UNIVERSIDADE ESTADUAL DE CAMPINAS

INSTITUTO DE BIOLOGIA

Catarina Baeta da Luz Bourgard

Analyses of the immunopathological and molecular mechanisms involved in cytoadherence of *Plasmodium vivax*

Análises dos mecanismos imunopatológicos e moleculares envolvidos no processo de citoaderência de *Plasmodium vivax*

CAMPINAS
2019

Catarina Baeta da Luz Bourgard

Analyses of the immunopathological and molecular mechanisms involved in cytoadherence of *Plasmodium vivax*

Análises dos mecanismos imunopatológicos e moleculares envolvidos no processo de citoaderência de *Plasmodium vivax*

Thesis presented to the Institute of Biology of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Genetics and Molecular Biology, in the Immunology area.

Tese apresentada ao Instituto de Biologia da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do Título de Doutora em Genética e Biologia Molecular, na área de Imunologia.

ESTE ARQUIVO DIGITAL CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELA ALUNA CATARINA BAETA DA LUZ BOURGARD E ORIENTADA PELO PROF. DR. FABIO TRINDADE MARANHÃO COSTA.

Orientador: Prof. Dr. Fabio Trindade Maranhão Costa

Co-Orientador: Profª. Drª. Letusa Albrecht

CAMPINAS
2019

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

<div style="border:1px solid">

B666a   Bourgard, Catarina Baeta da Luz, 1985-
       Analyses of the immunopathological and molecular mechanisms involved in
cytoadherence of *Plasmodium vivax* / Catarina Baeta da Luz Bourgard. –
Campinas, SP : [s.n.], 2019.

         Orientador: Fabio Trindade Maranhão Costa.
         Coorientador: Letusa Albrecht.
         Tese (doutorado) – Universidade Estadual de Campinas, Instituto de
Biologia.

         1. Malária. 2. *Plasmodium*. 3. *Plasmodium vivax*. 4. Fenótipo. 5.
Transcriptoma. I. Costa, Fabio Trindade Maranhão, 1972-. II. Albrecht, Letusa.
III. Universidade Estadual de Campinas. Instituto de Biologia. IV. Título.

</div>

Informações para Biblioteca Digital

Título em outro idioma: Análises dos mecanismos imunopatológicos e moleculares
envolvidos no processo de citoaderência de *Plasmodium vivax*
Palavras-chave em inglês:
Malaria
Plasmodium
Plasmodium vivax
Phenotype
Transcriptome
Área de concentração: Imunologia
Titulação: Doutora em Genética e Biologia Molecular
Banca examinadora:
Fabio Trindade Maranhão Costa [Orientador]
Cristiana Ferreira Alves de Brito
Fernanda Janku Cabral
Gerhard Wunderlich
Danilo Ciccone Miguel
Data de defesa: 01-03-2019
Programa de Pós-Graduação: Genética e Biologia Molecular

COMISSÃO EXAMINADORA

Prof. Dr. Fabio Trindade Maranhão Costa (Orientador)

Profª. Drª. Fernanda Janku Cabral

Prof. Dr. Danilo Ciccone Miguel

Profª. Drª. Cristiana Ferreira Alves de Brito

Prof. Dr. Gerhard Wunderlich

Os membros da Comissão Examinadora acima assinaram a Ata de defesa, que se encontra no processo de vida acadêmica do aluno.

To all family, friends and colleagues,
who go together with me through this adventurous

Aquarela

" (…)

E ali logo em frente, a esperar pela gente, o futuro está
E o futuro é uma astronave que tentamos pilotar,
Não tem tempo nem piedade, nem tem hora de chegar
Sem pedir licença muda nossa vida, depois convida a rir ou chorar

Nessa estrada não nos cabe conhecer ou ver o que virá
O fim dela ninguém sabe bem ao certo onde vai dar
Vamos todos numa linda passarela
De uma aquarela que um dia, enfim, descolorirá "

(in Aquarela *song from* Toquinho*,*
composed by Gabriela Luna Borges*)*

AGRADECIMENTOS

Na vã tentativa de incluir, neste pequeno espaço, a minha sincera e sentida gratidão a todos quantos verdadeiramente contribuíram e impactaram pessoal e profissionalmente nestes últimos cinco anos, começo por agradecer:

Ao Prof. Dr. Fabio T. M. Costa, que apostou e confiou nas minhas capacidades de adaptação ao desconhecido e de trabalho numa área de pesquisa tão desafiadora quanto motivante, como é a malária. Obrigada pela orientação e apoio sempre presentes, e além disso, por promover todo um conjunto de colaborações científicas, incentivando-me a buscá-las de forma independente, tanto no Brasil como por esse Mundo fora. Levo como lição o seu exemplo de capacidade de promoção da cooperação e divulgação da pesquisa científica, que espero continuar a fazer por muitos e bons anos.

À Profª. Drª. Letusa Albrecht, minha co-orientadora implacável, sempre exigindo mais e o melhor de mim, e acima de tudo uma verdadeira amiga, ensinando-me a ser positiva e a dar a volta por cima, principalmente quando "os ventos não estavam de feição".

À Profª. Drª. Stefanie Lopes, a minha segunda co-orientadora que me abriu as portas de Manaus e me ensinou a fazer trabalho de campo "como deve ser!". Minha amiga do coração que me "deu uma família adotiva" quando o meu coração saudoso mais precisou.

Muito obrigada aos três, por terem dado origem, promovido a continuidade e apoiarem até ao fim este projeto arriscado e ambicioso.

Aos membros da comissão da qualificação e defesa, Prof. Dr. Marco Aurélio Vinolo, Prof. Dr. Gerhard Wunderlich, Profª. Drª. Elisabeth Bilsland, Prof. Dr. Renato Vicentini dos Santos, Profª. Drª. Cristiana Brito e Profª. Drª. Fernanda Janku Cabral, pela importante discussão científica.

A todos os membros, passados e presentes, do Laboratório de Imuno-parasitologia Experimental, que tornaram os dias mais leves e divertidos: Carol, Catarina, Carla, Isabel, Juliana, João, João Luiz, Letícia, Liam, Letícia, Marcele, Najara, Tatyana, Natália, Ana, Gustavo, Kaira e Luis.

Ao Prof. Dr. Per Sunnerhagen, por ter compreendido em uma única apresentação, o desafio da análise de dados que estava por vir, e com isso ter-me aberto as portas do seu laboratório na Universidade de Gotemburgo, orientando-me a procurar colaborações e apoio no aprendizado nas áreas de bioinformática e genética de leveduras. Continuaremos juntos neste novo desafio da descoberta de novos antimaláricos contra a malária vivax.

A todos os membros do laboratório do Prof. Dr. Per Sunnerhagen com quem me cruzei, que me adotaram e em muito contribuíram para aquecer o ambiente, desafiando as temperaturas negativas que se faziam sentir no Inverno Sueco: Elena, Elisa, Heejung, Johanna, Hanna, Agata, Martin, Anaswara, Vianey, Hanne, Elisabet e Linde. Um abraço forte e especial para as minhas amigas Elena, Elisa e Agata.

Aos meus melhores e queridos amigos de Portugal, Germana, Inês e Afonso, reservo um caloroso e gigante abraço, do tamanho da distância que nos separa, mas que nunca impediu que estivessem sempre presentes no meu coração.

À minha família, que sempre desejou o melhor para mim, especialmente aos meus pais, que me apoiaram incondicionalmente nesta minha aventura, mesmo quando o coração apertava de saudade! E ao Philippe, que surgiu inesperadamente pra compor a minha vida, que lida comigo como ninguém e que me fez acreditar que o impossível está ali, à mão de semear.

ABSTRACT

Plasmodium vivax is the most prevalent, widespread and neglected human malaria parasite, currently placing billions of people at risk of infection, thus imposing major health and economic burdens. Worldwide, anti-malarial drug resistance emergence and severe clinical complications are of great concern. The mechanisms underlying the pathobiology of the neglected *P. vivax* are still little known. The lack of a reliable *in vitro P. vivax* long-term culture restricts its biology study in place and time, relegating researchers to work in malaria endemic field conditions, where successful *omics* applications are very challenging.

The capacity of *P. vivax* to remodel host reticulocyte membrane and promote adhesivity has been demonstrated, which is an important mechanism for host immune evasion. Functional studies have already reported that adhesion of *P. vivax* infected red blood cells (PvIRBCs) to the host endothelial cells, although in considerably lower rates, is as strong and stable as the verified for *P. falciparum* infections. Also, it has been reported adhesion of normocytes to the PvIRBCs is strong and results in stable rosette formation, which shows higher rates in vivax compared to falciparum malaria. More recently, it was reported that there is a correlation between rosette formation and altered membrane deformability of PvIRBCs, where the rosette-forming PvIRBCs are significantly more stiff and rigid than their non-rosetting equals. Mature staged parasites (schizonts) show a higher capacity for adherence than other asexual parasite stages both in cytoadherence and rosetting. The lower proportion of schizonts observed on the peripheral blood circulation of patients suggests that parasites could be sequestered on the host vascular endothelium. Rosette-forming PvIRBCs may also be the cause for this lower rate of schizonts in the patients' blood, contributing for parasite sequestration phenomena in the host microvasculature and/or spleen, and consequently, the rheopathological characteristics present in vivax malaria disease. Vivax malaria patient autopsies have shown accumulation of PvIRBCs in the lungs, spleen, liver and bone marrow. Additionally, it has been demonstrated that parasitemia underestimates total parasite biomass, which is greater in severe vivax malaria patients, and thus, capable of mediating systemic inflammatory pathology.

In this study, we aimed to understand the molecular mechanisms behind adherence phenotypes by identifying proteins, especially parasitic ligands, which might

be important in *P. vivax* adhesion capacity. Using RNA-seq coupled with parasite field sample enrichment, *ex vivo* maturation and cytoadherence assays, we have sequenced the whole transcriptome of parasite populations with distinct adhesive characteristics. Our expression profiles brings out the importance of membrane and membrane-associated proteins, with adhesin or adhesin-like properties, such as *Plasmodium* Interspersed repeats (PIR) and *Plasmodium* Helical Interspersed SubTelomeric (PHIST) proteins, which might pay a role in adherence phenotypes. Within those protein groups, we found a percentage of differentially expressed genes that traditionally are more expressed in sexual rather than asexual parasite stages, suggesting the relevance of rosette formation by *P. vivax* gametocytes. Importantly, we found host immune-related differentially expressed genes, of which several are associated with the human phagocytosis pathways. These data strongly suggest that rosetting can hamper leukocyte phagocytosis host immune response, as an effective mechanism of *P. vivax* immune evasion adaptation. Our results reflect the pathobiology of circulating Brazilian *P. vivax* populations, principally concerning its adhesive capacity as a possible source of the severe clinical manifestations reported. Furthermore, we hope that such achievements will further enable the investigations on the biology of *P. vivax* apicomplexan parasite, impacting considerably in vaccine and drug design, ultimately helping us achieve the future elimination of vivax malaria.

Resumo

Plasmodium vivax é o parasita causador malária humana mais prevalente, disseminado e negligenciado, colocando todos os anos bilhões de pessoas em risco de infecção, acarretando sérios problemas de saúde e econômicos. A emergência de resistência a antimaláricos e complicações clínicas graves são preocupantes. Pouco se sabe sobre os mecanismos envolvidos nas características patogénicas da biologia do negligenciado *P. vivax*. A impossibilidade de executar a cultura *in vitro* de isolados a longo prazo, limita os pesquisadores ao estudo da sua biologia no espaço e tempo, restringindo o trabalho experimental a áreas endêmicas de malária vivax, onde a execução bem-sucedida de aplicações ômicas é desafiadora.

A capacidade de *P. vivax* remodelar a membrana dos reticulócitos do hospedeiro e promover a sua adesividade foi já demonstrada, dendo um mecanismo importante de evasão ao sistema imunitário humano. Estudos funcionais têm reportado que a adesão de reticulócitos infectados por *P. vivax* (RTi-Pv) a células endoteliais do hospedeiro, apesar de em menor número, é forte e estável, como o verificado para eritrócitos infectados por *P. falciparum*. Também foi observado que a adesão de eritrócitos não infectados a RTi-Pv é forte e resulta na formação estável de rosetas, que apresenta uma maior taxa em malária vivax do que falciparum. Mais recentemente, foi publicado que existe uma correlação entre a formação de rosetas e a deformabilidade de RTi-Pv, na qual RTi-Pv que formam rosetas são significativamente mais rígidos. Estágios maduros de *P. vivax* (esquizontes) têm elevada capacidade de aderir relativa a outros estágios assexuais do parasita, em ambos fenótipos de citoadesão e roseteamento. A menor proporção de esquizontes na circulação sanguínea periférica de pacientes sugere que os parasitas podem sequestrar no endotélio vascular do hospedeiro. Os RTi-Pv que formam rosetas poderão ser a causa desta baixa taxa de esquizontes circulantes no sangue dos pacientes, contribuindo para o fenómeno de sequestro parasitário na microvasculatura e/ou baço do hospedeiro, e consequentemente contribuindo para características reopatológicas da malária vivax. Autópsias de pacientes com malária vivax mostraram a acumulação de RTi-Pv nos pulmões, baço, fígado e medula óssea. Ainda, foi demonstrado que a parasitemia subestima a biomassa total parasitária, que é elevada

em malária vivax severa, portanto, capaz de mediar a inflamação sistémica da patologia.

O objetivo deste estudo é a compreensão dos mecanismos moleculares envolvidos nos fenótipos adesivos, identificando proteínas, principalmente ligantes parasitários, que possam ser importantes na capacidade de aderência de *P. vivax*. Usando RNA-seq em conjunto com o enriquecimento, maturação *ex vivo* e ensaio funcional de adesão com amostras clínicas de *P. vivax*, foi sequenciado o transcriptoma de populações de parasitas com características adesivas distintas.

Os nossos perfis de expressão mostram a importância de diferentes grupos de proteínas membranares ou associadas a membranas, com propriedades de adesinas, tais como proteínas *Plasmodium* Interspersed repeats (PIR) e *Plasmodium* Helical Interspersed SubTelomeric (PHIST), que podem ter um papel importante no fenótipo adesivo de *P. vivax*. Dentro deste grupo de proteínas diferencialmente expressas foi verificado que muitas são tradicionalmente produzidas por parasitas em fase sexuada, sugerindo a importância da formação de rosetas por gametócitos de *P. vivax*. Adicionalmente, análise do perfil de expressão gênica humano permitiu a identificação de genes diferencialmente expressos associados a vias de fagocitose. Estes dados sugerem fortemente que o fenótipo de roseteamento pode impedir a fagocitose do parasita por leucócitos como resposta do sistema imune humano à infeção. Os resultados obtidos refletem as características patogénicas de populações brasileiras circulantes de *P. vivax*, principalmente no que diz respeito à sua capacidade de aderência como principal fonte das manifestações clínicas severas reportadas. Para além disso, esperamos que estes dados abram ainda mais as investigações sobre a biologia deste parasita apicomplexo, ajudando no desenho de vacinas e na descoberta de novos antimaláricos, promovendo o sucesso na eliminação da malária vivax no futuro.

**Palavras-chave:** malária, *Plasmodium* spp., *P. vivax*, fenótipo adesivo, transcriptoma

# INDEX

INTRODUCTION

VIVAX MALARIA WORLD SCENARIO

Plasmodium spp. belongs to a larger phylum of obligate intracellular parasites, the Apicomplexa, and is the causative agent of malaria. Worldwide, human malaria infections (Fig. 1) can be caused by five different *Plasmodium* species: *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae*, and *P. knowlesi* (Cowman et al. 2017). *P. falciparum* is by far the deadliest parasite, causing the most severe clinical outcomes. Therefore, it has received much attention in the past years.
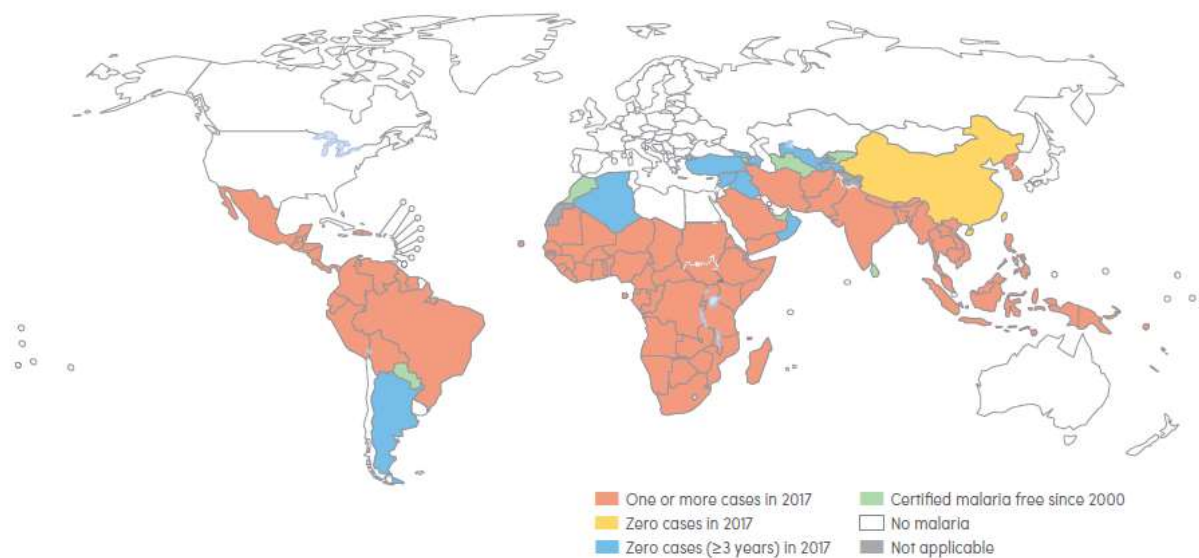
Figure 1. Status of 2017 malaria cases in countries assessed for malaria occurrence since 2000. All countries in the WHO European Region reported zero cases in 2016 and again in 2017. In 2017, both China and El Salvador reported zero cases. By definition, countries with zero malaria cases over at least the past 3 consecutive years are malaria free. Adapted from World Malaria Report 2018 (WHO 2018).

However, *P. vivax* is the most geographically spread human malaria parasite (WHO 2018, Gething et al. 2012), and recently, it has reemerged in regions formerly considered malaria free (Severini et al. 2004, Bitoh et al. 2011, Kim et al. 2009). *P. vivax* mortality and morbidity has been recently reassessed and are likely to have been significantly underestimated (Naing et al. 2014). Worldwide, about 2.85 billion people have been estimated to be in risk of being infected by *P. vivax* (Guerra et al. 2010, Gething et al. 2012, Price et al. 2007, Battle et al. 2012). This parasite is predominant in Central and Southeast Asia (Rahimi et al. 2014), Latin America (Guerra et al. 2010, Coura, Suarez-Mutis, and Ladeia-Andrade 2006, Ferreira and Castro 2016) and some African regions (Rosenberg 2007, Howes et al. 2015). Of those regions, like Central

and Southeast Asia, are densely populated, thus accentuating the social-economic burden caused by the disease (WHO 2018, Gething et al. 2012) (Fig. 2).
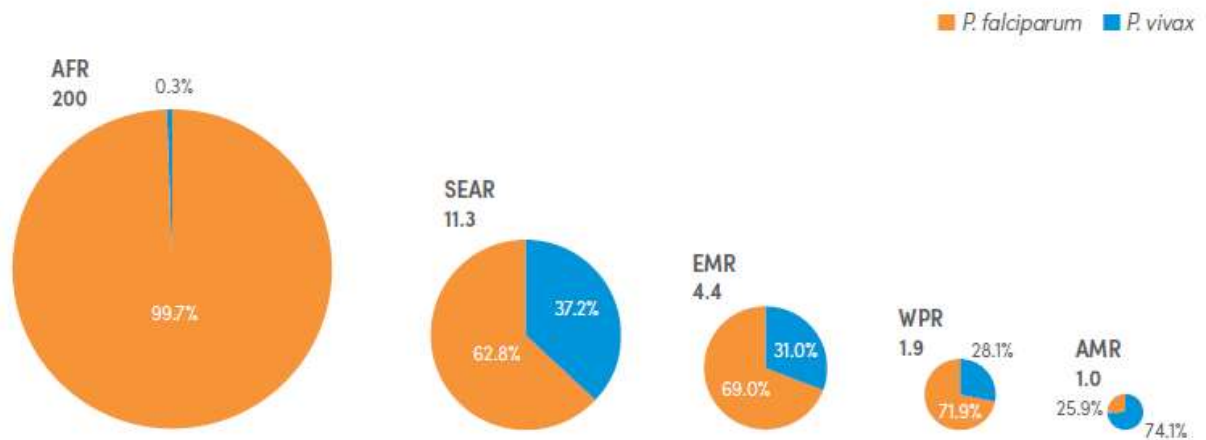

Figure 2. 2017 estimated (million) cases of malaria by WHO region. The circles' area reflects the percentage of the estimated number of malaria cases in each region. AFR: WHO African Region; AMR: WHO Region of the Americas; EMR: WHO Eastern Mediterranean Region; *P. falciparum*: *Plasmodium falciparum*; *P. vivax*: *Plasmodium vivax*; SEAR: WHO South-East Asia Region; WHO: World Health Organization; WPR: WHO Western Pacific Region. Adapted from World Malaria Report 2018 (WHO 2018).

The fact that human infecting *Plasmodium* spp is geographically widespread leads to infection occurrence in genetically distinct human populations. Clinic complications and heterogeneous resistance to antimalarial drugs observed are a direct result of different and diverse individual responses in a host-parasite relation in a context of *P. vivax* polyclonal infection (Guerra et al. 2010, Anstey et al. 2009). According to the WHO guidelines (WHO 2018), the first line of *P. vivax* chemotherapy is chloroquine (CQ) plus primaquine (PQ), the only approved drug targeting the latent parasite form. These drugs are extensively used in regions where resistance has not been verified, whereas in high transmission areas already presenting cases of drug resistance, an artemisinin based combination therapy (ACT) is recommended (WHO 2010). More recently, Tafenoquine (TAF), an analog of PQ, has been approved as an efficient single-dose treatment against vivax malaria relapse (Lacerda, Llanos-Cuentas et al. 2019).

In the last decade, an accentuated reduction of *P. falciparum* infection cases is occurring in falciparum and vivax malaria endemic areas (WHO 2018, Childs et al. 2006), as a direct consequence of the global advances accomplished to control malaria falciparum transmission and, more recently, with important breakthroughs in the generation of an effective malaria vaccine (Arama and Troye-Blomberg 2014, Lorenz, Karanis, and Karanis 2014). However, there are recent reports of a marked increase

of *P. vivax* resistance to CQ (Suwanarusk et al. 2007, Suwanarusk et al. 2008), for which the mechanisms of action remain undisclosed (Price et al. 2014), leading to a more accurate observation of the particular pattern of *P. vivax* transmission (Rahimi et al. 2014, Snounou and White 2004, Genton et al. 2008, Tjitra et al. 2008). We urge for a new therapeutic line of antimalarials that targets vivax malaria patients and an effective multi-species malaria vaccine for such endemic areas. Unfortunately, *P. vivax* malaria remains neglected in most aspects (Baird 2007, Mendis et al. 2001, Mueller et al. 2009, Price et al. 2007, Carlton, Sina, and Adams 2011).

The characteristics that permitted an easier control over *P. falciparum* infections allow studying and targeting *P. vivax*, which can make a difference in the future worldwide treatment and eradication of malaria. Following the decline of *P. falciparum* infections*, P. vivax* is now one of the dominant malaria species in several endemic regions (Gething et al. 2012, WHO 2018, Coura, Suarez-Mutis, and Ladeia-Andrade 2006, Hussain et al. 2013), especially on the region of the Americas (Fig. 1, 2 and 3). Last epidemiology studies reported by WHO (WHO 2018) from Latin America region, estimates that 138 million people are at risk, where *P. vivax* is responsible for 75% of malaria infections, mainly concentrated on the Amazonian region, where nine different *Anopheles* spp. vectors (*An. albimanus, An. albitarsis, An. aquasalis, An. braziliensis, An. darlingi, An. neivai, An. nuneztovari, An. pseudopunctipennis* and *An. punctimacula*) help spreading the disease (Fig. 3). Around 22% of all cases have been registered in Brazil, where since 2010, the main agent of malaria infection is *P. vivax* (WHO 2018).
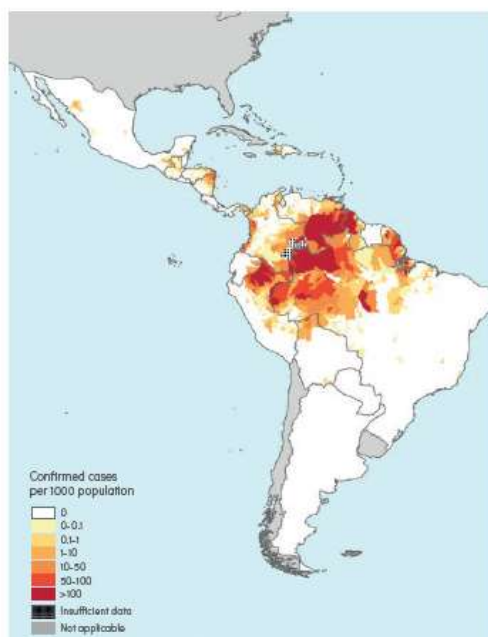
Figure 3. Confirmed malaria cases per 1000 population in 2017 (WHO 2018).

The direct transfer of *P. falciparum* control measures to *P. vivax* has turned out to be inadequate (Bockarie and Dagoro 2006, Luxemburger et al. 1994, Bousema and Drakeley 2011, Andrade et al. 2010, Baird 2010), a scenario that can be observed in some endemic regions (WHO 2018, Coura, Suarez-Mutis, and Ladeia-Andrade 2006). However, it is allowing the evaluation of *P. vivax* infection without *P. falciparum* influence on patients.

Several clinical complications that were normally associated with *P. falciparum* infections have been reported for vivax malaria (Rahimi et al. 2014, Barcus et al. 2007, Genton et al. 2008, Kochar, Mahajan, et al. 2009, Kochar, Das, et al. 2009, Kochar et al. 2005, Alexandre et al. 2010, Siqueira et al. 2010, Fernandez-Becerra, Pinazo, et al. 2009, Tjitra et al. 2008, Tan et al. 2008, Anstey et al. 2007, Anstey et al. 2002, Suratt and Parsons 2006, Poespoprodjo et al. 2008, McGready et al. 2004, Hutchinson and Lindsay 2006, Baird 2007, Erhart et al. 2004, Chung et al. 2008, Sharma et al. 1993, Saharan et al. 2009). More accurate and innovative diagnose tools are now able to give us a more accurate rates of *P. vivax* infections, mainly underreported in the past, which is challenging the pre-established view of *P. vivax* as a "benign" parasite (Mendis et al. 2001, Anstey et al. 2009, Mueller et al. 2009, Naing et al. 2014, Gething et al. 2012, Baird 2007). Furthermore, a constant increase and spread of anti-malarial resistance remains of great concern (Tjitra et al. 2008, Baird 2004, de Santana Filho et al. 2007, Poespoprodjo et al. 2008, Suwanarusk et al. 2007, Russell et al. 2008, Price, Douglas, and Anstey 2009, Price et al. 2014).

P. vivax is considered as an severe pathogen causing progressive anemia with repeated hemolysis (Collins, Jeffery, and Roberts 2003), and dyserythropoiesis (Wickramasinghe et al. 1989) episodes and other severe manifestations, showing a higher incidence in young children in endemic areas (Genton et al. 2008, Price et al. 2007, Tjitra et al. 2008, Marsh et al. 1995, Williams et al. 1997). This implies the urgent need to gather a profound knowledge about its biology, in general, and particularly its immunopathological mechanisms.

Nowadays, the available molecular tools (reviewed in (Escalante et al. 2015)) rise as a valuable option to better estimate the prevalence and incidence of vivax malaria, principally in low transmission (lower parasitemia) endemic regions and its dynamics in terms of differential contributions between host and vectors. Also, within the specific demographic and migration contexts of parasite populations (different patterns of gene flow observed under colonization of new areas and/or expansion) and infected

individuals, the efficacy of interventions on the infections (occurrence rates, complications and duration) can be carefully examined. Importantly, such tools can successfully distinguish cases of recrudescence, relapse and infection *de novo* and investigate the emergence of drug resistance.

PLASMODIUM VIVAX: AN OVERVIEW

The life cycle of *Plasmodium* spp. involves two different hosts: the definitive host or vector, which is the mosquito of the genus *Anopheles*, followed by a vertebrate intermediate host, such as a human (Fig. 4). As such, vivax malaria is a mosquito-borne infectious disease caused by the transmission of *P. vivax* protozoan parasites present on the salivary glands of infected female *Anopheles* spp. mosquitoes during the blood meal (Kiszewski et al. 2004).

In resume, upon entrance, the *P. vivax* sporozoites migrate to the liver where they either mature into an active schizont, going through several rounds of division, after which the resultant merozoites are released into the bloodstream, or developed into a latent form called hypnozoite (Krotoski 1985). The *P. vivax* erythrocytic cycle takes about 48 hours, where merozoites almost exclusively invade reticulocytes (RTs), altering some of their properties, in particular their deformability and enlargement (Suwanarusk et al. 2004), the formation of caveola-vesicle like complexes (CVC) and cytoplasmic cleft structures (Barnwell et al. 1990). Some parasites have the capacity of develop into round-shaped gametocytes, quite early after infection (Boyd 1937), which will be taken up by other *Anopheles* mosquitoes during the blood meal. Initiation of the sexual cycle occurs with the development of male and female gametocytes, subsequent fertilization, formation of motile ookinetes that migrate to the mosquito midgut epithelium and further differentiate into oocysts. Upon maturation, these oocysts release sporozoites, which travel and invade the salivary glands. This enables them to become injected into another host while the female mosquito feeds, perpetuating the complex *P. vivax* life cycle (Mueller et al. 2009) (Fig. 4).
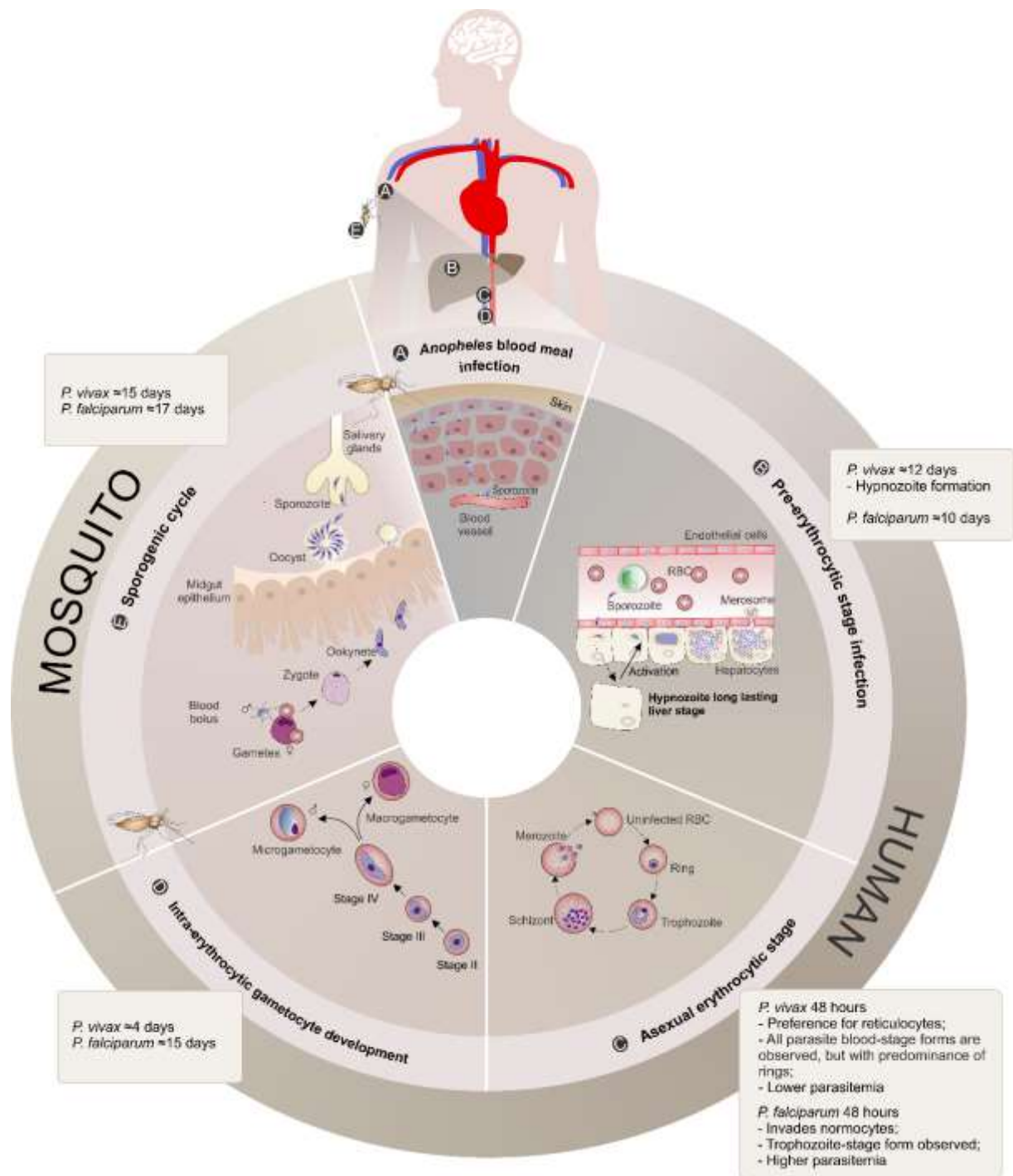
Figure 4. *Plasmodium vivax* and *Plasmodium falciparum* life cycle comparison. **(A)** *Anopheles* blood meal infection; **(B)** Pre-erythrocytic stage infection; **(C)** Asexual erythrocytic stage; **(D)** Intra-erythrocytic gametocyte development; **(E)** Mosquito Stage. Adapted from (Bourgard et al. 2018).

P. vivax had an evolutionary path distinct from *P. falciparum*, being more closely related to *P. cynomolgi*, a sister taxon that infects Asian macaque monkeys, and to other malaria parasites all belonging to a clade that infects Old World monkeys, Southeast Asia apes and West African wild apes (Liu et al. 2014, Luo, Sullivan, and Carlton 2015, Duval et al. 2010). In fact, in a context of close interaction, *P. cynomolgi* can infect humans (Ta et al. 2014, Ramasamy 2014). Probably, as a consequence of

such a unique evolutionary path, *P. vivax* has several features (Baird 2007, Mueller et al. 2009, Gething et al. 2012) of which three key biological properties are the source of the differences seen between *P. vivax* and *P. falciparum* (Fig. 4):

(1) Preference for invading RTs (Field 1956), increasing their deformability, size, fragility and permeability (Kitchen 1938, Suwanarusk et al. 2004, Handayani et al. 2009, Desai 2014). The ligand PvRBP2 was identified as the responsible for the *P. vivax* preferential binding to reticulocytes (Galinski et al. 1992). Some researchers have hypothesized that deformability would allow *P. vivax* to evade the host immune system, by safe passage of infected RTs through the spleen (del Portillo et al. 2004, Handayani et al. 2009, Suwanarusk et al. 2004) and also, because all blood-stage forms of the parasite are observed in the peripheral blood circulation, it was believed that *P. vivax* did not need to cytoadhere. However, studies have shown that *P. vivax* has the capacity to cytoadhere (Carvalho et al. 2010, Lopes et al. 2014, Field 1956, Chotinavich 2003, Rogerson, Mwapasa, and Meshnick 2007, Siqueira, Alexandre et al. 2010, Lacerda, Fragoso et al. 2012, Machado Siqueira, Lopes Magalhaes et al. 2012, Elizalde-Torrent, Val et al. 2018) and to form rosettes (Udomsanpetch et al. 1995) and some biopsy data have shown *P. vivax* sequestered in the spleen and the lungs (del Portillo et al. 2004, Anstey et al. 2007). Moreover, RTs represent 1-2% of blood stream circulating erythrocytes, maintaining the characteristic low biomass parasitemia, which could be an adaptation to limit the occurrence of hyperparasitemias and the associated virulence, or the necessity of a specific environment for *P. vivax* to grow and multiply (Mueller et al. 2009). Further studies are needed to clarify all these aspects.

(2) Earlier production of sexual stages during the infection (Boyd 1937) with a characteristic spherical shape seen in peripheral blood circulation, even before the beginning of clinical symptoms, which might function as a reservoir to promote successful transmission to the mosquitoes (Boyd 1937, Bousema and Drakeley 2011).

(3) the formation of dormant hypnozoites, which remain in the liver in a latent state (Krotoski et al. 1982, Baird, Schwartz, and Hoffman 2007, White 2011). Upon reactivation, hypnozoites, which are not eliminated by the majority of approved antimalarials, can restart the erythrocytic cycle of infection, causing subsequent relapses weeks or months after the first infection (Battle et al. 2014) and the likely transmission of the sexual gametocytes (White 2011, Krotoski et al. 1982, Betuela et al. 2012, Orjuela-Sanchez et al. 2009).

The mechanisms and triggering factors underlying hypnozoite activation are still unknown (Mueller et al. 2009). Environmental and host stress might have a role, and some reactivation patterns seems to be connected with *Anopheles* spp. seasonal peaks (Hulden, Hulden, and Heliovaara 2008), maintaining a dynamic genetic pool, available to continue the transmission and propagation of the parasite (Mueller et al. 2009, White et al. 2016, White 2011, Shanks and White 2013, Hulden and Hulden 2011, White et al. 2014). PQ is the only drug available that can effectively kill hypnozoites, but unfortunately, complications have been reported for glucose-6-phosphate dehydrogenase (G6PD) deficient patients (Beutler, Duparc, and Group 2007), thus limiting its mass administration in endemic regions (Bowman et al. 2004).

The lack of a reliable and reproducible *in vitro* system for long term culture of *P. vivax* and no easy access and use of monkey based studies with adapted strains of *P. vivax* (Beeson and Crabb 2007, Panichakul et al. 2007) has greatly restricted research in time and place and impose great limitations, both in the quantity and quality of functional assays performed in *ex vivo* short term cultures (Noulin et al. 2013). Thus, the *P. vivax* research community faces great hurdles to uncover the biology of this elusive parasite, posing huge hitches for control policies decision-making.

PLASMODIUM VIVAX GENETIC DIVERSITY

An important piece of evidence published by Orjuela-Sánchez, *P.* and collaborators (Orjuela-Sanchez et al. 2013), alerted the vivax malaria community to another level of complexity. A higher genetic diversity of *P. vivax* by microsatellite analysis was observed, in a context of infection polyclonality of Asian clinic isolates, compared to the same genetic diversity verified for *P. falciparum* Asian isolates at correspondent geographic locations (Orjuela-Sanchez et al. 2013). More recently, publications based on NGS methodologies support the high level of genetic diversity verified for this parasite (Luo, Sullivan, and Carlton 2015, Hupalo et al. 2016).

PLASMODIUM VIVAX ADHESION PHENOTYPES

Whereas some other aspects of *Plasmodium* spp. have been clarified, such as the core machinery driving invasion (reviewed in (Cowman et al. 2017)), the ligand-receptor interactions between the parasite and the host erythrocyte remain unidentified, but we know them to show significant molecular heterogeneity and evolution. For instance, *P. vivax* infecting humans have similar diversity as the chimp

and gorilla *P. vivax*-like parasites, suggesting that the emergence of *P. vivax* in the human population was an ancient event, which have led to a change and optimization of the parasite ligands from their ape counterparts to the human *P. vivax* (Liu, Li et al. 2014). The polymorphism of parasite ligands allows the parasite to colonize varied ecological niches in addition to evading the immune system. On the contrary, host erythrocyte variation can serve to curb the efficiency of *Plasmodium* invasion, a factor in limiting infections and pathogenesis. Little is known about these ligand-receptor interactions involved in the adhesive capacity of *P. vivax* (Bernabeu et al. 2012, Fernandez-Becerra, Yamamoto, et al. 2009) and its host cell tropism.

Cytoadhesion

In the 60's, a "hiding" capacity of *P. vivax* was proposed, in which a disproportionally low level of mature blood stages (schizonts) relative to young stages (trophozoites) was found in peripheral circulation, indicating that maturation of these parasites could occur, but deep in the human microvasculature (Field 1963). In 2010, we were the first group to demonstrate that *P. vivax*-infected erythrocytes (Pv-iE) collected from patients with malaria were capable to adhere *ex vivo* to placental cryosections, to human lung endothelial cells (HLEC), and to *Saimiri* brain endothelial cells, in static and flow conditions (Carvalho et al. 2010). Although the number of Pv-iE adhered in static conditions was 10 to 15 times lower than that of Pf-iE, the strength of interaction to the endothelium was similar for both parasite species (Carvalho et al. 2010). Likewise, in assays using transfected Chinese hamster ovary cells (CHO), the ligation of Pv-iE to CHO expressing ICAM-1 was 2 times higher than to non-transfected cells or to CHO expressing CD36 (CHO^{CD36}), suggesting ICAM-1 as a potential endothelium receptor for *P. vivax*. Also, Pv-iE were able to adhere to placental cryosections and the ligation to HLEC was inhibited by soluble chondroitin sulphate A (CSA), indicating the participation of this receptor (Carvalho et al. 2010). In a study, performed with 33 Asian-Pacific *P. vivax* isolates, adherence to immobilized CSA and Hyaluronic acid (HA) was shown, and previous incubation of all isolates with CSA and HA reversed parasite ligation, proving CSA involvement as an adhesion receptor in *P. vivax* (Chotivanich et al. 2012). Besides, we also confirmed the low frequency of circulating schizonts in peripheral circulation of patients infected with *P. vivax* (Lopes et al. 2014). Furthermore, we were able to demonstrate that *P. vivax* isolates matured

in vitro until the schizont stage have a higher adhesion capacity relative to the same isolates before maturation (Lopes et al. 2014). Thus, this data suggests that the low circulating schizont proportion seen in patients could be a consequence of parasite sequestration in another place (Lopes et al. 2014), which is related to parasite disappearance from the peripheral blood circulation. This sequestered *P. vivax* biomass is directly associated with poor clinical outcomes and disease severity (Barber et al. 2015).

Rosetting

A very important adhesive phenotype, associated with clinical complications in falciparum malaria is the capacity of Pf-iE rosette formation, characterized by the ligation of an infected erythrocyte to two or more healthy erythrocytes (Carlson et al. 1990). Several studies have investigated an association between the ABO blood type and in the rosetting process suggesting an important role (Rowe et al. 2007, Barragan et al. 2000). In *P. falciparum*, some scientific publications indicate that patients with A or B blood type showed a greater rosette presence than O blood type patients, which points at an enhanced protection against the occurrence of severe malaria outcomes (Rowe et al. 2007, Barragan et al. 2000). Although these phenomena have already been quite well described and studied for *P. falciparum*, little is known about the formation of rosettes in *P. vivax*. The first report about Pv-iE rosettes (Udomsanpetch et al. 1995) was published more than 20 years ago, but until now, few scientific experiments have explored and described the *P. vivax* rosetting phenotype (Udomsanpetch et al. 1995, Chotivanich, Pukrittayakamee, et al. 1998, Chotivanich, Udomsangpetch, et al. 1998, Chotivanich et al. 2012, Russell et al. 2011). Furthermore, in contrast to *P. falciparum*, the relation between *P. vivax* rosetting, disease severity, parasitemia and blood type is unknown (Russell et al. 2011, Udomsanpetch et al. 1995, Chotivanich, Pukrittayakamee, et al. 1998).

