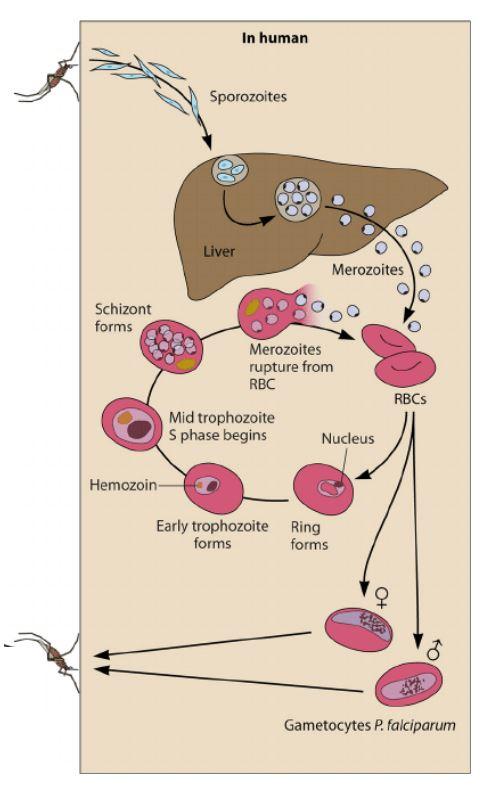
RNA-seq analysis of *Plasmodium falciparum* shows upregulated genes involved in merozoite egress & erythrocyte invasion

Abstract

Malaria is a febrile disease caused by parasites of the genus *Plasmodium*. *P. falciparum* causes the most severe cases of malaria, and is responsible for most of the deaths. Investigating transcriptional differences between severe and uncomplicated cases can shed light on malaria pathogenicity and virulence. In this study we used previously acquired publicly available RNA sequencing data from severe and uncomplicated cases of *P. falciparum* infection. Differential analysis of the *P. falciparum* transcriptome showed that a group of proteins involved in merozoite egress and erythrocyte invasion were significantly upregulated in severe malaria. The upregulated proteins are merozoite surface proteins 1 and 2, Rh5, and serine repeat antigen 5 (SERA5). Up-regulation of genes in this life cycle process provides insight into the cause of malaria severity. In addition, up-regulation of tyrosine kinase-like protein (TKL) in conjunction with SERA5 regulatory proteins suggests that the interaction between TKL and SERA5 may have a regulatory function.

Introduction

Malaria is a parasitic infection which affects approximately 500 million people annually14. The disease is caused by the Plasmodium species, which is transmitted through the saliva of a female Anopheles mosquito. There are five species of Plasmodium which infect humans, *P. vivax* and *P. falciparum* being the most common ones20. The parasite exists in several life cycle stages, in both mosquito and human hosts (Fig. 1). In this study we will be focusing on the life cycle stages that exist within the human host, specifically the sporozoite and merozoite stages. The parasite is in the sporozoite stage when introduced to the human body through Anopheles saliva. The sporozoites travel to the liver, where they infect hepatocytes (liver cells) and evade being filtered out of the bloodstream. After maturating into merozoites, the parasite egress the liver and enters into the blood stage by attacking erythrocytes. Once in the blood stage, the parasite either remains in the asexual stage (merozoite) to continue infectecting healthly erythrocytes or progresses to the sexual stage (gametocyte) waiting to infect healthy mosquitos. In order to progress through these stages, the Plasmodium must change expression of many different genes over the course of its life. The predominant life stage present in the samples from this study are the erythrocytic ring and trophozoite stages, which are a part of the blood cycle stages15.



When malaria is diagnosed, it is categorized as being either an uncomplicated or a severe case. Severe malaria is categorized by clinical manifestations meeting certain criteria, as published by the World Health Organization. Some examples of these symptoms include cerebral malaria, impaired consciousness, and severe anemia25. Most severe cases of malaria are caused by one species of Plasmodium in particular: *P. falciparum*. It is unknown why some cases of *P. falciparum* are severe while others are uncomplicated. Finding the cause of these clinical manifestations could help in developing treatments for severe malaria.

