combined three biological replicates proteomic data.

Figure 5.

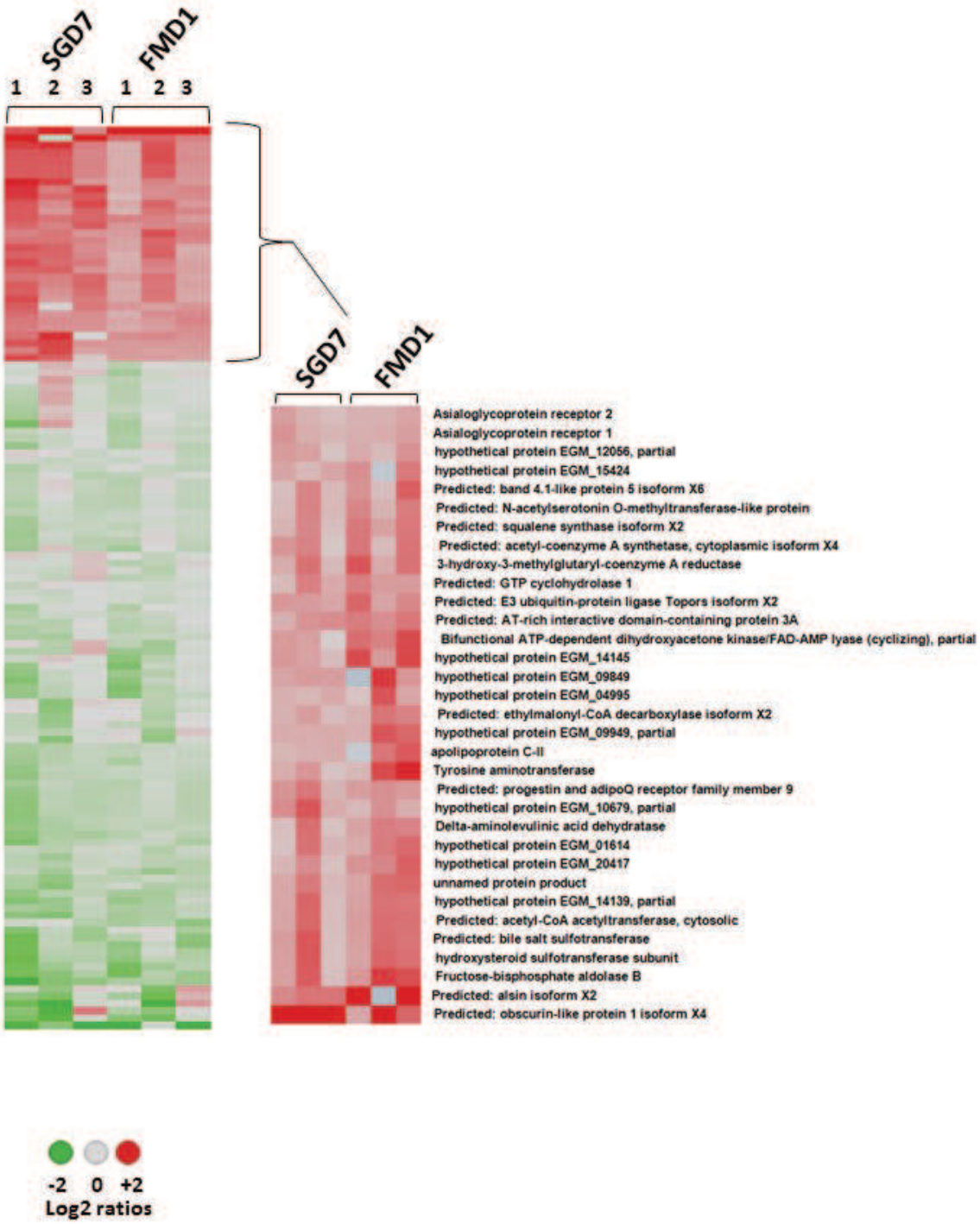

Figure 5. Hierarchical clustering of up- and down-regulated genes in FMD1 and SGD7 as compared to OMD7 for three biological replicates is shown. Red are up-regulated genes, white are non-affected genes and green are down-regulated genes.