More recently, Russell, B. and colleagues have demonstrated that, to accomplish *P. vivax* field isolate enrichment in parasite, it is necessary to use trypsin to disrupt the present rosettes (Russell et al. 2011), suggesting the existence of parasitic ligands involved in the process of *P. vivax* rosette formation. During the last years rosette formation traits have been explored (Zhang et al. 2016, Lee et al. 2014), drawing already some conclusions: (1) both frequency and rate of rosetting in patient samples

is higher in vivax than in falciparum malaria, (2) which occurs as soon as 20 hours after reticulocyte invasion by *P. vivax* asexual and sexual stages (Lee et al. 2014), (3) host ABO blood group, reticulocyte amount and *P. vivax* parasitemia do not significantly correlate with enhanced or diminished *P. vivax* rosetting capacity.

The rosette complex structure is based preferentially on mature erythrocytes (normocytes), where Glycophorin C receptor presence seems to have an important role (Lee et al. 2014). As it is known, Pv-iE have the rheological properties altered, principally in decrease of membrane elasticity, which enables them to avoid splenic clearance (Zhang et al. 2016). According to recent studies on deformability of Pv-iE, rosette-forming iE are distinctly more rigid than their non-rosetting counterparts, with long adhesion to normocytes, which suggests a high contribution of rosettes to the sequestration of schizonts Pv-iE in the host microvasculature and/or spleen (Zhang et al. 2016).

Genetic background and adhesion phenotype

The genetic basis and expression mechanisms of *P. vivax* adhesion phenotypes are almost completely unknown. The subtelomeric variant gene family of *P. vivax*, designated *vir,* was described 16 years ago (del Portillo et al. 2001), being characterized as having a structure and organization of extreme complexity. Initially, around 346 different genes where classified and grouped within this multigenic superfamily (del Portillo et al. 2001), subdivided in 12 distinct subfamilies (A to L), according to their sequence similarity and structure. Subsequently, a bioinformatic analysis study was published (Lopez et al. 2013), where all subtelomeric gene sequences were compared and examined considering their organization and structure characteristics, similar to those of *vir* genes (Lopez et al. 2013). In the end, this analysis revised the total number of *vir* genes to 314 (Lopez et al. 2013). It is estimated that *vir* genes represent 10 to 20% of *P. vivax* coding sequence (del Portillo et al. 2001). This great repertoire is expressed in *P. vivax* clinical isolates, in which different groups of subfamilies of *vir* genes can be transcribed simultaneously by each mature parasite (Fernandez-Becerra et al. 2005). Contrary to *var* genes from *P. falciparum*, which encode around 60 different PfEMP-1 proteins, only some *vir* genes (160 from 346) have motifs similar to those of exported proteins – "*Plasmodium* Export Element/Vacuolar Transport Signal (PEXEL/VTS)". The PEXEL/VTS molecules are responsible for protein export to the erythrocyte surface and cytosol (Marti et al. 2004).

In contrast to PfEMP-1, VIR antigens are not clonally expressed (Fernandez-Becerra et al. 2005) and can be found in the interior of reticulocytes, thus indicating different functions and locations (Fernandez-Becerra et al. 2005). Considering its variant nature and presence in the erythrocyte membrane, the role for VIR antigens in Pv-iE adhesion to endothelial cells was evaluated. Polyclonal antibodies capable to recognize VIR recombinant proteins of two different families (VIRE4 and VIRE5) inhibited *ex vivo* adhesion of Pv-iE to HLEC (Marti et al. 2004). Moreover, supporting our findings (Marti et al. 2004), Bernabeu, M., *et al.* evaluated the role of VIR proteins in *P. vivax* cytoadherence (Bernabeu et al. 2012). In this work, they transfected *P. falciparum* strain 3D7 with several recombined VIR proteins and proved the involvement of VIR antigens in adhesion to ICAM-1 (Bernabeu et al. 2012). Based on the evidences described above, there is no doubt that *P. vivax* has adhesive capacity. Still, it is not clear that this ability makes part of a strategy to avoid destruction of Pv-iE in the spleen (del Portillo et al. 2004), like it has been foreseen to *P. falciparum*, and/or what is the relation between adhesion and the observed pathological clinical complications reported for vivax malaria.

OMICS & MALARIA

P. vivax NGS: reference genome and transcriptome

The use of the reference genome *P. vivax* Salvador-1 (Sal-1) primate adapted strain (published 10 years ago (Carlton, Escalante, et al. 2008)) and the publication of recent WGS directly from *P. vivax* clinical isolates (Dharia et al. 2010, Chan et al. 2012, Menard et al. 2013, Neafsey et al. 2012, Luo, Sullivan, and Carlton 2015) has opened new possibilities for molecular biology studies (Feng et al. 2003, Forrester and Hall 2014, Carlton, Adams, et al. 2008). Importantly, a new enhanced *P. vivax* reference genome P01 was recently published (Auburn et al. 2016), providing an outstanding improvement on the current reference genome (*P. vivax* Sal-1) and its annotation, essential for all future NGS data analysis and revaluation on the existing published data. This new *P. vivax* P01 reference genome covers, in an enhanced way, the parasite subtelomeric and centromeric regions, where a high percentage of *vir* genes are located. As stated by the authors, "an extensive repertoire of over 1200 *Plasmodium* interspersed repeat (*pir*) genes were identified in PvP01 compared to 346 in Salvador-I, suggesting a vital role in parasite survival or development" (Auburn et al.

2016). As previously reported and rephrasing Auburn *et al.*, these variant genes might be involved in important molecular mechanisms for parasite survival and expansion, thus an enhanced reference sequence will help their identification.

However, transcriptome studies aiming to understand the molecular mechanisms underlying several aspects of the pathogenesis of *P. vivax*, are still scarce (Zhu et al. 2016, Westenberger et al. 2010, Bozdech et al. 2008, Boopathi et al. 2014, 2013). The data have provided a strong foundation for understanding *P. vivax* transcription profiles, but each one lacks information. The Bozdech *et al.* microarray experiment does not have a complete genome coverage, where genes not present in the initial *P. vivax* Sal-1 reference genome annotation do not have a transcriptional profile. The Westenberger et al. microarray experiment covers additional *unannotated P. vivax* genome sections, however abundance comparisons during the asexual life cycle are missing. Whole Transcriptome Shotgun Sequencing (WTSS), also known as RNA-seq, has grown significantly as a powerful approach. It is enabling researchers to produce unbiased transcript abundance data on the several stages of *P. vivax* parasites that is not limited to the specific probes used on a microarray chip. RNA-Seq also allows the definition of the boundaries of genes, such that it could uncover novel gene transcripts, alternative splicing events, validate and/or correct current gene models, and predict 5' and 3' untranslated regions. RNA-Seq therefore provides an opportunity to identify strain specific patterns of gene expression associated with parasite virulence and host-pathogen interactions. Moreover, comparative transcriptome analysis between the deadly *P. falciparum* and *P. vivax* would be key to access the biological and clinical differences between this human malaria parasite species (Carlton, Adams, et al. 2008, Mueller et al. 2009), impacting considerably in vaccine and drug design. Methodological achievements on the capacity to isolate, enrich and mature *ex vivo P. vivax* clinical isolates *(Russell et al. 2011, Udomsangpetch R 2001, Auburn et al. 2013)* are opening new avenues for high-throughput transcriptome sequencing analyses. Still, the very low parasitemias (Guerra et al. 2010) and the multi-clonality of the strains observed in a single patient (Orjuela-Sanchez et al. 2013), hamper the development of a reliable methodologies for RNA isolation with an appropriate quantity and purity suitable for WTSS and its latter analysis and data mining.

Transcriptomic approaches to host-pathogen interactions

The clinical manifestations that characterize malaria disease are caused by the sexual blood stage parasites and its interaction with the human host occurring in the (micro) vasculature and adjoining organs, such as spleen (review in (Cunnington, Walther, and Riley 2013, Gazzinelli et al. 2014, Buffet et al. 2011)). Thus, many features of host-pathogen interactions can be explored and evaluated through analyses of whole peripheral blood samples from malaria patients (comprising lymphocytes and erythrocytes), as sources of both pathogen and host cells for application of transcriptomic analyses, either considering one (Daily et al. 2004, Griffiths et al. 2005) or several cell types from the same sample (Yamagishi et al. 2014). These samples, even taking into account the low parasitemia verified in vivax malaria patients, can yield sufficient parasite RNA reads, allowing expression analysis when using the standard RNA-seq depth capacity (Adiconis et al. 2013, Haque et al. 2017, Malone and Oliver 2011, Conesa et al. 2016, Hrdlickova, Toloue, and Tian 2017). However, the lower the parasitemia, the more required is the depth of sequencing needed for meaningful biological results interpretation.

One important factor for the determination of the pathogenesis of many infectious diseases, including malaria, is the pathogen load, which is harder to quantify as a stimulus that elicit a systemic host immune response, especially when the pathogen is spread through multiple tissues (Cunnington 2015). In this scenario, assessment of host-parasite interactions in blood of (vivax) malaria using transcriptomics offers an excellent opportunity for understanding the role systemic host-parasite interactions in general, even if not always straightforward (Westermann et al. 2016, Westermann, Gorski, and Vogel 2012). One drawback comes from the fact that total RNA isolated from the parasite may only encompass a small proportion of the total RNA isolated from a specimen. To tackle this issue, both cell sorting using genetic engineered fluorescent strains, the specific enrichment and capture in parasite transcripts at the time of RNA extraction, and ribosomal RNA (rRNA) depletion to maximize sequencing of mRNA can be used. In brief, successful transcriptomics is highly reliant on the pathogen, the RNA content per pathogen, the sample type, and the pathogen load.

Several contributions have already been made that lead to the identification of host parasite interactions at cellular scale (extensively reviewed in (Lee et al. 2018)), including those related with regulation of invasion-associated effectors and virulence

genes by non-coding RNAs. There is the increased need to apply "dual RNA-seq" to *in vivo* infection experiments. Nevertheless, the present published studies report evidence that the parasite gene expression can vary within the host and can be driven by the host immune response. Particularly, eukaryotic pathogens, such as *Plasmodium*, contain a tightly packed and highly regulated transcriptome that allow their greater change in gene expression in response to their host.

AIMS

MAIN AIM

The major purpose of this research project is to understand the molecular mechanisms behind distinct immunopathogenesis of *P. vivax* adherence (Fig. 1). Specifically, we aim to

(i)   identify and focus our study the role of parasitic ligands, during the process of rosette formation and

(ii)   Pv-iE cytoadhesion to host endothelial receptors, using the potential of transcriptomics approaches.

The efforts of our research group concentrate in better understanding the expression patterns related to such pathogenicity characteristics of *P. vivax* malaria, in particular of adhesion phenotypic capacity, from Amazonian very low parasitemia clinical isolates, and thus go into the mysterious and mostly unknown biology of *P. vivax* apicomplexan parasite.
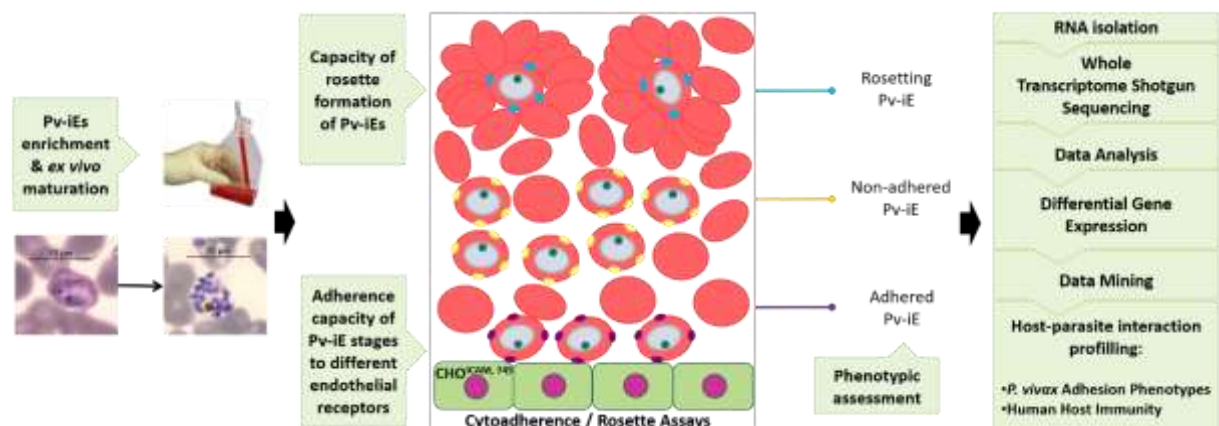
Figure 1. Flowchart showing the main and specific objectives of this study aiming to understand the molecular mechanisms behind the *P. vivax* adherence phenotype.

SPECIFIC AIMS: METHODOLOGICAL OVERVIEW OF THE ACCOMPLISHED RESEARCH

Having in mind this main goal, specific aims were accomplished in a stepwise manner, beginning by collection, characterization and WTSS (RNA-seq) of *P. vivax* samples, isolated from the clinic and enriched in adhesion phenotypes on the road to transcriptomics data analysis, differential gene expression (DE) profiling and data mining (Fig. 2, 3 and 4).

- *Plasmodium* spp. RNA isolation method: standardization and validation

(i) optimization of an RNA extraction method from a panel of *P. falciparum* serial diluted samples to determine limitations, accuracy and sensibility of the best protocols through NanoDrop® and quantitative real -time PCR;

*(ii)*the best methods for RNA extraction tested before were applied to several *P. vivax* clinical samples presenting very low parasitemia and validated towards RNA integrity, purity, quality and quantity suitable for WTSS using both BioAnalyser® and quantitative real -time PCR;

Design Experiment ❶
- Set up the experiment to address your questions
  P. vivax adherence phenotypic characterization

Sample Preparation ❷
- Functional assays and sample characterization and isolation
  P. vivax cytoadhesion and rosetting assays

RNA Isolation ❸
- Isolate and purify input RNA
  P. vivax isolation and total RNA extraction
  RNeasy® Micro kit (QIAGEN®)

Prepare Libraries ❹
- Convert the RNA to cDNA and add sequencing adapters
  P. vivax RNA conversion to cDNA through *SMART®-Seq V4 Ultra Low Input RNA kit* and indexed using the *Low Input Library Prep v2* (Clontech Laboratories, Inc. A Takara Bio Company)

Sequence Run ❺
- Sequence cDNAs using a sequencing platform
  P. vivax libraries were sequenced on Illumina® sequencer Hiseq 2500 in *Rapid Run mode*, with *HiSeq Rapid Cluster Kit v2 (100x100) Paired End kit* and *HiSeq Rapid SBS kit v2 (200 cycles)*

Data Analysis ❻
- Analyze the resulting short-read sequences
  P. vivax raw data analysis using bioinformatic RNA-seq pipeline

Figure 2. Methodological overview of the experimental design (Step ❶) for the execution of *P. vivax* RNA-seq experiment.

- *P. vivax* field work:

(i) effective and reproducible harvest, enrichment and *ex vivo* maturation of *P. vivax* field isolates to perform several different assays and experimental protocols;

(ii) standardization of an unbiased cryopreservation method by N_2 flash freezing of assayed *P. vivax* field isolates for sequencing experiments;

(iii) execution of *ex vivo* 48h culture and time-point preservation of Pv-iE for evaluation of temporal expression profile;

(iv) metadata collection and analysis;

- *P. vivax* adhesion assays:

(i) Accomplishment of several rosetting assays for evaluation of its rosette formation capacity through several maturation stages of different *P. vivax* field isolates;

(ii) optimization of a new protocol for cytoadhesion assessment and corresponding separation and recovery of Pv-iE populations with characteristically distinct ability to adhere to endothelial cells, constitutively expressing different human cell receptors;

(iii) execution of several cytoadhesion assays to isolate, recover and quantitate the Pv-iE populations with different capacities to bind endothelial receptors (CHO[ICAM] and CHO[745]);
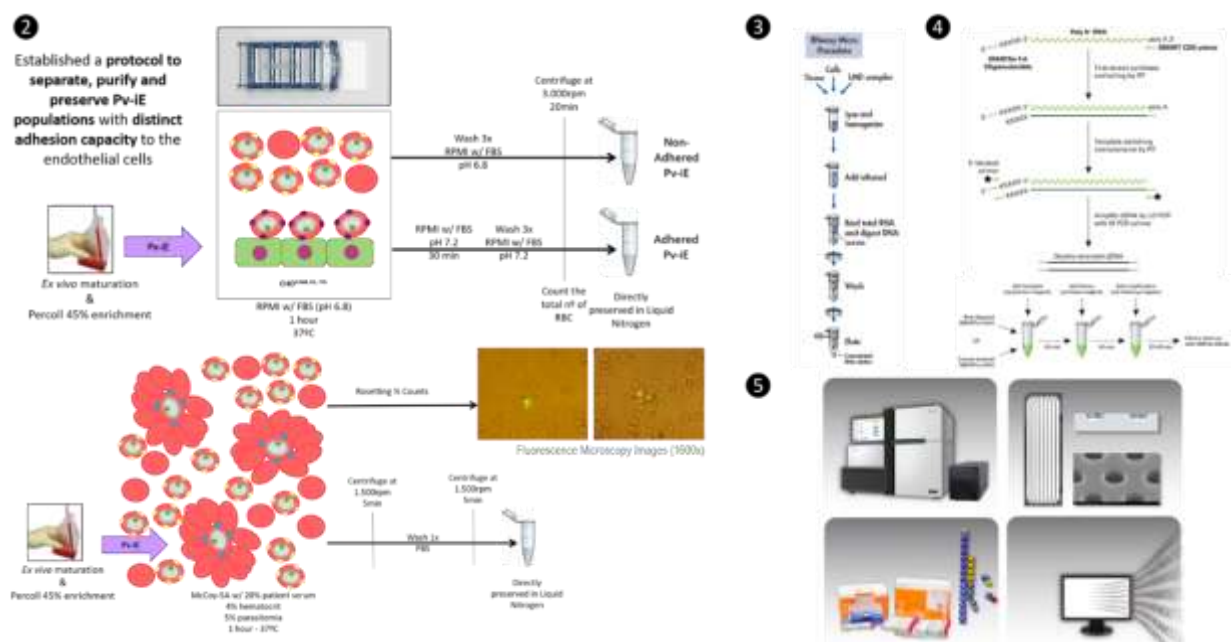
Figure 3. Protocol outlined for the several experimental steps previously designed and optimized for the execution of *P. vivax* RNA-seq experiment. Step ❷ – Sample preparation; Step ❸ – RNA Isolation; Step ❹ – Libraries preparation and Step ❺ – Sequence run.

- Whole Transcriptome Shotgun Sequencing (or RNA-seq):

(i) RNA-seq experimental design (Fig. 2) in accordance with our goal for *P. vivax* whole transcriptome sequencing from phenotypically characterized isolates enriched for parasite populations with cytoadherence or rosette formation capacity (Fig. 2, Step 1) from the malaria endemic region in the Amazons, where patients present a characteristic low parasitemia;

(ii) RNA-seq accomplished experimental steps:

a. Functional assays: isolation and stock of *P. vivax* parasite populations with adherence phenotypes well characterized through cytoadhesion and rosetting assays (Fig. 2 and Fig. 3, Step 2);

b. Isolation and purification of *P. vivax* total RNA (Fig. 2 and Fig. 3, Step 3) from the samples previously assayed;

c. Transcriptomic library preparation, with controlled cDNA synthesis, amplification and indexing for downstream sequencing (Fig. 2 and Fig. 3, Step 4);

d. Sequencing of our libraries using Illumina® platform (Fig. 2 and Fig. 3, Step 5).
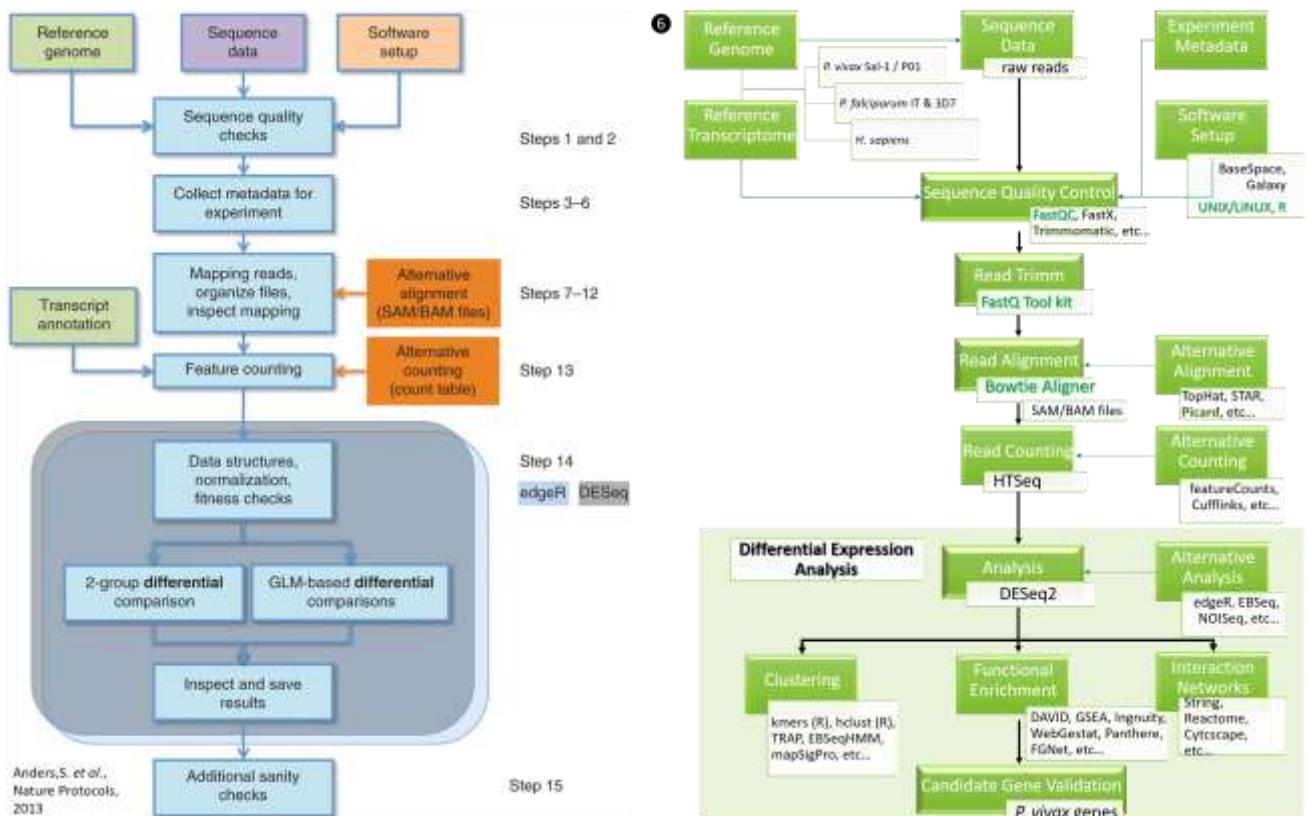
Figure 4. Bioinformatics workflow for RNA-seq data analysis (https://rnaseq.uoregon.edu/#analysis). Left panel shows the current overview for transcriptomics data analysis pipeline (Anders et al. 2013)

and the right panel shows a more detailed bioinformatic methodology overview of *P. vivax* transcriptomics data analysis.

(iii) RNA-seq data analysis (Fig. 2, Fig. 3, Step 6 and Fig. 4): the raw reads obtained were quality controlled and an overall evaluation of the sequencing run. Trimmed and quality accessed reads had the duplicates marked out, were aligned, indexed and mapped to the reference genomes of:

   a. *H. sapiens*, to analyze host immune response;
   b. *P. falciparum* IT and 3D7 for the two control samples;
   c. *P. vivax* reference genomes, Salvador-1 and P01.

(iv) Read count of the *P. vivax* and *H. sapiens* aligned data for RPKM (Reads Per Kilobase Million), FPKM (Fragments Per Kilobase Million) determination,

(v) Differential Gene Expression analysis (Fig. 4) and evaluation of the pool of gene with significant fold changes by sample or group of samples comparison;

(vi) Data mining: exploration of gene ontology enrichment landscapes concerning biological processes, molecular function and cellular components. Functional and metabolic pathways enrichment and interaction network analysis and visualization (Fig. 4).

CHAPTER 1

"*Plasmodium vivax* Biology:

Insights Provided by

Genomics, Transcriptomics and Proteomics"

Overview

In this review, we show the main achievements on *P. vivax* biology accomplished based on the published genome, transcriptome and proteome sequencing projects. In particular,

(1) the genome-wide comparative studies showing evolutionary relationships between parasites of the same genus,

(2) the broad genetic diversity landscape within the *P. vivax* populations reported as to understand specific selection pressures (environmental and host/vector related) acting presently on parasite populations,

(3) the more recent expression profile datasets and regulation mechanisms emerging from sensitive high-throughput WTS (RNA-seq) of *P. vivax* in different stages, and

(4) the attempts to identify parasite metabolic pathways and antigens as possible diagnosis biomarkers through mass spectrometry (MS) based proteomics analysis of parasites and human host profiling from vivax malaria patient samples.

Also, we present the main challenges encountered, remaining gaps and possible research avenues being explored and developed by the vivax malaria community.

CHAPTER 2

"A reliable RNA preparation methodology suitable for Whole Transcriptome Shotgun Sequencing harvested from *Plasmodium vivax*-infected patients with scarce parasitemia"

Preliminary version:

Bourgard C, Lopes SCV, Lacerda MVG, Albrecht L and Costa FTM

Overview

Herein, we present a methodological outline to successfully perform RNA-seq from *P. vivax* field isolates with scarce parasitemia from the low transmission Amazonian endemic region, with distinct cytoadherence capacity. We were able to overcome limitations concerning contaminations from gDNA from the parasite and from the human host by combining enrichment and *ex vivo* maturation known methods with the finest ribonucleic acid isolation procedures to obtain *P. vivax* RNA suitable and proceed for low input library preparation, achieving an unbiased transcriptome sequencing, which allow us insights into the transcriptomic profiling of parasites adhesive phenotypes.

We expect this work to promote further successful high-throughput sequencing application in future *P. vivax* transcriptomic analyses, particularly in malaria endemic regions. Such data will positively impact vivax research, improving knowledge of *P. vivax* biology and paving our way to malaria elimination.

PRELIMINARY VERSION

A reliable RNA preparation methodology suitable for Whole Transcriptome Shotgun Sequencing harvested from a *Plasmodium vivax*-infected patients with scarce parasitemia

Catarina Bourgard[1], Stefanie C V Lopes[2,3], Marcus V G Lacerda[3,2], Letusa Albrecht[4] and Fabio T M Costa[1*]

[1]Laboratory of Tropical Diseases – Prof. Dr. Luiz Jacintho da Silva, Department of Genetics, Evolution and Bioagents, University of Campinas – UNICAMP. Campinas, SP, Brazil.
[2]Instituto Leônidas e Maria Deane, Fundação Oswaldo Cruz – FIOCRUZ, Manaus, AM, Brazil
[3]Fundação de Medicina Tropical Dr. Heitor Vieira Dourado – FMT-HVD, Gerência de Malária, Manaus, AM Brazil.
[4]Instituto Carlos Chagas, Fundação Oswaldo Cruz - FIOCRUZ, Curitiba, PR, Brazil

* Correspondence:

Prof. Dr. Fabio T. M. Costa

University of Campinas – UNICAMP

Department of Genetics, Evolution and Bioagents

Institute of Biology

Laboratory of Tropical Diseases – Prof. Dr. Luiz Jacintho da Silva

13083-864, Campinas, SP, Brazil.

fabiotmc72@gmail.com

Abstract

Plasmodium vivax is one of the most world-threatening human malaria parasites, however knowledge regarding its biology remains elusive. The impossibility of a long-term *in vitro* culture, the tricky host contaminations, and especially, the so characteristic low parasite burden, makes RNA isolation from multi-clonal isolates, a challenging task. So far, obtaining *P. vivax* RNA in quantity and purity for the newest high throughput and sensible Whole Transcriptome Shotgun Sequencing technologies, especially after functional assays, are difficult to perform and furthermore analyze for biological significance.

Herein, we present a methodological outline to successfully perform RNA-seq from *P. vivax* field isolates with scarce parasitemia from the low transmission Amazonian endemic region, with distinct cytoadherence capacity. We were able to overcome limitations concerning contaminations from DNA parasite itself and from the human host by combining enrichment and *ex vivo* maturation known methods with the finest ribonucleic acid isolation procedures to obtain *P. vivax* RNA suitable and proceed for low input library preparation, achieving an unbiased transcriptome sequencing, which allow us insights into the transcriptomic profiling of parasites adhesive phenotypes.

We expect this work to promote further successful high-throughput sequencing application in future *P. vivax* transcriptomic analyses. Such data will positively impact vivax research, improving knowledge of *P. vivax* biology and paving our way to malaria elimination.

Introduction

Plasmodium vivax is the most prevalent malaria parasite outside Sub-Saharan Africa, and the most geographically widespread type of malaria, placing millions of people at risk of infection, therefore imposing major health and economic burdens [1, 2]. Infection occurs in genetically distinct populations with heterogeneous resistance to chloroquine, probably as a result of individual responses in a host-parasite relation [3-6]. Severe clinical complications, although scarce[7], have been of great concern [7]. Nevertheless, the lack of a reliable and reproducible *in vitro* system for long term *P. vivax* culture [8], restricts the study of its biology to endemic referral hospitals.

P. vivax Salvador-1 (Sal-1) primate adapted strain, sequenced 10 years ago [9], being currently used as reference genome, has opened new possibilities for molecular biology studies [10, 11]. Transcriptome studies aiming to understand the molecular mechanisms underlying several aspects of the pathogenesis of *P. vivax*, e.g. drug resistance, disease severity, the understanding of dormant liver-stage parasites (hypnozoites) are only now starting to be published by several research groups (reviewed in [12]). Most recently, publication of a new reference genome *P. vivax* P01 with improved scaffold assembly, specially of subtelomeric, telomeric and centromeric regions [13], will lead to a significant alignment, mapping and analysis into meaningful insights into *P. vivax* biology of the genomic and transcriptomic data. Moreover, comparative transcriptome analysis between the deadly *P. falciparum* and *P. vivax* would be key to access the biological and clinical differences between this human malaria parasite species [9], impacting considerably in drug and vaccine design. The capacity to isolate, enrich and mature *ex vivo P. vivax* clinical isolates [14-16] open new avenues for high-throughput transcriptome sequencing analyses [17-19]. However, very low parasitemias [1] verified in several vivax malaria endemic regions and the infections multi-clonality observed in a single patient [20], hindered further progress into *P. vivax* transcriptomics, in particular, the application of Whole Transcriptome Shotgun Sequencing (WTSS), also known as RNA-seq, as a powerful approach to identify strain specific patterns of gene expression associated with parasite virulence and host-pathogen interactions.

In the 60's, *a* "hidden" capacity of *P. vivax* was proposed, in that a disproportionally low level of mature blood stages (schizonts) relative to young stages (trophozoites) was found in peripheral circulation, indicating that maturation of these parasites could

occur, but deep in the human microvasculature [21]. In 2010, we demonstrate that *P. vivax* infected erythrocytes (Pv-iE) collected from patients with malaria were capable to adhere *ex vivo* to placental cryosections, to human lung endothelial cells (HLEC), and to *Saimiri* brain endothelial cells, in static and flow conditions [22]. Although the number of Pv-iE adhered in static conditions was 10 to 15 times lower than that of Pf-iE, the strength of interaction to the endothelium was similar for both parasite species [22]. Likewise, in assays using transfected Chinese hamster ovary cells (CHO), the ligation of Pv-iE to CHO expressing ICAM-1 was 2 times higher than to non-transfected cells or to CHO expressing CD36 (CHO^{CD36}), suggesting ICAM-1 as a potential endothelium receptor for *P. vivax*. Also, Pv-iE were able to adhere to placental cryosections and the ligation to HLEC was inhibited by soluble chondroitin sulphate A (CSA), indicating the participation of this receptor [22]. In a study, performed with 33 Asian-Pacific *P. vivax* isolates, adherence to immobilized CSA and Hyaluronic acid (HA) was shown, and previous incubation of all isolates with CSA and HA reversed parasite ligation, proving CSA involvement as an adhesion receptor in *P. vivax* [23]. Besides that, we also confirmed the low frequency of circulating schizonts in peripheral circulation of patients infected with *P. vivax* [24]. Furthermore, we were able to demonstrate that *P. vivax* isolates matured *in vitro* until the schizont stage have a higher adhesion capacity relative to the same isolates before maturation [24]. Thus, this data suggests that the low circulating schizont proportion seen in patients could be a consequence of parasite sequestration in another place [24], which is related to parasite disappearance from the peripheral blood circulation. This sequestered *P. vivax* biomass is directly associated with poor clinical outcomes and disease severity [25].

Taking in mind these difficulties, we developed a methodologic framework which would allow us successfully to achieve all steps from RNA isolation of a set of lower parasitemia *P. vivax* isolates from the Amazonian endemic region to an unbiased coverage WTSS, while studying the particular cytoadhesion phenotypic traits, further understanding the immuno-pathogenesis of this elusive parasite. Our expectation is that this work will help vivax malaria researchers to overcome some barriers to molecular biology in *P. vivax*. For instance, by promoting the generation of expression data from the world parasite populations, as well as, foster and improve the success rate of WTSS on clinical isolates in malaria endemic areas, consequently, instigate the investigations on the biology of *P. vivax* apicomplexan parasite.

Results

RNA isolation methods and quality control assessment

In the context of gene expression studies using transcriptome data, we tested several methods for RNA isolation in a panel of *P. falciparum* samples with a wide-ranging parasite density (10^2 to 10^7), which intended to simulate the conditions of those found in field isolates of *P. vivax*. Eight RNA isolation approaches were tested, covering several conditions: from the well-known single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction and a TRIzol based RNA extraction protocol, to the user-friendly kits based on the acid guanidinium thiocyanate-phenol-chloroform or TRIzol extraction, available on the market. To understand the transport options available for the isolates during the fieldwork, the influence on RNA of an initial sample preservation, frozen in a commercial stabilizer reagent or TRIzol was also evaluated (Figure 1). After completion of each RNA extraction, the sizing, quantity and quality of samples was estimated by using the known electrophoretic technology present in the BioAnalyser® platform (Supplemental Figure 1).

Next, to identify and evaluate the limitations of downstream reactions applied to this set of isolated RNAs, we amplified three housekeeping genes by qPCR (Figure 2). Between the three chosen housekeeping genes, average Ct for the amplification of the *DNA repair helicase* gene was higher than the *Gamma GCS* and *SER-tRNA ligase* (Figure 2). Generally, high Ct mean values indicate less efficiency of amplification reactions for the target fragment, which difficult the analysis of amplification in case of poorer amount RNA samples. Therefore, we looked more carefully to the amplifications observed for the other two genes (Figure 2A and C). The detection of contamination in the samples with gDNA from the *P. falciparum* itself, originated from an inefficient separation of RNA from DNA during the extraction procedure, was possible through the amplification of reverse transcriptase negative (RT-) controls. As expected, it was observed an inversely proportional relation between the initial quantities of parasites from which the RNA was extracted relatively to the mean Ct value obtained, consistent through all extraction methods. To evaluate which methods was the best candidate for RNA isolation in very small amounts, we discarded some cases where amplification was not successful, or we could detect some gDNA contamination. For instance, Ct values very high and/or coincident between the samples and its correspondent RT-controls where discarded as not successfully amplified samples. Contamination cases

where observed when RT- controls presented Ct values lower or similar to those obtained for gDNA amplification. We could verify that in general, we were able to extract small amounts of RNA more efficiently by applying the user-friendly columns base kits, instead of performing other large scale, manual protocols. RNA samples extracted using the commercial kits from biological material previous preserved, showed more reproducibility between replicates (see error bars). This is an advantage considering the future application of this extraction protocol to *P. vivax* field isolates that need to be preserved for transportation before RNA extraction.

Between all extraction procedures performed, the two most reliable combinations of preservation and extraction procedures, RNAlater®/RNeasy® Micro kit and TRI Reagent®/Direct-zol™ RNA MiniPrep, were chosen to extract RNA from *P. vivax* field isolates (Figure 3). After isolation, enrichment and short-term *ex vivo* maturation (Supplemental Figure 3), RNA from a set of samples with different number of parasites (10^3 to 10^6) was extracted. BioAnalyzer® quality controls are shown in Supplemental Figure 2.

The RNA extracted from these filed isolates was then amplified by qPCR for the same three housekeeping genes (Figure 4). We observed that, although the results were similar for both kits, Ct values were higher when the sample was preserved in TRI Reagent® and then extracted using the Direct-zol™ RNA MiniPrep. This may be a consequence of contamination, with organic compounds, common in protocols that use TRIzol and other phenols to separate the RNA from DNA, which inhibit enzymatic reactions, and thus, waning the efficacy of reverse transcription and amplification reactions. Like for *P. falciparum* samples, in *P. vivax* samples the gDNA contamination from the parasite was limited and, most importantly, host contamination was reduced and did not interfere in the qPCR reactions (Figure 4).

In the end, we determined that the RNeasy® Micro kit would be that most reliable to work with. However, instead of preserving the isolates in RNAlater® to transport the samples from one lab to another to continue with the sequencing set up, we opted for flash freezing in liquid nitrogen all our *P. vivax* isolates, prior to transport in controlled temperature. This decision considered two very important factors which could directly interfere greatly in our sequencing outputs. The use of preserving substances, like RNAlater®, is controversial. Not only we don't have complete and exact information about the content of such patented solution, as also, some researchers have alerted for incompatibilities verified when reading samples in spectrophotometry equipment for

RNA control assessment, as for instance, on the golden standard BioAnalyser®. We did saw some inconsistency of RNA Integrity Number (RIN) values, both in *P. falciparum* and *P. vivax* RNAs (Supplemental Figures 1 and 2). Given our samples have little RNA and the algorithm-based RIN values are calculated based on three features, quantity, quality and purity, of the samples, we couldn´t clearly sort out the source. Furthermore, RNA-seq is a sensible technique, that anything could affect with the final readouts, including the presence of traces of RNAlater®. The presence of such compounds could interfere and hinder our analysis of *P. vivax* expression patterns.

Evaluation of Brazilian P. vivax *isolates adherence phenotype*

To investigate the cytoadhesion capacity of different populations of Pv-iE from Brazilian Amazonian endemic field, we collected a total of 8 vivax malaria patient samples (Supplementary Table 1). The 3 female and 5 male patients have on average 29.8 years old and presented 500 to 10^4 Pv-iE per µL, counted on a thick smear by the specially trained microscopists from malaria diagnosis service at FMT-HVD. All samples were processed immediately after collection following the procedures for parasite isolation and enrichment to obtain parasitemias >50%, and thus, a total number of Pv-iE greater than 400,000 to enable us to proceed with cytoadhesion assays. Thin blood smears after Percoll 45% enrichment allowed us to choose isolates with a higher percentage of trophozoite to schizont staged parasites and control for host lymphocyte contamination. Each single Pv-iE isolate was assayed in duplicate against CHO^{ICAM} and CHO^{745} cell lines in one go. The complete count and estimations of the total number of Pv-iE adhered/non-adhered populations for each cytoadhesion assay are shown in Supplemental Table 2. For 4 of those isolates (92U15, 93U15, 101U15 and 105U15), 4 different adherent/non-adherent populations of Pv-iE were collected for later transcriptome sequencing. The comparison between all cytoadhesion assays against CHO^{ICAM} and CHO^{745} cell lines reveled a non-significant difference (paired t test, 2-tailed p-value=0.0595) on Pv-iE adherence capacity (Figure 5 A), even if Pv-iEs seem to adhere more to the CHO^{745} cell line, a confounding observation given our data shows a big variability between different isolates. When evaluating isolates for which we performed *P. vivax* transcriptomics (paired t test, 2-tailed p-value=0.1103) (Figure 5 B), there is no significantly difference between adherence to the different CHO cell lines.

P. vivax *low input cDNA synthesis, library preparation and sequencing*

In order to perform RNA-Seq, we followed a sequencing experimental design that we considered the most promising option, considering the particularities of *P. vivax* samples. RNA extractions of Pv-iE populations of adhered and non-adhered to CHO[745]/CHO[ICAM] for 4 different isolates were done using the RNeasy® Micro kit and its quantity and quality evaluated using the Bioanalyzer® platform (Supplementary Table 3). On average 21.201 pg/uL of RNA, ranging from 6.655 to 40.524 pg/uL, was obtained with an average 8.0 RIN (5.7-9.8). Given the low amounts of *P. vivax* RNA, we opted for the use of SMART® technology, which offers unparalleled sensitivity, unbiased amplification of cDNA transcripts from low input RNA samples (Supplemental Table 4), which is a tremendous advantage since the huge limitation imposed by low parasite burden in vivax malaria patients. Immediately after, the cDNA output was converted into sequencing templates suitable for cluster generation and high-throughput sequencing resulting into a sequencing-ready library for the Illumina® platform (Supplemental Table 5).

Whole Transcriptome Shotgun Sequencing data analysis

We obtained a total number of 7,506,734 raw reads. On average 595,231 paired end reads (100 bp) (Supplemental Figure 6) per sample from the 13 sample libraries were successfully sequenced (Supplemental Table 6), with an average 47.2% GC content. FastQC of the total number of raw reads obtained from all our libraries reveled good sequence quality and the necessary trimming steps only excluded a minor fraction of reads (Supplementary Table 6), in general, repetitive, not accurately determined and Illumina® adaptors run through sequences. Using *P. vivax* P01 reference genome, we were able to align and map on average half of the total number of trimmed reads obtained to annotated protein-coding genes (Table 2). Note that each sample is subdivided by 4 different assays, which reduce by a quarter the total number of reads available for downstream analysis. Sequences showing multiple or discordant alignments were excluded from the analysis.

Differential expression profiles associated to P. vivax *rosetting phenotype*

To proceed for RNA-seq of *P. vivax* iRBCs, we separate four groups of parasite population samples, with adhesion capacity to CHO[745] (92U15-2, 93U15-21, 105U15-

6 and 101U15-23) and CHO^{ICAM} (92U15-1, 105U15-5), and the correspondent non-adhered samples to CHO^{745} (92U15-4, 93U15-22,105U15-8 and 101U15-24) and CHO^{ICAM} (92U15-3, 105U15-7 and 101U15-20). Through RNA-seq data analysis, we accessed the differential gene expression profiles between samples of these two groups and dissect by data mining, possible differences that tentatively might explain *P. vivax* adhesion phenotype during the progress of vivax malaria disease. Although expression profiles were similar between samples within the same capacity level for rosette formation, our *P. vivax* populations isolated from malaria patients showed some degree of transcriptome heterogeneity. In any of the different groups analyzed we were able to see significantly expressed genes, when considering data normalization (q-values). However, a list of genes with significant p-values could be withdrawn and looked more carefully (Table 3). Especially considering the comparison group of parasites that adhered to CHO^{ICAM} versus CHO^{745}, we observed a total of 52 genes. A portion of those genes (7) codify conserved *Plasmodium* spp. proteins of yet unknown function. These genes showed a strong (2.3>log2 Fold Change>3.4) absolute expression, which might suggest their involvement in molecular processes important for adhesion phenotype. Future functional characterization of these proteins should clarify this possibility. From the pool of the other genes, 5 genes has caught our attention: one PIR protein (PVP01_0950000), one *Plasmodium* exported protein of unknown function (PVP01_1201600), the down-expression in adhered samples of the putative translocation protein SEC62 (PVP01_1268900) and two proteins with repetitive domains, the heptatricopeptide repeat-containing protein (PVP01_1416100) and the putative WD repeat-containing protein (PVP01_0905300). Further investigations are needed in order to understand the biological meaning of this results.