A study by Tonkin-Hill et al. has described the use of comparative transcriptomics to study the role of Plasmodium falciparum surface protein 1 (PfSP1) in malaria severity15. Our study uses the same dataset as Tonkin-Hill, but with several differences. The Tonkin-Hill et al. study only studied *var* genes, whereas we will be looking at the entire transcriptome, with a focus on genes relating to merozoite egress and erythrocyte invasion.

# Methods

All data for this study was collected by Tonkin-Hill et al from malaria patients in Papua, Indonesia15. A set of 44 transcriptomes was obtained from blood serum samples of 23 patients with severe malaria and 21 patients with uncomplicated malaria. The sequences were retrieved from the European Nucleotide Archive. The Galaxy cluster at useGalaxy.org was used to manipulate all transcriptomes and a pipeline for the assembly of the P. falciparum genome was adapted from Tonkin Hill et al15. Raw pair-end reads were trimmed with Trim Galore! to remove Illumina primers that were still attached to the ends of sequences. HISAT216 was then used to align the trimmed reads to three different reference genomes: *P.vivax* v10, the HG38 human genome, and *P. falciparum* 3D7. The human and *P. vivax* alignments were then converted to BED format.17 BAM filter was used to remove any reads that were mapped to either *P. vivax* or the human genome. After filtering, featureCounts18 was used to prepare a table containing counts of features such as genes, exons, non-coding regions, etc. These tables serve as the inputs for DEseq. The counted reads are then input to DEseq19, which outputs a list of genes which can be analyzed for significant differential expression. These genes were then categorized using the KEGG database. When no KEGG term was available for a gene, a Refseq name was used.

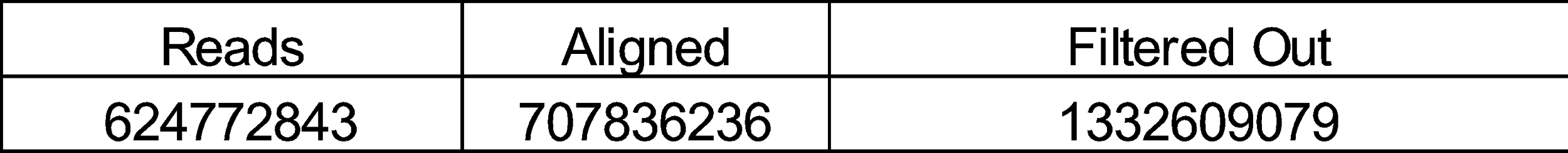
# 

# 

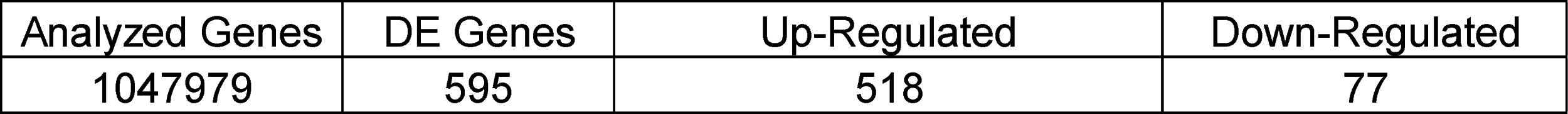
# Results and Discussion

### DEseq results show 595 upregulated genes in severe malaria

HISAT (Fig. 3) was used to map reads from NGS to reference genomes. Because the samples are from blood serum, it is probable that they are contaminated with other genetic material, such as the human transcriptome or other strains of plasmodium. *P. vivax* is common in Indonesia, and so this strain was included in the HISAT alignments. The reads that mapped to either the *P. vivax* genome or the human genome were then removed in a filtering process (Fig. 3). The filtering process showed that other transcripts were present, as shown by the large number of reads removed (Suppl Table 2).