Table 1. Fractionation of labeled peptides:

TMT labelling		
Monkey	**Sampl**	**TMT**
1	FMD1	126
	SGD7	127N
	OMD7	127C
2	FMD1	128N
	SGD7	128C
	OMD7	129N
3	FMD1	129C
	SGD7	130N
	OMD7	130C

Table 2. Pathway enrichment for up-regulated genes in Fresh monolayer day 1 (FMD1) compared to old monolayer day 7 (OMD7). Nine hundred and five up-regulated genes of FMD1 samples were analysed and significantly enriched pathways using GO and KEGG database. Selected pathways related to liver stage development and metabolism is shown (Go and KEGG pathways). Number of genes and the receptive *p*-value for each pathway in our dataset is also included (*p*-value < 0.05).

Pathway name in (FMD7 and OMD7)	**Number of genes**	***P*- value**
Metabolic pathways	99	2.6763E-39
cytosol	92	8.37E-13
extracellular exosome*	83	8.47146E-13
oxidation-reduction process	40	8.95E-16
mitochondrial matrix	38	8.29E-24
developmental	34	3.11E-02
endoplasmic reticulum membrane	28	3.46346E-05
aging	24	0.026212652
mitochondrial inner membrane	22	5.12E-07
intracellular membrane-bounded organelle	21	6.23758E-05
Peroxisome	14	1.86E-08
response to drug	13	1.42E-03
electron carrier activity	13	4.41E-09
receptor binding	13	4.17E-03
Carbon metabolism	13	4.93036E-06
Valine, leucine and isoleucine degradation	13	1.7696E-10
Chemical carcinogenesis	12	9.46E-07
liver development	6	4.06E-03

Table 3. Pathway enrichment for up-regulated genes in sandwich gel day 7 (SGD7) compared to old monolayer day 7 (OMD7). A hundred and nineteen up-regulated genes of SGD7 samples were examined and significantly enriched pathways were highlighted using GO and KEGG database. Pathways that were significantly enriched are shown here. Number of genes and the receptive *p*-value for each pathway in our dataset is also included (*p*- value < 0.05).

Pathway name in (SGD7/OMD7)	**Number of genes**	***P*- value**
extracellular space	6	0.007732
extracellular exosome*	8	0.037245
extracellular matrix structural constituent	3	0.001995
Complement Clr-like EGF domain	2	0.028201
Epidermal growth factor-like domain	3	0.0312
Pathways in cancer	5	0.009017
Small cell lung cancer	3	0.01434
ECM-receptor interaction	3	0.015331
Toxoplasmosis	3	0.025922
EGF	3	0.024899
Disulfide bond	7	0.001946

Table 4. List of hepatocyte surface receptors known for pre-erythrocytic stage of malaria infection.

Accession	Description	Log2 (FMD1/OMD7)	*P*-value	Log2 (SGD7/OMD7)	*P*-value
XP_005576959.1	CD81 antigen isoform X2	0.323	0.012	0.126	0.144
XP_005587354.1	Basigin isoform X1	0.063	0.245	0.048	0.345
XP_005572655.1	Scavenger receptor class B member 1 isoform X5	0.703	0.009 *	0.123	0.186
XP_005570076.1	Scavenger receptor cysteine-rich type 1 protein M130 isoform X4	-0.226	0.123	-0.207	0.023
XP_005563674.1	Fatty acid-binding protein, epidermal	-0.621	0.017 *	-0.453	0.023 **
XP_015303680.1	Ephrin type-A receptor 1 isoform X3	-0.333	0.118	-0.267	0.036
XP_005549301.1	Transferrin receptor protein 2 isoform X2 (CD71)	0.575	0.11	0.148	0.003

* Significantly differentially expressed in FMD1 samples (Cut off: greater than or less than 0.4, p-value < 0.05).

** Significantly differentially expressed in SGD7 samples (Cut off: greater than or less than 0.4, p-value < 0.05).