Discussion

P. vivax is one of the most widespread and world-threatening human malaria parasites, however researchers know little of its biology, principally in respect of the characteristics that distinguish it from the more known and studied closely related *P. falciparum*. Several factors have hampered *P. vivax* research, in special, the fact that it is still not possible to long-term culture it. Therefore, experiments involving *P. vivax* are restricted in time and place. Adding up, *P. vivax* infections characterize themselves by a log order lower parasitaemia, compared with those of *P. falciparum*. With low parasites available, very low amounts of ribonucleic acids are obtained during the

extraction process and substantial DNA/RNA contamination from the host, mainly from lymphocytes, is observed, which confounds data analysis and interpretation. Also important in such samples with reduced amount of RNA is DNA contamination from the parasite itself, which could lead to inaccurate gene expression level assessment. In the end, these technical hitches together with mapping short-read sequence data and attaining even coverage through the genome make sequencing *Plasmodium* parasites using high-throughput shotgun methodologies highly challenging.

In an attempt to overcome these difficulties, we tested several methods for sample preservation combined with RNA isolation in a panel of *P. falciparum* samples that could mimic some of the most important conditions found when dealing with *P. vivax* field isolates. The low parasitemias verified in peripheral blood samples drawn from the patients were simulated by the set of variable parasites density samples. Also, the different parasite amounts helped us to identify and draw with some interval of confidence, the maximum and minimum limits of RNA isolation (Figure 3) and successful downstream molecular reactions (Figure 4). By using qPCR, it was possible to amplify three different *Plasmodium spp.* constitutively expressed genes of RNA extracted from an initial very small number of Pv-iE, such as 10^3. Although present in some samples, *P. falciparum* or *P. vivax* genomic DNA amplification was reduced, could be easily distinguishable and did not interfered with downstream analysis. While lymphocytes are completely absent from *P. falciparum in vitro* culture, CF11 filtration efficiently removes the lymphocytes from the sample, as previously tested by others [16], which are reflected in our reduced amplification of our control host gene (Figure 6), not a limiting factor in qPCR reactions.

We were able to overcome the limitation of DNA contamination from the parasite itself and host DNA/RNA contamination, while successfully amplifying RNA isolated from scarce parasitemia multi-clonal single patient field isolates. Furthermore, there's no need for additional mRNA selection that could deplete even more our samples in RNA. Within cDNA synthesis, the presence of other RNA species, especially rRNA, act as an important source of mRNA sequencing corruption. So, for an RNA input, which was our case, not needing to execute RNA depletion techniques, is an important advantage when dealing with lower quantities of RNA. Also, ensuring that the final cDNA libraries contain the 5' end of the mRNA and maintain a true representation of the original mRNA transcripts, were factors considered, that are critical for transcriptome sequencing and gene expression analysis.

Conclusion

We hope that this methodology will help researchers to overcome the barriers to molecular biology in *P. vivax* by improving the success rate of WTSS on clinical isolates in malaria endemic areas and additionally opens further the investigations on the biology of *P. vivax* apicomplexan parasite.

Methods

Ethical Approval

Informed consent was sought and granted from all patients attending the *Fundação de Medicina Tropical Dr. Heitor Vieira Dourado* (FMT-HVD), Manaus, Amazonas State, Brazil. All procedures, including protocols and consent forms, were approved by the Ethics Review Board of FMT-HVD (processes CAAE-0044.0.114.000-11 and 54234216.0000.0005).

Study Area, Subjects and Sample Collection

Patient recruitment was implemented at FMT-FVD, a tertiary care center for infectious diseases in Manaus, Amazonas State, Brazil. Only adult patients were recruited. Exclusion criteria comprised severe malaria, patients under anti-malarial treatment, with *P. falciparum* malaria and/or *P. falciparum* and *P. vivax* mixed infections and pregnant women. After microscopic diagnosis of uncomplicated *P. vivax* malaria, determination of parasitaemia and before treatment was initiated, up to 8 mL of peripheral blood was collected in citrate-coated Vacutainer™ tubes (Becton-Dickinson). Complete information on patient profiles for *P. vivax* clinical isolates used in this work is on Supplementary Table 1. Then, patients were treated with Chloroquine and Primaquine according to the standard protocol recommended by the Brazilian Malaria Control Program. Subsequently, *P. vivax* mono-infection was confirmed by PCR analysis, as described elsewhere [26].

Parasite Isolation, Enrichment and ex vivo Maturation

All samples were immediately processed to obtain enriched *P. vivax* infected erythrocytes (Pv-iEs). First, the plasma and buffy coat layer was removed after separation by centrifugation at 400 x g for 5 min at room temperature. Pellet was ressuspended in an equal volume of RPMI parasite medium and then performed CF11

column filtration (Sigma®) to deplete white blood cells (WBC) [14, 16, 27]. After, completed parasite enrichment was achieved by Percoll 45% gradient as previously described [22]. Thin blood smears were prepared and stained with *Panótico Rápido* (Laborclin®) kit, before, during and after *ex vivo* short culture to control the extent of parasite maturation.

Chinese hamster ovary cells in vitro *culture*

Two different lines of Chinese hamster ovary (CHO) cells, the CHO[745] cell line depleted for known glycosaminoglycans [28-30] and CHO[ICAM] transfected cell line, expressing constitutively the with human intercellular adhesion molecule-1 (ICAM-1) [31, 32] were cultured following previous standardized methods using RPMI 1x (Sigma® #R4130) pH 6.8 with 10% Fetal Bovine Serum (FBS) and 40 mg/L Gentamycin media) and were periodically controlled for its receptor expression by flow cytometry (Gallios flow cytometer from Beckman-Coulter). We used primary antibody mouse anti-human CD54 (ICAM) (BD Cat. No. 559047) conjugate with chicken anti-mouse IgG (H+L) Alexa Fluor®488 (Cat. No. A21200) to monitor ICAM-1 receptor expression on our CHO[ICAM] cultures. A set of mouse anti-human primary antibodies to CD54 (ICAM), CD106 (VCAM) (BD Cat. No. 555645), CD36 (BD Cat. No. 555453), CD201 (EPCR) (BioLegend Cat. No 351902) and Chondroitin Sulfate (CSA) (BD Cat. No. 554275) were used to confirm the CHO[745] cell line status as depleted for these receptors. Mycoplasma PCR detection were routinely performed to exclude it as a contaminant of our cell lines.

Pv-iE cytoadhesion assays

After parasite enrichment, the number of Pv-iE per mL was determined using a Neubauer chamber before the assay (Supplemental Figure 3). We used 4-wells Lab-Tek previously prepared with a lane of confluent (>80%) and adhered ($\geq 1,0 \times 10^5$) CHO cells, 2-wells for each CHO[ICAM] and CHO[745]. Pv-iE ($\geq 1,0 \times 10^5$) in RPMI 1x (pH 6.8) with FBS medium were added to each Lab-Tek well and incubated for one hour at 37ºC. The Pv-iE that do not adhere to the CHO cells were carefully washed (3x) using the same media, collected by centrifugation at 3000 rpm, 20 min at room temperature and washed once with PBS 1x to clean from medium residues. Only on the duplicate well for each different CHO cell line under assay, we proceed with detachment of Pv-iE through 30 min incubation with RPMI 1x (pH 7.2) with FBS,

followed by 3x washes in 10 min intervals with the same media and making sure that the lane of CHO cells remained intact. Non-adhered Pv-iE were collected, washed with PBS 1x and pelleted down. All the pellets of adhered and non-adhered Pv-iE were immediately flash frozen in N$_2$. Quantification of the number of adhered Pv-iE relative to the initial total number of Pv-iE used in each assay was done by microscopy, after remove the Lab-Tek well structure from the glass slide and staining with *Panótico Rápido* kit.

P. falciparum *in vitro cultures*

P. falciparum FCR3 strain [33] was cultured accordingly to standard procedures previously described [34].

RNA Extraction and Quality Control

For RNA extraction of a set of *P. falciparum* culture dilutions (10^2 to 10^7), two different manual protocols and three different commercial kits were used (Figure 1). The *single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction* [35, 36] was directly used upon sample assembly, without any prior conservation by freezing or by adding a preservation reagent. This protocol was followed exactly as explained by the authors [35, 36], with some scale adaptations for the lower density parasite samples. The *reliable RNA preparation for Plasmodium falciparum* [34] was performed as described, also with some scale adaptations concerning the samples with the smallest amount of parasites and including an initial step of sample preservation in TRIzol at -80ºC, for no longer than 48 hours. The three commercial kits RNeasy® Micro (Qiagen), RNeasy® Micro Plus (Qiagen) and Direct-zol^{TM} RNA MiniPrep (Zymo Research) were used for purification of total RNA as per the manufacturer's protocol. For three sets of *P. falciparum* culture dilutions, the samples were preserved in RNAlater® RNA stabilization reagent (Qiagen) or TRI Reagent® (Zymo Research) before RNA extraction with each kit respectively. The total number of Pv-iEs from field isolates was determined before preservation in RNAlater® RNA stabilization reagent or TRI Reagent® at -80 ºC.RNA extraction was executed using the RNeasy® Micro kit (Qiagen) or Direct-zol^{TM} RNA MiniPrep (Zymo Research) (Figure 2). All RNA samples were carefully preserved precipitated in a solution of one-tenth volume of RNase-free 3M sodium acetate pH 5.2, 2.5 volumes of absolute ethanol and one volume of isopropanol, overnight at -20ºC and then stocked at -80ºC.

For RNA-seq, after Pv-iE quantification, the isolate was preserved by flash freezing the sample in liquid nitrogen. For RNA extraction, the RNeasy® Micro kit (Qiagen) was used according to the manufacturer instructions. Quality control was first accessed by running electrophoretically the extracted RNA samples in the Agilent 2100 Bioanalyzer instrument, using the Agilent RNA 6000 Pico Kit reagents and chips and analyzed on the 2100 Expert software, according to the ©Agilent Technologies recommendations.

cDNA synthesis and Quantitative Real Time PCR

The cDNA synthesis was carried out using 13.2µL of RNA sample for a total reaction volume of 20µL, using the High Capacity cDNA Reverse Transcription with RNase Inhibitor kit (Applied Biosystems®), with 10x RT buffer, 10x Random Primers, 25x (100mM) dNTP Mix, RNase Inhibitor and MultiScribe Reverse Transcriptase (5U.µL^{-1}), according to the manufacturer´s indications. A reverse transcriptase negative control (RT-) without the enzyme was performed for all different cDNA synthesis reactions. The cDNA synthesized from *P. falciparum* and *P. vivax* RNA extracted samples was used as template for the amplification of *seryl tRNA synthetase* (SER-tRNA S; PFIT_0715800 and PVX_1234480), *gamma-glutamylcysteine synthetase* (Gamma GCS; PFIT_0919000 and PVX_099360) and *DNA repair helicase* (DNA RH; PFIT_1409400 and PVX086025), three different *Plasmodium spp.* constitutively expressed genes. The following specific pairs of primers were used: Pf SER-tRNA S-Fw: GAGGAATTTTACGTGTTCATCAA; Pf SER-tRNA S-Rv: GATTACTTGTAGGAAAGAATCCTTC; Pv SER-tRNA S-Fw: AGGGATTGCTACGTGAGCACATT; Pv SER-tRNA S-Rv: GTTGCTGACTAGGTAGCCAGGCTTC; Pf Gamma GCS-Fw: TGCGAATATGGATGATGAAGG; Pf Gamma GCS-Rv: TAAGAGCAAGGAAAAGTGGT; Pv Gamma GCS-Fw: CAGCGACCTGGACGACGAGAA; Pv Gamma GCS-Rv: TTAGGGCTAAGAACAAAGGG; Pf DNA RH-Fw: GCCCTTTCTATGCTACGAGA; Pf DNA RH-Rv: TTTTCTAGTATGGTTAATGTAGCT; Pv DNA RH-Fw: GCCCCTTCTACGCCACGAGG; Pv DNA RH-Rv: TTGTCTAGCACAGTTAGTGTAGCT. To detect contamination in the *P. vivax* extracted RNA samples by human RNA and DNA, we also amplified the Toll-like receptor 9 (TLR9) gene with the specific primer pair, TLR9-F: ACGTTGGATGCAAAGGGCTGGCTGTTGTAG and TLR9-R:

ACGTTGGATGTCTACCACGAGCACTCATTC [16]. Power SYBR® Green PCR Master Mix (Applied Biosystems®) reagents were used to perform all qRT-PCR reactions in quadruplicate and run on the Applied Biosystems® 7500 Real-Time PCR System, following the company's RT-PCR guidelines. In brief, individual reactions were set-up using 1μL of cDNA previously synthetized, 2x Power SYBR® Green PCR Master Mix, 500nM each primer and RNase/DNase-free water to a final volume of 10μL. The run protocol comprised the following cycling conditions: initial holding stage of 50ºC/20 sec and 95ºC/10 min; 45x cycling stages of 95ºC/15 sec, 57ºC for *P. falciparum* primers pairs or 60ºC for *P. vivax* primers pairs/30 sec and 60ºC/1 min; and a 2x melting curve stage of 95ºC/15 sec and 60ºC/1 min. Applied Biosystems® 7500 Software v.2.0.6 was used for data analysis.

Low Input cDNA synthesis and Library Preparation for Whole Transcriptome Shotgun Sequencing

We used SMART-Seq V4 Ultra Low Input RNA kit for sequencing by the Clontech's patented SMART® (Switching Mechanism at 5' End of RNA Template) technology. cDNA quality, quantity and size range was evaluated through BioAnalyser platform from ©Agilent Technologies, Inc., using the Agilent High Sensitivity DNA Kit (cDNA, 5 to 500 pg/μL within a size range of 50 to 7000 bp), as per manufacturer instructions. Prior to generating the final library for Illumina® sequencing, the Covaris AFA system was used for controlled cDNA shearing, resulting in DNA fragments between 200 and 500 bp sizes. Instructions were followed as indicated in the SMART-Seq V4 Ultra Low Input RNA kit for sequencing user manual by Clontech Laboratories, Inc. A Takara Bio Company. cDNA output was then converted into sequencing templates suitable for cluster generation and high-throughput sequencing through the Low Input Library Prep v2 (Clontech Laboratories, Inc. A Takara Bio Company). Library quantification procedures using the Library Quantification kit (Clontech Laboratories, Inc. A Takara Bio Company) by the golden standard qPCR and Agilent's High Sensitivity DNA kit (Agilent Technologies, Inc.) were successfully completed before proceeding for the pool set-up (2 different pools of 12 samples differently indexed) at a final concentration of 2nM for direct sequencing. The library was sequenced on HiSeq 2500 sequencer on Rapid Run mode with the HiSeq Rapid Cluster Kit v2 (100x100) Paired End and HiSeq Rapid SBS Kit v2 (200 cycles) kit. The generated libraries were cluster amplified and sequenced on the Illumina platform using standard Illumina® reagents and

protocols for multiplexed libraries, by following their loading recommendations. The sequencing runs were performed on HiSeq 2500 sequencer on Rapid Run mode with the HiSeq Rapid Cluster Kit v2 (100x100) Paired End, HiSeq Rapid SBS Kit v2 (200 cycles) and HiSeq® Rapid Duo cBot v2 Sample Loading kits from Illumina®, Inc..

Transcriptomic data analysis

We used the EuPathDB-Galaxy free, interactive, web-based platform for our large-scale data analysis (https://eupathdb.globusgenomics.org/) [37]. The RNA-seq raw reads were checked for quality by running Fast Quality Control (FastQC - https://www.bioinformatics.babraham.ac.uk/projectY/fastqc/). The reads were then subjected to trimming using the Trimmomatic [38] Galaxy tool (v. 0.36.5; http://www.usadellab.org/cms/index.php?page=trimmomatic) and aligned using TopHat2 [39] (Galaxy Tool Version SAMTOOLS: 1.2; BOWTIE2: 2.1.0; TOPHAT2: 2.0.14), towards the *P. vivax* P01 PlasmoDB release 38 and the Homo sapiens UCSC hg38 (RefSeq & Gencode gene annotations embedded on HostDB release 29)[40] reference genomes. FPKM estimation of reference genes with htseq-count [41] (Galaxy Tool v. HTSEQ: default; SAMTOOLS: 1.2; PICARD: 1.134) allowed the final analysis of differentially expressed genes through DESeq2 [42] Galaxy Tool (v. 2.1.6.0). Differential gene expression between the different analysis groups was identified after a pairwise Wilcoxon test was used to compare the transcriptional profiles with the following cutoffs: p-value<0.05, q-value<0.5 and a log2 fold change > 1.5.

Acknowledgments

Author contributions statement

All authors conceived the experiments. CB did the field sample collection and cytoadhesion assays, designed and executed the RNA-seq experiments and data analysis. CB and LA went through data mining of the obtained results and draft the manuscript. All authors reviewed the manuscript.

References

1. Guerra, C.A., et al., *The international limits and population at risk of Plasmodium vivax transmission in 2009.* PLoS Negl Trop Dis, 2010. **4**(8): p. e774.
2. WHO, *World Malaria Report 2018.*, in *World Malar. Rep.* 2018: Geneva.
3. Baird, J.K., *Resistance to therapies for infection by Plasmodium vivax.* Clin Microbiol Rev, 2009. **22**(3): p. 508-34.
4. Price, R.N., N.M. Douglas, and N.M. Anstey, *New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance.* Curr Opin Infect Dis, 2009. **22**(5): p. 430-5.
5. Tjitra, E., et al., *Multidrug-resistant Plasmodium vivax associated with severe and fatal malaria: a prospective study in Papua, Indonesia.* PLoS Med, 2008. **5**(6): p. e128.
6. Mourao, L.C., et al., *Naturally acquired antibodies to Plasmodium vivax blood-stage vaccine candidates (PvMSP-1(1)(9) and PvMSP-3alpha(3)(5)(9)(-)(7)(9)(8) and their relationship with hematological features in malaria patients from the Brazilian Amazon.* Microbes Infect, 2012. **14**(9): p. 730-9.
7. Lacerda, M.V., et al., *Understanding the clinical spectrum of complicated Plasmodium vivax malaria: a systematic review on the contributions of the Brazilian literature.* Malar J, 2012. **11**: p. 12.
8. Mueller, I., et al., *Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite.* Lancet Infect Dis, 2009. **9**(9): p. 555-66.
9. Carlton, J.M., et al., *Comparative genomics of the neglected human malaria parasite Plasmodium vivax.* Nature, 2008. **455**(7214): p. 757-63.
10. Feng, X., et al., *Single-nucleotide polymorphisms and genome diversity in Plasmodium vivax.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8502-7.
11. Forrester, S.J. and N. Hall, *The revolution of whole genome sequencing to study parasites.* Mol Biochem Parasitol, 2014. **195**(2): p. 77-81.
12. Bourgard, C., et al., *Plasmodium vivax Biology: Insights Provided by Genomics, Transcriptomics and Proteomics.* Front Cell Infect Microbiol, 2018. **8**: p. 34.
13. Auburn, S., et al., *A new Plasmodium vivax reference sequence with improved assembly of the subtelomeres reveals an abundance of pir genes.* Wellcome Open Res, 2016. **1**: p. 4.
14. Russell, B., et al., *A reliable ex vivo invasion assay of human reticulocytes by Plasmodium vivax.* Blood, 2011. **118**(13): p. e74-81.
15. Udomsangpetch R, S.R., Williams J, Sattabongkot J., *Modified techniques to establish a continuous culture of Plasmodium vivax [abstract]. .* American Society of Tropical Medicine and Hygiene 51st Annual Meeting program; 2002 Nov 10-14; Denver, CO. , 2001.
16. Auburn, S., et al., *Effective preparation of Plasmodium vivax field isolates for high-throughput whole genome sequencing.* PLoS One, 2013. **8**(1): p. e53160.
17. Zhu, L., et al., *New insights into the Plasmodium vivax transcriptome using RNA-Seq.* Sci Rep, 2016. **6**: p. 20498.
18. Westenberger, S.J., et al., *A systems-based analysis of Plasmodium vivax lifecycle transcription from human to mosquito.* PLoS Negl Trop Dis, 2010. **4**(4): p. e653.
19. Bozdech, Z., et al., *The transcriptome of Plasmodium vivax reveals divergence and diversity of transcriptional regulation in malaria parasites.* Proc Natl Acad Sci U S A, 2008. **105**(42): p. 16290-5.
20. Orjuela-Sanchez, P., et al., *Higher microsatellite diversity in Plasmodium vivax than in sympatric Plasmodium falciparum populations in Pursat, Western Cambodia.* Exp Parasitol, 2013. **134**(3): p. 318-26.
21. Field, J.W., Sandosham, A.A., Fong, Y.L., *A Morphological Study of the Erythrocitic Parasites in Thick Blood Films, the Microscopic Diagnosis Of Human Malaria.* The Government Press, Kuala Lumpur, 1963.

22. Carvalho, B.O., et al., *On the cytoadhesion of Plasmodium vivax-infected erythrocytes.* J Infect Dis, 2010. **202**(4): p. 638-47.
23. Chotivanich, K., et al., *Plasmodium vivax adherence to placental glycosaminoglycans.* PLoS One, 2012. **7**(4): p. e34509.
24. Lopes, S.C., et al., *Paucity of Plasmodium vivax mature schizonts in peripheral blood is associated with their increased cytoadhesive potential.* J Infect Dis, 2014. **209**(9): p. 1403-7.
25. Barber, B.E., et al., *Parasite biomass-related inflammation, endothelial activation, microvascular dysfunction and disease severity in vivax malaria.* PLoS Pathog, 2015. **11**(1): p. e1004558.
26. Snounou, G. and B. Singh, *Nested PCR analysis of Plasmodium parasites.* Methods Mol Med, 2002. **72**: p. 189-203.
27. Sriprawat, K., et al., *Effective and cheap removal of leukocytes and platelets from Plasmodium vivax infected blood.* Malar J, 2009. **8**: p. 115.
28. Esko, J.D., T.E. Stewart, and W.H. Taylor, *Animal cell mutants defective in glycosaminoglycan biosynthesis.* Proc Natl Acad Sci U S A, 1985. **82**(10): p. 3197-201.
29. Esko, J.D., et al., *Inhibition of chondroitin and heparan sulfate biosynthesis in Chinese hamster ovary cell mutants defective in galactosyltransferase I.* J Biol Chem, 1987. **262**(25): p. 12189-95.
30. Andrews, K.T., et al., *Adherence of Plasmodium falciparum infected erythrocytes to CHO-745 cells and inhibition of binding by protein A in the presence of human serum.* Int J Parasitol, 2005. **35**(10): p. 1127-34.
31. Hasler, T., et al., *An improved microassay for Plasmodium falciparum cytoadherence using stable transformants of Chinese hamster ovary cells expressing CD36 or intercellular adhesion molecule-1.* Am J Trop Med Hyg, 1993. **48**(3): p. 332-47.
32. Buffet, P.A., et al., *Plasmodium falciparum domain mediating adhesion to chondroitin sulfate A: a receptor for human placental infection.* Proc Natl Acad Sci U S A, 1999. **96**(22): p. 12743-8.
33. Kraemer, S.M., et al., *Patterns of gene recombination shape var gene repertoires in Plasmodium falciparum: comparisons of geographically diverse isolates.* BMC Genomics, 2007. **8**: p. 45.
34. Kirsten Moll, I.L., Hedvig Perlmann, Artur Scherf, Mats Wahlgren., *Methods in Malaria Research.* Malaria Research and Reference Reagent Resource Center (MR4) and American Type Culture Collection 10801 University Boulevard, Manassas, VA 20110-2209, 2008.
35. Chomczynski, P. and N. Sacchi, *Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction.* Anal Biochem, 1987. **162**(1): p. 156-9.
36. Chomczynski, P. and N. Sacchi, *The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on.* Nat Protoc, 2006. **1**(2): p. 581-5.
37. Aurrecoechea, C., et al., *EuPathDB: the eukaryotic pathogen genomics database resource.* Nucleic Acids Res, 2017. **45**(D1): p. D581-D591.
38. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.
39. Trapnell, C., L. Pachter, and S.L. Salzberg, *TopHat: discovering splice junctions with RNA-Seq.* Bioinformatics, 2009. **25**(9): p. 1105-11.
40. Rosenbloom, K.R., et al., *The UCSC Genome Browser database: 2015 update.* Nucleic Acids Res, 2015. **43**(Database issue): p. D670-81.
41. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.
42. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

Figures
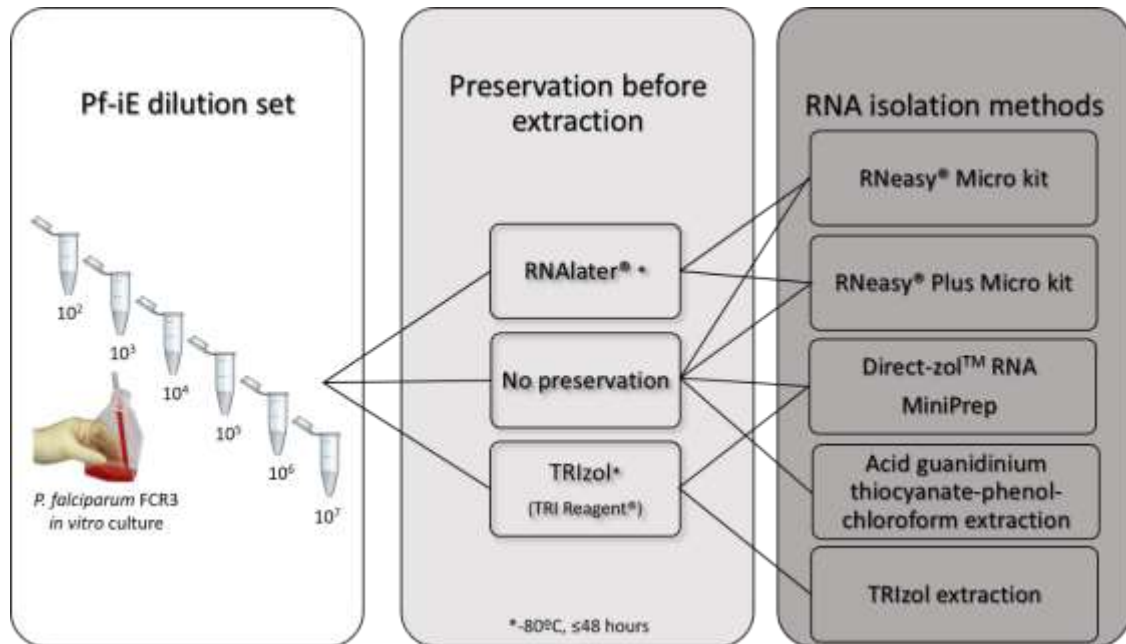
Figure 1. *P. falciparum in vitro* culture and total RNA isolation methods. RNA was extracted from a set of *P. falciparum* infected erythrocytes (Pf-iE) culture dilutions (10^2 to 10^7 in triplicate). The *single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction* [35, 36] was directly used upon sample assembly. An initial step of sample preservation in TRIzol at -80⁰C, for no longer than 48h, was executed accordingly to the *reliable RNA preparation for Plasmodium falciparum* protocol [34]. The RNeasy® Micro, RNeasy® Micro Plus and Direct-zol™ RNA MiniPrep kits were applied as per manufacturer's protocol. Two sets of Pf-iE dilutions were preserved in RNAlater® stabilization reagent or TRI Reagent® at -80⁰C, for no longer than 48h.
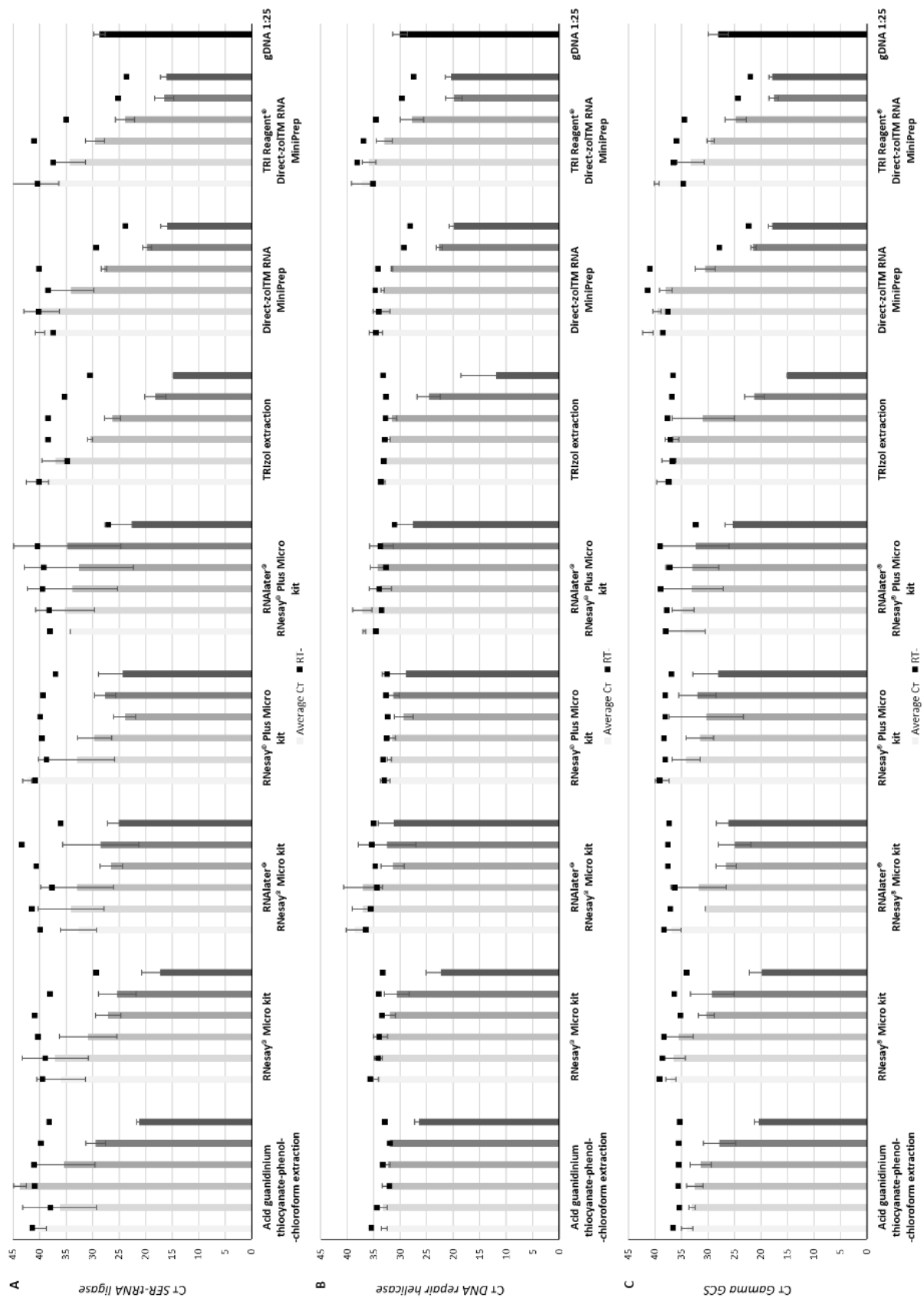
Figure 2. qPCR of *P. falciparum* total RNA samples. qPCR amplification of housekeeping genes *seryl tRNA synthetase* (A), *DNA repair helicase* (B) and *gamma-glutamylcysteine synthetase* (C), from the RNA samples of Pf-iE culture dilutions (10^2 in light gray to 10^7 in dark gray), extracted by each different method. Bars and error bars represent the mean Ct and standard deviations, respectively. Black square

marks represent mean Ct of RT- reactions, when amplification was observed before the last (45th) cycle stage.
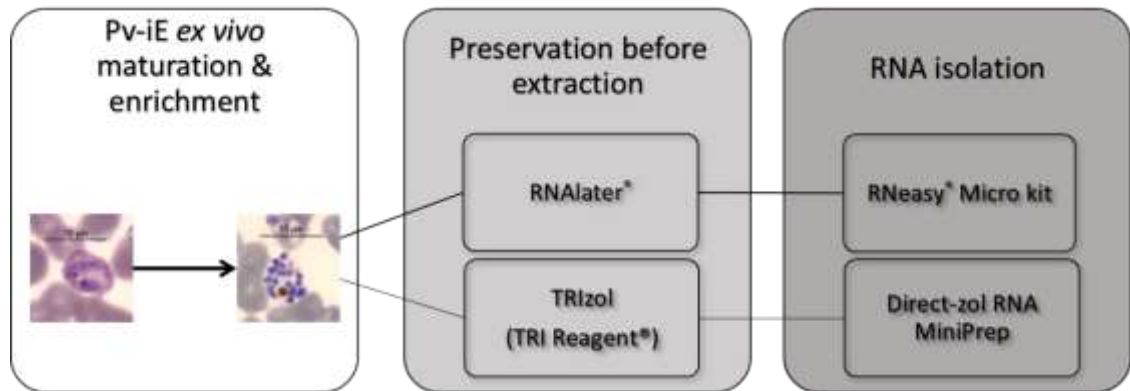
Figure 3. *P. vivax* isolation, short-term *ex vivo* maturation, enrichment and total RNA extraction. Blood samples were immediately processed. White blood cells were depleted by CF11 filtration [16]. The early blood staged parasites were *ex vivo* short-term cultured to allow maturation followed by Percoll gradient enrichment [14]. Total number of Pv-iEs was accessed before preservation in RNAlater® or TRI Reagent® at -80ºC. RNA extraction was executed using the RNeasy® Micro kit.
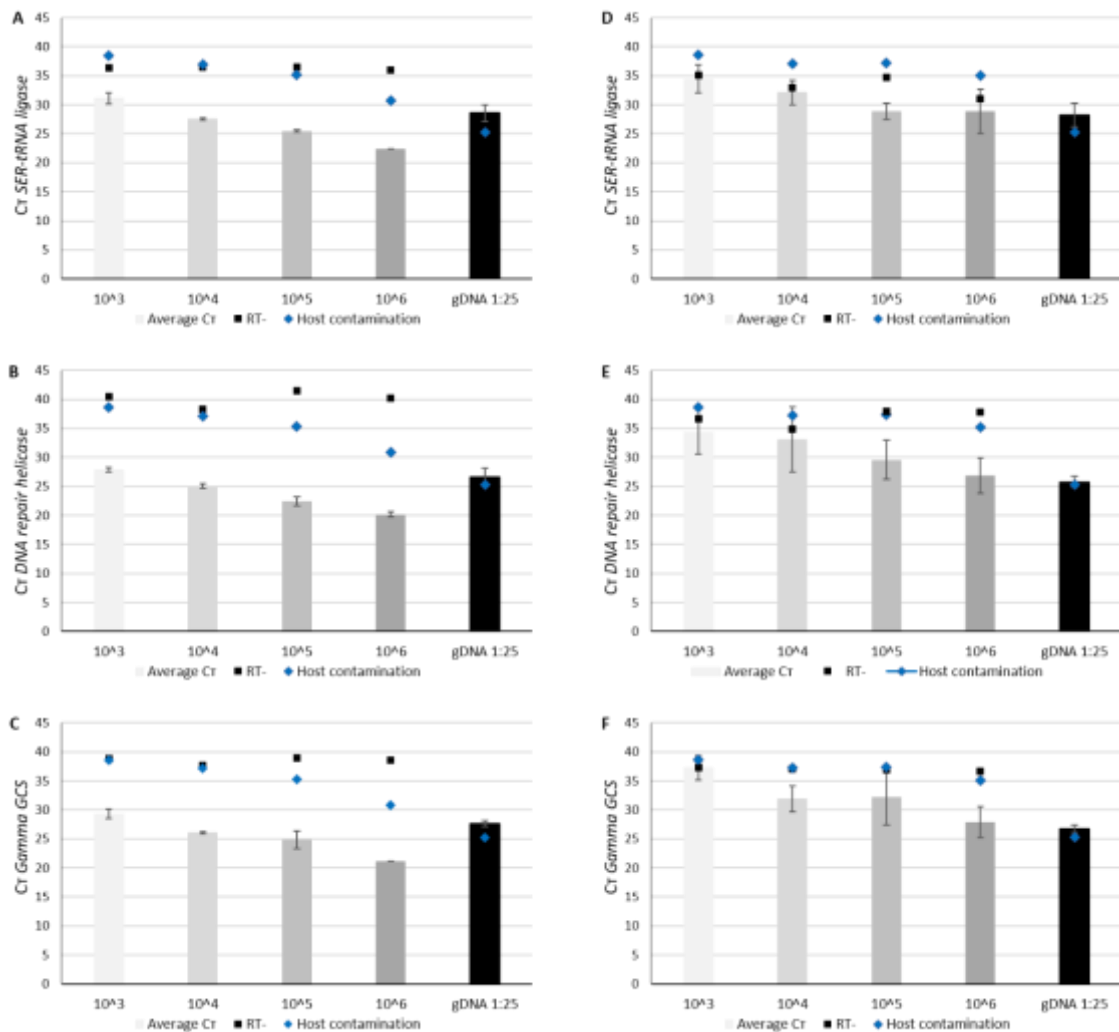
Figure 4. qPCR of *P. vivax* total RNA samples. qPCR amplification of housekeeping genes *seryl tRNA synthetase* (A and D), *DNA repair helicase* (B and E) and *gamma-glutamylcysteine synthetase* (C and F), from the RNA samples of Pv-iE field-isolates (10^3 in light gray to 10^6 in dark gray), preserved in RNAlater® or TRI Reagent® and extracted by RNeasy® Micro kit (A-C) or Direct-zol™ RNA MiniPrep (D-F), respectively. Bars and error bars represent the mean Ct and standard deviations, respectively. Black square marks represent average Ct of RT- reactions, when amplification was observed before the last (45th) cycle stage. Blue diamond marks represent mean Ct amplification of human TLR9 gene for host contamination detection.
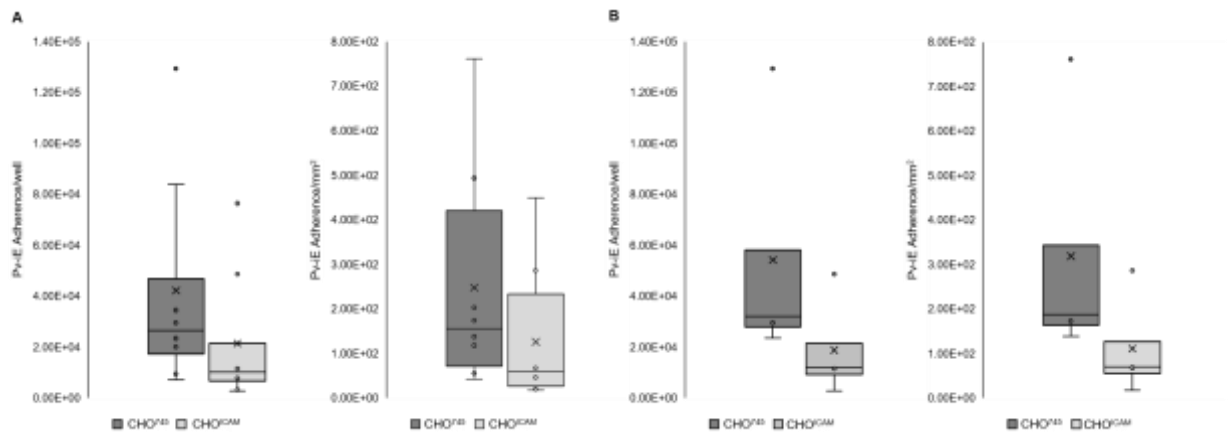
Figure 5. Pv-iE adherence per well and per mm² considering all cytoadhesion assays against CHO[745] and CHO[ICAM] cell lines (A) (paired t test, 2-tailed p-value=0.0595) and only the isolates which the distinct adhered and non-adhered Pv-iE populations were transcriptome sequenced (B) (paired t test, 2-tailed p-value=0.1103).

Tables

Table 1. Sample profiles for *P. vivax* clinical isolates

Sample Code	Initial smear stages (h)	Pv-iE stages after Percoll (h)	Parasitemia after Percoll (%)	Total nº of erythrocytes /mL	Total nº of Pv-iE/mL	Volume of Pv culture (µL/well)	Total nº of Pv-iE/well	Cell lines /Adhesion assay	Sample for RNA-seq
92U15	trophozoites (11-15) and some schizonts	trophozoites (11-15) and some schizonts	50,4%	7,10E+06	3,58E+06	250	8,95E+05	CHO^{745} CHO^{ICAM}	**Y**
93U15	trophozoites (11-15)	trophozoites (11-15)	94,0%	3,88E+06	3,65E+06	250	9,12E+05	CHO^{745} CHO^{ICAM}	**Y**
96U15.1	trophozoites (15)	trophozoites (15-17)	80,7%	7,24E+06	5,84E+06	250	1,46E+06	CHO^{745} CHO^{ICAM}	N
101U15	rings (few), trophozoites (15-19) and schizonts	trophozoites (mature >15) and schizonts	54,0%	2,70E+06	1,46E+06	250	3,65E+05	CHO^{745} CHO^{ICAM}	**Y**
102U15	trophozoites (young 11-15)	trophozoites (young 11-15)	98,2%	5,90E+07	5,80E+07	125	7,25E+06	CHO^{745} CHO^{ICAM}	N
103U15	rings (9), trophozoites (11-15) and schizonts	trophozoites (young 11-15) and gametocytes	97,0%	1,74E+06	1,69E+06	130	2,19E+05	CHO^{745} CHO^{ICAM}	N
104U15	some rings and trophozoites (young 11-15)	(some) rings, trophozoites (11-15) and gametocytes	65,3%	2,81E+06	1,83E+06	250	4,59E+05	CHO^{745} CHO^{ICAM}	N
105U15	rings (few), trophozoites (15-20) and schizonts	trophozoites (15-20)	81,6%	5,20E+06	4,24E+06	250	1,06E+06	CHO^{745} CHO^{ICAM}	**Y**

Summary of sample information during parasite enrichment of clinical isolates and on cytoadhesion assays performed during field work at malaria vivax endemic area, adequate for RNA-seq downstream application. Y – Yes; N – No; All samples were preserved through N_2 flash freezing after processing and during transportation.

Table 2. Alignment and mapping summary report for *P. vivax* clinical isolates RNA sequenced.