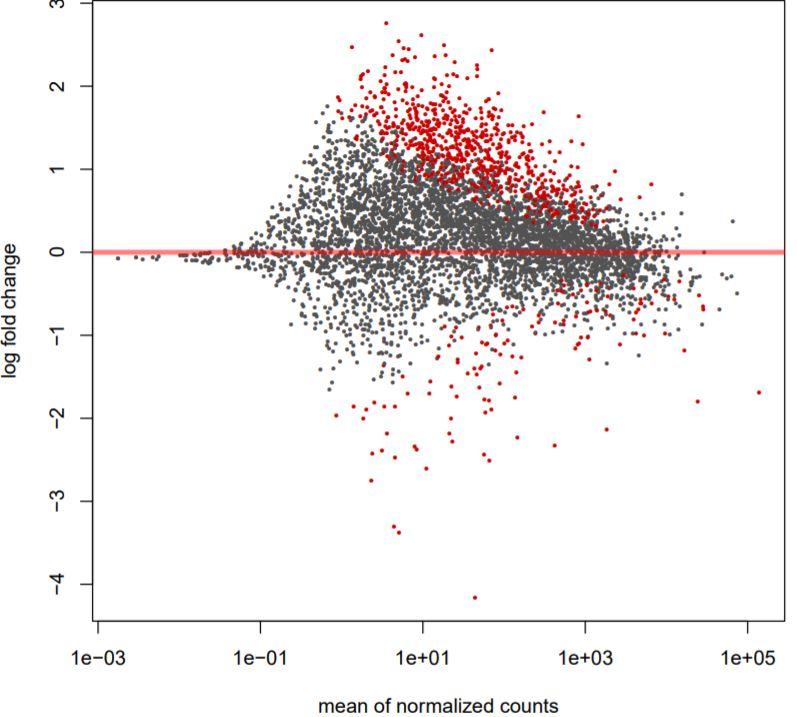


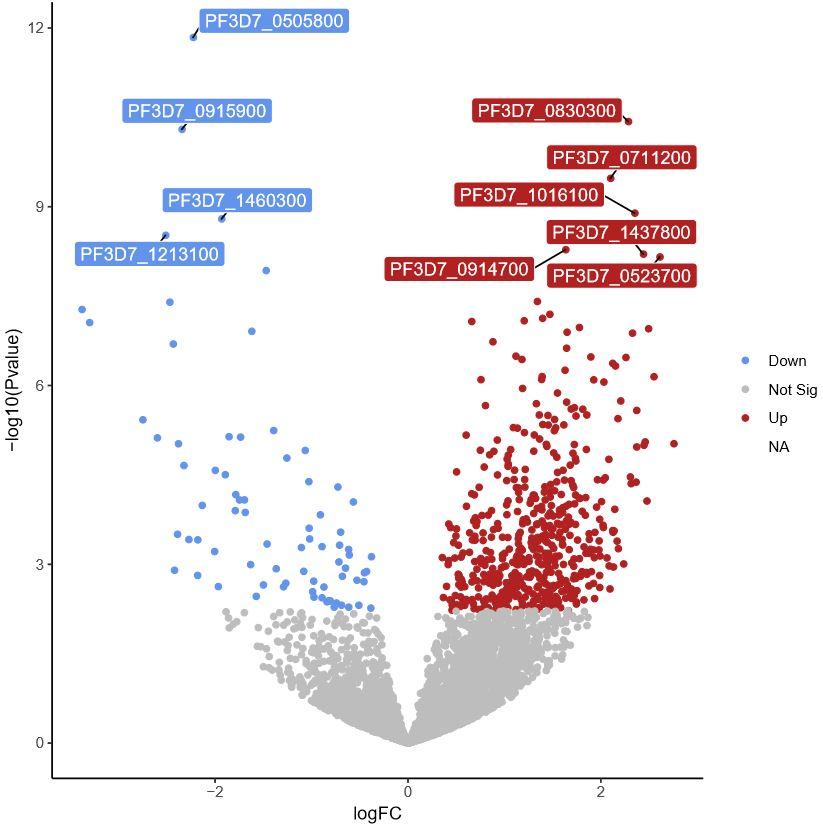
Filtering is important to prevent the contaminants from skewing DEseq analysis later on. Featurecounts (Fig. 2) counts the features present in the assembled genome, such as genes and exons. DEseq2 (Fig. 2) was used to perform the differential analysis itself. After completion, the dispersion plot was analyzed to ensure that the assumptions of DEseq were met (Fig. 5). DEseq shows that five hundred sixty-one genes were found to have a significant change in regulation. This is shown in the default MA plot (Fig. 6), as well as in a volcano plot (Fig. 7) created to better visualize the data. The MA plot shows the magnitude of up-regulation by distance from the center line on a log base 2 scale. Significantly differentially expressed genes 

are shown in red. The volcano plot shows the significance of the difference in expression versus the magnitude of the change in expression. The ten genes with the most significant changes in expression were selected to have their IDs shown, to guide analysis. There are more genes which are upregulated than down-regulated, by a factor of 6 (Suppl. Table 3).

# 

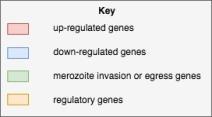
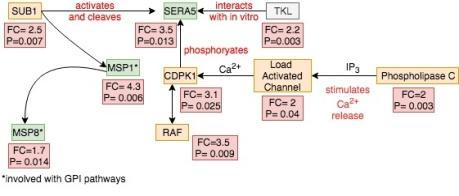








After all of the genes with a significant change were categorized by biological function with either a KEGG or Refseq name, a literature search was conducted to determine the function of these genes in P. falciparum. This comprehensive review allowed patterns to be found, and thus reveal the biological significance of the differences in expression.



### Proteins involved in erythrocyte egress found to be upregulated

Subtilisin-like protease 1 (SUB1) (PF3D7\_0507500) is upregulated by a factor of 2.5 (P=0.007) between severe and uncomplicated malaria blood samples(Suppl. Table 1). SUB1 is a regulatory protease, which is believed to remove regulatory domains from other proteinswhen secreted into the parasitophorous vacuole1. Specifically, it activates serine repeat antigen 5 (SERA5), SERA6, merozoite surface protein 1 (MSP1), MSP6, and MSP75. These proteins are involved in egress from the infected erythrocytes and invasion of uninfected erythrocytes1. Of these targets for subtilisin 1, SERA5 and MSP1 have been found to be upregulated in cases of severe malaria.