	Sample Code (RNA-seq ID)						
	92U15 (1)	**92U15** (2)	**92U15** (3)	**92U15** (4)	**93U15** (21)	**93U15** (22)	**105U15** (5)
Left reads							
Input	345888	1070871	1003388	219624	475794	214949	663014
Mapped (% of input)	2584 (0.7%)	54770 (5.1%)	9455 (0.9%)	18629 (8.5%)	5239 (1.1%)	18875 (8.8%)	89193 (13.5%)
Multiple alignments (% of mapped)	1504 (58.2%)	3019 (5.5%)	2594 (27.4%)	1455 (7.8%)	1118 (21.3%)	1108 (5.9%)	3460 (3.9%)
Multiple alignments (>20)	72	43	47	48	27	20	65
Right reads							
Input	345888	1070871	1003388	219624	475794	214949	663014
Mapped	3415 (1.0%)	55627 (5.2%)	11637 (1.2%)	19160 (8.7%)	6457 (1.4%)	19727 (9.2%)	91575 (13.8%)
Multiple alignments	2329 (68.2%)	4863 (8.7%)	4505 (38.7%)	1959 (10.2%)	2053 (31.8%)	2171 (11.0%)	5589 (6.1%)
Multiple alignments (>20)	72	43	47	49	27	20	68
Overall read mapping rate	0.90%	5.20%	1.10%	8.60%	1.20%	9.00%	13.60%
Aligned pairs	618	44480	4288	15033	2914	14723	76226
Multiple alignments (% of aligned pairs)	255 (41.3%)	815 (1.8%)	302 (7.0%)	352 (2.3%)	177 (6.1%)	279 (1.9%)	1263 (1.7%)
Discordant alignments (% of aligned pairs)	168 (27.2%)	501 (1.1%)	270 (6.3%)	199 (1.3%)	159 (5.5%)	215 (1.5%)	697 (0.9%)
Concordant pair alignment rate	0.10%	4.10%	0.40%	6.80%	0.60%	6.70%	11.40%

	105U15 (6)	**105U15** (7)	**105U15** (8)	**101U15** (20)	**101U15** (23)	**101U15** (24)
Left reads						
Input	670600	1288122	188117	226387	754181	475848
Mapped (% of input)	97757 (14.6%)	68891 (5.3%)	21157 (11.2%)	26956 (11.9%)	47834 (6.3%)	51521 (10.8%)
Multiple alignments (% of mapped)	4391 (4.5%)	4283 (6.2%)	1305 (6.2%)	1810 (6.7%)	2019 (4.2%)	2624 (5.1%)
Multiple alignments (>20)	58	71	33	55	38	57
Right reads						
Input	670600	1288122	188117	226387	754181	475848
Mapped	96131 (14.3%)	70237 (5.5%)	21050 (11.2%)	26585 (11.7%)	48540 (6.4%)	51689 (10.9%)
Multiple alignments	5524 (5.7%)	7053 (10.0%)	1727 (8.2%)	2256 (8.5%)	3527 (7.3%)	4078 (7.9%)
Multiple alignments (>20)	58	71	33	55	38	57
Overall read mapping rate	14.50%	5.40%	11.20%	11.80%	6.40%	10.80%
Aligned pairs	82207	54997	17065	21599	39428	41269
Multiple alignments (% of aligned pairs)	1332 (1.6%)	1111 (2.0%)	384 (2.3%)	676 (3.1%)	551 (1.4%)	769 (1.9%)
Discordant alignments (% of aligned pairs)	697 (0.8%)	744 (1.4%)	227 (1.3%)	393 (1.8%)	416 (1.1%)	580 (1.4%)
Concordant pair alignment rate	12.20%	4.20%	9.00%	9.40%	5.20%	8.60%

Table 3. Number of genes differentially expressed analyzed in several comparison groups

DE Comparison groups	Number of genes DE
Adhered vs non-adhered	107
Adhered vs non-adhered to CHO[ICAM]	137
Adhered vs non-adhered to CHO[745]	75
Adhered to CHO[ICAM] vs adhered to CHO[745]	52

Summary of the number of genes verified to be differentially expressed (DE) (p-value<0.05) not considering further normalization of read counts against the total number of reads per sample.

Table 4. List of differential expressed gene in the adhered to CHO^{ICAM} vs adhered to CHO^{745} comparison group

Gene ID	mean counts	Log2Fold Change	SD	Wald stat	p-value	q-value	Transcript Length	Gene Symbol	Product Description
PVP01_0419200	2.249161	-3.234168928	1.443578	-2.24038	0.025066	0.998539	3120	RPN1	26S proteasome regulatory subunit RPN1, putative
PVP01_0612900	9.095748	-2.045834326	1.042515	-1.9624	0.049716	0.998539	1161	RPL3	60S ribosomal protein L3, putative
PVP01_1252300	2.690943	-3.31818665	1.440708	-2.30316	0.02127	0.998539	1422	6PGD	6-phosphogluconate dehydrogenase, decarboxylating, putative
PVP01_0530800	74.28897	-2.240479019	1.124194	-1.99296	0.046265	0.998539	1353	null	alpha tubulin 2, putative
PVP01_0313300	6.145463	-2.572203722	1.243332	-2.0688	0.038565	0.998539	1590	CDPK4	calcium-dependent protein kinase 4, putative
PVP01_0408100	5.392482	2.256093649	1.132427	1.992264	0.046342	0.998539	1521	null	conserved Plasmodium protein, unknown function
PVP01_0506800	2.279292	-3.151362815	1.448797	-2.17516	0.029618	0.998539	4998	null	conserved Plasmodium protein, unknown function
PVP01_0933800	8.945893	2.468248713	1.150215	2.145902	0.031881	0.998539	90	null	conserved Plasmodium protein, unknown function
PVP01_1128400	2.84275	-3.422256378	1.422272	-2.40619	0.01612	0.998539	4200	null	conserved Plasmodium protein, unknown function
PVP01_1205300	2.251303	-3.21605919	1.437719	-2.23692	0.025292	0.998539	8064	null	conserved Plasmodium protein, unknown function
PVP01_1323500	3.311216	2.526861583	1.275671	1.98081	0.047613	0.998539	2340	null	conserved Plasmodium protein, unknown function
PVP01_1443200	2.39271	-2.981006304	1.470583	-2.02709	0.042653	0.998539	4914	null	conserved Plasmodium protein, unknown function
PVP01_0602200	5.299142	-3.063963135	1.208347	-2.53566	0.011223	0.998539	3291	null	DNA polymerase delta catalytic subunit, putative
PVP01_1310900	4.100622	-3.473190249	1.438152	-2.41504	0.015734	0.998539	2277	MCM5	DNA replication licensing factor MCM5, putative
PVP01_0114000	5.324549	-3.072061846	1.262611	-2.4331	0.01497	0.998539	2571	null	DnaJ protein, putative

PVP01_1323000	9.373775	-2.417507984	1.222703	-1.97718	0.048021	0.998539	1833	null	dynein intermediate chain, putative
PVP01_0924600	2.719771	-3.26828004	1.446023	-2.26019	0.02381	0.998539	1404	ERO1	endoplasmic reticulum oxidoreductin, putative
PVP01_0909400	8.954836	-2.218764339	1.097648	-2.02138	0.04324	0.998539	1035	ERC	endoplasmic reticulum-resident calcium binding protein, putative
PVP01_0816000	19.21719	-2.728331165	1.19935	-2.27484	0.022915	0.998539	1341	ENO	enolase, putative
PVP01_0833800	6.276239	-2.724353955	1.167899	-2.3327	0.019664	0.998539	1968	null	FAD-dependent glycerol-3-phosphate dehydrogenase, putative
PVP01_1244000	47.80517	-2.084812785	0.963461	-2.16388	0.030474	0.998539	1014	GAPDH	glyceraldehyde-3-phosphate dehydrogenase, putative
PVP01_1420800	2.247696	-3.235434643	1.435285	-2.25421	0.024183	0.998539	3924	PRP2	golgi protein 1, putative
PVP01_1416100	2.246649	-2.936955808	1.466663	-2.00248	0.045234	0.998539	2883	null	heptatricopeptide repeat-containing protein, putative
PVP01_0823800	6.368487	-2.694417238	1.159689	-2.3234	0.020158	0.998539	1125	PPase	inorganic pyrophosphatase, putative
PVP01_1022200	4.675448	-2.841341159	1.305358	-2.17668	0.029505	0.998539	1869	INO1	inositol-3-phosphate synthase, putative
PVP01_1229400	4.985151	-2.924554163	1.282764	-2.27988	0.022615	0.998539	1005	null	lactate dehydrogenase, putative
PVP01_1229700	22.06872	-3.093375026	1.153024	-2.68284	0.0073	0.998539	951	LDH	L-lactate dehydrogenase
PVP01_0704800	3.729869	-3.459672135	1.436295	-2.40875	0.016007	0.998539	1482	null	Maf-like protein, putative
PVP01_1132100	2.52518	-2.915734306	1.480095	-1.96996	0.048842	0.998539	1194	TOM40	mitochondrial import receptor subunit TOM40, putative
PVP01_1256700	57.76185	5.453513457	1.344569	4.055958	4.99E-05	0.092369	2409	CPR	NADPH--cytochrome P450 reductase, putative
PVP01_1207600	2.927111	-3.289399089	1.447057	-2.27317	0.023016	0.998539	1251	NT1	nucleoside transporter 1
PVP01_1303000	3.732117	-2.644459966	1.298276	-2.0369	0.04166	0.998539	1047	NAPL	nucleosome assembly protein, putative
PVP01_0929800	13.20492	-2.405356196	1.012372	-2.37596	0.017503	0.998539	1290	PV1	parasitophorous vacuolar protein 1, putative
PVP01_0721000	10.5437	-2.95607265	1.121929	-2.63481	0.008418	0.998539	1251	PGK	phosphoglycerate kinase, putative
PVP01_0950000	2.286886	2.281440402	1.131166	2.016892	0.043707	0.998539	1389	null	PIR protein
PVP01_1201600	15.80213	-2.003063043	0.94768	-2.11365	0.034545	0.998539	897	null	Plasmodium exported protein, unknown function

PVP01_0315800	6.720583	1.831032209	0.920029	1.990189	0.04657	0.998539	993	ESF2	pre-rRNA-processing protein ESF2, putative
PVP01_0111900	3.485238	-3.611668991	1.417048	-2.54873	0.010812	0.998539	2061	null	regulator of chromosome condensation, putative
PVP01_0308900	3.910606	-2.770995611	1.233291	-2.24683	0.024651	0.998539	3426	RPA1	replication protein A1, large subunit, putative
PVP01_1020200	8.949359	-3.138646877	1.164765	-2.69466	0.007046	0.998539	2196	PSOP1 2	secreted ookinete protein, putative
PVP01_0933000	2.418975	-3.102077405	1.460756	-2.12361	0.033703	0.998539	1632	null	T-complex protein 1 subunit alpha, putative
PVP01_1447600	2.915036	-3.469798639	1.423434	-2.43762	0.014784	0.998539	1635	CCT3	T-complex protein 1 subunit gamma, putative
PVP01_1140200	14.84691	-2.139734954	0.972114	-2.20112	0.027728	0.998539	1632	CCT6	T-complex protein 1 subunit zeta, putative
PVP01_0722300	7.180133	-2.710393719	1.286114	-2.10743	0.03508	0.998539	1641	TRXR	thioredoxin reductase, putative
PVP01_0924800	4.174081	-2.72011548	1.311654	-2.07381	0.038097	0.998539	1137	null	thioredoxin-like protein, putative
PVP01_1268900	2.934532	-3.103016632	1.466193	-2.11638	0.034313	0.998539	1131	null	translocation protein SEC62, putative
PVP01_MIT00500	3.556286	3.153000571	1.314067	2.399422	0.016421	0.998539	207	null	unspecified product
PVP01_MIT01200	51.17361	3.425556543	1.197268	2.861144	0.004221	0.998539	196	null	unspecified product
PVP01_1412900	3.040085	-3.513129173	1.423222	-2.46843	0.013571	0.998539	1836	vapA	V-type proton ATPase catalytic subunit A, putative
PVP01_0113600	2.686498	-3.131490422	1.458399	-2.14721	0.031777	0.998539	2946	null	V-type proton ATPase subunit a, putative
PVP01_0305500	2.895724	-3.378843987	1.436515	-2.35211	0.018667	0.998539	1485	null	V-type proton ATPase subunit B, putative
PVP01_0905300	2.737354	-3.150100879	1.459302	-2.15863	0.030879	0.998539	1851	null	WD repeat-containing protein, putative

Supplemental Figures and Tables
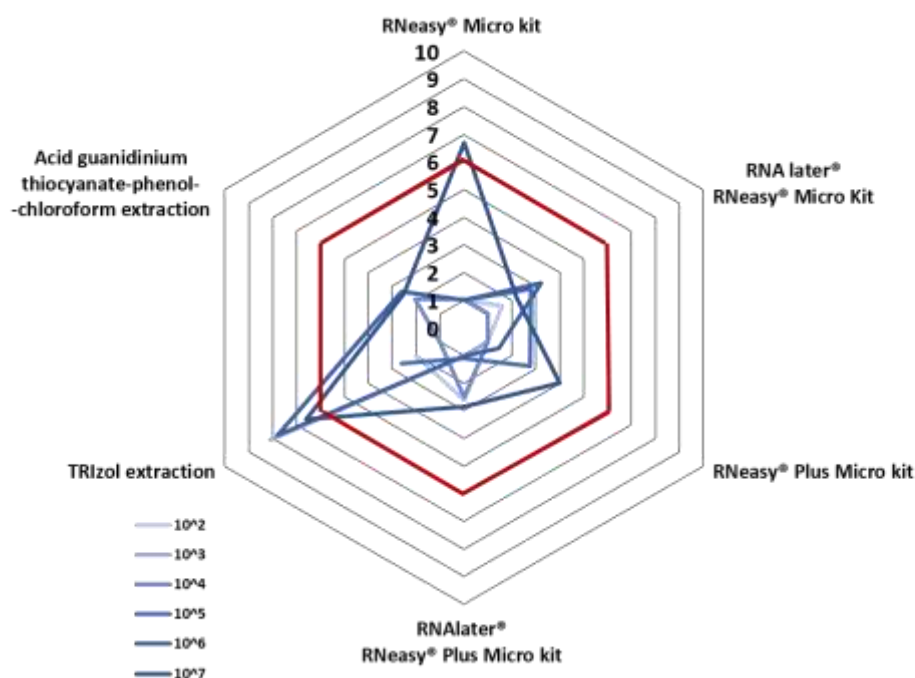
Figure 1. *P. falciparum* total RNA sample analysis. Average RNA Integrity Number (RIN) for all RNA samples extracted by each method obtained on the ©Agilent Bioanalyzer platform. RNA with RIN ≥ 6.0 is considered suitable for RNA-seq.
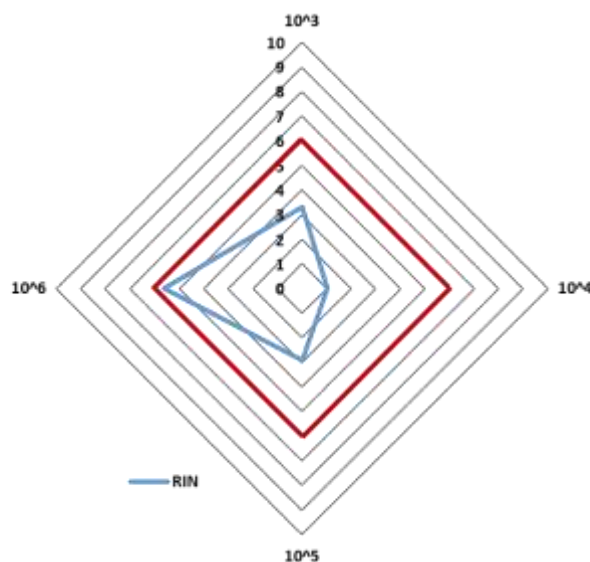
Figure 2. *P. vivax* total RNA sample analysis. Average RNA Integrity Number (RIN) for all RNA samples preserved in RNAlater® and extracted by RNeasy® Micro kit, obtained on the ©Agilent Bioanalyzer platform. RNA with RIN ≥ 6.0 is considered suitable for RNA-seq.
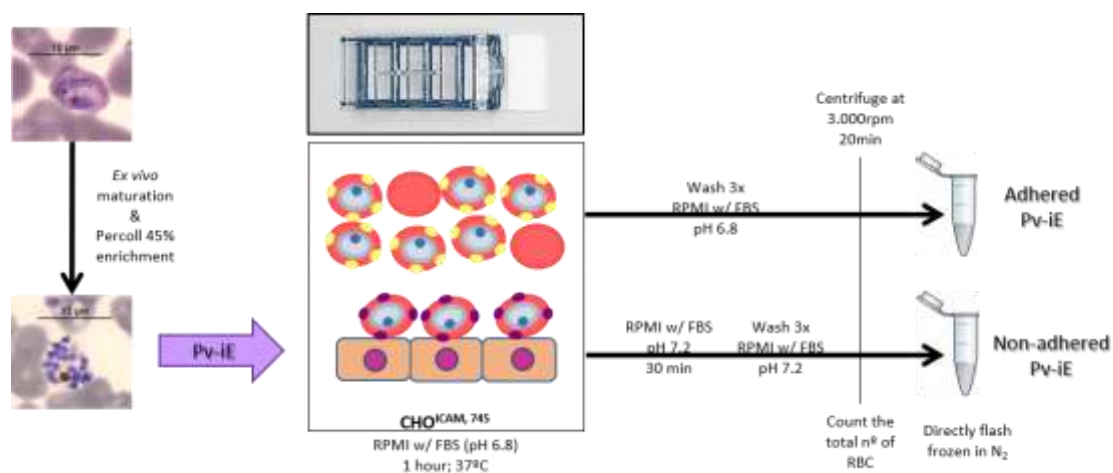
Figure 3. Methodologic scheme of cytoadhesion assays to determine the adherence capacity of distinct populations of Pv-iE from vivax malaria patient isolates.
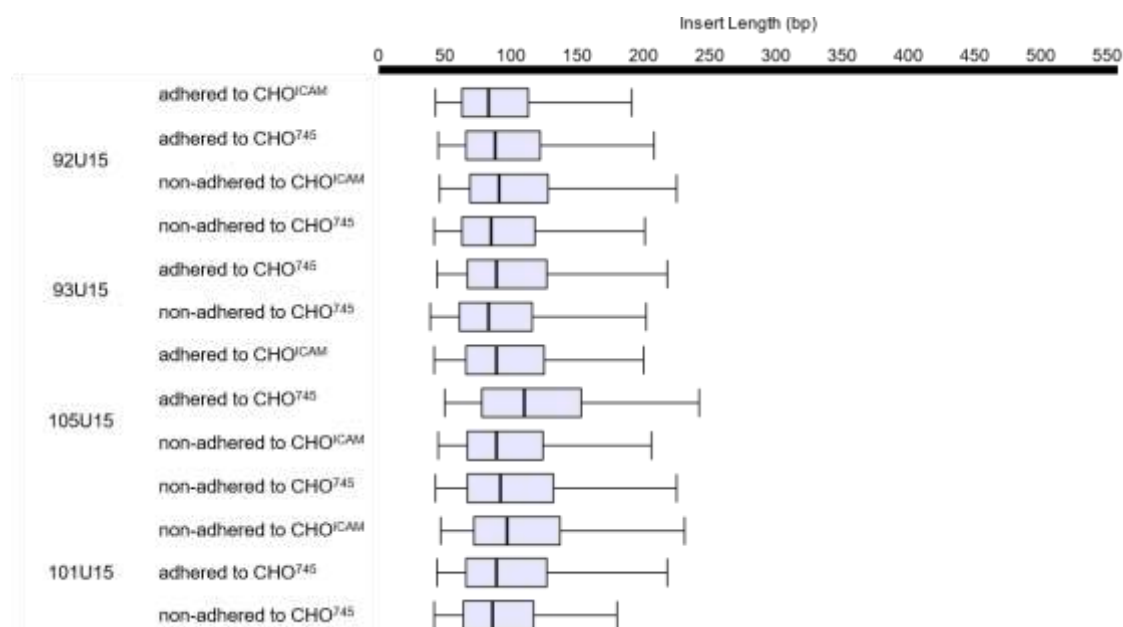
Figure 4. The insert length distribution boxplot summarizing the insert length distribution of paired-end reads of RNA sequenced libraries.

Table 1. Patient profiles for *P. vivax* clinical isolates

Collection date	Patient code	Sex	Age	Parasitemia*	Pv-iE/200 leucocyt	WBC (x10³/µl)	RBC (x10⁶/µL)	HGB (g/dL)	HCT (%)	MCV (fL)	MCH (pg)	MCHC (g/dL)	PLT (x10³/µL)	LYM (x10³/µL)	MXD (x10³/µL)	NEUT (x10³/µL)	RDW_SD (fL)	RDW_VC (%)	MPV (fL)
7/12/15	**92U15**	M	25	2+	278	3.8	4.54	12.5	39.6	87.2	27.5	31.6	66	0.9	F2* 1,1	F3* 1,8	40.6	12.8	8.5
7/12/15	**93U15**	M	44	2+	331	3.4	4.43	11.8	39.4	88.9	26.6	29.9	45	F1* 0,5	F2* 0,6	2.3	41.6	12.8	9.3
9/12/15	**96U15**	F	28	2+	530	3.8	3.49	10.3	31.5	90.3	29.5	32.7	AG 58	1.1	0.4	2.3	42.5	13.6	9.3
14/12/15	**101U15**	M	32	2+	193	4.6	5.16	15	47.4	91.9	29.1	31.6	93	0.8	0.1	3.7	49.6	14.8	8
15/12/15	**102U15**	F	23	2+	200	6.4	4.69	11.6	38.1	81.2	24.7	30.4	AG 117	1	0.9	4.5	43.1	14.9	10
15/12/15	**103U15**	M	24	2+	270	WL* 3	4.77	13.4	41.9	87.8	28.1	32	AG 23	WL* 0,6	WL* 0,2	WL* 2,2	42.8	13.6	
15/12/15	**104U15**	F	25	2+	287	WL* 7,9	4.37	12.2	39.9	91.3	27.9	30.6	AG 130	WL* 1,8	WL* 1,1	WL* 5	42.4	11.8	9.7
15/12/15	**105U15**	M	38	2+	250	6.9	5.38	15.1	48	89.2	28.1	31.5	AG 47	0.5	0.6	5.8	42.7	12.5	10.7

Summary of patients and clinical isolates information after collection during field work at vivax malaria endemic area. Date: dd/mm/yy; M: Male; F: Female; *Parasitemia estimation and **initial counts of the number of Pv-iE per 200 leucocytes on the slide by thick smear made available by the specially trained microscopists from malaria diagnosis service from FMT-HVD. Two crosses (2+) equals to 2 to 20 parasites counted by microscope field and an average of 500 to 10^4 parasites per µL. WBC: White Blood Cells; RBC: Red Blood Cells; HGB: hemoglobin; HCT: Hematocrit; MCV: Mean Corpuscular Volume; MCH: Mean Corpuscular Hemoglobin; MCHC: mean corpuscular hemoglobin concentration; PLT: Platelets; LYM: Lymphocytes; MXD: mixed cells; NEUT: Neutrophils; RDW: Red Cell Distribution Width SD: Standard Deviation and VC: Variance Coefficient; MPV: Mean Platelet Volume. For information on hemogram flags please see https://www.sysmex.co.za/fileadmin/media/f112/SEED/English/Sysmex_SEED_6_2013_Haematology_Results_Interferences__Flagging_and_Interpretation_-_Part_II_EN.pdf.

Table 2. Cytoadhesion assays for *P. vivax* clinical isolates.

Sample code - cell line assay	Total nº of Pv-iE/well*	Nº lines	Pv-iE counts**	Pv-iE /nº lines	Pv-iE /mm²	Adhered Pv-iE /well	Non-adhered Pv-iE /well	Total nº of Pv-iE /well	Total nº of Pv-iE /mm²	Adhered Pv-iE /well (%)	Non-adhered Pv-iE /well (%)	Adhered Pv-iE /mm²	Non-adhered Pv-iE /mm²
92U15-CHO[745]	8.95E+05	2	637	318.5	202.8	3.45E+04	8.61E+05	8.95E+05	5.26E+03	3.85	96.15	203	5062
92U15-CHO[ICAM]	8.95E+05	8	211	26.4	16.8	2.85E+03	8.92E+05	8.95E+05	5.26E+03	0.32	99.68	17	5248
93U15-CHO[745]	9.12E+05	1	1195	1195.0	760.8	1.29E+05	7.83E+05	9.12E+05	5.36E+03	14.18	85.82	761	4604
93U15-CHO[ICAM]	9.12E+05	1	449	449.0	285.8	4.86E+04	8.63E+05	9.12E+05	5.36E+03	5.33	94.67	286	5079
96U15-CHO[745]	1.46E+06	1	185	185.0	117.8	2.00E+04	1.44E+06	1.46E+06	8.59E+03	1.37	98.63	118	8470
96U15-CHO[ICAM]	1.46E+06	1	72	72.0	45.8	7.79E+03	1.45E+06	1.46E+06	8.59E+03	0.53	99.47	46	8542
101U15-CHO[745]	3.65E+05	3	649	216.3	137.7	2.34E+04	3.42E+05	3.65E+05	2.15E+03	6.41	93.59	138	2009
101U15-CHO[ICAM]	3.65E+05	5	574	114.8	73.1	1.24E+04	3.53E+05	3.65E+05	2.15E+03	3.40	96.60	73	2074
102U15-CHO[745]	7.25E+06	1	775	775.0	493.4	8.39E+04	7.17E+06	7.25E+06	4.26E+04	1.16	98.84	493	42154
102U15-CHO[ICAM]	7.25E+06	1	706	706.0	449.5	7.64E+04	7.17E+06	7.25E+06	4.26E+04	1.05	98.95	449	42198
103U15-CHO[745]	2.19E+05	6	520	86.7	55.2	9.38E+03	2.10E+05	2.19E+05	1.29E+03	4.28	95.72	55	1233
103U15-CHO[ICAM]	2.19E+05	9	273	30.3	19.3	3.28E+03	2.16E+05	2.19E+05	1.29E+03	1.50	98.50	19	1269
104U15-CHO[745]	4.59E+05	8	525	65.6	41.8	7.10E+03	4.52E+05	4.59E+05	2.70E+03	1.55	98.45	42	2658
104U15-CHO[ICAM]	4.59E+05	7	559	79.9	50.8	8.64E+03	4.50E+05	4.59E+05	2.70E+03	1.88	98.12	51	2649
105U15-CHO[745]	1.06E+06	2	545	272.5	173.5	2.95E+04	1.03E+06	1.06E+06	6.24E+03	2.78	97.22	173	6062
105U15-CHO[ICAM]	1.06E+06	5	531	106.2	67.6	1.15E+04	1.05E+06	1.06E+06	6.24E+03	1.08	98.92	68	6168

Data table with all parasite counts before and after each cytoadhesion assay. *Total nº of Pv-iE per well was estimated using a Neubauer chamber counts; **Pv-iE were counted from duplicate experiment wells from the same Lab-Tek 4-well slide strained with *Panótico Rápido* kit. Calculation of the final number of adhered per mm^2 and per well took into consideration the dimensions of well from a 4 Lab-Tek 4-well slide (width = 17 mm X h = 10 mm), where h corresponds to 50 fields on the used microscope (diameter microscopic fields = 0.2mm). Area (A) of the microscopic fields = 3.1416 x 0.01 = 0,031416 mm^2. A of each counted field line on the microscopic = 0,031416 x 50 = 1,5708 mm^2; A = h x l = 10 x 17 mm = 170 mm^2.

Table 3. Bioanalyzer measurements after RNA extraction with RNeasy® Micro kit.

Sample Code	Cytoadhesion assay description of Pv-iE collected samples	RNA [pg/uL]	RIN
92U15	adhered to CHO^{745}	12.457	9.1
	non-adhered to CHO^{745}	36.245	9.3
	adhered to CHO^{ICAM}	7.935	8.9
	non-adhered to CHO^{ICAM}	21.608	8.9
93U15	adhered to CHO^{745}	12.437	N/A
	non-adhered to CHO^{745}	17.819	9.3
101U15	adhered to CHO^{745}	9.405	4.3
	non-adhered to CHO^{745}	65.81	N/A
	adhered to CHO^{ICAM}	10.465	N/A
	non-adhered to CHO^{ICAM}	16.529	N/A
105U15	adhered to CHO^{745}	6.655	9.8
	non-adhered to CHO^{745}	26.412	N/A
	adhered to CHO^{ICAM}	12.533	6.8
	non-adhered to CHO^{ICAM}	40.524	5.7

RIN: RNA Integrity Number; N/A: undetermined.

Table 4. Bioanalyzer quality and quantity controls after cDNA library generation and amplification.

Sample Code	Cytoadhesion assay description of Pv-iE collected samples	SMART-Seq v4 Ultra Low Input RNA kit for Sequencing							Shearing and library amplification		
		cDNA [pg/uL]	cDNA [ng/uL]	V to 20ng (µL)	Total cDNA [ng/75uL]	V_T to 20ng/reaction from 75µL after shearing (µL)	V Elution Buffer (µL)	Total cDNA [ng/reaction]	Nº of PCR cycles for Low Input Library amplification	Fragment average size (bp)	Molarity [pmol/L]
92U15	adhered to CHO[745]	40817.2	40.817	0.49	3061.290	3.7	6.3	20	6	358	9.07E+04
	non-adhered to CHO[745]	1036.28	1.036	19.30	77.721	12.0	-	0.156	11	333	1.83E+03
	adhered to CHO[ICAM]	24790.7	24.791	0.81	1859.300	6.1	3.9	20	6	386	7.52E+02
	non-adhered to CHO[ICAM]	72193.1	72.193	0.28	5414.482	2.1	7.9	20	6	369	6.50E+05
93U15	adhered to CHO[745]	26400.6	26.401	0.76	1980.044	5.7	4.3	20	6	388	4.50E+05
	non-adhered to CHO[745]	4153.18	4.153	4.82	311.489	12.0	-	0.66	8	259	4.23E+02
101U15	adhered to CHO[745]	290901	290.901	0.07	21817.556	0.5	9.5	20	6	364	1.71E+05
	non-adhered to CHO[745]	45498.6	45.499	0.44	3412.397	3.3	6.7	20	6	361	5.37E+04
	adhered to CHO[ICAM]	2627.92	2.628	7.61	197.094	12.0	-	0.42	9	589	4.36E+03
	non-adhered to CHO[ICAM]	65064	65.064	0.31	4879.796	2.3	7.7	20	6	376	3.04E+03
105U15	adhered to CHO[745]	48301.3	48.301	0.41	3622.595	3.1	6.9	20	6	322	2.34E+03
	non-adhered to CHO[745]	3329.65	3.330	6.01	249.724	12.0	-	0.533	8	452	6.21E+03
	adhered to CHO[ICAM]	1213.14	1.213	16.49	90.986	12.0	-	0.194	10	334	2.95E+03
	non-adhered to CHO[ICAM]	28025.7	28.026	0.71	2101.928	5.4	4.6	20	6	351	8.14E+04

V: Volume; V_T: Total Volume; bp: base pairs.

Table 5. Library quantification by qPCR and sample pool for HiSeq 2500 load and run.

Sample Code	Cytoadhesion assay description of Pv-iE collected samples	Dilution	Cq Mean	Cq Error	Mean [pM]	Error [pM]	Total [pM]	Total [nM]	*447bp from DNA fragments	Library Concentration [nM]	Dilution	V at 2nM (µL)	V ddH$_2$O (µL)	V$_T$/sample (7µL)
92U15	adhered to CHO^{745}	1.00E-06	10.99	0.59	4.59E+00	2.07E+00	4.59E+06	4590.0	2.05E+09	5731.09	dil -3	5.73109	2.44	4.56	7.00
	non-adhered to CHO^{745}	1.00E-04	15.88	0.33	1.34E-01	3.37E-02	1.34E+03	1.34	6.00E+05	1.80277	dil 0	1.80277	3.33	3.67	7.00
	adhered to CHO^{ICAM}	1.00E-03	15.14	0.97	2.58E-01	1.52E-01	2.58E+02	0.26	1.15E+05	0.29889	dil 0	0.29889	20.07	N/A	9.00
	non-adhered to CHO^{ICAM}	1.00E-07	15.97	0.08	1.26E-01	6.99E-03	1.26E+06	1.26E03	5.62E+08	1523.92	dil -2	15.2392	0.92	6.08	7.00
93U15	adhered to CHO^{745}	1.00E-09	16.03	0.79	1.43E-01	6.84E-02	1.43E+08	1.42E05	6.37E+10	164284	dil -4	16.4284	0.85	6.15	7.00
	non-adhered to CHO^{745}	1.00E-04	14.77	0.34	2.99E-01	7.21E-02	2.99E+03	2.99	1.34E+06	5.15862	dil 0	5.15862	2.71	4.29	7.00
101U15	adhered to CHO^{745}	1.00E-09	11.79	0.43	2.66E+00	8.74E-01	2.66E+09	2.66E06	1.19E+12	3266538	dil -6	3.26654	4.29	2.71	7.00
	non-adhered to CHO^{745}	1.00E-09	12.81	0.74	1.16E+00	5.44E-01	1.16E+09	1.16E06	5.19E+11	1436343	dil -5	14.3634	0.97	6.03	7.00
	adhered to CHO^{ICAM}	1.00E-04	20.52	0.07	5.75E-03	2.55E-04	5.75E+01	0.06	2.57E+04	0.04365	dil 0	0.04365	320.71	N/A	N/A
	non-adhered to CHO^{ICAM}	1.00E-04	17.56	0.42	4.11E-02	1.07E-02	4.11E+02	0.41	1.84E+05	0.48873	dil 0	0.48873	28.65	N/A	9.00
105U15	adhered to CHO^{745}	1.00E-04	14.89	0.85	3.03E-01	1.89E-01	3.03E+03	3.03	1.35E+06	4.20347	dil 0	4.20347	3.33	3.67	7.00
	non-adhered to CHO^{745}	1.00E-05	14.80	0.60	3.06E-01	1.34E-01	3.06E+04	30.63	1.37E+07	30.2912	dil -1	3.02912	4.62	2.38	7.00
	adhered to CHO^{ICAM}	1.00E-04	11.73	0.88	3.03E+00	2.29E+00	3.03E+04	30.30	1.35E+07	40.5512	dil -1	4.05512	3.45	3.55	7.00
	non-adhered to CHO^{ICAM}	1.00E-06	11.42	0.22	3.22E+00	4.74E-01	3.22E+06	3220.0	1.44E+09	4100.68	dil -3	4.10068	3.41	3.59	7.00

V: Volume; V$_T$: Total Volume; bp: base pairs.

Table 6. Raw data (reads) description output from RNA-sequencing.

Sample ID	Lane	Index	Sample code	Description	File name	Total Sequences (FastQ)	Total nº of reads after trimming (%)	Sequence length	%GC	Sample Ref
1	1	ATCACG		adhered to CHO^{ICAM}	1_S1_L001_R1_001.fastq.gz 1_S1_L001_R2_001.fastq.gz	348167 348167	345888 345888 (88.3%)	101 101	48 49	P. vivax Sal-1/P01
2	1	CGATGT		adhered to CHO^{745}	2_S2_L001_R1_001.fastq.gz 2_S2_L001_R2_001.fastq.gz	1082819 1082819	1070871 1070871 (98.9%)	101 101	48 48	P. vivax Sal-1/P01
3	1	TTAGGC	92U15	non-adhered to CHO^{ICAM}	3_S3_L001_R1_001.fastq.gz 3_S3_L001_R2_001.fastq.gz	1021141 1021141	1003388 1003388 (98.3%)	101 101	48 48	P. vivax Sal-1/P01
4	2	ATCACG		non-adhered to CHO^{745}	4_S12_L001_R1_001.fastq.gz 4_S12_L001_R2_001.fastq.gz	223979 223979	219624 219624 (98.1%)	101 101	49 50	P. vivax Sal-1/P01
21	1	CAGATC		adhered to CHO^{745}	21_S7_L001_R1_001.fastq.gz 21_S7_L001_R2_001.fastq.gz	481872 481872	475794 475794 (98.7%)	101 101	49 49	P. vivax Sal-1/P01
22	1	TAGCTT	93U15	non-adhered to CHO^{745}	22_S10_L001_R1_001.fastq.gz 22_S10_L001_R2_001.fastq.gz	225383 225383	214949 214949 (95.4%)	101 101	43 45	P. vivax Sal-1/P01
5	2	CGATGT		adhered to CHO^{ICAM}	5_S13_L001_R1_001.fastq.gz 5_S13_L001_R2_001.fastq.gz	683478 683478	663014 663014 (97.0%)	101 101	43 44	P. vivax Sal-1/P01
6	1	TGACCA		adhered to CHO^{745}	6_S4_L001_R1_001.fastq.gz 6_S4_L001_R2_001.fastq.gz	680245 680245	670600 670600 (98.6%)	101 101	44 45	P. vivax Sal-1/P01
7	1	ACAGTG	105U15	non-adhered to CHO^{ICAM}	7_S5_L001_R1_001.fastq.gz 7_S5_L001_R2_001.fastq.gz	1312365 1312365	1288122 1288122 (98.2%)	101 101	47 47	P. vivax Sal-1/P01
8	1	GGCTAC		non-adhered to CHO^{745}	8_S11_L001_R1_001.fastq.gz 8_S11_L001_R2_001.fastq.gz	192240 192240	188117 188117 (97.9%)	101 101	48 49	P. vivax Sal-1/P01

20	1	GCCAAT		non-adhered to CHO[ICAM]	20_S6_L001_R1_001.fastq.gz 20_S6_L001_R2_001.fastq.gz	231269 231269	226387 226387 (97.9%)	101 101	46 47	P. vivax Sal-1/P01
23	1	ACTTGA	**101U15**	adhered to CHO[745]	23_S8_L001_R1_001.fastq.gz 23_S8_L001_R2_001.fastq.gz	761780 761780	754181 754181 (99.0%)	101 101	48 49	P. vivax Sal-1/P01
24	1	GATCAG		non-adhered to CHO[745]	24_S9_L001_R1_001.fastq.gz 24_S9_L001_R2_001.fastq.gz	493265 493265	475848 475848 (96.5%)	101 101	47 48	P. vivax Sal-1/P01

CHAPTER 3

"*Plasmodium vivax* and Rosette Formation:

what can transcriptomics tell us?"

Overview

In this study, we aim to understand the molecular mechanisms responsible for rosette formation by identifying possible molecules, especially parasitic ligands, which might be important in *P. vivax* adhesion capacity. Using high-throughput RNA-seq technology coupled with parasite field sample enrichment, *ex vivo* maturation and rosetting assays, we have sequenced the whole transcriptome of parasite populations. Differential expression analysis presented here was accessed by comparison between rosette-forming *P. vivax* samples against their non-forming counterpart.

Taken together, our results point out to the importance of membrane and membrane-associated proteins, whose sequence predict them to be adhesin or adhesin like, and thus important in rosetting phenotype. We expect the results to reflect this parasite pathobiology, principally concerning its adhesive capacity, possible the source of the severe clinical manifestations globally reported.

PRELIMINARY VERSION

Plasmodium vivax and rosette formation:

what can transcriptomics tell us?

Catarina Bourgard[1], Stefanie C. V. Lopes[3,4], Letusa Albrecht[1,2]* and Fabio T. M. Costa[1]*

[1]Laboratory of Tropical Diseases – Prof. Dr. Luiz Jacintho da Silva, Department of Genetics, Evolution, Microbiology and Immunology, University of Campinas – UNICAMP. Campinas, SP, Brazil
[2]Instituto Carlos Chagas, Fundação Oswaldo Cruz - FIOCRUZ, Curitiba, PR, Brazil
[3]Instituto Leônidas e Maria Deane, Fundação Oswaldo Cruz – FIOCRUZ, Manaus, AM, Brazil
[4]Fundação de Medicina Tropical Dr. Heitor Vieira Dourado – FMT-HVD, Gerência de Malária,Manaus, AM Brazil.

* Correspondence:

Prof. Dr. Fabio T. M. Costa

University of Campinas – UNICAMP

Department of Genetics, Evolution and Bioagents

Institute of Biology

Laboratory of Tropical Diseases – Prof. Dr. Luiz Jacintho da Silva

13083-864, Campinas, SP, Brazil.

fabiotmc72@gmail.com

or

Prof. Dr. Letusa Albrecht

Instituto Carlos Chagas, Fundação Oswaldo Cruz

Apicomplexan Research Laboratory

81350-010, Curitiba, PR, Brazil

letusa.albrecht@fiocruz.br or letusaa@gmail.com

CB: catbourgard@gmail.com
SCVL: stefaniecplopes@gmail.com

Abstract

The mechanisms underlying the pathobiology of the neglected *P. vivax* involved in the severe manifestations observed for vivax malaria patients worldwide are still little known. The ability of *P. vivax* to promote the host reticulocyte deformability has been demonstrated and might be the principle behind splenic clearance avoidance. Functional studies have already reported that adhesion of normocytes to the *P. vivax* infected red blood cells (PvIRBCs) is strong and results in stable rosette formation. Mature staged parasites (schizonts) were also reported presenting a higher capacity for rosetting than younger ones. More recently, it was shown that there is a correlation between rosette formation and altered membrane deformability of PvIRBCs, where the rosette-forming PvIRBCs are significantly more stiff and rigid than their non-rosetting equals. Thus, rosette-forming PvIRBCs may be the cause for the lower rates of schizonts peripheral circulation, contributing for parasite sequestration phenomena in the host microvasculature and/or spleen, and consequently, the rheopathological characteristics present in vivax malaria disease. In this study we aim to understand the molecular mechanisms responsible for rosette formation by identifying possible molecules, especially parasitic ligands, which might be important in *P. vivax* adhesion capacity. Using high-throughput RNA-seq technology coupled with parasite field sample enrichment, *ex vivo* maturation and rosetting assays, we have sequenced the whole transcriptome of parasite populations. Differential expression analysis presented here was accessed by comparison between rosette-forming *P. vivax* samples against their non-forming counterpart. Taken together, our results point out to the importance of membrane and membrane-associated proteins, whose sequence predict them to be adhesin or adhesin like, and thus important in rosetting phenotype. We expect the results to reflect this parasite pathobiology, principally concerning its adhesive capacity, possible the source of the severe clinical manifestations globally reported.

Introduction

Plasmodium vivax is the most prevalent malaria parasite outside Sub-Saharan Africa and the most geographically widespread type of malaria, placing billions of people at risk of infection, thus imposing major health and economic burdens [1]. Infection occurs in genetically distinct populations with heterogeneous resistance to chloroquine [2-5]. Severe clinical complications, although scarce [6], have been of great concern [6] and could be directly associated to parasite adherence [7, 8].

One very important adhesive phenotype associated with clinical complications in falciparum malaria is the capacity of PfiRBCs rosette formation, characterized by the ligation of an infected erythrocyte to two or more healthy erythrocytes [9]. Several studies have investigated an association between the ABO blood type and the rosetting process suggesting an important role [10, 11]. In *P. falciparum*, some scientific publications indicate that patients with A blood type showed a greater rosette presence than O blood type patients, which points for an enhanced protection against the occurrence of severe malaria outcomes [10, 11].

Although these phenomena have already been quite well described and studied for *P. falciparum*, little is known about the formation of rosettes in *P. vivax* or any other *Plasmodium* spp.. In falciparum malaria, the *P. falciparum* Erythrocyte Membrane Protein 1 (PfEMP1) has been identified as the main protein responsible for rosette formation by biding to different healthy erythrocyte receptors, such as CD35, heparan sulfate and blood group antigens. However, orthologs in other species have yet to be identified. Other PfiRBCs surface proteins, such as RIFIN proteins, codified by a subtelomeric polymorphic multigene family [12], contribute to the parasite antigenic variability and can also be used as immune inhibitory receptors to achieve immune evasion [13].

The first report about PviRBCs rosettes [14] was published more than 20 years ago, but until now, few scientific experiments have explored and described the *P. vivax* rosetting phenotype [14-18]. In contrast to *P. falciparum*, the relation between *P. vivax* rosetting, disease severity, parasitemia and blood type is unknown [14, 15, 18].

Russell, B. and colleagues have demonstrated that, to accomplish *P. vivax* field isolate enrichment in parasite, it is necessary to use trypsin to disrupt the present rosettes [18], suggesting the existence of parasitic ligands involved in the process of *P. vivax*

rosette formation. During the last years and in cooperation with Russel, B and colleagues, we have explored rosette formation traits [19, 20], drawing already some conclusions. Both incidence and percentage of rosetting in patient samples is more common in vivax than in falciparum malaria, which occurs as soon as 20h after reticulocyte invasion by *P. vivax* asexual and sexual stages [20]. Some traits, such as host ABO blood group, reticulocyte amount and *P. vivax* parasitemia do not significantly correlate with enhanced or diminished *P. vivax* rosetting capacity. The rosette complex structure is based preferentially on mature erythrocytes (normocytes), where Glycophorin C receptor presence seems to have an important role [20]. As it is known, PviRBCs have the rheological properties altered, principally in decrease of membrane elasticity, which enables them to avoid splenic clearance [19]. According to recent studies on deformability of PviRBCs, rosette-forming iRBCs are distinctly more rigid than their non-rosetting counterparts, with long adhesion to normocytes, which suggests a high contribution of rosettes to the sequestration of schizonts PviRBCs in the host microvasculature and/or spleen [19].