### Ca2+ pathways and other erythrocyte egress regulatory proteins found to be upregulated

SERA5 is involved with erythrocyte egress during the *Plasmodium* asexual blood stage. PF3D7\_0207600 is the gene which encodes for this antigen, is upregulated by a factor of 3.5. (p=0.013) in severe malariaS1. In addition, a regulator of Sera5, CDPK1, is also upregulated, by a factor of 3.1 (P=0.025) (Suppl. Table 1). The mechanism of SERA5 action is unclear, however studies have shown its activity in merozoite egress 1,2. A study in which the gene (PF3D7\_0207600) encoding for SERA5 was removed by DiCre-mediated excision showed a 50% reduction in intracellular population and invasion rate3. This knockout study further provides evidence of the relationship between erythrocyte egress and higher virulence. In addition to SERA5, upstream Ca2+ regulatory genes have been found to be upregulated in *P. falciparum* from severe malaria cases (Suppl. Table 1). One such gene is PF3D7\_1013500, which encodes phosphoinositide-specific phospholipase C (PLC), was found to be upregulated by a factor of 2 (P=0.0027)(Suppl. Table 1). PLC is responsible for hydrolyzing diacylglycerol to form inisitol triphosphate (IP3). The upregulation of PLC is significant, as IP3 stimulates the release of Ca2+ ions from the endoplasmic reticulum, through load activated Ca2+ channels2. The expression of the gene (PF3D7\_1362300) encoding for Ca2+ channels is also upregulated in cases of severe malaria, by a factor of 2 (P=0.04)(Suppl. Table 1). The endoplasmic reticulum stores Ca2+ in plasmodium, to set up a potential difference across its membrane. The release of Ca2+ can then activate or deactivate regulatory proteins, and thus cell function. Both the reputed serine protease (SERA5), and its regulatory pathway are significantly upregulated severe malaria. Another gene related to SERA5 that is upregulated is tyrosine kinase like protein (TKL). The protein is upregulated by a factor of 2.2 (p-0.0027)F8. The function of TKL is unknown, but it has been shown to interact with SERA5 by a sterile alpha motif (SAM) and has the ability to bind ATP.13 The fact that SERA5, as well as its regulatory proteins are upregulated suggests that the increased virulence seen in some cases of *P. falciparum* may be due to the increased ability of the parasite to exit the erythrocyte and invade new host cells.

### Merozoite Surface Proteins are upregulated and is involved in erythrocyte egress

Merozoite surface protein 1 (MSP1) is associated with erythrocyte egress by the parasite7. After post translational processing by subtilisin to yield its active form, the protein is attached to the cell membrane using a GPI anchor 1,8. MSP1 (PF3D7\_1216500) is upregulated by a factor of 4.3 (p=0.006)(Suppl. Table 1). Merozoite surface protein 8 (MSP8) is upregulated by a factor of 1.7 (p=0.014) (Suppl. Table 1). MSP8 is not processed by subtilisin. MSP8 is attached to the cell surface via GPI, as is MSP1. MSP8 has also been implicated in erythrocyte egress9.

### Upregulation of Rh5 suggests increased erythrocyte invasion

Rh5 (PF3D7\_0323400) is a protein expressed during erythrocyte invasion. Rh5 first binds to the p113 protein on the surface of the *plasmodium*, and then to the basigin receptors on the surface of the targeted erythrocytes. This interaction works to draw the erythrocyte closer to the *plasmodium*, in preparation for the formation of the moving junction11. It is upregulated in severe malaria by a factor of 3.6 (p=0.0026)(Suppl. Table 1). While p113 is not upregulated, it has a higher base level of expression11. Rh5 upregulation, when considered together with MSP1 and SERA5, presents an interesting observation, as they all serve a role in erythrocyte invasion.

# Conclusion

### DEseq results show 595 upregulated genes in severe malaria

DEseq analysis was crucial to our study and revealed many genes involved in erythrocyte and merozoite egress are upregulated in severe malaria. We made sure that data we were analyzing was just the *Plasmodium falciparum* transcriptome, but removing possible contaminants such as other strains of plasmodium. The MA and volcano plots allowed us to visualize the magnitude of expression change, and to see whether the genes were upregulated or downregulated. There were six times more upregulated genes than there were downregulated genes, and we paid special attention to the genes that were involved in any type of invasion mechanism when categorizing the genes. Once we categorized the genes, we discovered many genes related to erythrocyte invasion and merozoite egress (and their regulatory genes) were upregulated and this lead us to our literature search of the function of these genes.