Here, our efforts concentrate in better understanding the expression patterns that could suggest which parasite ligands or metabolic pathways might be involved in the reshape of normocytes cell membrane and properties that are related to *P. vivax* rosetting capacity. For that, we performed several rosetting assays on Amazonian low parasitemia clinical isolates, in order to separate two groups of parasite population samples, with and without rosetting capacity. From whole transcriptome sequencing data analysis, we accessed the differential gene expression profiles between these two groups of samples to dissect by data mining, possible differences that tentatively might explain *P. vivax* rosetting phenotype during the progress of vivax malaria disease.

Results

Evaluation of Brazilian P. vivax *isolates rosetting phenotype*

To investigate the rosetting capacity of different populations of PviRBCs from Brazilian Amazonian endemic field, we collected a total of 26 vivax malaria patient samples (Supplementary Table 1). The patient pool was mainly constituted by male

individuals (2.25 F:M ratio) with an average 38.6 years old and presented a range of 500 to 10^4 parasites per µL, counted on a thick smear by the specially trained microscopists from malaria diagnosis service at FMT-HVD (Supplementary Table 1). All samples were processed immediately after collection following the procedures for parasite isolation and enrichment to obtain parasitemias >50%, and thus, the greatest total number of PviRBCs to enable us to proceed with a total of 27 rosetting assays (Table 1). Taking into consideration the stage of maturation of *P. vivax* parasites predominant on the clinical isolates, 10 underwent *ex vivo* culture to allow maturation of the parasites from younger to more mature stages (in general trophozoites and schizonts) before proceeding for rosetting assays (Table 1). By subdividing our assays according to the predominant parasite stages used as young trophozoites (n=10), trophozoites (n=8) and schizonts (n=9) (Table 1), we observed a statistically significant enhanced capacity for rosetting formation of more mature PviRBCs compared with more young staged samples (ANOVA statistical test p-value=0.0004; Figure 1).

P. vivax *low input cDNA synthesis, library preparation and sequencing*

Considering the particularities of *P. vivax* samples, we followed an experimental design that allowed the transcriptome sequencing of a set of 8 PviRBCs pre-selected samples, that could reflect expression profiles characteristic of *P. vivax* rosetting phenotype. RNA extractions of PviRBC populations were done using the RNeasy® Micro kit and its quantity and quality evaluated using the Bioanalyzer® platform (Supplementary Table 2). On average 21.939 pg/uL of RNA, ranging from 4.925 to 65.531 pg/uL, was obtained with an average 8.0 RIN (6.3-9.8). Given the low amounts of *P. vivax* RNA, we opted for the use of SMART® technology, which offers unparalleled sensitivity, unbiased amplification of cDNA transcripts from low input RNA samples (Supplementary Table 2), which is a tremendous advantage since the huge limitation imposed by low parasite burden in vivax malaria patients. Immediately after, the cDNA output was converted into sequencing templates suitable for cluster generation and high-throughput sequencing resulting into a sequencing-ready library for the Illumina® platform (Supplemental Table 3). Same procedures were performed for *P. falciparum* FCR3 S1.2 and S20 strains.

Whole Transcriptome Shotgun Sequencing data analysis

We obtained a total number of 4,248,722 raw reads. On average 455,253 paired end reads (100 bp) (Supplemental Figure 1) per sample from the 8 sample libraries were successfully sequenced (Supplemental Table 4), with an average 46.4% GC content. FastQC of the total number of raw reads obtained from all our libraries reveled good sequence quality and the necessary trimming steps only excluded a minor fraction of reads (Supplementary Table 4), in general, repetitive, not accurately determined and Illumina® adaptors run through sequences. Using *P. vivax* P01 reference genome, we were able to align and map on average half of the total number of trimmed reads obtained to annotated protein-coding genes (Table 2). Sequences showing multiple or discordant alignments were excluded from the analysis. The same analysis pipeline was followed for the analysis of *P. falciparum* FCR3 S1.2 and S20 raw data, using *P. falciparum* IT and 3D7 reference genomes (Table 3).

Differential expression profiles associated to *P. vivax* rosetting phenotype

To proceed for RNA-seq of *P. vivax* iRBCs, we separate two groups of parasite population samples, with high (9-29% rosetting rates; 62U15, 66U15.2, 69U15 and 106U16) and low rosetting capacity (3-6% rosetting rates; 63U15, 65U15, 73U15 and 109U15). Through RNA-seq data analysis, we accessed the differential gene expression profiles between samples of these two groups and dissect by data mining, possible differences that tentatively might explain *P. vivax* rosetting phenotype during the progress of vivax malaria disease. Although expression profiles were similar between samples within the same capacity level for rosette formation, our *P. vivax* populations isolated from malaria patients showed some degree of transcriptome heterogeneity. Nevertheless, analysis between the high (69U15 and 106U16) against the low (73U15 and 109U16) rosetting samples revealed a group of 94 differential expressed genes (q-value<0.05, Supplementary Table 5). A big portion of those genes (31) codify conserved *Plasmodium* spp. proteins of yet unknown function. With the exception of only two genes (PVP01_1231000 and PVP01_0944300) with negative log2 fold changes (Bozdech, Mok et al. 2008), all genes showed a high expression (1.8>log2 Fold Change>3.5) in our group

of high rosetting samples, which might suggest their involvement in molecular processes important for rosetting or the adhesion binding itself. Future functional characterization of these proteins should clarify this possibility. Numerous other genes (38) are enzyme-coding proteins responsible for DNA/cremation organization packaging and assembly or other cell metabolic processes, including 5 upregulated kinases, some have/are being characterized (NEK3, MAPK2 and RKIP). Interestingly, we could identify an important group of 20 membrane or membrane-associated proteins (Table 4). Within this group, we found two 6-cystein proteins (P47 and P48/45), 2 LCCL (CCp2 and LAP5), one PH domain-containing protein and one WD-repeat domain-containing protein. Also upregulated were the PhL1 interacting proteins PIP2 and 3, the early transcribed membrane protein (ETRAMP), MSP7-like protein and 3 genes from CPW-WPC protein family. Only one gene described as a leucine-rich repeat protein showed a downregulated expression (Log2 Fold Change=2.5). Importantly, 4 glideosome-associated protein also appear to be highly expressed in our rosetting isolates. Curiously, our analysis was able to pick up the significant downregulated expression of the *P. vivax* macrophage migration inhibitory factor (MIF) gene.

Among the upregulated genes, four were predicted to be adhesins or adhesin-like, as MSP7 and three CPW-WPC genes (Table 5).

Discussion

To have a better understanding of *P. vivax* expression patterns could be involved in rosetting capacity, we performed several rosetting assays on Amazonian low parasitemia clinical isolates. In accordance with previously published data [19], we could observe and confirm a proportional relation between PviRBCs maturation and its enhanced capacity to form rosettes, where young staged parasites showed a low percentage of rosetting (<6%) that progressively augments towards schizont parasites, which can rosette to a 50% rate or more (Figure 1). This result is indicative of the fact that *P. vivax* must be expressing proteins responsible for the reshape of normocytes cell membrane and properties, which might be directly involved in rosette formation.

Through RNA-seq of *P. vivax* iRBCs, we evaluated the differential gene expression profiles between two groups with different rosetting capacity (high against low) and identified expressed genes that could explain *P. vivax* rosetting during the progress of vivax malaria disease. Considering the confounding expression variability expected between different *P. vivax* clinical isolates, our analysis between the double sample high and low rosetting samples revealed a group of 94 differentially expressed genes.

As anticipated, a large number of genes have not yet been characterized for their protein function, but the conservation of their sequence throughout *Plasmodium* spp. may be indicative of their importance on parasite rosetting phenotype. Furthermore, majority of this conserved *Plasmodium* showed high upregulated expression, suggesting their involvement in molecular processes important for erythrocyte binding. These proteins functional characterization should further elucidate this possibility.

Within the group of differentially expressed enzymes, 3 upregulated kinases have caught our attention, since kinases are classic targets for discovery of new therapeutic drugs. NEK3 has been reported has essential for mitosis progression in *P. berghei* blood-stage development [21], MAPK2 seems to play an important role in stress response in *Toxoplasma gonddi* [22] and RKIP affects activity of another kinase, the calcium-dependent protein kinase 1 [23], which regulates several important *Plasmodium* spp. reliant on calcium metabolic processes. Studies on calcium homeostasis have reported that parasitized RBCs show an increased influx of calcium when compared to the decreased efflux of unparasitized RBCs. Calcium content has been localized in the *Plasmodium* spp. compartment. For the RBC invasion by the merozoite, extracellular calcium is needed, as well as the subsequently parasite development and maturation inside of the erythrocyte [24, 25].

P. vivax field isolates are often characterized by asynchronous populations of parasites in different stages of develop and/or maturation. Although our samples were chosen with the aim to access the transcriptomic profiles of trophozoites and/or early schizont parasites, we could also catch the expression of some interesting gametocyte membrane surface genes, reflecting the importance of the study of mechanisms of *P. vivax* transmission. Together with P48/45 surface protein [26], P47 is one of such proteins,

having been reported as requires for optimal fertilization in *P. berghei* and for mosquito immune evasion, showing a strong signature of natural selection and population structure in the *P. falciparum* and *P. vivax* genomes [27].

Also, we verified differential expression of two LCCL lectin domain adhesive-like proteins (LAPs), a family of conserved six modular proteins, conserved through the apicomplexan genus, which are expressed in sexual stages of *Plasmodium* parasites and reported to be involved in the formation of protein complexes required for successful *P. berghei* sporogony [28, 29].

One of the most important aspects of erythrocyte infection by *P. vivax* is the dramatic morphogenesis of the erythrocyte membrane, driven by a network of microtubules (MT) sustained by the inner membrane complex (IMC). As expected, genes codifying the actin and tubulin backbone molecules of MTs were found to be upregulated in our rosetting samples, together with a group of MT motor enzymes, IMC proteins such as PhIL1 (PIP2 and 3-)integrating proteins, which are critical in various processes such as signal transduction and intracellular and membrane trafficking [30, 31].

Furthermore, we were able to catch the overexpression of an early-transcribed membrane protein (ETRAMP). ETRAMPs are important proteins present on the membrane of intracellular parasites of *Plasmodium* species, formed during erythrocyte invasion as an invagination of the iE cell surface during the asexual blood stage parasites. Recent studies showed that ETRAMPs have been localized on the intracellular membranes of immature schizont and at the apical organelles of newly *formed P. vivax* merozoites of mature schizont and have the capacity to elicit high antibody titers capable of recognizing parasites of vivax malaria patients [32].

Together with other 3 expressed genes from CPW-WPC surface protein family, another important membrane protein found differentially expressed in our study was the *P. vivax* (MSP7)-like protein. Merozoite surface proteins belong to families of proteins often involved in complex *Plasmodium* invasion processes. Pf MSP7 interactions with host P-selectin receptors have been demonstrated [33], which in consequence block interactions between host P-selectin and leukocyte ligands and could underlie the mechanism for the known immunomodulatory effects of both MSP7 and P-selectin in malaria infection models. Although *msp7* in *P. vivax* has not yet been functionally

characterized, there is evidence this protein is under selection and thus, being functionally important in *P. vivax* [34].

In addition, 4 glideosome-associated proteins are observed to be differentially expressed in our high rosetting parasites. The capacity to bind, reorient and invade new host cells is mainly powered by the "glideosome" proteins [35]. The glideosome is a macromolecular complex comprising proteins with adhesive properties. These proteins are released apically on the parasite membrane and translocated to the opposite pole of the parasite through the actomyosin system anchored in the IMC.

Finally, we have detected the expression of macrophage migration inhibitory factor (MIF) gene, one of the first cytokines described, which has a broad range of pro-inflammatory properties. It has been reported that expressed PfMIF protein localizes to the Maurer's cleft during asexual blood stage parasites. PfMIF *in vitro* treatment of human monocytes inhibited their random migration and reduced the surface expression of toll like receptor (TLR) 2, TLR4 and CD86, indicating that its release potentially modulates the host monocytes functions during *Plasmodium* acute infection [36]. In accordance with this data, our analysis performed in isolates from non-severe vivax malaria patients reported a significant downregulated expression of the *P. vivax mif* gene in parasite populations showing rosetting phenotypes.

Conclusion

Taken together, these results point out to the importance of membrane and membrane associate proteins in rosetting phenotype. Functional assays might further clarify of these proteins allow the parasites to adhere to the surface of host cells, such as health erythrocytes, maintaining them anchored in order to create the characteristic rosette of surrounding erythrocytes, which enable the parasite to evade from the host immune system.

Methods

Ethical Approval

Informed consent was sought and granted from all patients attending the *Fundação de Medicina Tropical Dr. Heitor Vieira Dourado* (FMT-HVD), Manaus, Amazonas State, Brazil. All procedures, including protocols and consent forms, were approved by the Ethics Review Board of FMT-HVD (process CAAE-0044.0.114.000-11 and 54234216.0000.0005).

Study Area, Subjects and Sample Collection

Adult patients were recruited at FMT-FVD Manaus, Amazonas State, Brazil, a tertiary care centre for infectious diseases following a pipeline of microscopic diagnosis of uncomplicated *P. vivax* malaria and determination of parasitaemia, after which the patients are treated with Chloroquine and Primaquine according to the standard protocol recommended by the Brazilian Malaria Control Program. Before initiation of treatment, up to 8 mL of peripheral blood was collected from each patient in citrate-coated Vacutainer^{TM} tubes (Becton-Dickinson). Severe malaria, patients under anti-malarial treatment, with *P. falciparum* malaria and/or *P. falciparum* and *P. vivax* mixed infections, and pregnant women were excluded from this study. *P. vivax* mono-infection was confirmed by PCR analysis, as described elsewhere [37].

Parasite Isolation, Enrichment and ex vivo Maturation

To obtain enriched *P. vivax* infected erythrocytes (PviRBCs), samples were immediately processed. Plasma and buffy coat layer were removed after centrifugation at 400 x g for 5 min at room temperature. The pellet was ressuspended in an equal volume of RPMI parasite medium and then performed CF11 column filtration (Sigma) to deplete white blood cells [18, 38, 39], followed by parasite enrichment through Percoll 45% gradient protocol as previously described [40]. Depending on the stage of parasite maturation, the early blood staged parasites were cultured for 18-22 hours to allow them to mature to late trophozoites and/or schizonts [18, 41]. Thin blood smears were prepared

and stained with *Panótico Rápido* (Laborclin®) kit, before, during and after *ex vivo* short culture to control the extent of parasite maturation.

Rosetting rates assessment

Twenty µL of PvRBCs at 2.5-5% parasitemia and 2.5-5% hematocrit were incubated for 40 min at 37°C in rosetting media (McCoy's 5A medium supplemented with 20% of patient autologous plasma). Triplicate aliquots of each sample were stained with 45 µg/ml of acridine orange and examined by direct light and fluorescence microscopy (Nikon Eclipse 50i, filter 96311 B-2E/C). Rosetting rates were access by counting 200 Pv-iRBCs in triplicate. As our criteria, the rosette complex was defined by the binding of two or more uninfected erythrocytes to an Pv-iE.

P. falciparum *in vitro cultures*

The Brazilian field isolate S20 (resistant to artemisinin) and *P. falciparum* FCR3 S1.2 [42] strains were cultured according to standard procedures previously described [43] and the last was enriched up to ≥ 80% of rosetting. For both strains, about 10^6 PfiRBCs were isolated for RNA extraction and downstream transcriptomic sequencing.

RNA Extraction and Quality Control

After PviRBCs quantification, the isolate was preserved by flash freezing the sample in liquid nitrogen. RNA extractions were accomplished by using the RNeasy® Micro kit (Qiagen) according to the manufacturer instructions. Before attempting RNA-seq, quality control was done by running electrophoretically the extracted RNA samples in the Agilent 2100 Bioanalyzer instrument using the Agilent RNA 6000 Pico Kit reagents and chips and analyzed on the 2100 Expert software, according to the ©Agilent Technologies recommendations.

Low Input cDNA synthesis and Library Preparation for Whole Transcriptome Shotgun Sequencing

SMART-Seq V4 Ultra Low Input RNA kit was used for cDNA libraries generation. cDNA quality, quantity and size range were evaluated through BioAnalyser using the

Agilent High Sensitivity DNA Kit (cDNA, 5 to 500 pg/µL within a size range of 50 to 7000 bp), as per manufacturer instructions. Covaris AFA system was used for controlled cDNA shearing, resulting in DNA fragments between 200 and 500 bp sizes. Instructions were followed as indicated in the SMART-Seq V4 Ultra Low Input RNA kit for sequencing user manual by Clontech Laboratories, Inc. A Takara Bio Company. cDNA output was then converted into sequencing templates suitable for cluster generation and high-throughput sequencing through the Low Input Library Prep v2 (Clontech Laboratories, Inc. A Takara Bio Company). Library quantification procedures using the Library Quantification kit (Clontech Laboratories, Inc. A Takara Bio Company) by the golden standard qPCR and Agilent's High Sensitivity DNA kit were successfully completed before proceeding for the pool set-up) at a final concentration of 2nM for direct sequencing. The generated libraries were cluster amplified and sequenced on the Illumina® platform using standard Illumina® reagents and protocols for multiplexed libraries by following their loading recommendations. Sequencing runs were performed on HiSeq 2500 sequencer on Rapid Run mode with the HiSeq Rapid Cluster Kit v2 (100x100) Paired End, HiSeq Rapid SBS Kit v2 (200 cycles) and HiSeq® Rapid Duo cBot v2 Sample Loading kits from Illumina®, Inc..

Raw reads Alignment and Mapping

Our data was analysed using the EuPathDB-Galaxy hub (https://eupathdb.globusgenomics.org/) free, interactive, web-based platform for large-scale data analysis by assembling a new workflow adapted for our RNA-seq experimental design on Galaxy platform [44] and using the PlasmoDB [45] pre-loaded *P. vivax* reference genome [46]. In summary, raw reads were checked for quality by running Fast Quality Control (Galaxy Tool Version FASTQC: 0.11.3; https://www.bioinformatics.babraham.ac.uk/projectY/fastqc/), a java quality control tool for high throughput sequencing data. Illumina® adaptors were trimmed through Trimmomatic (Galaxy Tool Version 0.36.5) on our Illumina® paired-end data [47], and read alignment and mapping was performed with TopHat2 [48] (Galaxy Tool Version SAMTOOLS: 1.2; BOWTIE2: 2.1.0; TOPHAT2: 2.0.14), towards the *Plasmodium vivax* P01 reference genome from PlasmoDB release 38 known transcripts and splice junctions. FPKM

estimation by count of the number of aligned reads matching the annotated reference genes was executed with htseq-count [49] (Galaxy Tool v. HTSEQ: default; SAMTOOLS: 1.2; PICARD: 1.134) . Final differential gene expression was performed with DESeq2 [50] Galaxy Tool (v. 2.1.6.0). Differential gene expression between the different analysis groups was identified after a pairwise Wilcoxon test was used to compare the transcriptional profiles with the following cutoffs: p-value<0.05, q-value<0.5 and a log2 fold change > 1.5 (Supplementary Figure 2).

Differential expression profiles, Gene Ontology enrichment, Metabolic Pathways and MAAP analysis

From the final list of genes computed, we selected those that could be informative about the host expression differences between high against low rosetting *P. vivax* infecting populations, focusing on the relative presence or absence within the sample comparison. For that, we considered the total pool of genes with a p-value < 0.05 and a q-value 0 – 1.0 and performed Gene Ontology enrichment on PlasmoDB (release 40) [45] to find all GO terms associated to biological processes, cellular components and molecular functions (genes computed and curated for a p-value cutoff of 0.05. towards *P. vivax* reference). By using REViGO [51], a web server to remove redundant terms, we visualized in semantic similarity-based scatterplots, interactive graphs and tag clouds, the relevant GO terms. We also searched for Metabolic Pathways based on PlasmoDB link using KEGG (https://www.genome.jp/kegg/) and MetaCyc (https://metacyc.org/) pathway resources (p-value cutoff of 0.05. towards *P. vivax* reference). The Malarial Adhesins and Adhesin-like proteins predictor (MAAP) was used to score the group of differentially expressed membrane proteins regarding its adhesin sequence and structure.

Acknowledgments

Author contributions statement

All authors conceived the experiments. CB did the field sample collection and rosetting assays, designed and executed the RNA-seq experiments and data analysis. CB and LA went through data mining of the obtained results and draft the manuscript. All authors reviewed the manuscript.

References

1. Guerra, C.A., et al., *The international limits and population at risk of Plasmodium vivax transmission in 2009.* PLoS Negl Trop Dis, 2010. **4**(8): p. e774.
2. Baird, J.K., *Resistance to therapies for infection by Plasmodium vivax.* Clin Microbiol Rev, 2009. **22**(3): p. 508-34.
3. Price, R.N., N.M. Douglas, and N.M. Anstey, *New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance.* Curr Opin Infect Dis, 2009. **22**(5): p. 430-5.
4. Tjitra, E., et al., *Multidrug-resistant Plasmodium vivax associated with severe and fatal malaria: a prospective study in Papua, Indonesia.* PLoS Med, 2008. **5**(6): p. e128.
5. Mourao, L.C., et al., *Naturally acquired antibodies to Plasmodium vivax blood-stage vaccine candidates (PvMSP-1(1)(9) and PvMSP-3alpha(3)(5)(9)(-)(7)(9)(8) and their relationship with hematological features in malaria patients from the Brazilian Amazon.* Microbes Infect, 2012. **14**(9): p. 730-9.
6. Lacerda, M.V., et al., *Understanding the clinical spectrum of complicated Plasmodium vivax malaria: a systematic review on the contributions of the Brazilian literature.* Malar J, 2012. **11**: p. 12.
7. Costa, F.T., et al., *On cytoadhesion of Plasmodium vivax: raison d'etre?* Mem Inst Oswaldo Cruz, 2011. **106 Suppl 1**: p. 79-84.
8. del Portillo, H.A., et al., *Variant genes and the spleen in Plasmodium vivax malaria.* Int J Parasitol, 2004. **34**(13-14): p. 1547-54.
9. Carlson, J., et al., *Antibodies to a histidine-rich protein (PfHRP1) disrupt spontaneously formed Plasmodium falciparum erythrocyte rosettes.* Proc Natl Acad Sci U S A, 1990. **87**(7): p. 2511-5.
10. Rowe, J.A., et al., *Blood group O protects against severe Plasmodium falciparum malaria through the mechanism of reduced rosetting.* Proc Natl Acad Sci U S A, 2007. **104**(44): p. 17471-6.
11. Barragan, A., et al., *Blood group A antigen is a coreceptor in Plasmodium falciparum rosetting.* Infect Immun, 2000. **68**(5): p. 2971-5.
12. Bultrini, E., et al., *Revisiting the Plasmodium falciparum RIFIN family: from comparative genomics to 3D-model prediction.* BMC Genomics, 2009. **10**: p. 445.
13. Saito, F., et al., *Immune evasion of Plasmodium falciparum by RIFIN via inhibitory receptors.* Nature, 2017. **552**(7683): p. 101-105.

14. Udomsanpetch, R., et al., *Rosette formation by Plasmodium vivax.* Trans R Soc Trop Med Hyg, 1995. **89**(6): p. 635-7.

15. Chotivanich, K.T., et al., *Characteristics of Plasmodium vivax-infected erythrocyte rosettes.* Am J Trop Med Hyg, 1998. **59**(1): p. 73-6.

16. Chotivanich, K.T., et al., *Rosetting characteristics of uninfected erythrocytes from healthy individuals and malaria patients.* Ann Trop Med Parasitol, 1998. **92**(1): p. 45-56.

17. Chotivanich, K., et al., *Plasmodium vivax adherence to placental glycosaminoglycans.* PLoS One, 2012. **7**(4): p. e34509.

18. Russell, B., et al., *A reliable ex vivo invasion assay of human reticulocytes by Plasmodium vivax.* Blood, 2011. **118**(13): p. e74-81.

19. Zhang, R., et al., *Rheopathologic Consequence of Plasmodium vivax Rosette Formation.* PLoS Negl Trop Dis, 2016. **10**(8): p. e0004912.

20. Lee, W.C., et al., *Glycophorin C (CD236R) mediates vivax malaria parasite rosetting to normocytes.* Blood, 2014. **123**(18): p. e100-9.

21. Tewari, R., et al., *The systematic functional analysis of Plasmodium protein kinases identifies essential regulators of mosquito transmission.* Cell Host Microbe, 2010. **8**(4): p. 377-87.

22. Huang, H., et al., *Molecular cloning and characterization of mitogen-activated protein kinase 2 in Toxoplasma gondii.* Cell Cycle, 2011. **10**(20): p. 3519-26.

23. Kugelstadt, D., et al., *Raf kinase inhibitor protein affects activity of Plasmodium falciparum calcium-dependent protein kinase 1.* Mol Biochem Parasitol, 2007. **151**(1): p. 111-7.

24. Krishna, S. and L. Squire-Pollard, *Calcium metabolism in malaria-infected erythrocytes.* Parasitol Today, 1990. **6**(6): p. 196-8.

25. Leida, M.N., J.R. Mahoney, and J.W. Eaton, *Intraerythrocytic plasmodial calcium metabolism.* Biochem Biophys Res Commun, 1981. **103**(2): p. 402-6.

26. van Dijk, M.R., et al., *A central role for P48/45 in malaria parasite male gamete fertility.* Cell, 2001. **104**(1): p. 153-64.

27. Molina-Cruz, A., G.E. Canepa, and C. Barillas-Mury, *Plasmodium P47: a key gene for malaria transmission by mosquito vectors.* Curr Opin Microbiol, 2017. **40**: p. 168-174.

28. Tremp, A.Z., et al., *LCCL protein complex formation in Plasmodium is critically dependent on LAP1.* Mol Biochem Parasitol, 2017. **214**: p. 87-90.

29. Pradel, G., et al., *A multidomain adhesion protein family expressed in Plasmodium falciparum is essential for transmission to the mosquito.* J Exp Med, 2004. **199**(11): p. 1533-44.

30. Parkyn Schneider, M., et al., *Disrupting assembly of the inner membrane complex blocks Plasmodium falciparum sexual stage development.* PLoS Pathog, 2017. **13**(10): p. e1006659.

31. Ebrahimzadeh, Z., A. Mukherjee, and D. Richard, *A map of the subcellular distribution of phosphoinositides in the erythrocytic cycle of the malaria parasite Plasmodium falciparum.* Int J Parasitol, 2018. **48**(1): p. 13-25.

32. Cheng, Y., et al., *Characterization of Plasmodium vivax Early Transcribed Membrane Protein 11.2 and Exported Protein 1.* PLoS One, 2015. **10**(5): p. e0127500.

33. Perrin, A.J., S.J. Bartholdson, and G.J. Wright, *P-selectin is a host receptor for Plasmodium MSP7 ligands.* Malar J, 2015. **14**: p. 238.

34. Castillo, A.I., M. Andreina Pacheco, and A.A. Escalante, *Evolution of the merozoite surface protein 7 (msp7) family in Plasmodium vivax and P. falciparum: A comparative approach.* Infect Genet Evol, 2017. **50**: p. 7-19.

35. Keeley, A. and D. Soldati, *The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa.* Trends Cell Biol, 2004. **14**(10): p. 528-32.

36. Cordery, D.V., et al., *Characterization of a Plasmodium falciparum macrophage-migration inhibitory factor homologue.* J Infect Dis, 2007. **195**(6): p. 905-12.

37. Snounou, G. and B. Singh, *Nested PCR analysis of Plasmodium parasites.* Methods Mol Med, 2002. **72**: p. 189-203.

38. Auburn, S., et al., *Effective preparation of Plasmodium vivax field isolates for high-throughput whole genome sequencing.* PLoS One, 2013. **8**(1): p. e53160.

39. Sriprawat, K., et al., *Effective and cheap removal of leukocytes and platelets from Plasmodium vivax infected blood.* Malar J, 2009. **8**: p. 115.

40. Carvalho, B.O., et al., *On the cytoadhesion of Plasmodium vivax-infected erythrocytes.* J Infect Dis, 2010. **202**(4): p. 638-47.

41.   Udomsangpetch R, S.R., Williams J, Sattabongkot J., *Modified techniques to establish a continuous culture of Plasmodium vivax [abstract]. .* American Society of Tropical Medicine and Hygiene 51st Annual Meeting program; 2002 Nov 10-14; Denver, CO. , 2001.

42.   Kraemer, S.M., et al., *Patterns of gene recombination shape var gene repertoires in Plasmodium falciparum: comparisons of geographically diverse isolates.* BMC Genomics, 2007. **8**: p. 45.

43.   Kirsten Moll, I.L., Hedvig Perlmann, Artur Scherf, Mats Wahlgren., *Methods in Malaria Research.* Malaria Research and Reference Reagent Resource Center (MR4) and American Type Culture Collection 10801 University Boulevard, Manassas, VA 20110-2209, 2008.

44.   Afgan, E., et al., *The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update.* Nucleic Acids Res, 2018. **46**(W1): p. W537-W544.

45.   Aurrecoechea, C., et al., *PlasmoDB: a functional genomic database for malaria parasites.* Nucleic Acids Res, 2009. **37**(Database issue): p. D539-43.

46.   Auburn, S., et al., *A new Plasmodium vivax reference sequence with improved assembly of the subtelomeres reveals an abundance of pir genes.* Wellcome Open Res, 2016. **1**: p. 4.

47.   Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data.* Bioinformatics, 2014. **30**(15): p. 2114-20.

48.   Trapnell, C., L. Pachter, and S.L. Salzberg, *TopHat: discovering splice junctions with RNA-Seq.* Bioinformatics, 2009. **25**(9): p. 1105-11.

49.   Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.

50.   Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

51.   Supek, F., et al., *REVIGO summarizes and visualizes long lists of gene ontology terms.* PLoS One, 2011. **6**(7): p. e21800.

Figures and Tables
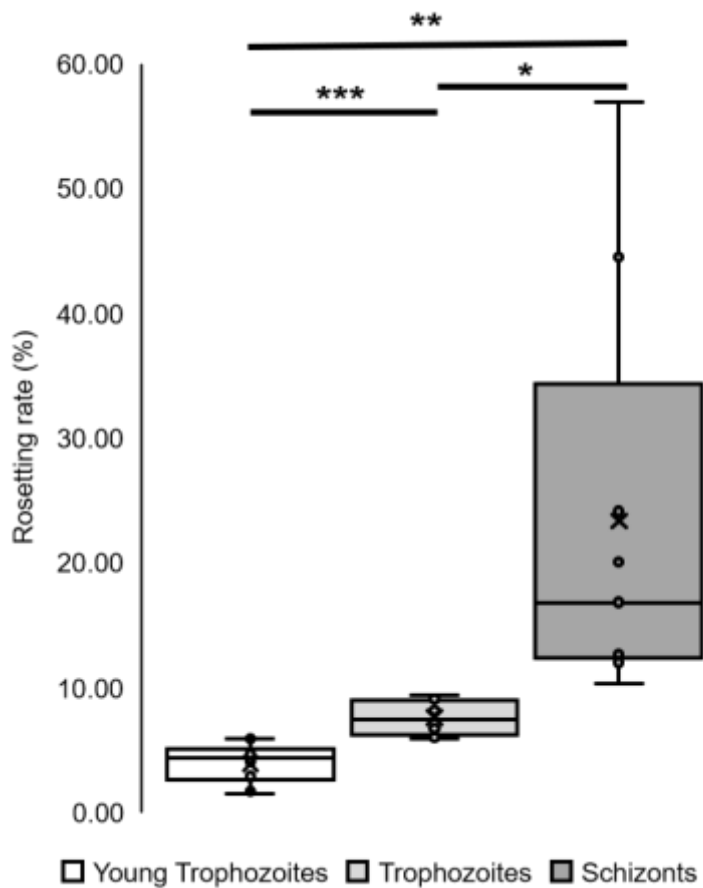
Figure 1. Pv-iE isolate rosetting rates, normalized against the sample's total parasitemia, represented on whiskers plots as percentages of rosette formation (%) according to the most prevalent parasite stages into 3 groups, young trophozoites (white; n=10), trophozoites (light grey; n=8) and schizonts (grey; n=9). Nonparametric one-ways ANOVA (Kruskal-Wallis) statistical test gave statistical significance (p-value<0.0001) between the 3 PviRBCs staged groups. Unpaired t test with Welch's correction 2-tails between young trophozoites - trophozoites (***p-value=0.0001), trophozoites – schizonts (*p-value=0.0206) and young trophozoites – schizonts (**p-value=0.0075) groups.
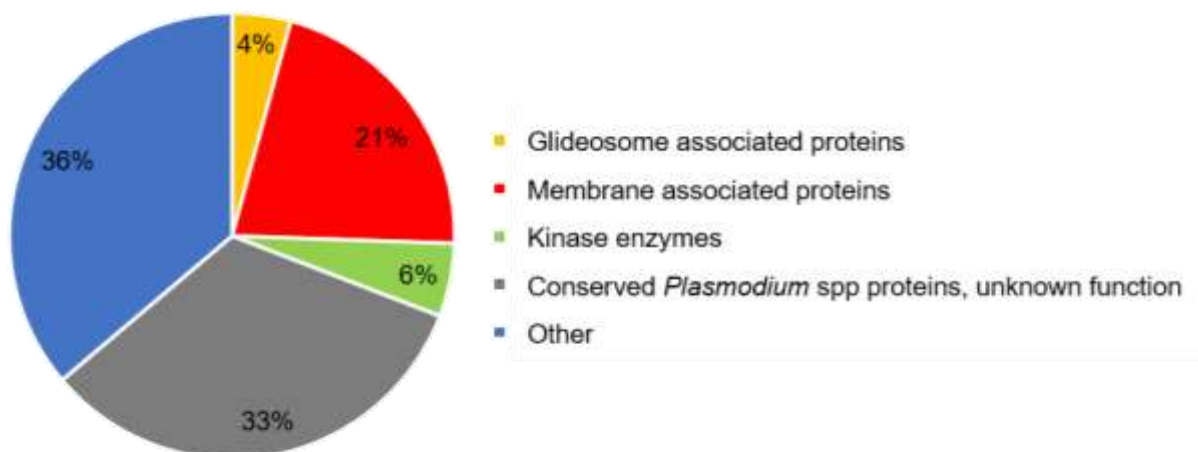
Figure 2. Pie chart showing the 94 differentially expressed genes grouped by protein function in glideosome associated proteins (4 genes; yellow), membrane associated proteins (20 genes; red), kinase enzymes (5 genes; green), conserved *Plasmodium* spp. proteins of unknown function (31 genes; grey) and other (34 genes; blue).

Table 1. Sample profiles for *P. vivax* clinical isolates

Sample Code	Pv-iE/200 leucocytes*	Initial smear stages (h)	*ex vivo* culture /maturation time (h)	Smear after *ex vivo* culture and/or Percoll 45%	Parasitemia after Percoll (%)	Rosetting (%)			
						All rosetting assays	Young Trophozoites	Trophozoites	Schizonts
59U15.1	438	trophozoites (15-22) and schizonts (≥22)	N	trophozoites (15-22) and schizonts (≥22)	87.50%	56.88			56.88
61U15	185	trophozoites (>20) and some schizonts (≥22)	N	schizonts (≥22)	75.80%	44.49			44.49
62U15	335	trophozoites (15-20)	N	trophozoites (15-20)	96.60%	9.37		9.37	
63U15	400	young trophozoites (11-13)	N	young trophozoites (11-13)	84.40%	5.91	5.91		
65U15	428	trophozoites young (12-15)	N	trophozoites (12-15)	84.40%	5.91	5.91		
66U15.1	340	rings (5) and some schizonts (≥22)	Y/16h	young trophozoites (13-18)	97.30%	4.67	4.67		
66U15.2	340	rings (5) and some schizonts	Y/16h	trophozoites (15-20) and some schizontes (≥22)	97.00%	20.08			20.08
67U17	42	young trophozoites (11-12)	Y/16h	trophozoites (18-20)	50.00%	6.00		6.00	
68U15	250	trophozoites (15-22) and some	N	trophozoites (15-22) and some	88.50%	8.76		8.76	

		schizonts (≥22)		schizonts (≥22)					
69U15	430	schizonts (21-26)	N	schizonts (21-26)	98.90%	24.13			24.13
70U15	135	schizonts (≥22)	N	schizonts (22-26)	80.00%	12.00			12.00
71U15	210	rings (5-7)	Y/16h	trophozoites (15-20)	74.50%	8.20		8.20	
72U15.1	1050	rings (4-5), trophozoites (15) and schizonts	Y/16h	young trophozoites	34.70%	1.49	1.49		
72U15.2	1050	rings (4-5) and trophozoites (15)	N	young trophozoites (11-15)	56.70%	1.70	1.70		
73U15	ND	rings (4-6) and few mature trophozoites /schizonts	N	trophozoites (11-17)	83.20%	4.66	4.66		
76U15	270	young trophozoites (11-15)	N	young trophozoites (11-15)	53.30%	5.86		5.86	
78U15.1	326	rings (few), mature trophozoites (15-22) and schizonts (≥22)	N	trophozoites (15-21) and schizonts (≥22)	93.00%	13.02			13.02
82U15.2	430	mature trophozoites and schizonts (≥15)	Y/16h	mature trophozoites and schizonts (≥15)	64.20%	10.27			10.27
84U15.1	210	trophozoites (11-15) and some schizonts	Y/1h	young trophozoites (11-15)	96.40%	2.89	2.89		

84U15.2	210	trophozoites (11-15) and some schizonts	Y/5h	trophozoites (>15)	96.40%	6.75		6.75
87U15	220	trophozoites young (11-15) and (some) schizonts	N	trophozoites (15-22	56.80%	9.09		9.09
94U15	230	rings and young trophozoites (11-13)	Y/18h	trophozoites (11-15)	67.10%	4.03	4.03	
95U15	120	young trophozoites young (11-15)	N	young trophozoites (11-15)	62.80%	4.84	4.84	
98U15.1	370	trophozoites (11-15) and some schizonts	N	mature trophozoites (≥15)	67.70%	6.77		6.77
98U15.2	370	trophozoites (11-15) and schizonts	Y/17h	schizonts (≥22)	99.00%	16.83			16.83
106U16	230	mature trophozoites (15-22)	N	mature trophozoites (15-22) and schizonts (≥22)	42.20%	12.66			12.66
109U16	480	rings, young trophozoites (11-15)	N	young trophozoites (11-15)	50.00%	3.00	3.00	

Summary of sample information during parasite enrichment and *ex vivo* maturation of clinical isolates and on rosetting assays performed during field work at malaria vivax endemic area, adequate for RNA-seq downstream application. Y – Yes; N – No; ND – Not Determined; All samples were preserved through N_2 flash freezing after processing and during transportation. RNA sequenced samples are indicated in light grey.

Table 2. Alignment and mapping summary report for *P. vivax* clinical isolates RNA sequenced.

	Sample Code							
	62U15	**66U15**	**69U15**	**73U15**	**106U16**	**109U16**	**63U15**	**65U15**
Left reads								
Input	546571	393248	436832	457871	726393	244626	550489	218478
Mapped (% of input)	278618 (51.0%)	185087 (47.1%)	220515 (50.5%)	268217 (58.6%)	241490 (33.2%)	153683 (62.8%)	331462 (60.2%)	127874 (58.5%)
Multiple alignments (% of mapped)	7705 (2.8%)	2873 (1.6%)	4044 (1.8%)	4963 (1.9%)	4997 (2.1%)	2479 (1.6%)	6071 (1.8%)	3103 (2.4%)
Multiple alignments (>20)	89	15	27	30	55	17	32	20
Right reads								
Input	546571	393248	436832	457871	726393	244626	550489	218478
Mapped	277677 (50.8%)	184911 (47.0%)	219709 (50.3%)	267472 (58.4%)	241068 (33.2%)	153046 (62.6%)	330121 (60.0%)	127278 (58.3%)
Multiple alignments	8896 (3.2%)	3363 (1.8%)	4561 (2.1%)	5608 (2.1%)	6237 (2.6%)	2776 (1.8%)	6813 (2.1%)	3376 (2.7%)
Multiple alignments (>20)	90	15	27	30	57	17	31	20
Overall read mapping rate	**50.90%**	**47,0%**	**50.40%**	**58.50%**	**33.20%**	**62.70%**	**60.10%**	**58.40%**
Aligned pairs	**232635**	**169102**	**198574**	**241964**	**210760**	**139372**	**301369**	**114319**
Multiple alignments (% of aligned pairs)	5007 (2.2%)	2021 (1.2%)	2864 (1.4%)	3680 (1.5%)	2816 (1.3%)	1955 (1.4%)	4825 (1.6%)	2259 (2.0%)
Discordant alignments (% of aligned pairs)	7591 (3.3%)	1297 (0.8%)	3479 (1.8%)	3272 (1.4%)	3360 (1.6%)	2646 (1.9%)	4114 (1.4%)	1374 (1.2%)
Concordant pair alignment rate	**41.20%**	**42.70%**	**44.70%**	**52.10%**	**28.60%**	**55.90%**	**54.00%**	**51.70%**

Data table summarizing the alignment and mapping results against the *P. vivax* P01 reference genome obtained using TopHat2 on our raw data (reads) for *P. vivax* clinical isolates RNA sequenced. Both pair of reads (left and right) we mapped, checked for multiple alignments and concordance of pair alignment.

Table 3. Alignment and mapping summary report for *P. falciparum* FCR3 S1.2 and S20 parasite cultures RNA sequenced.

P. falciparum Reference Genome	*Pf* IT		*Pf* 3D7	
P. falciparum cultures	*Pf* FCR3 S1.2	*Pf* S20	*Pf* FCR3 S1.2	*Pf* S20
Left reads				
Input	537931	379684	537931	379684
Mapped (% of input)	299290 (55.6%)	216060 (56.9%)	291552 (54.2%)	222644 (58.6%)
Multiple alignments (% of mapped)	33175 (11.1%)	13834 (6.4%)	53244 (18.3%)	19376 (8.7%)
Multiple alignments (>20)	237	71	326	161
Right reads				
Input	537931	379684	537931	379684
Mapped	301231 (56.0%)	215955 (56.9%)	293297 (54.5%)	222754 (58.7%)
Multiple alignments	35211 (11.7%)	13229 (6.1%)	54969 (18.7%)	18833 (8.5%)
Multiple alignments (>20)	239	69	328	160
Overall read mapping rate	**55.80%**	**56.90%**	**54.40%**	**58.70%**
Aligned pairs	**270875**	**197266**	**261576**	**206532**
Multiple alignments (% of aligned pairs)	27722 (10.2%)	11210 (5.7%)	46279 (17.7%)	16871 (8.2%)
Discordant alignments (% of aligned pairs)	7248 (2.7%)	3454 (1.8%)	3588 (1.4%)	2483 (1.2%)
Concordant pair alignment rate	**49.00%**	**51.00%**	**48.00%**	**53.70%**

Data table summarizing the alignment and mapping results against the *P. falciparum* IT and 3D7 reference genomes obtained using TopHat2 on our raw data (reads) for *P. falciparum* FCR3 S1.2 and S20 cultures that were sequenced. Both pair of reads (left and right) we mapped, checked for multiple alignments and concordance of pair alignment.

Table 4. List of differentially expressed genes codifying membrane and glideosome associated proteins.

Gene ID	mean counts	Log2 Fold Change	SD	Wald stat	p-value	q-value	Transcript Length	Gene Symbol	Product Description
PVP01_0616000	112.7691	3.297695	0.641278	5.142383	2.71E-07	0.000241	711	P28	ookinete surface protein P28, putative
PVP01_0616100	252.2767	2.822779	0.57479	4.910976	9.06E-07	0.000268	660	P25	ookinete surface protein P25
PVP01_0702600	121.3386	2.292091	0.602623	3.803521	0.000143	0.008636	915	PH	PH domain-containing protein, putative
PVP01_0705200	23.38039	-2.54479	0.816246	-3.11768	0.001823	0.031738	1245	LRR9	leucine-rich repeat protein
PVP01_0734800	47.52981	2.437732	0.688854	3.538825	0.000402	0.012534	564	ETRAMP	early transcribed membrane protein
PVP01_0820000	38.75789	2.309619	0.774088	2.983663	0.002848	0.040146	1683		CPW-WPC family protein
PVP01_0904300	29.44514	2.21377	0.753047	2.93975	0.003285	0.040824	1758		CPW-WPC family protein
PVP01_0905300	25.05779	3.023211	0.852766	3.545184	0.000392	0.012534	1851		WD repeat-containing protein, putative
PVP01_0942600	54.05728	2.203111	0.6829	3.226109	0.001255	0.025799	1551	IMC1b	inner membrane complex protein 1b, putative
PVP01_1020200	34.351	2.280086	0.741552	3.074748	0.002107	0.034015	2196	PSOP12	secreted ookinete protein, putative
PVP01_1020200	110.9312	2.230126	0.629647	3.541867	0.000397	0.012534			
PVP01_1137600	38.75146	2.054745	0.70281	2.923613	0.00346	0.041519	2187	BTP1	basal complex transmembrane protein 1, putative
PVP01_1208000	150.0577	2.3366	0.582524	4.011162	6.04E-05	0.004127	1302	P47	6-cysteine protein
PVP01_1208100	82.41245	2.148073	0.731291	2.937372	0.00331	0.040824	1353	P48/45	6-cysteine protein
PVP01_1219900	50.53231	2.237958	0.753953	2.9683	0.002995	0.04029	1236		MSP7-like protein
PVP01_1223200	72.71893	1.911415	0.654919	2.918551	0.003517	0.041637	1446		CPW-WPC family protein
PVP01_1251100	39.43236	3.389008	0.742545	4.564045	5.02E-06	0.000743	4824	CCp2	LCCL domain-containing protein
PVP01_1251100	152.1176	2.738815	0.64522	4.244776	2.19E-05	0.002159			
PVP01_1255400	54.03757	2.419967	0.728522	3.32175	0.000895	0.020904	2625	LAP5	LCCL domain-containing protein, putative
PVP01_1318300	93.48751	2.420415	0.721748	3.353546	0.000798	0.019148	1053	PIP2	PhIL1 interacting protein PIP2, putative
PVP01_1318600	100.5654	2.447863	0.675563	3.623441	0.000291	0.010756	807	PIP3	PhIL1 interacting protein PIP3, putative
PVP01_1439700	61.58366	1.950321	0.64478	3.024786	0.002488	0.036824	1692	IMC1h	inner membrane complex protein 1h, putative
PVP01_0716400	214.8741	2.048356	0.550386	3.721671	0.000198	0.008784	1188	GAP50	glideosome-associated protein 50, putative
PVP01_0532000	21.49608	2.492458	0.804617	3.097695	0.00195	0.033305	1116	GAPM2	glideosome associated protein with multiple membrane spans 2, putative
PVP01_1018200	75.41593	1.939599	0.627571	3.090644	0.001997	0.033463	1395	GAP40	glideosome-associated protein 40, putative

| PVP01_1230900 | 16.81091 | 2.446957 | 0.874358 | 2.798575 | 0.005133 | 0.049011 | 912 | GAPM1 | glideosome associated protein with multiple membrane spans 1, putative |

Table 5. MAAP analysis

Gene ID	Product description	MAAP score
PVP01_1219900	merozoite surface protein 7 (MSP7), putative	0.734
PVP01_1223200	CPW-WPC family protein, putative	0.766
PVP01_0820000	CPW-WPC family protein, putative	0.797
PVP01_0904300	CPW-WPC family protein, putative	1.132

Scores were calculated using the Malarial Adhesins and Adhesin-like Proteins predictor

Supplemental Material
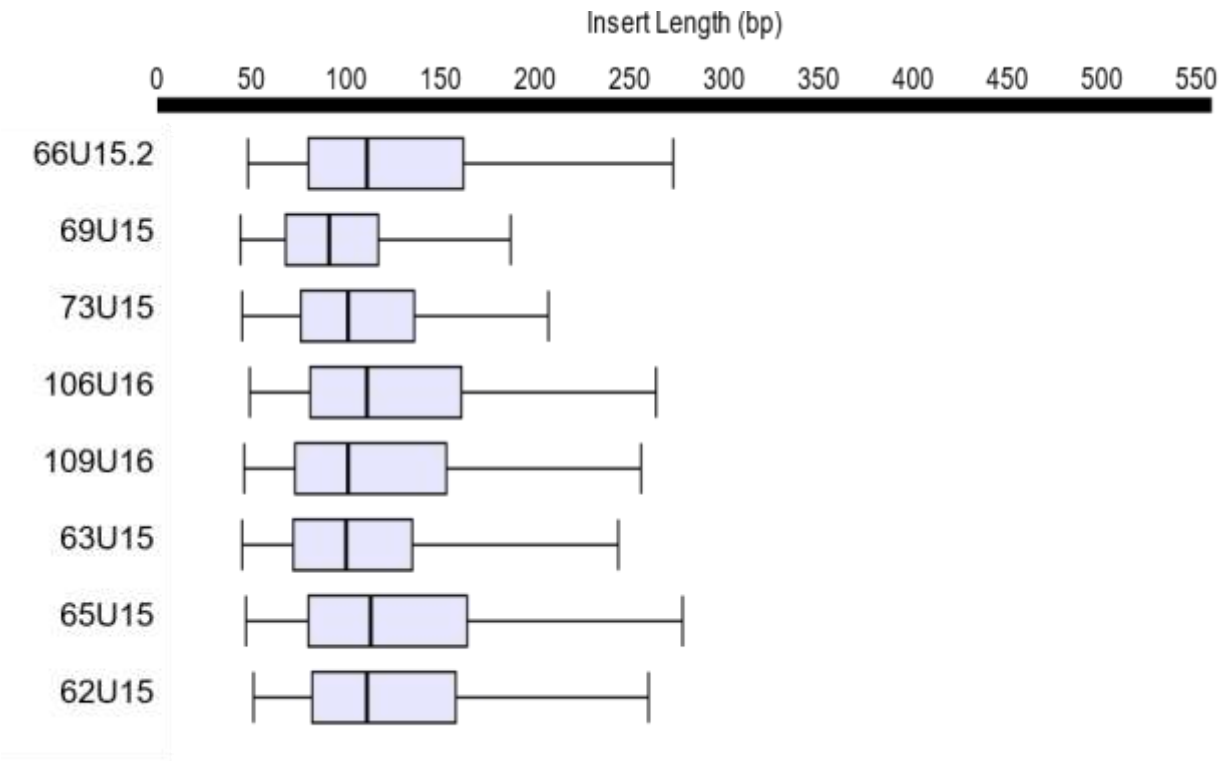
Figure 1. The insert length distribution boxplot summarizing the insert length distribution of paired-end reads of RNA sequenced libraries.
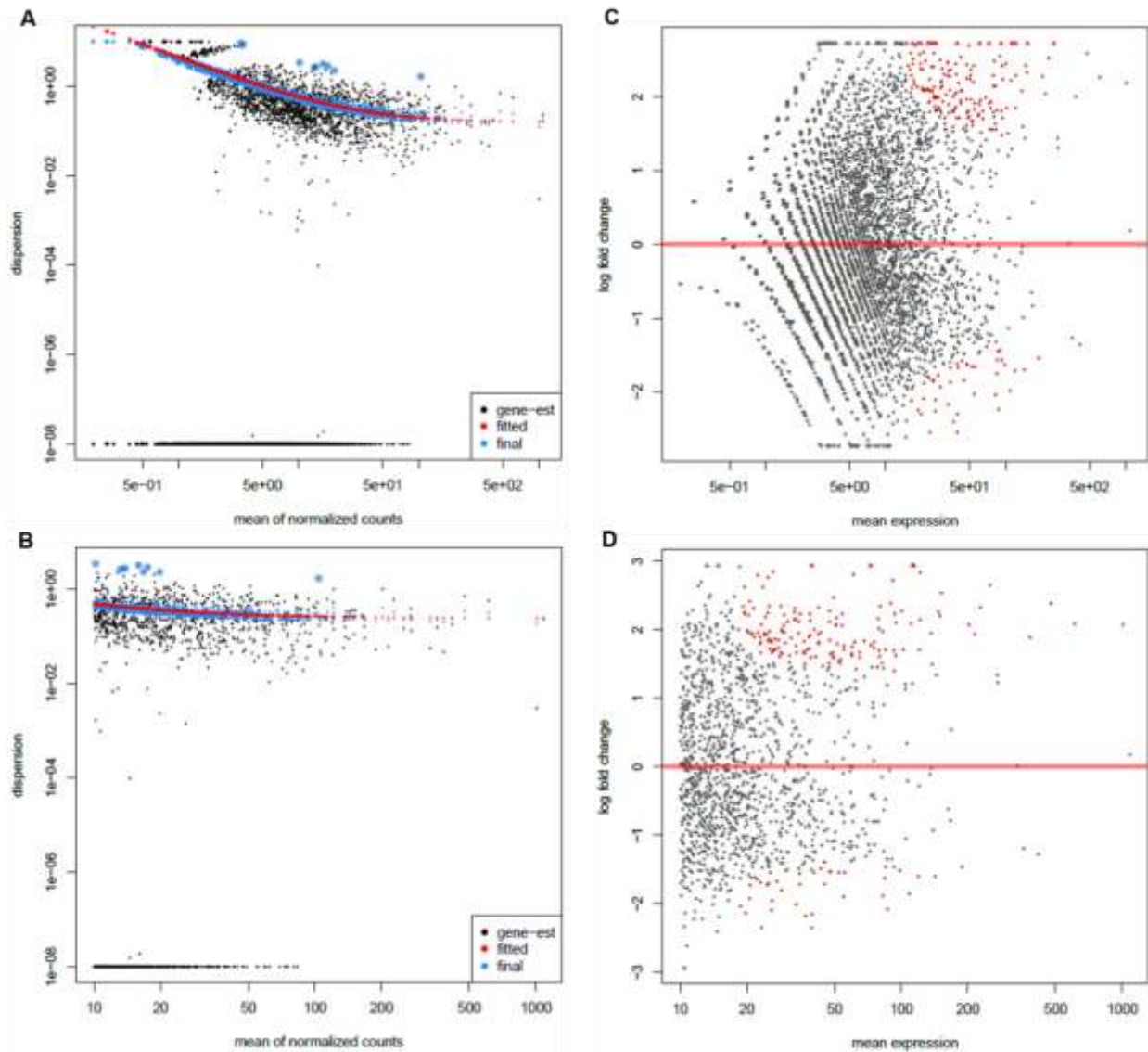
Figure 2. DESeq2 report analysis. Plots of dispersion estimates against the mean of normalized counts for complete (A) and filtered (B) data, showing the dispersion estimate for each gene as obtained by considering the information from each gene separately (black dots), the fitted estimates showing the dispersions' dependence on the mean (red line), the final dispersion estimates shrunk from the gene-wise estimates towards the fitted estimates (blue dots) and genes which have high gene-wise dispersion estimates and are hence labelled dispersion outliers and not shrunk toward the fitted trend line. The significantly differentially expressed genes (highlighted in red) are shown in MA scatter plots for both complete (C) and filtered (B) data.

Table 1. Patient profiles for *P. vivax* clinical isolates

Collection date	Patient code	Sex	Age	Parasitemia*	Pv-iE/200 leucocytes*	WBC (x10³/µl)	RBC (x10⁶/µL)	HGB (g/dL)	HCT (%)	MCV (fL)	MCH (pg)	MCHC (g/dL)	PLT (x10³/µL)	LYM (x10³/µL)	MXD (x10³/µL)	NEUT (x10³/µL)	RDW_SD (fL)	RDW_VC (%)	MPV (fL)
23/11/15	59U15	M	41	2+	448	5.9	5.07	14.1	45.8	90.3	27.8	30.8	43	0.6	0.5	4.8	48.7	13.8	11.1
24/11/15	61U15	F	42	2+	185	3.8	4.66	13	42.4	91	27.9	30.7	132	F1*1.1	F2*0.7	2	45.4	12.9	9.6
25/11/15	62U15	F	32	2+	335	5.5	4.04	11	34.6	85.6	27.2	31.8	42	1.2	0.5	3.8	43.3	13.8	10.4
25/11/15	63U15	F	36	2+	400	3.6	4.82	12.7	41.7	86.5	26.3	30.5	79	1.1	0.6	1.9	44.9	13.5	10
24/11/15	64U15	M	46	2+	365	4.4	4.27	12.1	38.2	89.5	28.3	31.7	68	0.5	0.5	3.2	45	13.3	9
25/11/15	65U15	M	49	2+	428	4.7	4.6	13	41.7	90.7	28.3	31.2	95	1.9	0.4	2.7	49	13.7	9.8
25/11/15	66U15	M	45	3+	340	11.2	5.17	14	45	87.8	27.1	30.8	260	0.6	1.2	9.4	44.7	13.1	6.8
26/11/15	67U15	F	36	1+	42	1.6	3.97	10.7	36.9	92.9	27	29	122	0.6	0.3	0.7	52.6	14.6	9.4
26/11/15	68U15	M	ND	2+	250	6.2	4.13	11.3	36.6	88.6	27.4	30.9	81	0.7	0.3	5.2	45.5	14	10.5
27/11/15	69U15	M	56	2+	430	3	4.39	13.3	42.2	96.1	30.3	31.5	43	0.3	0.1	2.6	48.6	13.1	8.7
26/11/15	70U15	F	25	2+	135	2.3	5.49	15.5	49.8	90.7	28.2	31.1	67	0.6	0.2	1.5	46.8	13.7	8.5
27/11/15	71U15	M	36	2+	210	3.3	4.95	13	42.8	86.5	26.3	30.4	42	0.5	0.5	2.3	46.5	13.8	8.9
27/11/15	72U15	F	36	3+	1050	3.6	4.17	8	29	69.5	19.2	27.6	35	F1* 0.6	F2* 0.5	2.5	41.8	16.8	ND
28/11/15	73U15	M	23	3+	ND	ND	ND	3.3	3.76	9.9	33	87.8	26.3	30	39	0.4	0.2	2.7	77.3
30/11/15	76U15	M	24	2+	270	ND	ND	5.1	4.97	12.1	40.2	80.9	24.3	30.1	131	F1*1.3	F2*0.7	3.1	42.7
1/12/15	78U15	M	26	3+	326	6.7	4.58	11.5	37.9	82.8	25.1	30.3	AG 76	1.2	0.3	5.2	47.1	15.4	9.4
2/12/15	82U15	M	26	2+	430	8.1	4.97	13.9	45.9	92.4	28	30.3	AG 196	1.5	0.7	5.9	50.7	14.2	8.6
2/12/15	84U15	M	48	2+	210	WL* 6.1	5.1	14.1	44.7	87.6	27.6	31.5	AG 67	WL* 1.9	WL* 0.5	WL* 3.7	43.8	13.4	10.2
3/12/15	87U15	M	29	2+	220	WL* 4	5.48	15.7	48.3	88.1	28.6	32.5	AG-36	WL* 0.5	WL* 1	WL* 2.5	45.1	13.8	ND
7/12/15	94U15	F	43	2+	230	4.8	3.87	10.6	32.9	85	27.4	32.2	PL* 11	1	0.1	3.7	43.9	14.9	ND
9/12/15	95U15	M	59	2+	120	5	5.08	13.9	42.6	83.9	27.4	32.6	AG 83	0.9	0.2	3.9	46.7	16.7	8.8
9/12/15	96U15	F	28	2+	530	3.8	3.49	10.3	31.5	90.3	29.5	32.7	AG 58	1.1	0.4	2.3	42.5	13.6	9.3
9/12/15	97U15	M	26	2+	170	3.2	4.68	13.3	41.8	89.3	28.4	31.8	101	1	0.3	1.9	40.1	13	9.6

9/12/15	98U15	M	47	2+	370	5.7	4.99	13.9	44.3	88.8	27.9	31.4	116	ND	ND	ND	46.3	13.8	9.2
11/1/16	106U16	M	57	2+	230	7.3	5.09	14.2	46.3	91	27.9	30.7	94	0.7	0.6	6	44.4	12.4	10.8
12/1/16	109U16	M	50	2+	480	1.3	4.64	11.7	39.7	85.6	25.6	29.5	47	0.3	0	1	50.2	16.1	7.5

Summary of patients and clinical isolates information after collection during field work at vivax malaria endemic area. Date: dd/mm/yy; M: Male; F: Female; ND – Not Determined; *Parasitemia estimation and **initial counts of the number of Pv-iE per 200 leucocytes on the slide by thick smear made available by the specially trained microscopists from malaria diagnosis service from FMT-HVD. Two crosses (2+) equals to 2 to 20 parasites counted by microscope field and an average of 500 to 10^4 parasites per µL. WBC: White Blood Cells; RBC: Red Blood Cells; HGB: hemoglobin; HCT: Hematocrit; MCV: Mean Corpuscular Volume; MCH: Mean Corpuscular Hemoglobin; MCHC: mean corpuscular hemoglobin concentration; PLT: Platelets; LYM: Lymphocytes; MXD: mixed cells; NEUT: Neutrophils; RDW: Red Cell Distribution Width SD: Standard Deviation and VC: Variance Coefficient; MPV: Mean Platelet Volume. For information on hemogram flags please see https://www.sysmex.co.za/fileadmin/media/f112/SEED/English/Sysmex_SEED_6_2013_Haematology_Results_Interferences__Flagging_and_Interpretation_-_Part_II_EN.pdf.

Table 2. Bioanalyzer measurements after RNA extraction with RNeasy® Micro kit and quality and quantity controls after cDNA library generation and amplification.

Sample Code	RNA [pg/uL]	RIN	cDNA [pg/uL]	cDNA [ng/uL]	V to 20ng (µL)	Total cDNA [ng/75uL]	V_T to 20ng/reaction from 75µL after shearing (µL)	V Elution Buffer (µL)	Total cDNA [ng/reaction]	Nº of PCR cycles for Low Input Library amplification	Fragment average size (bp)	Molarity [pmol/L]
62U15	65.531	8.4	2149.92	2.150	9.30	161.244	12.0	-	0.344	9	376	4.52E+04
63U15	8.914	9.8	124671	124.671	0.16	9350.300	1.2	8.8	20	6	366	1.97E+05
65U15	21.499	9.7	53184.4	53.184	0.38	3988.827	2.8	7.2	20	6	416	1.51E+05
66U15.2	522	N/A	46864.8	46.865	0.43	3514.859	3.2	6.8	20	6	356	1.01E+05
69U15	13.529	6.6	137982	137.982	0.14	10348.684	1.1	8.9	20	6	396	1.85E+05
73U15	4.925	7.4	249417	249.417	0.08	18706.261	0.6	9.4	20	6	310	5.69E+04
106U16	17.234	6.3	10061.3	10.061	1.99	754.595	12.0	-	1.609	6	369	7.98E+05
109U16	1.87	N/A	413707	413.707	0.05	31028.009	0.4	9.6	20	6	393	8.56E+04
Pf FCR3 S1.2	3.264	N/A	246122	246.122	0.08	18459.169	0.6	9.4	20	6	378	8.87E+04
Pf S20	1.164	N/A	2566.13	2.566	7.79	192.460	12.0	-	0.41	9	351	4.19E+04

P. falciparum FCR3 S1.2 strain (~10^6 Pf-iE) enriched for >80% rosetting formation and the Brazilian *P. falciparum* S20 field isolate (~10^6 Pf-iE) resistant to Artemisinin with no rosette formation were used as positive and negative controls for rosetting phenotype. RIN: RNA Integrity Number; N/A: undetermined; V: Volume; V_T: Total Volume; bp: base pairs.

Table 3. Library quantification by qPCR and sample pool for Illumina® HiSeq 2500 load and run.

Sample Code	Dilution	Cq Mean	Cq Error	Mean [pM]	Error [pM]	Total [pM]	Total [nM]	*447bp from DNA fragments	Library Concentration [nM]	Dilution		V at 2nM (µL)	V ddH$_2$O (µL)
62U15	1.00E-06	11.6	0.63	3.03E+00	1.48E+00	3.03E+06	3030.00	1.35E+09	3602.154	dil -3	3.6021543	3.89	3.11
63U15	1.00E-09	13.39	0.58	8.79E-01	2.97E-01	8.79E+08	878500.00	3.93E+11	1072922	dil -5	10.729221	1.30	5.70
65U15	1.00E-09	13.83	0.77	6.72E-01	3.24E-01	6.72E+08	672100.00	3.00E+11	722184.4	dil -5	7.2218438	1.94	5.06
66U15.2	1.00E-09	14.25	0.84	5.29E-01	3.66E-01	5.29E+08	528800.00	2.36E+11	663970.8	dil -5	6.6397079	2.11	4.89
69U15	1.00E-09	15.13	0.75	2.73E-01	1.74E-01	2.73E+08	272800.00	1.22E+11	307933.3	dil -5	3.0793333	4.55	2.45
73U15	1.00E-09	12.55	1.40	2.09E+00	1.88E+00	2.09E+09	2090000.00	9.34E+11	3013645	dil -6	3.0136452	4.65	2.35
106U16	1.00E-10	17.23	0.23	5.40E-02	8.57E-03	5.40E+08	540000.00	2.41E+11	654146.3	dil -5	6.5414634	2.14	4.86
109U16	1.00E-09	14.44	0.33	4.03E-01	9.00E-02	4.03E+08	403000.00	1.80E+11	458374	dil -5	4.5837405	3.05	3.95
Pf FCR3 S1.2	1.00E-09	18.12	2.45	7.05E-02	8.30E-02	7.05E+07	70450.00	3.15E+10	83309.92	dil-4	8.3309921	1.68	5.32
Pf S20	1.00E-07	13.20	0.90	1.09E+00	5.77E-01	1.09E+07	10900.00	4.87E+09	13881.2	dil-3	13.881197	1.01	5.99

P. falciparum FCR3 S1.2 strain (~10^6 Pf-iE) enriched for >80% rosetting formation and the Brazilian *P. falciparum* S20 field isolate (~10^6 Pf-iE) resistant to Artemisinin with no rosette formation were used as positive and negative controls for rosetting phenotype. RIN: RNA Integrity Number; N/A: undetermined; V: Volume; V$_T$/sample = 7.0µL; bp: base pairs.

Table 4. Raw data (reads) description output from RNA-sequencing.

Sample ID	Lane	Index	Sample Code	Filename	Total Sequences (FastQ)	Total nº of reads after trimming (%)	Sequence length	%GC	Sample Ref
9	2	TTAGGC	62U15	9_S14_L001_R1_001.fastq.gz 9_S15_L001_R2_001.fastq.gz	567222 567222	546571 546571 (96.4%)	101 101	42 43	P. vivax Sal-1/P01
10	2	GCCAAT	66U15.2	10_S17_L001_R1_001.fastq.gz 10_S17_L001_R2_001.fastq.gz	396869 396869	393248 393248 (99.1%)	101 101	49 49	P. vivax Sal-1/P01
11	2	CAGATC	69U15	11_S18_L001_R1_001.fastq.gz 11_S18_L001_R2_001.fastq.gz	444674 444674	436832 436832 (93.8%)	101 101	46 47	P. vivax Sal-1/P01
12	2	ACTTGA	73U15	12_S19_L001_R1_001.fastq.gz 12_S19_L001_R2_001.fastq.gz	465530 465530	457871 457871 (98,4%)	101 101	45 46	P. vivax Sal-1/P01
13	2	ACAGTG	106U16	13_S16_L001_R1_001.fastq.gz 13_S16_L001_R2_001.fastq.gz	741000 741000	726393 726393 (98.0%)	101 101	49 49	P. vivax Sal-1/P01
14	2	GATCAG	109U16	14_S20_L001_R1_001.fastq.gz 14_S20_L001_R2_001.fastq.gz	247006 247006	244626 244626 (99.0%)	101 101	46 46	P. vivax Sal-1/P01
15	2	TAGCTT	63U15	15_S21_L001_R1_001.fastq.gz 15_S21_L001_R2_001.fastq.gz	556333 556333	550489 550489 (98.9%)	101 101	47 47	P. vivax Sal-1/P01
16	2	GGCTAC	65U15	17_S23_L001_R1_001.fastq.gz 17_S23_L001_R2_001.fastq.gz	223390 223390	218478 218478 (97.8%)	101 101	45 46	P. vivax Sal-1/P01

| 17 | 1 | CTTGTA | Pf FCR3 S1.2 | 16_S22_L001_R1_001.fastq.gz
16_S22_L001_R2_001.fastq.gz | 553823
553823 | 537931
537931
(97.1%) | 101
101 | 29
30 | P. falciparum IT/3D7 |
| 18 | 2 | TGACCA | Pf S20 | 18_S15_L001_R1_001.fastq.gz
18_S15_L001_R2_001.fastq.gz | 383309
383309 | 379684
379684
(99.1%) | 101
101 | 30
30 | P. falciparum IT/3D7 |

Tabela 5. Complete diferencial gene expression list

Gene ID	mean counts	Log2 Fold Change	SD	Wald stat	p-value	q-value	Transcript Length	Gene Symbol	Product Description
PVP01_0616000	112.7691	3.297695	0.641278	5.142383	2.71E-07	0.000241	711	P28	ookinete surface protein P28, putative
PVP01_0616100	252.2767	2.822779	0.57479	4.910976	9.06E-07	0.000268	660	P25	ookinete surface protein P25
PVP01_0702600	121.3386	2.292091	0.602623	3.803521	0.000143	0.008636	915	PH	PH domain-containing protein, putative
PVP01_0705200	23.38039	-2.54479	0.816246	-3.11768	0.001823	0.031738	1245	LRR9	leucine-rich repeat protein
PVP01_0734800	47.52981	2.437732	0.688854	3.538825	0.000402	0.012534	564	ETRAMP	early transcribed membrane protein
PVP01_0820000	38.75789	2.309619	0.774088	2.983663	0.002848	0.040146	1683		CPW-WPC family protein
PVP01_0904300	29.44514	2.21377	0.753047	2.93975	0.003285	0.040824	1758		CPW-WPC family protein
PVP01_0905300	25.05779	3.023211	0.852766	3.545184	0.000392	0.012534	1851		WD repeat-containing protein, putative
PVP01_0942600	54.05728	2.203111	0.6829	3.226109	0.001255	0.025799	1551	IMC1b	inner membrane complex protein 1b, putative
PVP01_1020200	34.351	2.280086	0.741552	3.074748	0.002107	0.034015	2196	PSOP12	secreted ookinete protein, putative
PVP01_1020200	110.9312	2.230126	0.629647	3.541867	0.000397	0.012534			
PVP01_1137600	38.75146	2.054745	0.70281	2.923613	0.00346	0.041519	2187	BTP1	basal complex transmembrane protein 1, putative
PVP01_1208000	150.0577	2.3366	0.582524	4.011162	6.04E-05	0.004127	1302	P47	6-cysteine protein
PVP01_1208100	82.41245	2.148073	0.731291	2.937372	0.00331	0.040824	1353	P48/45	6-cysteine protein
PVP01_1219900	50.53231	2.237958	0.753953	2.9683	0.002995	0.04029	1236		MSP7-like protein
PVP01_1223200	72.71893	1.911415	0.654919	2.918551	0.003517	0.041637	1446		CPW-WPC family protein
PVP01_1251100	39.43236	3.389008	0.742545	4.564045	5.02E-06	0.000743	4824	CCp2	LCCL domain-containing protein
PVP01_1251100	152.1176	2.738815	0.64522	4.244776	2.19E-05	0.002159			
PVP01_1255400	54.03757	2.419967	0.728522	3.32175	0.000895	0.020904	2625	LAP5	LCCL domain-containing protein, putative
PVP01_1318300	93.48751	2.420415	0.721748	3.353546	0.000798	0.019148	1053	PIP2	PhIL1 interacting protein PIP2, putative
PVP01_1318600	100.5654	2.447863	0.675563	3.623441	0.000291	0.010756	807	PIP3	PhIL1 interacting protein PIP3, putative
PVP01_1439700	61.58366	1.950321	0.64478	3.024786	0.002488	0.036824	1692	IMC1h	inner membrane complex protein 1h, putative

PVP01_0716400	214.8741	2.048356	0.550386	3.721671	0.000198	0.008784	1188	GAP50	glideosome-associated protein 50, putative
PVP01_0532000	21.49608	2.492458	0.804617	3.097695	0.00195	0.033305	1116	GAPM2	glideosome associated protein with multiple membrane spans 2, putative
PVP01_1018200	75.41593	1.939599	0.627571	3.090644	0.001997	0.033463	1395	GAP40	glideosome-associated protein 40, putative
PVP01_1230900	16.81091	2.446957	0.874358	2.798575	0.005133	0.049011	912	GAPM1	glideosome associated protein with multiple membrane spans 1, putative
PVP01_1463200	79.29981	2.152384	0.609051	3.533994	0.000409	0.012534	1131	ACT1	actin, putative
PVP01_0215500	23.68609	2.847756	0.826299	3.4464	0.000568	0.016274	294		dynein light chain, putative
PVP01_05190000	16.75521	3.047203	1.005411	3.030802	0.002439	0.03671			
PVP01_0607900	20.39195	2.373058	0.807591	2.938439	0.003299	0.040824	1035		phospholipid scramblase, putative
PVP01_0517400	94.53938	1.716136	0.589713	2.910122	0.003613	0.041873	300	HMGB2	high mobility group protein B2, putative
PVP01_0530800	30.68695	2.303592	0.800011	2.87945	0.003984	0.044779	1353		alpha tubulin 2, putative
PVP01_0530800	609.8787	2.261229	0.664683	3.401966	0.000669	0.018246			
PVP01_1270600	19.23439	2.578164	0.912845	2.824316	0.004738	0.047275	1521		tubulin epsilon chain, putative
PVP01_0905800	1007.859	2.19129	0.495873	4.419055	9.91E-06	0.001258	312	H4	histone H4, putative
PVP01_0905900	383.0346	2.000833	0.526814	3.797986	0.000146	0.008636	357	H2B	histone 2B, putative
PVP01_0106900	26.27094	3.098066	0.834422	3.712829	0.000205	0.008784	2628	MSH2-2	DNA mismatch repair protein MSH2, putative
PVP01_0716900	34.89807	2.374967	0.719649	3.300174	0.000966	0.021451	2270	TR	telomerase RNA
PVP01_0315400	39.4368	-2.30529	0.72849	-3.16447	0.001554	0.029354	411	RPS10	40S ribosomal protein S10, putative
PVP01_1231500	108.8071	-1.99161	0.637346	-3.12484	0.001779	0.031595	573		60S ribosomal protein L6, putative
PVP01_0312600	54.89049	-2.03154	0.665866	-3.05097	0.002281	0.035506	984	EIF3I	eukaryotic translation initiation factor 3 subunit I, putative
PVP01_1225200	61.51756	1.953305	0.667764	2.925142	0.003443	0.041519	1602		chromatin assembly factor 1 P55 subunit, putative
PVP01_1114900	33.45494	-2.17093	0.75858	-2.86184	0.004212	0.046375	1332		elongation factor 1-alpha, putative
PVP01_1334600	119.6808	-1.71801	0.612209	-2.80625	0.005012	0.048722	660		60S ribosomal protein L10, putative

PVP01_0108700	142.2223	-1.69775	0.574609	-2.95462	0.003131	0.040824	2247	HSP90	heat shock protein 90, putative
PVP01_0816000	188.2958	-1.54743	0.547881	-2.82438	0.004737	0.047275	1341	ENO	enolase, putative
PVP01_0614700	86.80803	-2.24092	0.677097	-3.3096	0.000934	0.021273	1092	ADA	adenosine deaminase, putative
PVP01_1013000	82.33313	-2.04289	0.72857	-2.80397	0.005048	0.048722	1434	SAHH	adenosylhomocysteinase, putative
PVP01_1253500	22.46193	3.016197	0.810966	3.719266	0.0002	0.008784	3543	NNT	NAD(P) transhydrogenase, putative
PVP01_1312300	202.3929	2.277458	0.761236	2.991789	0.002773	0.039723	2547		ribonucleoside-diphosphate reductase large subunit, putative
PVP01_1448300	92.65471	1.814316	0.643397	2.819903	0.004804	0.047398	1098	TLAP1	thioredoxin-like associated protein 1, putative
PVP01_1408700	25.86374	-2.26161	0.795743	-2.84214	0.004481	0.046726	1329		FHA domain-containing protein, putative
PVP01_1229400	42.25079	2.535802	0.775752	3.268829	0.00108	0.022833	1005		lactate dehydrogenase, putative
PVP01_1229400	113.716	3.525322	0.71681	4.918068	8.74E-07	0.000268			
PVP01_1229400	227.9158	2.525019	0.688415	3.667872	0.000245	0.009872			
PVP01_0917100	43.27545	2.088327	0.666576	3.132919	0.001731	0.031595	1356	FT2	folate transporter 2, putative
PVP01_1412100	476.0866	2.592285	0.698835	3.709436	0.000208	0.008784	1488		meiosis-specific nuclear structural protein 1, putative
PVP01_1452800	28.53532	2.508619	0.746829	3.359026	0.000782	0.019148	3297	ULG8	upregulated in late gametocytes ULG8, putative
PVP01_1467200	147.4798	2.422746	0.643959	3.762265	0.000168	0.008784	7743	G377	osmiophilic body protein G377, putative
PVP01_0833800	42.01345	1.951736	0.703105	2.775881	0.005505	0.050142	1968		FAD-dependent glycerol-3-phosphate dehydrogenase, putative
PVP01_0315200	42.39439	1.91849	0.698976	2.744717	0.006056	0.052214	1302	CDC50A	LEM3/CDC50 family protein, putative
PVP01_0904800	17.00196	2.513344	0.905969	2.774206	0.005534	0.050142	1854		phenylalanine--tRNA ligase beta subunit, putative
PVP01_1220500	26.03336	-2.09938	0.762184	-2.75442	0.00588	0.051187	1569		chaperone binding protein, putative
PVP01_1447500	39.41054	-2.55487	0.775925	-3.29268	0.000992	0.021494	351	MIF	macrophage migration inhibitory factor, putative
PVP01_1300900	121.6468	3.121743	0.770805	4.049976	5.12E-05	0.004127	891	NEK3	NIMA related kinase 3, putative
PVP01_0914700	49.36319	2.704811	0.795931	3.3983	0.000678	0.018246	1728	MAPK2	mitogen-activated protein kinase 2, putative
PVP01_0724800	19.61031	2.877202	0.893319	3.220799	0.001278	0.025799	2853		protein kinase, putative

PVP01_1441300	20.65536	2.601104	0.831047	3.129912	0.001749	0.031595	1323		cAMP-dependent protein kinase regulatory subunit, putative
PVP01_1438100	15.34182	2.419945	0.866682	2.792195	0.005235	0.049212	576	RKIP	raf kinase inhibitor, putative
PVP01_0317200	34.32217	2.092318	0.735009	2.846656	0.004418	0.046706	2226		conserved Plasmodium protein, unknown function
PVP01_0414700	25.24296	2.59296	0.909784	2.850084	0.004371	0.046706	6750		conserved Plasmodium protein, unknown function
PVP01_0419500	29.54548	2.378209	0.771529	3.082464	0.002053	0.03376	1998		conserved Plasmodium protein, unknown function
PVP01_0505700	50.85069	2.298813	0.832034	2.762883	0.005729	0.050750	3135		conserved protein, unknown function
PVP01_0508600	78.83837	2.299477	0.610533	3.766344	0.000166	0.008784	897		conserved Plasmodium protein, unknown function
PVP01_0526400	80.26688	2.483304	0.605028	4.104448	4.05E-05	0.003599	960		conserved Plasmodium protein, unknown function
PVP01_0526400	33.02295	2.480573	0.733532	3.381683	0.00072	0.018816			
PVP01_0530400	18.53342	3.184352	0.891261	3.572862	0.000353	0.012534	225		conserved Plasmodium protein, unknown function
PVP01_0723600	24.91472	2.416044	0.790179	3.057592	0.002231	0.035381	1278		conserved Plasmodium protein, unknown function
PVP01_0816200	17.60976	2.696404	0.885229	3.045996	0.002319	0.035506	717		conserved Plasmodium protein, unknown function
PVP01_0822100	89.50047	2.552484	0.635803	4.014585	5.96E-05	0.004127	651		conserved Plasmodium protein, unknown function
PVP01_0913300	36.88973	-2.39957	0.842584	-2.84788	0.004401	0.046706	879		conserved protein, unknown function
PVP01_0932300	79.43851	1.801584	0.60573	2.974236	0.002937	0.04029	777		conserved Plasmodium protein, unknown function
PVP01_1018500	28.87016	2.119632	0.746608	2.839015	0.004525	0.046726	1815		conserved Plasmodium protein, unknown function
PVP01_1020900	84.8206	1.823611	0.644594	2.829083	0.004668	0.047275	816		conserved Plasmodium protein, unknown function
PVP01_1136600	47.4886	1.96798	0.677522	2.904675	0.003676	0.041873	366		conserved Plasmodium protein, unknown function
PVP01_1140300	94.8669	2.238768	0.612655	3.654209	0.000258	0.00996	1368		conserved Plasmodium protein, unknown function
PVP01_1145000	38.62408	2.451921	0.844171	2.904533	0.003678	0.041873	1128		conserved Plasmodium protein, unknown function

PVP01_1216000	51.19263	2.365249	0.746088	3.170202	0.001523	0.029354	1977	conserved Plasmodium protein, unknown function
PVP01_1231000	19.8116	-2.29708	0.823219	-2.79037	0.005265	0.049212	1305	conserved Plasmodium protein, unknown function
PVP01_1326800	32.84243	2.296351	0.720483	3.187235	0.001436	0.028345	936	conserved Plasmodium protein, unknown function
PVP01_1335900	33.41938	2.354546	0.823133	2.860468	0.00423	0.046375	1323	conserved Plasmodium protein, unknown function
PVP01_1344600	60.71377	3.015201	0.697504	4.322843	1.54E-05	0.00171	399	conserved Plasmodium protein, unknown function
PVP01_1344600	32.80864	2.457339	0.815522	3.013208	0.002585	0.037631		
PVP01_1345600	91.62995	2.908091	0.612192	4.750293	2.03E-06	0.000361	1077	conserved Plasmodium protein, unknown function
PVP01_1421800	40.48658	2.190182	0.793415	2.760449	0.005772	0.050750	1137	conserved Plasmodium protein, unknown function
PVP01_1424300	16.24633	2.592357	0.881688	2.940218	0.00328	0.040824	4605	conserved Plasmodium protein, unknown function
PVP01_1433600	19.57315	2.389032	0.80832	2.955552	0.003121	0.040824	330	conserved Plasmodium protein, unknown function
PVP01_1433600	47.38325	2.258372	0.669807	3.371676	0.000747	0.018956		
PVP01_1457500	75.18845	1.843871	0.620267	2.972708	0.002952	0.04029	576	conserved Plasmodium protein, unknown function
PVP01_1465500	72.66478	3.514997	0.737318	4.767275	1.87E-06	0.000361	888	conserved Plasmodium protein, unknown function
PVP01_1467500	136.6578	2.080068	0.597305	3.482423	0.000497	0.014708	1950	conserved Plasmodium protein, unknown function
PVP01_0944300	42.06959	2.084861	0.75081	2.776818	0.005489	0.050142	975	conserved Plasmodium protein, unknown function
PVP01_1218300	19.48892	2.456623	0.887304	2.768639	0.005629	0.050491	1137	conserved Plasmodium protein, unknown function

"*Plasmodium vivax* rosetting

impacts on

host immune response"

Overview

Here, the present study used a range of *ex vivo* approaches to dissect the role of rosetting in vivax malaria. *P. vivax* rosetting rates were enhanced by autologous plasma and total immunoglobulin M levels correlated with rosetting frequency. Moreover, rosetting was also correlated with parasitemia, IL-6 and IL-10 in infected patients. Transcriptomic analysis of peripheral leukocytes from *P. vivax*-infected patients with low or moderated rosetting frequency allowed identification of differentially expressed actin-related and immunoglobulin genes, known to be associated with the human host phagocytosis pathway. In addition, a smaller phagocytic index was found for the high rosetting group, and phagocytosis functional assays demonstrated that rosetting integrity interfered with the phagocytic index. Collectively, these results showed that rosette formation plays a role in host immune response by hampering leukocyte phagocytosis. Thus, these findings suggested that rosetting is a novel and effective *P. vivax* immune evasion adaptation.

Important Note: To comply with scientific journal publication rules to which we have submitted this work, the draft version presented here only contains results and discussion of RNA-seq data analysis, my main contribution for the preparation of this manuscript within this thesis project.

Abstract

Plasmodium vivax is the most prevalent cause of malaria outside of Africa. Although *P. vivax* pathogenesis is poorly understood, it is known that infected reticulocytes cytoadhere readily to noninfected normocytes, forming rosettes. Despite the high prevalence of *P. vivax* rosetting, the biological purpose of this phenomenon is unknown. Here, the present study used a range of *ex vivo* approaches to dissect the role of rosetting in vivax malaria. *P. vivax* rosetting rates were enhanced by autologous plasma and total immunoglobulin M levels correlated with rosetting frequency. Moreover, rosetting was also correlated with parasitemia, IL-6 and IL-10 in infected patients. Transcriptomic analysis of peripheral leukocytes from *P. vivax*-infected patients with low or moderated rosetting frequency allowed identification of differentially expressed actin-related and immunoglobulin genes, known to be associated with the human host phagocytosis pathway. In addition, a smaller phagocytic index was found for the high rosetting group, and phagocytosis functional assays demonstrated that rosetting integrity interfered with the phagocytic index. Collectively, these results showed that rosette formation plays a role in host immune response by hampering leukocyte phagocytosis. Thus, these findings suggested that rosetting is a novel and effective *P. vivax* immune evasion adaptation.

Keywords: malaria, rosetting, *P. vivax*, immune evasion, phagocytosis

Author Summary: The role of *Plasmodium vivax* rosettes was investigated using several *ex vivo* experiments. Higher rosetting parasites were less phagocyted by THP-1 cells, suggesting that this mechanism can protect the parasite from the host immune system.

Introduction

(…)

Results

(…)

Differential gene expression in peripheral blood mononuclear cells from P. vivax infection and its relationship with rosetting

Considering differences in cytokine levels between high and low rosetting isolates, transcriptome analysis was performed on peripheral leukocytes from *P. vivax* infected patients (Table 1A) with very low or moderate rosetting. Eight genes that showed a log2 fold change higher than 1.5 were identified (Table 1B). Three out of eight genes were in the Fc gamma receptor (FCGR)-dependent phagocytosis pathway. Interestingly, immunoglobulin kappa constant (IGKC) and immunoglobulin heavy constant gamma 1 (IGHG1) were upregulated and actin-related protein 2/3 complex subunit 2 (ARPC2) was downregulated in individuals with moderate rosetting compared to patients with low rosetting.

Table 1 – RNAseq analysis of PBMCs from patients infected with *P. vivax*.

A. Summarized metadata description of patients infected with *P. vivax* analyzed by RNAseq. **B.** Graph of log2 fold change of a selected pool of genes from PBMC of *P. vivax* infected patients between moderate and low rosetting isolates data analysis. Gene lists using p-value<0.05, q-value<0.5 and log2(fold change)>1.5 cut-offs obtained from RNAseq differential gene expression analysis.