### Proteins involved in erythrocyte egress found to be upregulated

Efficient egress from the infected erythrocyte, and subsequent invasion of new erythrocytes is an important component of the Plasmodium life cycle. Merozoite egress and erythrocyte invasion are especially important with regards to the severity of malaria. The blood stage of the Plasmodium is the cause of all clinical symptoms27. As such, an increase in the ability of the parasite to persist in the blood stages would be expected to increase the severity of the disease. Consistent with this hypothesis, we have observed increased transcription of four out of at least eight critical proteins associated with erythrocyte egress and subsequent erythrocyte invasion in severe cases of P. falciparum. It is important to note that we have not been able to confirm an increase in protein abundance or function. These upregulated genes are merozoite surface protein 1, merozoite protein 8, and serine repeat antigen 5 (SERA5) (Suppl. Table 1). These genes being upregulated provide evidence that there is an increase in malaria virulence when the ability to invade new healthy erythrocytes after infected erythrocyte egress.

### Ca2+ pathways and other erythrocyte egress regulatory proteins found to be upregulated

In addition, up-regulation of the regulatory proteins of SERA5 has been shown in severe cases of P. falciparum (Suppl. Table 1). This data show that malaria severity is correlated with the expression of these genes. This finding is consistent with previous studies which have been done to characterize the function and importance of these genes in P. falciparum. 1,2,7,8,11 In addition, CDPK1, Ca2+ Channels, PLC, and RAF kinase inhibitor are also upregulated (Suppl. Table 1). These proteins are regulators of SERA5. The fact that they are upregulated is of special note, because TKL is upregulated with them. Much research has been done to characterise the proteins involved in merozoite egress and invasion. A previous study has shown that TKL interacts with SERA5 in vitro, although the exact function of TKL is not known13. Up-regulation of TKL along with other regulators of SERA5 provides evidence that TKL is a regulator of SERA5, and plays a role in merozoite egress. In light of this evidence, biochemical research characterizing the interaction between TKL and SERA5 is needed to more fully understand the critical process of merozoite egress. This could be a future direction for other research groups to investigate in order to confirm our findings.

### Proteins involved with erythrocyte invasion and merozoite egress found to be upregulated

A number of merozoite surface proteins are found to be upregulated (Suppl. Table 1) in severe cases of malaria, and is important to the process of erythrocyte egress. These proteins are MSP1 and MSP8. The upregulation of these proteins provides further evidence that the upregulation of genes involved with invasion mechanisms correspond to higher disease virulence. Through this upregulation, the parasite is more effective in egressing infected erythrocytes and infecting healthy ones. Other proteins were found to be upregulated which are involved in erythrocyte invasion, such as Rh5. Studies have found that the role of Rh5 in erythrocyte invasion is to draw the erythrocyte closer to the surface of the plasmodium with the aid of p113 protein.11 All of these proteins are involved in erythrocyte invasion and merozoite egress in some capacity, and while some of their exact functions are unknown, we can see that they result in higher disease virulence. This study laid the groundwork for using RNA sequencing analysis to discover the genetic basis of malaria virulence, but further work needs to be done to study the parasite *in vitro*. A future direction for further studies could be studying the biochemical mechanisms of these proteins further to discover their exact role in erythrocyte invasion and merozoite egress.