A.

Group Analysis	Sample Code	Rosette formation (%)
Moderate rosetting	106U16	30.0%
	69U15	24.4%
Very Low rosetting	109U16	6.0%
	73U15	5.6%

B.

Gene ID	Gene name	Gene description	Gene locus	Nº of counted reads		Log2 (Fold Change)	q-value
				Very low rosetting	Moderate rosetting		
ENSG00000100325.14	ASCC2	Activating Signal Cointegrator 1 Complex Subunit 2	chr22:29788607-29838304	1422.06	111.412	-3.67	0.01
ENSG00000163466.15	ARPC2	Actin Related Protein 2/3 Complex Subunit 2	chr2:218217093-218254356	1134.05	274.937	-2.04	0.09
ENSG00000196998.16	WDR45	WD Repeat Domain 45	chrX:49071155-49101170	482.955	123.784	-1.96	0.05
ENSG00000166710.17	B2M	***Beta-2-Microglobulin***	chr15:44711476-44718877	3709.93	12757.5	1.78	0.23
ENSG00000105701.15	FKBP8	FK506 Binding Protein 8	chr19:18531612-18544077	164.806	569.29	1.79	0.17
ENSG00000211592.7	IGKC	Immunoglobulin Kappa Constant	chr2:88811185-89245596	2195.89	9705.52	2.14	0.29
ENSG00000211896.7	IGHG1	Immunoglobulin Heavy Constant Gamma 1	chr14:105664632-106538344	336.287	2065.88	2.62	0.20
ENSG00000124098.9	FAM210B	Family with Sequence Similarity 210 Member B	chr20:56358914-56368663	65.2525	411.436	2.66	0.22

(…)

Discussion

(…)

The host transcriptomic profile directly compared patient isolates with low versus moderate rosetting, which were quantified immediately after blood sample collection, when most parasites were young [1]. Because rosetting is mostly formed by mature parasites, this constituted an experimental limitation of the present study. Additionally, a low quantity of host sample input effectively transcribed resulted in few statistically significant pools of differentially expressed genes. Nevertheless, three out of eight genes showing more than 1.5-fold change in expression were found, and their functions were relevant in the Fc gamma phagocytic pathway. The present transcriptome findings, supported by our results showing that plasma from patients with high rosetting levels inhibit the phagocytic capacity of THP-1 cells, support the idea that *P. vivax* rosettes strongly inhibit the phagocytosis pathway.

(…)

Taken together, these data indicated that vivax malaria rosetting is an evasion mechanism that allows the parasite to escape from the host immune system (Figure 6), and this generates more questions regarding this common phenotype. The fact that rosetting is a frequent feature in *P. vivax* late stages and that the high prevalence of rosettes may indicate that this phenotype is an important advantage for the parasite by conferring significant protection from the host immune system. Therefore, the understanding of *P. vivax* rosettes may help the scientific community to develop new strategies for malaria control.

Material and Methods

Ethical Approval

Informed consent was granted from all patients attending the *Fundação de Medicina Tropical Dr. Heitor Vieira Dourado* (FMT-HVD; Manaus, Amazonas State, Brazil). All

procedures, including protocols and consent forms, were approved by the Ethics Review Board of FMT-HVD (process CAAE-0044.0.114.000-11).

Blood sample collection

Blood samples were obtained from malaria patients presenting at the clinic of the Hospital de Medicina Tropical Heitor Vieira Dourado (Manaus, Amazonas State, Brazil). Prior to sample collection, patients granted informed consent. Blood samples were collected using BD Vacutainer® tubes with sodium citrate anticoagulant. A thin blood smear was prepared from each sample to determine species of malaria parasites involved, parasitemia and the predominant erythrocytic stage of the parasite. Blood cell count was performed immediately after blood collection using a Sysmex KX21N (Sysmex Corporation-Roche, Japan). After collection of peripheral blood, patients received treatment following national guidelines with chloroquine plus primaquine.

Parasite isolation and enrichment

Once microscopic diagnosis of uncomplicated vivax malaria was made and before the treatment was initiated, 8 mL of blood was collected into citrate-coated Vacutainer® tubes (BD). Blood was immediately processed to obtain enriched Pv-iEs. Immediately after collection, the RBCs containing trophozoites and schizonts were separated from the younger forms on a 45% Percoll™ (GE Healthcare) gradient as previously described [2].

Rosetting Assay

IEs (20 μL) at 2.5-5% parasitemia and 2.5-5% hematocrit were incubated for 40 min at 37°C in rosetting medium (McCoy's 5A medium supplemented with 20% of autologous plasma). Duplicated samples were stained with 45 μg/ml acridine orange and examined by direct light and fluorescence microscopy (Nikon Eclipse 50i, filter 96311 B-2E/C). Rosetting was assessed by counting 200 IEs, in duplicate. A rosette was determined by the binding of two or more uninfected erythrocytes to an IE. To assess the involvement of

plasma factors in the rosette formation, the plasma in rosetting medium was substituted for 0,5% of Albumax II. For heat-inactivated plasma, samples were heat inactivated for 30 min at 56°C.

(…)

RNAseq of Peripheral Blood Mononuclear Cells (PBMCs)

mRNA from PBMCs isolated from total blood of patients infected with *P. vivax* with high (n=2) or low (n=2) rosetting was analyzed by RNAseq. For this specific approach, moderate rosetting isolates were considered as 20% rosettes, and low rosetting isolates were considered as less than 10% rosettes. RNA extractions were performed a RNeasy® Micro Kit (Qiagen) according to the manufacturer's instructions. Before conducting WTSS, RNA quality was accessed using an Agilent 2100 Bioanalyzer as well as Agilent RNA 6000 Pico Kit reagents and chips, and it was analyzed using 2100 Expert software according to the recommendations of Agilent Technologies. A SMART-Seq V4 Ultra Low Input RNA Kit was used for sequencing via Clontech's patented Switching Mechanism at 5' End of RNA Template (SMART®) technology. cDNA quality, quantity and size range were evaluated using the BioAnalyzer platform from Agilent Technologies, Inc. using the Agilent High Sensitivity DNA Kit (cDNA, 5 to 500 pg/µL within a size range of 50 to 7000 bp) following the manufacturer's instructions. Prior to generating the final library for Illumina® sequencing, the Covaris AFA system was used for controlled cDNA shearing. cDNA output was then converted into sequencing templates suitable for cluster generation and high-throughput sequencing through the Low Input Library Prep v2 (Clontech Laboratories, Inc.; Takara Bio Company). Library quantification procedures were performed by qPCR using the Library Quantification Kit (Clontech Laboratories, Inc.; Takara Bio Company) and Agilent's High Sensitivity DNA Kit (Agilent Technologies, Inc.), and they were successfully completed before proceeding to the pool setup at a final concentration of 2 nM for direct sequencing. The library was sequenced on a HiSeq 2500 sequencer on Rapid Run mode with the HiSeq Rapid Cluster Kit v2 (100x100) Paired End and HiSeq Rapid SBS Kit v2 (200 cycles). The generated libraries were cluster amplified and sequenced on the Illumina® platform using standard Illumina® reagents and protocols for multiplexed libraries following the manufacturer's loading recommendations. On

average, approximately 474.553 paired end reads were obtained per sample from the 4 sample libraries sequenced. The RNAseq raw reads were checked for quality by running Fast Quality Control (FastQC - https://www.bioinformatics.babraham.ac.uk/projectY/fastqc/). The reads were then subjected to a RNAseq alignment v1.1.1 workflow from BaseSpace platform for cloud-based genomics analysis and storage, and they were integrated with Illumina® sequencers. This workflow allowed trimming of Illumina® adaptors, read mapping using TopHat 2 (Bowtie 2) aligner towards the *Homo sapiens* UCSC hg38 (RefSeq & Gencode gene annotations)[3], and FPKM estimation of reference genes and transcripts using Cufflinks 2. Final assembly and analysis of differentially expressed reference transcripts were performed with Cuffdiff 2 within the Cufflinks Assembly & DE pipeline v2.1.0. Differential gene expression between moderate and low rosetting groups was identified after a pairwise Wilcoxon test was used to compare the transcriptional profiles with the following cutoffs: p-value<0.05, q-value<0.5 and a log2 fold change > 1.5.

(…)

Author's contributions

LA, SCPL, ABIES, RA, AS, JAL, CB and ACAVK performed the experiments. LA, SCPL, CB, BR, LR, and FTMC participated in the data analyses and helped to draft the manuscript. ISS and MVGL contributed with reagents/materials/analysis tools. LA and FTMC conceived the study and wrote the final version of the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

References

1. Lopes SC, Albrecht L, Carvalho BO, Siqueira AM, Thomson-Luque R, Nogueira PA, et al. Paucity of Plasmodium vivax mature schizonts in peripheral blood is associated with their increased cytoadhesive potential. J Infect Dis. 2014;209(9):1403-7. Epub 2014/01/15. doi: jiu018 [pii]10.1093/infdis/jiu018. PubMed PMID: 24415786.
2. Carvalho BO, Lopes SC, Nogueira PA, Orlandi PP, Bargieri DY, Blanco YC, et al. On the cytoadhesion of Plasmodium vivax-infected erythrocytes. The Journal of infectious diseases. 2010;202(4):638-47. doi: 10.1086/654815. PubMed PMID: 20617923.
3. Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, et al. The UCSC Genome Browser database: 2015 update. Nucleic acids research. 2015;43(Database issue):D670-81. doi: 10.1093/nar/gku1177. PubMed PMID: 25428374; PubMed Central PMCID: PMC4383971.

(…)

Figure captions

Figure 1- Features of *Plasmodium vivax* rosettes.
Figure 2 - Peripheral parasitemia, hematocrit and platelet levels of patients.
Figure 3 – The correlation of **total immunoglobulin and naturally acquired antibodies to merozoite proteins relationship with rosetting.**
Figure 4 – Cytokine profile of vivax malaria patients and its correlation with rosetting.
Figure 5- **Non-rosetting parasites are more likely to be phagocytosed by THP-1 cells.**
Figure 6 – Rosetting model in vivax malaria. Proposed model of rosetting role in vivax malaria.

Supporting information captions

SI Fig 1 – Rosetting frequency.
SI Fig 2 – Rosetting frequency and acquired immune response to MSP-1 and AMA-1.
SI Fig 3 – Correlation of rosetting and other parameters.

CHAPTER 5

"Micro RNAs in the Host-Apicomplexan Parasites Interactions: A Review of Immunopathological Aspects"

Overview

MicroRNAs (miRNAs), a class of small non-coding regulatory RNAs, have been detected in a variety of organisms ranging from ancient unicellular eukaryotes to mammals. They have been associated with numerous molecular mechanisms involving developmental, physiological and pathological changes of cells and tissues. Despite the fact that miRNA-silencing mechanisms appear to be absent in some Apicomplexan species, an increasing number of studies have reported a role for miRNAs in host-parasite interactions, which we summarized in this review. Host miRNA expression can change following parasite infection and the consequences can lead, for instance, to parasite clearance. In this context, the immune system signaling appears to have a crucial role.

DISCUSSION

Sequencing the *P. vivax* transcriptome is a significant challenge. From experimental design, beginning with blood sample collection of vivax malaria patients under ethical consent, processing samples keeping enrichment in mind, both in quantity and adhesion phenotypes, to finally sequencing and obtaining a useful amount of raw data. To increase in value the sequencing aspect of this project, an extensive literature review on *P. vivax* "omics" field and applications was published last year (Chapter 1,(Bourgard et al. 2018)), gathering the state of the art information on *Plasmodium* spp. biology and host-parasite interactions. This review particularly focused on key discoveries already achieved in the *P. vivax* sequencing field, and the developments, hurdles, and limitations currently faced by the research community, as well as future perspectives on vivax malaria research (Chapter 1,(Bourgard et al. 2018)).

The characteristically low parasitemia of Brazilian vivax malaria patients and the impossibility of long-term *in vitro* culture (Noulin et al. 2013) defies all the lower thresholds of workable methodologies currently, available (Chapter 1,(Bourgard et al. 2018)). All kits were chosen and tested (when possible) for parasitemias as low as 10^3 parasites, with the objective to optimize the whole processes from collection until extraction of RNA to guarantee successful downstream application of RNA-seq (Chapter 2). The true bottlenecks were sample *ex vivo* maturation, which are historically low (Noulin et al. 2013). The parasitemia greatly impacts quality and purity of the isolated RNA for sequencing (on the order of nano to picograms). Only samples with a viable number of cytoadhesion and rosetting enriched phenotypes will provide good quality and purity of RNA.

As an example, although several short-term *ex vivo* maturation time-lapse experiments were performed successfully, attempts to extract RNA from those samples failed. The bulk part of the sample was made up mainly by healthy erythrocytes, which formed a biological mass too large to extract total RNA from the relatively few parasites present. Overcoming this problem required a complete set of Percoll gradients, which do not exist previously. It required considerable time and effort to establish and optimize this protocol. With adherence phenotypic assessment as my priority, I did not follow this

experimental path. Such experiments would have been important to determine a "safe window" of parasite maturation, in which the *P. vivax* produce and exports parasitic ligand proteins to the surface membrane of the erythrocyte, allowing its reshaping and promoting the capacity of adhesion. This is the basis for both cytoadhesion and rosetting phenotypes. On the other hand, a maturation timing in which we still could pick up the transcription profiles of the unknown genes that codify the target parasitic ligands. However, at the time, three different studies had sequenced the whole-transcriptome of *P. vivax* isolates through all erythrocytic stages (Zhu et al. 2016, Bozdech et al. 2008, Westenberger et al. 2010), which in combination with our most recent published data on cytoadhesion (Lopes et al. 2014) and rosetting (Zhang et al. 2016) phenotypes, gave us insights in the optimal parasite stages to investigate.

On one hand, rosetting assays have been well described and established (Aims Fig. 3. Step 3, personal information shared by Albrecht, L., (Zhang et al. 2016),(Lee et al. 2014) Chapter 3 and 4), and excluding the small aliquot necessary for rosetting rates count, the bulk of the parasites present in the sample could be preserved for RNA isolation. This gave us the opportunity to gather a bigger pool of data to reflect the rosetting phenotypic characteristics present on the currently circulating *P. vivax* populations of the Amazonian endemic area (Chapter 3).

On the other hand, optimization of cytoadhesion assays was required with the objective to recover a proportion of parasites that have adhered to a layer of different CHO cells (Chapter 2). The detachment of the adhered parasites was achieved by altering incubation times followed by careful slide washes with higher pH media than that used during the adhesion assay itself and assuring that the lane of CHO cells remained intact (Fig. 3, Step 2). Duplicate assays were performed to allow the counting of adhered parasites by staining. Only with the duplicate well for each different CHO cell line under assay did we detach and recover the parasites. Since we wanted to evaluate the adherence capacity of *P. vivax* populations against CHO cells expressing constitutively different endothelial receptors, known to be involved in *P.* falciparum binding, a control with CHO[745] cells had to be done accordingly. In summary, all samples were processed immediately after collection following the procedures for parasite isolation and enrichment to obtain parasitemias at least greater than 50%. Thus, a total minimum number of Pv-iE

greater than 400,000 to enable us to proceed with cytoadhesion assays. Given these restrictions, both the low rates of successful *ex vivo* maturation along with the inability to evaluate the cytoadhesion capacity of the parasites against more than two CHO cells lines at the time, fewer assays provided samples that could be further sequenced (Chapter 2).

Low-input kits and single-cell kits for library generation are the only ones available to tackle such difficult samples. By first-hand experience, the majority of technology and reagents applied here are set up for mammalian cell lineages, thus several criteria had to be considered when choosing the kits for RNA-seq. These decisions included between low input or single-cell kits to non-human cell lines, mRNA *versus* total RNA (specially the appeal for an ncRNA setup), depletion of rRNA vs mRNA cDNA targeted synthesis, cDNA shearing, size of read amplification, paired- or non-paired-end, stranded or un-stranded protocol, index management of all samples, and others.

We were able to sequence the whole transcriptome of *P. vivax* from clinical isolates enriched in adhesion phenotypes from the Brazilian Amazon malaria endemic region. A sufficient quantity of data (~2.8M raw reads) was generated using the Illumina® NGS platform (Aims Fig. 3.; view Chapters 2 to 4). In general, although the total number of reads obtained was satisfactory for this experiment (~2.8M reads), we verified that the level of clustering on the flowcell (~1.6M reads on lane 1 and ~1.2M reads on lane 2) executed on cBot was lower than average, possibly because we were dealing with low input sample sizes. In our case, clustering on cBot was performed before and independently of the sequencing run on the HiSeq 2500. As we were cautious to not reach both upper and lower thresholds for accurate clustering we chose the libraries pooled samples as we better judged, given the fact that each combination of machines (cBot and sequencer) has an optimum concentration sample input to load on cBot. Based on an approximate calculation within the big range recommended by Illumina® (8 to 16 µM) and previous experience, we loaded our pooled libraries at an average concentration of 12µM. We now know that in case of library resequencing, we should load as much as 14-15µM of our pooled sample (still within the maximum range recommended by Illumina®), if we follow the same experimental setup.

Nevertheless, the amplification of each library was successful, generating reads of outstanding quality (view Chapters 2 to 4), exceptions being made for one cDNA library

that failed amplification before we attempted sequencing. As estimated according to the concentration of PhiX DNA control loaded for the experimental procedure, about 18% of our reads were automatically excluded from our data. The per-base sequence quality and quality scores (QS≥30) were good and GC content agreed with the expected for *P. vivax* or *P. falciparum* control samples. Reads with N content were low, the sequence length distribution was 100% 100bp for all samples (1bp is added by the sequencer as to sequence 100bp for each pair-ended reads – R1 and R2), and sequence duplication levels were also low, although accounting for some overrepresented sequences detected as primers/adaptors sequencing-through. We could satisfactorily conclude that sequencing itself by the HiSeq 2500 sequencer after cBot clustering went smooth. Even though we were careful in measuring and pooling together each library at the same concentration, samples coming from cytoadhesion assays had fewer parasites (note that from one patient, the sample was divided into amongst four assays), less RNA and thus, less sequencing power and fewer reads obtained. By the nature of rosetting assays, the sample wasn't subdivided and thus, more enriched in parasites, RNA and in final number of reads obtained for the parasite. RNA from *P. falciparum* strains was isolated from ~10^6 parasites to a level that was expected for *P. vivax* sample isolates and not unbalance too much the final pool of libraries.

The RNA-seq data analysis process (Aims Fig. 4) that follows can be technically daunting and difficult to interpret. We must keep in mind that *P. vivax* reference genomes are yet to be fully annotated and the few RNA-seq studies published (review on Chapter 1 (Bourgard et al. 2018)) have already shown transcriptome species-specific properties, sometimes difficult to detect, analyse and interpret using the current bioinformatic tools. To overcome this challenge, I learned and developed several bioinformatic skills: UNIX/LINUX command line, R environment (R studio) and the execution in R of several packages of NGS data analysis in Java, Pearl and Python languages. User friendly platforms such as the BaseSpace from Illumina® and the EuPathDB-Galaxy online workflows were tailored and used in a complementary way. For instance, all the applications for read count and differential gene expression analysis offered on BaseSpace platform do not yet support other genomes than human and a selected few othe model organisms. Furthermore, the available Bowtie Aligner (BWA aligner) at the

BaseSpace platform used previously to align our samples (consisting of FASTQ files) using the BWA-MEM aligner to the *P. vivax* reference genomes (both *P. vivax* P01 or Sal-1, created from imported FASTA files using FASTA Upload App), do not have enough power to properly align our *P. vivax* raw reads. In parallel with application of all bioinformatic basics and programing language skills necessary to perform the next steps of analysis into a UNIX/LINUX command line and R environment, I invested into exploring other platforms for *Plasmodium* spp. data analysis. The EuPathDB-Galaxy hub (https://eupathdb.globusgenomics.org/) offered by Globus Genomics, is a free, interactive, web-based platform for large-scale data analysis that joins all the best characteristics of EuPath database (https://eupathdb.org/eupathdb/), specially from PlasmoDB (http://plasmodb.org/plasmo/) and HostDB (http://hostdb.org/hostdb/), into the well-known Galaxy platform. Galaxy hub (https://usegalaxy.org/) allows NGS data analysis in interactive workflows that can be easily customized or newly created from the several applications pre-loaded on the online platform or in UNIX/LINUX command line, manageable by a bioinformatician without the need for in-deep informatic programming skills. Combined with all EuPathDB pre-loaded reference genomes, the availability of some pre-configured workflows and the possibility of visualization of the results in GBrowse, EuPathDB-Galaxy was extremely useful for delving in into our *P. vivax* RNA-seq data.

The first step into NGS data analysis is the quality assessment of the raw reads obtained directly from the sequencer. For that, Fast Quality Control (FastQC - https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), java software package was executed in UNIX/LINUX command line or accessed through FastQC BaseSpace App or EuPathDB-Galaxy hub. FastQC is a quality control tool for high throughput sequence data, which generates a HTML report for further evaluation. It provides a modular set of analyses which were used to have a quick impression of whether our data had any obvious problems of which we should be aware, before doing any further analysis. Quality control analysis did not reveal any serious problems with the data we obtained, other than some readthrough of Illumina® adaptors, with all our samples having passed the test.

After sequence quality control check, the reads were "trimmed" (cut or excluded) successively considering some important criteria. Low quality reads (*e.g.* reads with *n*' N

bases) were immediately excluded. Good quality reads were trimmed for primer content, by aligning our obtained raw reads with the specific sequences primers from the kit used to synthesize the libraries. Next, the Illumina® adaptor sequences present in our raw reads, originated from sequence read-through the adaptors were filtered out. Afterwards, the long-A/T sequences derived from mRNA poly-A tails, for which our sample preparation was enriched given the fact that we synthesized libraries selective for mRNA, were cut. The trimming process was performed using Trimmomatic (at EuPathDB-Galaxy hub) or the Fast Quality Tool kit (App from BaseSpace based on TagCleaner and Trimmomatic software packs). The trimming process involved the manipulation of FASTQ files, including adapter trimming, quality trimming, length filtering, format conversions and down-sampling. From the statistical summary after trimming (view Chapters 2 to 4), we could observe a difference on the percentage of reads obtained to the different samples, which match accordingly with each sample quantity of initial RNA and consequently the amount of each library for sequencing. Overall, a good quality was verified and only a minimal number of sequences (always < 10%) were excluded during this process, demonstrating that the entire process from RNA extraction until sequencing at the Illumina® platform was successful (view Chapters 2 to 4).

The following step is the alignment of these quality accessed reads to the genomes and/or transcriptomes of reference, a procedure in which duplicated reads, singletons and other secondary aligned sequences and sequences with multiple and/or pairwise discordant alignment were filtered out. The alignment process was altogether applied to the respective reference genomes and transcriptomes. Reference transcriptomes, even if more specific and accurate are tremendously biased by the amount of available, or in our case, the lack of information deposited on the databanks, greatly restricting our downstream analysis. We chose to use reference genomes that gave more coverage and confidence to straighten the differential expression analysis ahead.

Considering our experimental data, read alignment of our *P. vivax* samples was performed to *P. vivax* Sal-1(Carlton, Adams, et al. 2008) reference genome and our reads for *P. falciparum* S20 and FCR3 S1.2 control samples were aligned to the *P. falciparum IT* and *3D7* reference genomes from PlasmoDB(Bahl et al. 2003). Meanwhile, a new enhanced *P. vivax* reference genome P01 was published (Auburn et al. 2016), to which

we   had   access   latter   on   PlasmoDB
(http://plasmodb.org/plasmo/showXmlDataContent.do?name=XmlQuestions.News#plas
modb02_17_release). Given this fact, efforts were focused into aligning these data to this
new *P. vivax P01* reference genome, which betters covers the parasite subtelomeric
regions, where a high percentage of the variant genes (*vir*) are located. As previously
reported, these variant genes might be involved in the molecular mechanisms of parasite
survival and expansion, thus an enhanced assembly reference genome, which also is
better annotated, helped identify them in our transcriptome data. By using TopHat2
(Bowtie aligner) (https://ccb.jhu.edu/software/tophat/index.shtml) our reads were aligned
and mapped against the *P. vivax* P01 reference genome, ranging from 33.2% to 62.7%
of overall read mapping rates of our samples. Additionally, the mapping results were also
analysed to identify splice junctions between exons, as well as insertions and deletions.

In parallel, alignment of our trimmed reads was also performed against *H. sapiens*
reference genome (sequence data was imported from the NCBI SRA using the SRA
Import App) to understand the possible level of host contamination. As expected, some
level of human host contamination was present, which greatly varies from sample to
sample, ranging from 2.6% to 38.7%. Note that, even a minimal contamination with
lymphocytes, which express more RNA than a parasite (~1000x), has serious
repercussions in sequence data output. The bioinformatic tool used to perform this task
was Bowtie Aligner (BWT Aligner) at BaseSpace. This application aligns samples
(consisting of FASTQ files) using the BWA-MEM aligner to a reference genome, including
a custom reference genome created from imported FASTA files. The aligner comprises
the Burrow-Wheeler Aligner (BWA, https://github.com/lh3/bwa) for pairwise alignment (as
proper tool for our sequencing method), SAMtools (http://www.htslib.org/) for BAM file
conversion to SAM and Picard (http://broadinstitute.github.io/picard/) to mark duplicates,
index the files and map them to the reference genome.

Following the data analysis pipeline, the aligned data independent on which reference
genome was used, was subjected to a read count for RPKM (Reads Per Kilobase Million),
FPKM (Fragments Per Kilobase Million) determination and data normalization by
executing the Python code pack HTSeq (http://htseq.readthedocs.io/en/release_0.9.1/),
and finally, differential expression analysis followed with application of the DEseq pack

(http://bioconductor.org/packages/release/bioc/html/DESeq.html) in an R (Bioconductor) environment utilizing the experimental design and biological replicated sequenced.

Overall, several alternative RNA-Seq data analysis workflows at EuPathDB-Galaxy and Illumina® BaseSpace online platforms were intensively applied for our *P. vivax* and *H. sapiens* reads. We executed a total of 6 different pre-configured EuPathDB-Galaxy workflows for Illumina® paired-end RNA-seq analysis of our *P. vivax* aligned raw data. From this preliminary analysis, we have narrowed down the 4 choices of data analysis pipelines (Fig. 1) to single best and more efficient workflow for our RNA-seq data (Fig 1. Workflow 3).
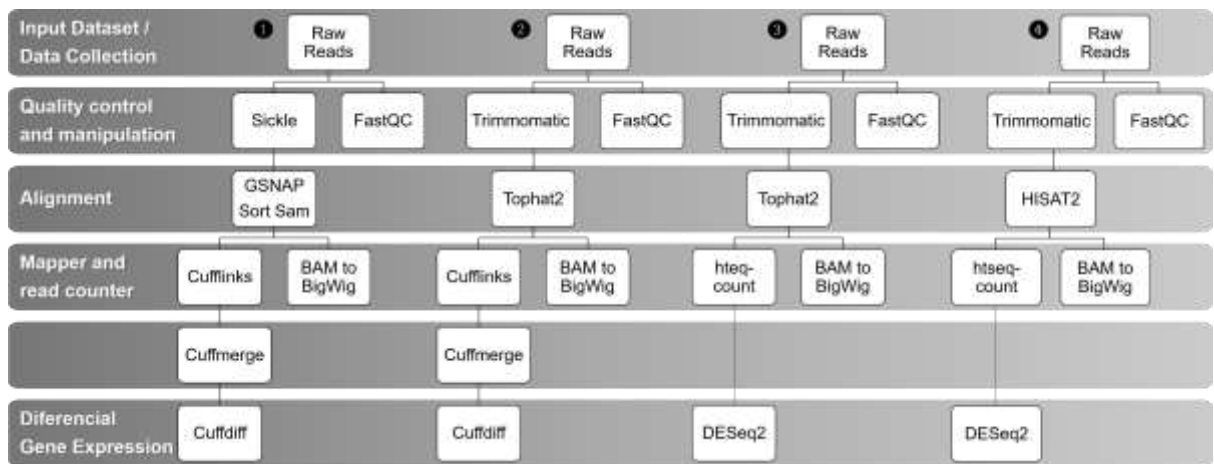
Figure 1. EuPathDB-Galaxy bioinformatics workflows for RNA-seq data analysis tested.

In summary, our RNA-seq raw reads were checked for quality by running Fast Quality Control (FastQC). This workflow allowed the trimming of Illumina® adaptors. Read mapping using Trimmomatic and alignment was done using TopHat2 towards the *Plasmodium* spp. deposited on PlasmoDB. HTseq-cont allowed FPKM estimation of reference genes and transcripts and differential expression of reference transcripts was performed with DESeq2 (Fig. 1, Workflow 3).

One of the interesting points of this RNA-Seq data was analyzing the reads which aligned for *H. sapiens* reference genome. Although the experiment was designed and executed to have as little human host RNA contaminant possible, there is always a small percentage of lymphocytes carrying as much as 10x much RNA per lymphocyte compared to a single Pv-iE that were picked up and sequenced. The generally better-established

platforms for human data analysis allowed considerably faster progress in RNA-seq data analysis, enabling us to obtain gene differential expression results from the few percentages of reads aligning with the human genome reference. By using the user-friendly interface application BaseSpace online platform from Illumina® (https://www.illumina.com/products/by-type/informatics-products/basespace-sequence-hub.html), we explored several different workflows (Fig. 2), evaluating at each step whose option was giving better outputs and general results until we reach the conclusion that the workflow using RNA-Seq Alignment using Tophat2 aligner, coupled with Cufflinks Assembly and Differential Expression was the most appropriate.
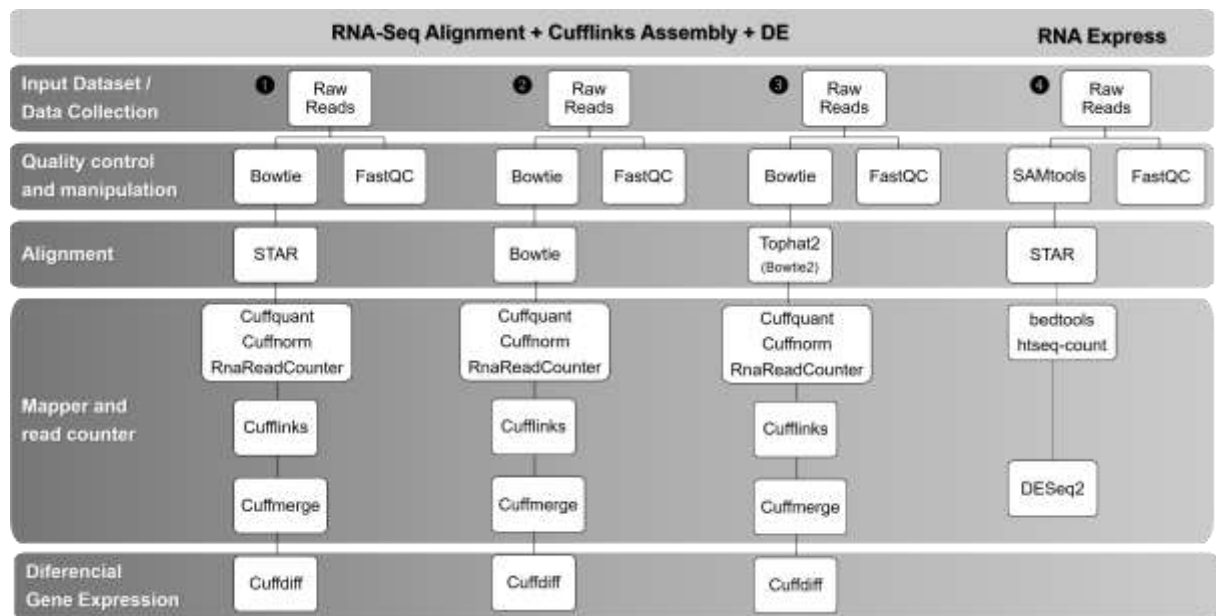
Figure 2. BaseSpace Illumina® sequencing hub bioinformatics workflows for RNA-seq data analysis.

In summary, our RNA-seq raw reads were checked for quality by running Fast Quality Control (FastQC) and subsequently were run through RNA-seq Alignment application from BaseSpace platform for cloud-based genomics analysis and storage, integrated by Illumina® sequencers. This workflow allowed trimming of Illumina® adaptors, read mapping using TopHat2 (Bowtie 2) aligner towards the *Homo sapiens* UCSC hg38 (RefSeq & Gencode gene annotations) (Speir et al. 2016, Rosenbloom et al. 2015), FPKM estimation of reference genes and transcripts using Cufflinks2. Final assembly and differential expression of reference transcripts was performed with Cuffdiff2 within the Cufflinks assembly & DE pipeline (Fig 2., Workflow 3).

For all the several group lists of differential expressed genes, we have analysed Gene Ontology (GO) Enrichment on PlasmoDB (release 37) (Aurrecoechea et al. 2009) to find all GO terms associated to biological processes, cellular components and molecular functions (genes computed and curated for a p-value cutoff of 0.05. towards the reference genomes). Additionally, by using REViGO (Supek et al. 2011), a web server to remove redundant terms, we were able to visualize in semantic similarity-based scatterplots, interactive graphs and tag clouds, the relevant GO terms. Also, were performed searches for Metabolic Pathways Enrichment based on PlasmoDB linked with the well-known KEGG (https://www.genome.jp/kegg/) and MetaCyc (https://metacyc.org/) pathway resources (p-value cutoff of 0.05) were performed.

It is important to stress that Illumina® sequencing technology is based on the generation of short-sized reads. This characteristic does bring the possibility to whole sequence species at higher rates in an easy, fast and less expensive way. However, this short-reads output data also comes at a great cost when taking into consideration the capacity of the currently available software algorithms to perform alignment and mapping tasks, which greatly impact the following read count and differential expression steps. Raw short-sized reads increase the processing difficulty when their sequence is not sufficiently specific to the organism in question (in this case *P. vivax* or *P. falciparum*). For instance, reads correspondent to highly repetitive sequences or with low sequence complexity will easily align with whichever genome/transcriptome we input into the system, especially when reference genomes are still precariously assembled incomplete annotation. Although extensive progress has been made during the last couple of years to have well annotated *P. falciparum* and *P. vivax* reference genomes, the aligned and mapped data shown here were based on the well annotated protein-coding transcripts and corresponding alternatively spliced variants.

Considering the still "under construction" reference genome, the fact that *P. vivax* a have highly repetitive genome particularly on centromeric, telomeric and subtelomeric chromosome regions and the fact that a fair number of reads weren't aligned and mapped accordingly, we will subject them to a *de novo* assembly protocol. The most used *de novo* assemblers are Velvet/Oases (http://www.ebi.ac.uk/~zerbino/velvet/ and http://www.ebi.ac.uk/~zerbino/oases/) and Trinity (trinityrnaseq.sourceforge.net).

Velvet/Oases together use a sophisticated set of algorithms that construct *de Bruijn* graphs and correct for errors and repeats based on information given by paired-end reads (our case). Post-processing of Velvet assemblies, performed by Oases, can be done by defining different k-mer sizes and contig coverage in order to explore and generate a final merged assembly result from all k-mer lengths to reflect a consensus across multiple assemblies and overcome the transcriptome inherently non-uniformity coverage. Trinity produces a transcriptome assembly through three steps, which correspond to discrete modules named "Inchworm", "Chrysalis", and "Butterfly". Inchworm builds initial contigs by finding paths through k-mer graphs, each contig assembly seeded by k-mers in order of k-mer abundance. Chrysalis groups these contigs together given sufficient k-mer overlap and builds *de Bruijn* graphs for these groups, in which the overlaps are nodes and the k-mers connecting edges. Butterfly simplifies the graphs when possible, then reconciles the graphs with original reads (or read pairs) to output individual contigs representative of unique splice variants and paralogous transcripts. In general terms, *de novo* assembly is used when there is no availability of a reference genome or when the one available is not well annotated, is incomplete or, as it can be our case, does not mirror the actual parasite population sequence circulating right now. We know that the *P. vivax* Sal-1 (Carlton, Escalante, et al. 2008) genome reference was adapted to primates, since it was first isolated in the 70's, and *P. vivax* P01 (Auburn et al. 2016) was assembled from 3 clinical Asian isolates. These tools can help to explore the remaining fraction of reads into new transcripts and splice variants, contributing to ameliorate the *P. vivax* reference genome.

To proceed for RNA-seq of *P. vivax* iRBCs, we separate four groups of parasite population samples, with adhesion capacity to CHO^{745} (92U15-2, 93U15-21, 105U15-6 and 101U15-23) and CHO^{ICAM} (92U15-1, 105U15-5), and the correspondent non-adhered samples to CHO^{745} (92U15-4, 93U15-22,105U15-8 and 101U15-24) and CHO^{ICAM} (92U15-3, 105U15-7 and 101U15-20) (Chapter 2). Through RNA-seq data analysis, we accessed the differential gene expression profiles between samples of these two groups and dissect by data mining, possible differences that tentatively might explain *P. vivax* adhesion phenotype during the progress of vivax malaria disease. Although expression profiles were similar between samples within the same capacity level for rosette formation,

our *P. vivax* populations isolated from malaria patients showed some degree of transcriptome heterogeneity. In any of the different groups analyzed we were able to see significantly expressed genes, when considering data normalization (q-values). However, a list of genes with significant p-values could be withdrawn and looked more carefully (Chapter 2, Table 3). Especially considering the comparison group of parasites that adhered to CHO^{ICAM} versus CHO^{745}, we observed a total of 52 genes. A portion of those genes (7) codify conserved *Plasmodium* spp. proteins of yet unknown function. These genes showed a strong (2.3>log2 Fold Change>3.4) absolute expression, which might suggest their involvement in molecular processes important for adhesion phenotype. Future functional characterization of these proteins should clarify this possibility. From the pool of the other genes, 5 genes has caught our attention: one PIR protein (PVP01_0950000), one Plasmodium exported protein of unknown function (PVP01_1201600), the down-expression in adhered samples of the putative translocation protein SEC62 (PVP01_1268900) and two proteins with repetitive domains, the heptatricopeptide repeat-containing protein (PVP01_1416100) and the putative WD repeat-containing protein (PVP01_0905300). Further investigations are needed in order to understand the biological meaning of this results.

To have a better understanding if *P. vivax* expression patterns could be functionally related to rosetting capacity, we performed several rosetting assays on Amazonian low parasitemia clinical isolates. In accordance with previously published data (Zhang et al. 2016), we could observe and confirm a proportional relation between PviRBCs maturation and its enhanced capacity to form rosettes, where young staged parasites showed a low percentage of rosetting (<6%) that progressively augments towards schizont parasites, which can rosette at a rate of 50% rate or more (Chapter 3, Fig. 1). This result is indicative of the fact that *P. vivax* must be expressing proteins responsible for the reshaping of cell membrane of normocytes, which might be directly involved in rosette formation.

Through RNA-seq of *P. vivax* iRBCs, we evaluated the differential gene expression profiles between two groups of samples with different (high against low) rosetting capacity and identified expressed genes that tentatively can explain *P. vivax* rosetting phenotype during the progression of vivax malaria disease. Considering the confounding expression

variability expected between different *P. vivax* clinical isolates, our analysis between the double sample high and low rosetting samples revealed a group of 94 differentially expressed genes (Chapter 3, Table 4).

As anticipated, a large number of genes have not yet been characterized for their protein function, but the conservation of their sequence throughout *Plasmodium* spp. may be indicative of their importance to the parasite rosetting phenotype. Furthermore, the majority of these conserved *Plasmodium* genes showed highly upregulated expression, suggesting their involvement in molecular processes important for erythrocyte binding. These proteins functional characterization should further elucidate this possibility.

Within the group of differentially expressed enzymes, three upregulated kinases have caught our attention, since kinases are classic targets for discovery of new therapeutic drugs. NEK3 has been reported has essential for mitosis progression in *P. berghei* blood-stage development (Tewari et al. 2010), MAPK2 seems to play an important role in stress response in *Toxoplasma gondii* (Huang et al. 2011), and RKIP affects the activity of another kinase, the calcium-dependent protein kinase 1 (Kugelstadt et al. 2007) which regulates several important *Plasmodium* spp. reliant on calcium metabolic processes. Studies on calcium homeostasis have reported that parasitized RBCs show an increased influx of calcium when compared to the decreased efflux of unparasitized RBCs. Calcium content has been localized in the Plasmodium spp. compartment. For the RBC invasion by the merozoite, extracellular calcium is needed, as well as the subsequently parasite development and maturation inside of the erythrocyte (Krishna and Squire-Pollard 1990, Leida, Mahoney, and Eaton 1981).

P. vivax field isolates are often characterized by asynchronous populations of parasites in different stages of development and/or maturation. Although our samples were chosen with the aim to access the transcriptomic profiles of trophozoites and/or early schizont parasites, we could also observe the expression of some interesting gametocyte membrane surface genes, reflecting the importance of the study of mechanisms of *P. vivax* transmission. Together with P48/45 surface protein (van Dijk et al. 2001), P47 is one such protein, reported as required for optimal fertilization in *P. berghei* and for mosquito

immune evasion. This showed a strong signature of natural selection and population structure in the *P. falciparum* and *P. vivax* genomes (Molina-Cruz, Canepa, and Barillas-Mury 2017).

Also, we verified differential expression of two LCCL lectin domain adhesive-like proteins (LAPs), a family of conserved six modular proteins, conserved through the apicomplexan genus, which are expressed in sexual stages of *Plasmodium* parasites and reported to be involved in the formation of protein complexes required for successful *P. berghei* sporogony (Tremp et al. 2017, Pradel et al. 2004).

One of the most important aspects of erythrocyte infection by *P. vivax* is the dramatic morphogenesis of the erythrocyte membrane, driven by a network of microtubules (MT) sustained by the inner membrane complex (IMC). As expected, genes encoding the actin and tubulin backbone molecules of MTs were found to be upregulated in our rosetting samples, together with a group of MT motor enzymes, IMC proteins such as PhIL1 (PIP2 and 3-)integrating proteins, which are critical in various processes such as signal transduction and intracellular and membrane trafficking (Parkyn Schneider et al. 2017, Ebrahimzadeh, Mukherjee, and Richard 2018).

Furthermore, we were able to catch the overexpression of an early-transcribed membrane protein (ETRAMP). ETRAMPs are important proteins present on the membrane of intracellular parasites, such as *Plasmodium* species, formed during erythrocyte invasion as an invagination of the iE cell surface during the asexual blood stage parasites. Recent studies showed that ETRAMPs have been localized on the intracellular membranes of immature schizont and at the apical organelles of newly *formed P. vivax* merozoites of mature schizont and have the capacity to elicit high antibody titers capable of recognizing parasites of vivax malaria patients (Cheng et al. 2015).