**References**

1. Withers-Martinez, C., Suarez, C., Fulle, S., Kher, S., Penzo, M., Ebejer, J.-P., … Blackman, M. J. (2012). Plasmodium subtilisin-like protease 1 (SUB1): insights into the active-site structure, specificity and function of a pan-malaria drug target. International Journal for Parasitology, 42(6), 597–612. https://doi.org/10.1016/j.ijpara.2012.04.005
2. Iyer, G. R., Singh, S., Kaur, I., Agarwal, S., Siddiqui, M. A., Bansal, A., … Malhotra, P. (2018). Calcium-dependent phosphorylation of Plasmodium falciparum serine repeat antigen 5 triggers merozoite egress. The Journal of Biological Chemistry, 293(25), 9736–9746. <https://doi.org/10.1074/jbc.RA117.001540>
3. Collins, C. R., Hackett, F., Atid, J., Tan, M. S. Y., & Blackman, M. J. (2017). The Plasmodium falciparum pseudoprotease SERA5 regulates the kinetics and efficiency of malaria parasite egress from host erythrocytes. PLOS Pathogens, 13(7), e1006453. https://doi.org/10.1371/journal.ppat.1006453
4. Lourido, S., & Moreno, S. N. J. (2015). The calcium signaling toolkit of the Apicomplexan parasites Toxoplasma gondii and Plasmodium spp. Cell Calcium, 57(3), 186–193. <https://doi.org/10.1016/j.ceca.2014.12.010>
5. Withers-Martinez, Chrislaine et al. “Plasmodium subtilisin-like protease 1 (SUB1): insights into the active-site structure, specificity and function of a pan-malaria drug target.” International journal for parasitology vol. 42,6 (2012): 597-612. [https://doi:10.1016/j.ijpara.2012.04.005](https://www.ncbi.nlm.nih.gov/pubmed/22543039)
6. Naik, R. S., Branch, O. H., Woods, A. S., Vijaykumar, M., Perkins, D. J., Nahlen, B. L., … Gowda, D. C. (2000). Glycosylphosphatidylinositol anchors of Plasmodium falciparum: molecular characterization and naturally elicited antibody response that may provide immunity to malaria pathogenesis. The Journal of Experimental Medicine, 192(11), 1563–1576. <https://doi.org/10.1084/jem.192.11.1563>
7. Das, S., Hertrich, N., Perrin, A. J., Withers-Martinez, C., Collins, C. R., Jones, M. L., … Blackman, M. J. (2015). Processing of Plasmodium falciparum Merozoite Surface Protein MSP1 Activates a Spectrin-Binding Function Enabling Parasite Egress from RBCs. Cell Host & Microbe, 18(4), 433–444. <https://doi.org/10.1016/j.chom.2015.09.007>
8. Garzón-Ospina, D., Buitrago, S. P., Ramos, A. E., & Patarroyo, M. A. (2018). Identifying Potential Plasmodium vivax Sporozoite Stage Vaccine Candidates: An Analysis of Genetic Diversity and Natural Selection. Frontiers in Genetics, 9, 10. <https://doi.org/10.3389/fgene.2018.00010>
9. Drew, D. R., Sanders, P. R., & Crabb, B. S. (2005). Plasmodium falciparum merozoite surface protein 8 is a ring-stage membrane protein that localizes to the parasitophorous vacuole of infected erythrocytes. Infection and Immunity, 73(7), 3912–3922. <https://doi.org/10.1128/IAI.73.7.3912-3922.2005>
10. Annoura, T., Van Schaijk, B. C. L., Ploemen, I. H. J., Sajid, M., Lin, J. W., Vos, M. W., … Khan, S. M. (2014). Two Plasmodium 6-Cys family-related proteins have distinct and critical roles in liver-stage development. FASEB Journal, 28(5), 2158–2170. <https://doi.org/10.1096/fj.13-241570>
11. Galaway, F., Drought, L. G., Fala, M., Cross, N., Kemp, A. C., Rayner, J. C., & Wright, G. J. (2017). P113 is a merozoite surface protein that binds the N terminus of Plasmodium falciparum RH5. Nature Communications, 8, 14333. <https://doi.org/10.1038/ncomms14333>
12. Green, J. L., Rees-Channer, R. R., Howell, S. A., Martin, S. R., Knuepfer, E., Taylor, H. M., … Holder, A. A. (2008). The motor complex of Plasmodium falciparum: phosphorylation by a calcium-dependent protein kinase. The Journal of Biological Chemistry, 283(45), 30980–30989. <https://doi.org/10.1074/jbc.M803129200>
13. Gnangnon, B., Fréville, A., Cailliau, K., Leroy, C., De Witte, C., Tulasne, D., … Pierrot, C. (2019). Plasmodium pseudo-Tyrosine Kinase-like binds PP1 and SERA5 and is exported to host erythrocytes. Scientific Reports, 9(1), 8120. <https://doi.org/10.1038/s41598-019-44542-3>
14. World Health Organization, Malaria Factsheet ; WHO, March 27, 2019 https://www.who.int/news-room/fact-sheets/detail/malariaAccessed 6/6/2019
15. Tonkin-Hill, G. Q., Trianty, L., Noviyanti, R., Nguyen, H. H. T., Sebayang, B. F., Lampah, D. A., … Maccallum, P. (2018). The Plasmodium falciparum transcriptome in severe malaria reveals altered expression of genes involved in important processes including surface antigen–encoding var genes. <https://doi.org/10.1371/journal.pbio.2004328>
16. Kim, Daehwan and Langmead, Ben and Salzberg, Steven L (2015). HISAT: a fast spliced aligner with low memory requirements. In Nature Methods, 12 (4), pp. 357â360. [doi:10.1038/nmeth.3317][Link]
17. Quinlan, Aaron R. and Hall, Ira M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. In *Bioinformatics, 26 (6), pp. 841â842.* [[doi:10.1093/bioinformatics/btq033](https://doi.org/10.1093/bioinformatics/btq033)][[Link](http://dx.doi.org/10.1093/bioinformatics/btq033)]
18. Liao, Y. and Smyth, G. K. and Shi, W. (2013). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. In *Bioinformatics, 30 (7), pp. 923â930.* [[doi:10.1093/bioinformatics/btt656](https://doi.org/10.1093/bioinformatics/btt656)][[Link](http://dx.doi.org/10.1093/bioinformatics/btt656)]
19. Love, Michael I and Huber, Wolfgang and Anders, Simon (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. In *Genome Biology, 15 (12).* [[doi:10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8)][[Link](http://dx.doi.org/10.1186/s13059-014-0550-8)]
20. Nadjm B, Behrens RH (2012). "Malaria: An update for physicians". Infectious Disease Clinics of North America. 26 (2): 243–59. [doi](https://en.wikipedia.org/wiki/Digital_object_identifier):[10.1016/j.idc.2012.03.010](https://doi.org/10.1016%2Fj.idc.2012.03.010). [PMID](https://en.wikipedia.org/wiki/PubMed_Identifier) [22632637](https://www.ncbi.nlm.nih.gov/pubmed/22632637).
21. de Macedo CS, Schwarz RT, Todeschini AR, Previato JO, Mendonça-Previato L. Overlooked post-translational modifications of proteins in Plasmodium falciparum: N- and O-glycosylation—a review. Mem Inst Oswaldo Cruz. 2010;105:949–56.
22. Von Itzstein M, Plebanski M, Cooke BM, Coppel RL. Hot, sweet and sticky: the glycobiology of Plasmodium falciparum. Trends Parasitol. 2008;24:210–8.
23. Cova, M., Rodrigues, J. A., Smith, T. K., & Izquierdo, L. (2015). Sugar activation and glycosylation in Plasmodium. Malaria Journal, 14(1), 427. <https://doi.org/10.1186/s12936-015-0949-z>
24. Lee, A. H., Symington, L. S., & Fidock, D. A. (2014). DNA Repair Mechanisms and Their Biological Roles in the Malaria Parasite Plasmodium falciparum. Microbiology and Molecular Biology Reviews, 78(3), 469–486. <https://doi.org/10.1128/MMBR.00059-13>
25. Severe Malaria Section 1: Epidemiology of severe falciparum malaria. (2014). <https://doi.org/10.1111/tmi.12313>
26. <https://www.cdc.gov/malaria/about/disease.html>

Supplemental Information

1. <https://github.com/Clementse90/Malaria-Study>
2. <https://github.com/Clementse90/Malaria-Study/blob/master/Sample%20Reads.xlsx>
3. <https://github.com/Clementse90/Malaria-Study/blob/master/Up_vs_Down_Regulation_Falciparum.html>

*Last modified: 2019-07-09 13:35*