Together with two other proteins from CPW-WPC surface protein family, another important membrane protein found differentially expressed in our study was the *P. vivax* (MSP7)-like protein. Merozoite surface proteins belong to families of proteins often involved in complex *Plasmodium* invasion processes. Pf MSP7 interactions with host P-lectin receptors have been previously demonstrated (Perrin, Bartholdson, and Wright 2015), which consequently block interactions between host P-selection and leukocyte ligands and could underlie the mechanism for the known immunomodulatory effects of

both MSP7 and P-selectin in malaria infection models. Although MSP7 in *P. vivax* has not yet been functionally characterized, there is evidence this protein is under selection, and thus functionally important in *P. vivax* (Castillo, Andreina Pacheco, and Escalante 2017). In addition, 4 glideosome-associated proteins are observed to be differentially expressed in our high rosetting parasites. The capacity to bind, reorient and invade new host cells is mainly powered by the "glideosome" proteins (Keeley and Soldati 2004). The glideosome is a macromolecular complex comprising proteins with adhesive properties. These proteins are released apically on the parasite membrane and translocated to the opposite pole of the parasite through the actomyosin system anchored in the IMC,

Finally, we detected the expression of the macrophage migration inhibitory factor (MIF) gene, one of the first cytokines described, which has a broad range of pro-inflammatory properties. It has been reported that expressed PfMIF protein localizes to the Maurer's cleft during asexual blood stage parasites. PfMIF *in vitro* treatment of human monocytes inhibited their random migration and reduced the surface expression of toll like receptor (TLR) 2, TLR4 and CD86, indicating that its release potentially modulates the host monocytes functions during *Plasmodium* acute infection (Cordery et al. 2007). In accordance with this data, our analysis performed in isolates from non-severe vivax malaria patients reported a significantly downregulated expression of the *P. vivax mif* gene in parasite populations showing rosetting phenotypes.

Together, these results point out to the importance of membrane and membrane associate proteins in rosetting phenotype. Functional assays might further clarify that these proteins allow the parasites to adhere to the surface of host cells, such as health erythrocytes, maintaining them anchored in order to create the characteristic rosette of surrounding erythrocytes, which enable the parasite to evade from the host immune system.

Considering differences in cytokine levels observed between moderate and low rosetting isolates, differential gene expression from the few reads of human peripheral blood mononuclear cells sequenced was performed. Our aim was to evaluate possible relationships between leukocytes expression profiling from *P. vivax* infected patients (Chapter 4, Table 1A) with a rosetting phenotype. Identified were eight genes that showed a log2 fold change higher than 1.5 (Chapter 4, Table 1B). Three out of eight genes were

in the Fc gamma receptor (FCGR)-dependent phagocytosis pathway. Remarkably, immunoglobulin kappa constant (IGKC) and immunoglobulin heavy constant gamma 1 (IGHG1) were upregulated and actin-related protein 2/3 complex subunit 2 (ARPC2) was downregulated in individuals with moderate rosetting compared to patients with low rosetting. The host transcriptomic profile directly compared patient isolates with low versus moderate rosetting, which were quantified immediately after blood sample collection, when most parasites were still young trophozoites (Lopes et al. 2014). Because rosetting is mostly formed by mature parasites (schizonts), this constituted an experimental limitation of this study. Additionally, even with the high quality of sequenced reads obtained in our RNA-seq, the low quantity of host sample input effectively transcribed resulted in few statistically significant pools of differentially expressed genes. Nevertheless, three out of eight genes showing more than 1.5-fold change in expression have relevant functions in the Fc gamma phagocytic pathway, demonstrating the importance of this and other future RNA-seq projects tailored to catch host and parasite transcriptomes with regards to its interactions during vivax malaria disease progression.

The present transcriptome findings, supported by our results showing that plasma from patients with high rosetting levels inhibit the phagocytic capacity of THP-1 cells, support the idea that *P. vivax* rosettes strongly inhibit the phagocytosis pathway. Because transcriptional analysis indicated that a phagocytosis pathway may be affected by rosetting, functional assays were conducted to investigate the role of this adhesive phenomenon. As rosetting is dependent on plasma factors, the capacity of plasma from individuals with different rates of rosetting to facilitate or inhibit phagocytosis to some extent was investigated. The ability of *P. vivax* isolates to form rosettes at different degrees and phagocytose was inversely correlated with rosetting levels. The most likely explanation for this observation is that the noninfected erythrocytes in the rosette formation shielded the iE from antibodies and provided a physical barrier to restrict contact with phagocytes and other effector cells of the immune system. However, additional studies are needed to search for plasma factors that can inhibit the phagocytosing capacity of macrophages.

Taken together, these data indicated that vivax malaria rosetting is an evasion mechanism that allows the parasite to shield itself from the host immune system (Chapter

4, Fig. 6), and this generates more questions regarding this common phenotype. The fact that rosetting is a frequent feature in *P. vivax* late stages and that the high prevalence of rosettes may indicate that this phenotype is an important advantage for the parasite by conferring significant protection from the host immune system. An understanding of *P. vivax* rosettes may help the scientific community to develop new strategies for malaria control.

Although the *P. vivax* RNA-seq project described here only targets protein-coding transcripts, it is known that *Plasmodium* spp. express a relatively high amount of non-coding RNAs (ncRNAs) (Zhu et al. 2016), which have important, albeit unknown, regulatory functions. One important class of small ncRNAs are the microRNAs (miRNAs), which have been detected in a variety of organisms ranging from ancient unicellular protozoans to mammals. miRNAs have been associated with numerous molecular mechanisms, such as those involved in developmental, physiological and pathological changes of cells and tissues. Even though miRNA-silencing mechanisms appear to be absent in a number of protozoan parasites, including *Plasmodium* spp., an increasing number of studies have reported a role for miRNAs in host-parasite interactions. Host miRNA expression can change following parasite infection and the consequences can lead, for instance, to parasite clearance. In this context, the immune system signaling appears to play a crucial role. As such, future research avenues should focus on the sequencing, identification and functional study of the different ncRNAs through parasite infection and host-parasite interactions during disease progress (Chapter 5, (Judice et al. 2016)). For instance, further miRNA research may (i) uncover new biomarkers related with disease progression, (ii) determine miRNA target genes that can clarify the miRNAs role in the pathogenesis and finally (iii) aid in the discovery of new therapeutic targets.

CONCLUSION

Recent observations of increasingly severe vivax malaria cases are rising awareness of the fact that *P. vivax* is one of most important, widespread and neglected human malaria parasites. It is responsible for millions of malaria cases per year, thus having a worldwide strong social impact. Unfortunately, little is known concerning this parasite biology and the molecular mechanisms involved in vivax malaria immuno-pathogenicity.

Transcriptomes of *P. vivax* clinical isolates are not routinely explored, and at the time of this writing, few published RNA-Seq datasets exist (Chapter 1, (Bourgard et al. 2018)). Our study aims to benefit the scientific community in several ways. Firstly, the study provides a guide for processing *P. vivax* clinical isolates from collection in the field through Illumina® library construction (Chapter 2). Secondly, it provides publicly available transcriptome sequencing datasets for *P. vivax* clinical isolates, which will be used for the improvement of reference gene models, such as 5' and 3' untranslated regions, novel gene transcripts, and alternative splicing events. Finally, the transcriptomes will inform our understanding of *P. vivax* invasion and host evasion mechanisms by identifying genes that are up-regulated in mature parasites and enabling inter-isolate comparisons of expression patterns, including the evaluation of multi-gene families known to be important for erythrocyte rheopathological properties.

Particularly, we believe that the experiments performed by us, calling upon the new technological tools currently available, and further bioinformatically analysed (Chapter 1, (Bourgard et al. 2018) and 5, (Judice et al. 2016)), will allow the mapping of the potential factors behind *P. vivax* cytoadhesion (Chapter 2) and rosette formation (Chapter 3) and identify the parasitic ligands, human host endothelial receptors and, most importantly, the pathways involved in these host-pathogen interactions (Chapters 2 and 4). Such knowledge could be of paramount importance for the development of vaccine (Annex 1, (Bittencourt et al. 2018)) and chemotherapeutic strategies (Annex 2) by identifying new molecular targets. Our expectation is that such knowledge of the vast biology of *P. vivax* (Chapter 5, (Judice et al. 2016)) will open new avenues on the development of more efficacious treatments for vivax malaria patients in the near future.

SCIENTIFIC IMPACT

As above mentioned on this brief introduction, both in the "omics" data generation field and new platforms for drug screening, there are outstanding technological achievements to aid the *P. vivax* research community in drug and vaccine discovery by getting a better knowledge of the parasite biology. Most importantly, data integration in *Plasmodium* spp. made available on databases such as PlasmoDB (Bahl et al. 2003), can aid us into understanding host-parasite interactions, by using the powerful bioinformatic tools and/or following very well thought-through *in silico* workflows and/or pipelines for data analysis. Today, such methodological approaches certainly open further the investigations on the biology of the *P. vivax* apicomplexan parasite, to succeed on the now recognized vivax malaria burden, and thus, steadily walk to accomplish the set agenda of malaria eradication.

FINANCIAL AIDS, COLLABORATIONS AND INSTITUTIONAL SUPPORT

My supervisor, Prof. Dr. Fabio T. M. Costa, is responsible and leader of our research group in Immuno-parasitology and Experimental Malaria at Tropical Diseases Laboratory (LDT), Biology Institute – UNICAMP, Campinas – SP, where we have an excellent infrastructure and also, financial and scientific support for a sustained development of this project. Within the Genetics and Molecular Biology PhD program, I too have enjoyed from CAPES funding to attend scientific events and the availability of Biology Institute infrastructure support, as well as a continuously scientific academic offer.

Our research group maintain direct collaboration with Prof. Dr. Marcus V. G. Lacerda from Fundação de Medicina Tropical Dr. Heitor Vieira Dourado, Manaus – AM. The two entities, LDT and FMT-HVD, gathered all necessary conditions and support for the complete execution of periods of experimental field work at malaria vivax Amazonian endemic region, indispensable to this project.

During my PhD project, several technological platforms have been used, between them, all the ones directly related to the quality and quantity analysis of genetic material (RNA and DNA) (*BioAnalyser*, ©Agilent Technologies, Inc.), execution of all protocols for transcriptomic library construction for WTSS of our samples in suitable equipment, following Illumina® high-throughput platform. For that, our research group has a very interactive collaboration with Prof. Dr. Letusa Albrecht (my PhD co-supervisor), principal investigator at Instituto Carlos Chagas (Fiocruz-Paraná), Curitiba-PR and scientist collaborator at Genetics, Evolution and Microbiology and Immunology department at Biology Institute – UNICAMP. We sequenced all our samples at the USP-ESALQ Sequencing Facility, Piracicaba-SP, coordinated by Prof. Dr. Luiz Lehmann Coutinho and executed by the lab technician Dr. Ricardo Brassaloti, which gave us full support to

accomplish the RNA-seq runs, as well as the possibility to do all follow-up of Illumina$^{®}$ sequencing platform.

Within existing Brazil-Sweden collaborations and backed-up by Prof. Dr. Per Sunnerhagen for University of Gothenburg (UGOT), we have in place a project (support #2853) to the Sweden Bioinformatics Infrastructure for Life Sciences (BILS/SciLifeLab) institute for RNA-seq data analysis support and my tutoring on the bioinformatic research field. The support of BILS/SciLifeLab during the one year internship at UGOT has been indispensable for this project, both by the tutoring and platforms (e.g. servers) offered at the University, network between bioinformatician Swedish community from BILS and SciLifeLab, and with several researchers who tutored me during EuPathDB, EMBL and EMBL-EBI courses, which gave me a solid know-how for present and future endeavors into "omics" fields.

References

Adiconis, X., D. Borges-Rivera, R. Satija, D. S. DeLuca, M. A. Busby, A. M. Berlin, A. Sivachenko, D. A. Thompson, A. Wysoker, T. Fennell, A. Gnirke, N. Pochet, A. Regev, and J. Z. Levin. 2013. "Comparative analysis of RNA sequencing methods for degraded or low-input samples." *Nat Methods* no. 10 (7):623-9. doi: 10.1038/nmeth.2483.

Alexandre, M. A., C. O. Ferreira, A. M. Siqueira, B. L. Magalhaes, M. P. Mourao, M. V. Lacerda, and Md Alecrim. 2010. "Severe Plasmodium vivax malaria, Brazilian Amazon." *Emerg Infect Dis* no. 16 (10):1611-4. doi: 10.3201/eid1610.100685.

Anders, S., D. J. McCarthy, Y. Chen, M. Okoniewski, G. K. Smyth, W. Huber, and M. D. Robinson. 2013. "Count-based differential expression analysis of RNA sequencing data using R and Bioconductor." *Nat Protoc* no. 8 (9):1765-86. doi: 10.1038/nprot.2013.099.

Andrade, B. B., A. Reis-Filho, S. M. Souza-Neto, J. Clarencio, L. M. Camargo, A. Barral, and M. Barral-Netto. 2010. "Severe Plasmodium vivax malaria exhibits marked inflammatory imbalance." *Malar J* no. 9:13. doi: 10.1186/1475-2875-9-13.

Anstey, N. M., T. Handojo, M. C. Pain, E. Kenangalem, E. Tjitra, R. N. Price, and G. P. Maguire. 2007. "Lung injury in vivax malaria: pathophysiological evidence for pulmonary vascular sequestration and posttreatment alveolar-capillary inflammation." *J Infect Dis* no. 195 (4):589-96. doi: 10.1086/510756.

Anstey, N. M., S. P. Jacups, T. Cain, T. Pearson, P. J. Ziesing, D. A. Fisher, B. J. Currie, P. J. Marks, and G. P. Maguire. 2002. "Pulmonary manifestations of uncomplicated falciparum and vivax malaria: cough, small airways obstruction, impaired gas transfer, and increased pulmonary phagocytic activity." *J Infect Dis* no. 185 (9):1326-34. doi: 10.1086/339885.

Anstey, N. M., B. Russell, T. W. Yeo, and R. N. Price. 2009. "The pathophysiology of vivax malaria." *Trends Parasitol* no. 25 (5):220-7. doi: 10.1016/j.pt.2009.02.003.

Arama, C., and M. Troye-Blomberg. 2014. "The path of malaria vaccine development: challenges and perspectives." *J Intern Med* no. 275 (5):456-66. doi: 10.1111/joim.12223.

Auburn, S., U. Bohme, S. Steinbiss, H. Trimarsanto, J. Hostetler, M. Sanders, Q. Gao, F. Nosten, C. I. Newbold, M. Berriman, R. N. Price, and T. D. Otto. 2016. "A new Plasmodium vivax reference sequence with improved assembly of the subtelomeres reveals an abundance of pir genes." *Wellcome Open Res* no. 1:4. doi: 10.12688/wellcomeopenres.9876.1.

Auburn, S., J. Marfurt, G. Maslen, S. Campino, V. Ruano Rubio, M. Manske, B. Machunter, E. Kenangalem, R. Noviyanti, L. Trianty, B. Sebayang, G. Wirjanata, K. Sriprawat, D. Alcock, B. Macinnis, O. Miotto, T. G. Clark, B. Russell, N. M. Anstey, F. Nosten, D. P. Kwiatkowski, and R. N. Price. 2013. "Effective preparation of Plasmodium vivax field isolates for high-throughput whole genome sequencing." *PLoS One* no. 8 (1):e53160. doi: 10.1371/journal.pone.0053160.

Aurrecoechea, C., J. Brestelli, B. P. Brunk, J. Dommer, S. Fischer, B. Gajria, X. Gao, A. Gingle, G. Grant, O. S. Harb, M. Heiges, F. Innamorato, J. Iodice, J. C. Kissinger, E. Kraemer, W. Li, J. A. Miller, V. Nayak, C. Pennington, D. F. Pinney, D. S. Roos, C. Ross, C. J. Stoeckert, Jr., C. Treatman, and H. Wang. 2009. "PlasmoDB: a functional genomic database for malaria parasites." *Nucleic Acids Res* no. 37 (Database issue):D539-43. doi: 10.1093/nar/gkn814.

Bahl, A., B. Brunk, J. Crabtree, M. J. Fraunholz, B. Gajria, G. R. Grant, H. Ginsburg, D. Gupta, J. C. Kissinger, P. Labo, L. Li, M. D. Mailman, A. J. Milgram, D. S. Pearson, D. S. Roos, J. Schug, C. J. Stoeckert, Jr., and P. Whetzel. 2003. "PlasmoDB: the Plasmodium genome resource. A database integrating experimental and computational data." *Nucleic Acids Res* no. 31 (1):212-5.

Baird, J. K. 2004. "Chloroquine resistance in Plasmodium vivax." *Antimicrob Agents Chemother* no. 48 (11):4075-83. doi: 10.1128/AAC.48.11.4075-4083.2004.

Baird, J. K. 2007. "Neglect of Plasmodium vivax malaria." *Trends Parasitol* no. 23 (11):533-9. doi: 10.1016/j.pt.2007.08.011.

Baird, J. K. 2010. "Eliminating malaria--all of them." *Lancet* no. 376 (9756):1883-5. doi: 10.1016/S0140-6736(10)61494-8.

Baird, J. K., E. Schwartz, and S. L. Hoffman. 2007. "Prevention and treatment of vivax malaria." *Curr Infect Dis Rep* no. 9 (1):39-46.

Barber, B. E., T. William, M. J. Grigg, U. Parameswaran, K. A. Piera, R. N. Price, T. W. Yeo, and N. M. Anstey. 2015. "Parasite biomass-related inflammation, endothelial activation, microvascular dysfunction and disease severity in vivax malaria." *PLoS Pathog* no. 11 (1):e1004558. doi: 10.1371/journal.ppat.1004558.

Barcus, M. J., H. Basri, H. Picarima, C. Manyakori, Sekartuti, I. Elyazar, M. J. Bangs, J. D. Maguire, and J. K. Baird. 2007. "Demographic risk factors for severe and fatal vivax and falciparum malaria among hospital admissions in northeastern Indonesian Papua." *Am J Trop Med Hyg* no. 77 (5):984-91. doi: 10.4269/ajtmh.2007.77.984.

Barnwell, J. W., P. Ingravallo, M. R. Galinski, Y. Matsumoto, and M. Aikawa. 1990. "Plasmodium vivax: malarial proteins associated with the membrane-bound caveola-vesicle complexes and cytoplasmic cleft structures of infected erythrocytes." *Exp Parasitol* no. 70 (1):85-99.

Barragan, A., P. G. Kremsner, M. Wahlgren, and J. Carlson. 2000. "Blood group A antigen is a coreceptor in Plasmodium falciparum rosetting." *Infect Immun* no. 68 (5):2971-5.

Battle, K. E., P. W. Gething, I. R. Elyazar, C. L. Moyes, M. E. Sinka, R. E. Howes, C. A. Guerra, R. N. Price, K. J. Baird, and S. I. Hay. 2012. "The global public health significance of Plasmodium vivax." *Adv Parasitol* no. 80:1-111. doi: 10.1016/B978-0-12-397900-1.00001-3.

Battle, K. E., M. S. Karhunen, S. Bhatt, P. W. Gething, R. E. Howes, N. Golding, T. P. Van Boeckel, J. P. Messina, G. D. Shanks, D. L. Smith, J. K. Baird, and S. I. Hay. 2014. "Geographical variation in Plasmodium vivax relapse." *Malar J* no. 13:144. doi: 10.1186/1475-2875-13-144.

Beeson, J. G., and B. S. Crabb. 2007. "Towards a vaccine against Plasmodium vivax malaria." *PLoS Med* no. 4 (12):e350. doi: 10.1371/journal.pmed.0040350.

Bernabeu, M., F. J. Lopez, M. Ferrer, L. Martin-Jaular, A. Razaname, G. Corradin, A. G. Maier, H. A. Del Portillo, and C. Fernandez-Becerra. 2012. "Functional analysis of Plasmodium vivax VIR proteins reveals different subcellular localizations and cytoadherence to the ICAM-1 endothelial receptor." *Cell Microbiol* no. 14 (3):386-400. doi: 10.1111/j.1462-5822.2011.01726.x.

Betuela, I., A. Rosanas-Urgell, B. Kiniboro, D. I. Stanisic, L. Samol, E. de Lazzari, H. A. Del Portillo, P. Siba, P. L. Alonso, Q. Bassat, and I. Mueller. 2012. "Relapses contribute significantly to the risk of Plasmodium vivax infection and disease in Papua New Guinean children 1-5 years of age." *J Infect Dis* no. 206 (11):1771-80. doi: 10.1093/infdis/jis580.

Beutler, E., S. Duparc, and G. Pd Deficiency Working Group. 2007. "Glucose-6-phosphate dehydrogenase deficiency and antimalarial drug development." *Am J Trop Med Hyg* no. 77 (4):779-89.

Bitoh, T., K. Fueda, H. Ohmae, M. Watanabe, and H. Ishikawa. 2011. "Risk analysis of the re-emergence of Plasmodium vivax malaria in Japan using a stochastic transmission model." *Environ Health Prev Med* no. 16 (3):171-7. doi: 10.1007/s12199-010-0184-8.

Bittencourt, N. C., J. A. Leite, Abie Silva, T. S. Pimenta, J. L. Silva-Filho, G. C. Cassiano, S. C. P. Lopes, J. C. K. Dos-Santos, C. Bourgard, H. I. Nakaya, A. M. R. da Silva Ventura, M. V. G. Lacerda, M. U. Ferreira, R. L. D. Machado, L. Albrecht, and F. T. M. Costa. 2018. "Genetic sequence characterization and naturally acquired immune response to Plasmodium vivax Rhoptry Neck Protein 2 (PvRON2)." *Malar J* no. 17 (1):401. doi: 10.1186/s12936-018-2543-7.

Bockarie, M. J., and H. Dagoro. 2006. "Are insecticide-treated bednets more protective against Plasmodium falciparum than Plasmodium vivax-infected mosquitoes?" *Malar J* no. 5:15. doi: 10.1186/1475-2875-5-15.

Boopathi, P. A., A. K. Subudhi, S. Garg, S. Middha, J. Acharya, D. Pakalapati, V. Saxena, M. Aiyaz, B. Chand, R. C. Mugasimangalam, S. K. Kochar, P. Sirohi, D. K. Kochar, and A. Das. 2013. "Revealing natural antisense transcripts from Plasmodium vivax isolates: evidence of genome regulation in complicated malaria." *Infect Genet Evol* no. 20:428-43. doi: 10.1016/j.meegid.2013.09.026.

Boopathi, P. A., A. K. Subudhi, S. Garg, S. Middha, J. Acharya, D. Pakalapati, V. Saxena, M. Aiyaz, B. Chand, R. C. Mugasimangalam, S. K. Kochar, P. Sirohi, D. K. Kochar, and A. Das. 2014. "Dataset of natural antisense transcripts in P. vivax clinical isolates derived using custom designed strand-specific microarray." *Genom Data* no. 2:199-201. doi: 10.1016/j.gdata.2014.06.024.

Bourgard, C., L. Albrecht, Acav Kayano, P. Sunnerhagen, and F. T. M. Costa. 2018. "Plasmodium vivax Biology: Insights Provided by Genomics, Transcriptomics and Proteomics." *Front Cell Infect Microbiol* no. 8:34. doi: 10.3389/fcimb.2018.00034.

Bousema, T., and C. Drakeley. 2011. "Epidemiology and infectivity of Plasmodium falciparum and Plasmodium vivax gametocytes in relation to malaria control and elimination." *Clin Microbiol Rev* no. 24 (2):377-410. doi: 10.1128/CMR.00051-10.

Bowman, Z. S., J. E. Oatis, Jr., J. L. Whelan, D. J. Jollow, and D. C. McMillan. 2004. "Primaquine-induced hemolytic anemia: susceptibility of normal versus glutathione-depleted rat erythrocytes to 5-hydroxyprimaquine." *J Pharmacol Exp Ther* no. 309 (1):79-85. doi: 10.1124/jpet.103.062984.

Boyd, M. F., Kitchen, S. F. 1937. "On the infectiousness of patients infected with Plasmodium vivax and Plasmodium falciparum." *Am J Trop Med Hyg* no. 17:253-62.

Bozdech, Z., S. Mok, G. Hu, M. Imwong, A. Jaidee, B. Russell, H. Ginsburg, F. Nosten, N. P. Day, N. J. White, J. M. Carlton, and P. R. Preiser. 2008. "The transcriptome of Plasmodium vivax reveals divergence and diversity of transcriptional regulation in malaria parasites." *Proc Natl Acad Sci U S A* no. 105 (42):16290-5. doi: 10.1073/pnas.0807404105.

Buffet, P. A., I. Safeukui, G. Deplaine, V. Brousse, V. Prendki, M. Thellier, G. D. Turner, and O. Mercereau-Puijalon. 2011. "The pathogenesis of Plasmodium falciparum malaria in humans: insights from splenic physiology." *Blood* no. 117 (2):381-92. doi: 10.1182/blood-2010-04-202911.

Carlson, J., G. Holmquist, D. W. Taylor, P. Perlmann, and M. Wahlgren. 1990. "Antibodies to a histidine-rich protein (PfHRP1) disrupt spontaneously formed Plasmodium falciparum erythrocyte rosettes." *Proc Natl Acad Sci U S A* no. 87 (7):2511-5.

Carlton, J. M., J. H. Adams, J. C. Silva, S. L. Bidwell, H. Lorenzi, E. Caler, J. Crabtree, S. V. Angiuoli, E. F. Merino, P. Amedeo, Q. Cheng, R. M. Coulson, B. S. Crabb, H. A. Del Portillo, K. Essien, T. V. Feldblyum, C. Fernandez-Becerra, P. R. Gilson, A. H. Gueye, X. Guo, S. Kang'a, T. W. Kooij, M. Korsinczky, E. V. Meyer, V. Nene, I. Paulsen, O. White, S. A. Ralph, Q. Ren, T. J. Sargeant, S. L. Salzberg, C. J. Stoeckert, S. A. Sullivan, M. M. Yamamoto, S. L. Hoffman, J. R. Wortman, M. J. Gardner, M. R. Galinski, J. W. Barnwell, and C. M. Fraser-Liggett. 2008. "Comparative genomics of the neglected human malaria parasite Plasmodium vivax." *Nature* no. 455 (7214):757-63. doi: 10.1038/nature07327.

Carlton, J. M., A. A. Escalante, D. Neafsey, and S. K. Volkman. 2008. "Comparative evolutionary genomics of human malaria parasites." *Trends Parasitol* no. 24 (12):545-50. doi: 10.1016/j.pt.2008.09.003.

Carlton, J. M., B. J. Sina, and J. H. Adams. 2011. "Why is Plasmodium vivax a neglected tropical disease?" *PLoS Negl Trop Dis* no. 5 (6):e1160. doi: 10.1371/journal.pntd.0001160.

Carvalho, B. O., S. C. Lopes, P. A. Nogueira, P. P. Orlandi, D. Y. Bargieri, Y. C. Blanco, R. Mamoni, J. A. Leite, M. M. Rodrigues, I. S. Soares, T. R. Oliveira, G. Wunderlich, M. V. Lacerda, H. A. del Portillo, M. O. Araujo, B. Russell, R. Suwanarusk, G. Snounou, L. Renia, and F. T. Costa. 2010. "On the cytoadhesion of Plasmodium vivax-infected erythrocytes." *J Infect Dis* no. 202 (4):638-47. doi: 10.1086/654815.

Castillo, A. I., M. Andreina Pacheco, and A. A. Escalante. 2017. "Evolution of the merozoite surface protein 7 (msp7) family in Plasmodium vivax and P. falciparum: A comparative approach." *Infect Genet Evol* no. 50:7-19. doi: 10.1016/j.meegid.2017.01.024.

Chan, E. R., D. Menard, P. H. David, A. Ratsimbasoa, S. Kim, P. Chim, C. Do, B. Witkowski, O. Mercereau-Puijalon, P. A. Zimmerman, and D. Serre. 2012. "Whole genome sequencing of field isolates provides robust characterization of genetic diversity in Plasmodium vivax." *PLoS Negl Trop Dis* no. 6 (9):e1811. doi: 10.1371/journal.pntd.0001811.

Cheng, Y., F. Lu, S. K. Lee, D. H. Kong, K. S. Ha, B. Wang, J. Sattabongkot, T. Tsuboi, and E. T. Han. 2015. "Characterization of Plasmodium vivax Early Transcribed Membrane Protein 11.2 and Exported Protein 1." *PLoS One* no. 10 (5):e0127500. doi: 10.1371/journal.pone.0127500.

Childs, D. Z., I. M. Cattadori, W. Suwonkerd, S. Prajakwong, and M. Boots. 2006. "Spatiotemporal patterns of malaria incidence in northern Thailand." *Trans R Soc Trop Med Hyg* no. 100 (7):623-31. doi: 10.1016/j.trstmh.2005.09.011.

Chotinavich, K. et al. 2003. "The adhesion receptors of P. vivax-infected red cells." *Exp. Parasitol.* no. 105:33-34.

Chotivanich, K. T., S. Pukrittayakamee, J. A. Simpson, N. J. White, and R. Udomsangpetch. 1998. "Characteristics of Plasmodium vivax-infected erythrocyte rosettes." *Am J Trop Med Hyg* no. 59 (1):73-6.

Chotivanich, K. T., R. Udomsangpetch, B. Pipitaporn, B. Angus, Y. Suputtamongkol, S. Pukrittayakamee, and N. J. White. 1998. "Rosetting characteristics of uninfected erythrocytes from healthy individuals and malaria patients." *Ann Trop Med Parasitol* no. 92 (1):45-56.

Chotivanich, K., R. Udomsangpetch, R. Suwanarusk, S. Pukrittayakamee, P. Wilairatana, J. G. Beeson, N. P. Day, and N. J. White. 2012. "Plasmodium vivax adherence to placental glycosaminoglycans." *PLoS One* no. 7 (4):e34509. doi: 10.1371/journal.pone.0034509.

Chung, B. H., S. W. Lee, S. E. Lee, T. J. Hwang, and H. S. Shin. 2008. "Predictors of Plasmodium vivax malaria-induced nephropathy in young Korean men." *Nephron Clin Pract* no. 110 (3):c172-7. doi: 10.1159/000167023.

Collins, W. E., G. M. Jeffery, and J. M. Roberts. 2003. "A retrospective examination of anemia during infection of humans with Plasmodium vivax." *Am J Trop Med Hyg* no. 68 (4):410-2.

Conesa, A., P. Madrigal, S. Tarazona, D. Gomez-Cabrero, A. Cervera, A. McPherson, M. W. Szczesniak, D. J. Gaffney, L. L. Elo, X. Zhang, and A. Mortazavi. 2016. "A survey of best practices for RNA-seq data analysis." *Genome Biol* no. 17:13. doi: 10.1186/s13059-016-0881-8.

Cordery, D. V., U. Kishore, S. Kyes, M. J. Shafi, K. R. Watkins, T. N. Williams, K. Marsh, and B. C. Urban. 2007. "Characterization of a Plasmodium falciparum macrophage-migration inhibitory factor homologue." *J Infect Dis* no. 195 (6):905-12. doi: 10.1086/511309.

Coura, J. R., M. Suarez-Mutis, and S. Ladeia-Andrade. 2006. "A new challenge for malaria control in Brazil: asymptomatic Plasmodium infection--a review." *Mem Inst Oswaldo Cruz* no. 101 (3):229-37.

Cowman, A. F., C. J. Tonkin, W. H. Tham, and M. T. Duraisingh. 2017. "The Molecular Basis of Erythrocyte Invasion by Malaria Parasites." *Cell Host Microbe* no. 22 (2):232-245. doi: 10.1016/j.chom.2017.07.003.

Cunnington, A. J. 2015. "The importance of pathogen load." *PLoS Pathog* no. 11 (1):e1004563. doi: 10.1371/journal.ppat.1004563.

Cunnington, A. J., M. Walther, and E. M. Riley. 2013. "Piecing together the puzzle of severe malaria." *Sci Transl Med* no. 5 (211):211ps18. doi: 10.1126/scitranslmed.3007432.

Daily, J. P., K. G. Le Roch, O. Sarr, X. Fang, Y. Zhou, O. Ndir, S. Mboup, A. Sultan, E. A. Winzeler, and D. F. Wirth. 2004. "In vivo transcriptional profiling of Plasmodium falciparum." *Malar J* no. 3:30. doi: 10.1186/1475-2875-3-30.

de Santana Filho, F. S., A. R. Arcanjo, Y. M. Chehuan, M. R. Costa, F. E. Martinez-Espinosa, J. L. Vieira, Md Barbosa, W. D. Alecrim, and Md Alecrim. 2007. "Chloroquine-resistant Plasmodium vivax, Brazilian Amazon." *Emerg Infect Dis* no. 13 (7):1125-6. doi: 10.3201/eid1307.061386.

del Portillo, H. A., C. Fernandez-Becerra, S. Bowman, K. Oliver, M. Preuss, C. P. Sanchez, N. K. Schneider, J. M. Villalobos, M. A. Rajandream, D. Harris, L. H. Pereira da Silva, B. Barrell, and M. Lanzer. 2001. "A superfamily of variant genes encoded in the subtelomeric region of Plasmodium vivax." *Nature* no. 410 (6830):839-42. doi: 10.1038/35071118.

del Portillo, H. A., M. Lanzer, S. Rodriguez-Malaga, F. Zavala, and C. Fernandez-Becerra. 2004. "Variant genes and the spleen in Plasmodium vivax malaria." *Int J Parasitol* no. 34 (13-14):1547-54. doi: 10.1016/j.ijpara.2004.10.012.

Desai, S. A. 2014. "Why do malaria parasites increase host erythrocyte permeability?" *Trends Parasitol* no. 30 (3):151-9. doi: 10.1016/j.pt.2014.01.003.

Dharia, N. V., A. T. Bright, S. J. Westenberger, S. W. Barnes, S. Batalov, K. Kuhen, R. Borboa, G. C. Federe, C. M. McClean, J. M. Vinetz, V. Neyra, A. Llanos-Cuentas, J. W. Barnwell, J. R. Walker, and E. A. Winzeler. 2010. "Whole-genome sequencing and microarray analysis of ex vivo Plasmodium vivax reveal selective pressure on putative drug resistance genes." *Proc Natl Acad Sci U S A* no. 107 (46):20045-50. doi: 10.1073/pnas.1003776107.

Duval, L., M. Fourment, E. Nerrienet, D. Rousset, S. A. Sadeuh, S. M. Goodman, N. V. Andriaholinirina, M. Randrianarivelojosia, R. E. Paul, V. Robert, F. J. Ayala, and F. Ariey. 2010. "African apes as reservoirs of Plasmodium falciparum and the origin and diversification of the Laverania subgenus." *Proc Natl Acad Sci U S A* no. 107 (23):10561-6. doi: 10.1073/pnas.1005435107.

Ebrahimzadeh, Z., A. Mukherjee, and D. Richard. 2018. "A map of the subcellular distribution of phosphoinositides in the erythrocytic cycle of the malaria parasite Plasmodium falciparum." *Int J Parasitol* no. 48 (1):13-25. doi: 10.1016/j.ijpara.2017.08.015.

Elizalde-Torrent, A., F. Val, I. C. C. Azevedo, W. M. Monteiro, L. C. L. Ferreira, C. Fernandez-Becerra, H. A. Del Portillo and M. V. G. Lacerda (2018). "Sudden spleen rupture in a Plasmodium vivax-infectedpatient undergoing malaria treatment." Malar J **17**(1): 79.

Erhart, L. M., K. Yingyuen, N. Chuanak, N. Buathong, A. Laboonchai, R. S. Miller, S. R. Meshnick, R. A. Gasser, Jr., and C. Wongsrichanalai. 2004. "Hematologic and clinical indices of malaria in a semi-immune population of western Thailand." *Am J Trop Med Hyg* no. 70 (1):8-14. doi: 10.4269/ajtmh.2004.70.8.

Escalante, A. A., M. U. Ferreira, J. M. Vinetz, S. K. Volkman, L. Cui, D. Gamboa, D. J. Krogstad, A. E. Barry, J. M. Carlton, A. M. van Eijk, K. Pradhan, I. Mueller, B. Greenhouse, M. A. Pacheco, A. F. Vallejo, S.

Herrera, and I. Felger. 2015. "Malaria Molecular Epidemiology: Lessons from the International Centers of Excellence for Malaria Research Network." *Am J Trop Med Hyg* no. 93 (3 Suppl):79-86. doi: 10.4269/ajtmh.15-0005.

Feng, X., J. M. Carlton, D. A. Joy, J. Mu, T. Furuya, B. B. Suh, Y. Wang, J. W. Barnwell, and X. Z. Su. 2003. "Single-nucleotide polymorphisms and genome diversity in Plasmodium vivax." *Proc Natl Acad Sci U S A* no. 100 (14):8502-7. doi: 10.1073/pnas.1232502100.

Fernandez-Becerra, C., O. Pein, T. R. de Oliveira, M. M. Yamamoto, A. C. Cassola, C. Rocha, I. S. Soares, C. A. de Braganca Pereira, and H. A. del Portillo. 2005. "Variant proteins of Plasmodium vivax are not clonally expressed in natural infections." *Mol Microbiol* no. 58 (3):648-58. doi: 10.1111/j.1365-2958.2005.04850.x.

Fernandez-Becerra, C., M. J. Pinazo, A. Gonzalez, P. L. Alonso, H. A. del Portillo, and J. Gascon. 2009. "Increased expression levels of the pvcrt-o and pvmdr1 genes in a patient with severe Plasmodium vivax malaria." *Malar J* no. 8:55. doi: 10.1186/1475-2875-8-55.

Fernandez-Becerra, C., M. M. Yamamoto, R. Z. Vencio, M. Lacerda, A. Rosanas-Urgell, and H. A. del Portillo. 2009. "Plasmodium vivax and the importance of the subtelomeric multigene vir superfamily." *Trends Parasitol* no. 25 (1):44-51. doi: 10.1016/j.pt.2008.09.012.

Ferreira, M. U., and M. C. Castro. 2016. "Challenges for malaria elimination in Brazil." *Malar J* no. 15 (1):284. doi: 10.1186/s12936-016-1335-1.

Field, J., Shute, P. 1956. "Plasmodium vivax. In The microscopuic diagnosis of human malaria. II. A morphological study of the erythrocytic parasites, Institute for Medical Research." *Studies from the Institute for Medical Research, Federated Malay States, Kuala Lumpur: Govt. Press.* no. 2 (24):251.

Field, J.W., Sandosham, A.A., Fong, Y.L. 1963. "A Morphological Study of the Erythrocitic Parasites in Thick Blood Films, the Microscopic Diagnosis Of Human Malaria." *The Government Press, Kuala Lumpur.*

Forrester, S. J., and N. Hall. 2014. "The revolution of whole genome sequencing to study parasites." *Mol Biochem Parasitol* no. 195 (2):77-81. doi: 10.1016/j.molbiopara.2014.07.008.

Galinski, M. R., C. C. Medina, P. Ingravallo, and J. W. Barnwell. 1992. "A reticulocyte-binding protein complex of Plasmodium vivax merozoites." *Cell* no. 69 (7):1213-26.

Gazzinelli, R. T., P. Kalantari, K. A. Fitzgerald, and D. T. Golenbock. 2014. "Innate sensing of malaria parasites." *Nat Rev Immunol* no. 14 (11):744-57. doi: 10.1038/nri3742.

Genton, B., V. D'Acremont, L. Rare, K. Baea, J. C. Reeder, M. P. Alpers, and I. Muller. 2008. "Plasmodium vivax and mixed infections are associated with severe malaria in children: a prospective cohort study from Papua New Guinea." *PLoS Med* no. 5 (6):e127. doi: 10.1371/journal.pmed.0050127.

Gething, P. W., I. R. Elyazar, C. L. Moyes, D. L. Smith, K. E. Battle, C. A. Guerra, A. P. Patil, A. J. Tatem, R. E. Howes, M. F. Myers, D. B. George, P. Horby, H. F. Wertheim, R. N. Price, I. Mueller, J. K. Baird, and S. I. Hay. 2012. "A long neglected world malaria map: Plasmodium vivax endemicity in 2010." *PLoS Negl Trop Dis* no. 6 (9):e1814. doi: 10.1371/journal.pntd.0001814.

Griffiths, M. J., M. J. Shafi, S. J. Popper, C. A. Hemingway, M. M. Kortok, A. Wathen, K. A. Rockett, R. Mott, M. Levin, C. R. Newton, K. Marsh, D. A. Relman, and D. P. Kwiatkowski. 2005. "Genomewide analysis of the host response to malaria in Kenyan children." *J Infect Dis* no. 191 (10):1599-611. doi: 10.1086/429297.

Guerra, C. A., R. E. Howes, A. P. Patil, P. W. Gething, T. P. Van Boeckel, W. H. Temperley, C. W. Kabaria, A. J. Tatem, B. H. Manh, I. R. Elyazar, J. K. Baird, R. W. Snow, and S. I. Hay. 2010. "The international limits and population at risk of Plasmodium vivax transmission in 2009." *PLoS Negl Trop Dis* no. 4 (8):e774. doi: 10.1371/journal.pntd.0000774.

Handayani, S., D. T. Chiu, E. Tjitra, J. S. Kuo, D. Lampah, E. Kenangalem, L. Renia, G. Snounou, R. N. Price, N. M. Anstey, and B. Russell. 2009. "High deformability of Plasmodium vivax-infected red blood cells under microfluidic conditions." *J Infect Dis* no. 199 (3):445-50. doi: 10.1086/596048.

Haque, A., J. Engel, S. A. Teichmann, and T. Lonnberg. 2017. "A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications." *Genome Med* no. 9 (1):75. doi: 10.1186/s13073-017-0467-4.

Howes, R. E., R. C. Reiner, Jr., K. E. Battle, J. Longbottom, B. Mappin, D. Ordanovich, A. J. Tatem, C. Drakeley, P. W. Gething, P. A. Zimmerman, D. L. Smith, and S. I. Hay. 2015. "Plasmodium vivax Transmission in Africa." *PLoS Negl Trop Dis* no. 9 (11):e0004222. doi: 10.1371/journal.pntd.0004222.

Hrdlickova, R., M. Toloue, and B. Tian. 2017. "RNA-Seq methods for transcriptome analysis." *Wiley Interdiscip Rev RNA* no. 8 (1). doi: 10.1002/wrna.1364.

Huang, H., Y. F. Ma, Y. Bao, H. Lee, M. P. Lisanti, H. B. Tanowitz, and L. M. Weiss. 2011. "Molecular cloning and characterization of mitogen-activated protein kinase 2 in Toxoplasma gondii." *Cell Cycle* no. 10 (20):3519-26. doi: 10.4161/cc.10.20.17791.

Hulden, L., and L. Hulden. 2011. "Activation of the hypnozoite: a part of Plasmodium vivax life cycle and survival." *Malar J* no. 10:90. doi: 10.1186/1475-2875-10-90.

Hulden, L., L. Hulden, and K. Heliovaara. 2008. "Natural relapses in vivax malaria induced by Anopheles mosquitoes." *Malar J* no. 7:64. doi: 10.1186/1475-2875-7-64.

Hupalo, D. N., Z. Luo, A. Melnikov, P. L. Sutton, P. Rogov, A. Escalante, A. F. Vallejo, S. Herrera, M. Arevalo-Herrera, Q. Fan, Y. Wang, L. Cui, C. M. Lucas, S. Durand, J. F. Sanchez, G. C. Baldeviano, A. G. Lescano, M. Laman, C. Barnadas, A. Barry, I. Mueller, J. W. Kazura, A. Eapen, D. Kanagaraj, N. Valecha, M. U. Ferreira, W. Roobsoong, W. Nguitragool, J. Sattabonkot, D. Gamboa, M. Kosek, J. M. Vinetz, L. Gonzalez-Ceron, B. W. Birren, D. E. Neafsey, and J. M. Carlton. 2016. "Population genomics studies identify signatures of global dispersal and drug resistance in Plasmodium vivax." *Nat Genet*. doi: 10.1038/ng.3588.

Hussain, M. M., M. Sohail, K. Abhishek, and M. Raziuddin. 2013. "Investigation on Plasmodium falciparum and Plasmodium vivax infection influencing host haematological factors in tribal dominant and malaria endemic population of Jharkhand." *Saudi J Biol Sci* no. 20 (2):195-203. doi: 10.1016/j.sjbs.2013.01.003.

Hutchinson, R. A., and S. W. Lindsay. 2006. "Malaria and deaths in the English marshes." *Lancet* no. 367 (9526):1947-51. doi: 10.1016/S0140-6736(06)68850-8.

Judice, C. C., C. Bourgard, A. C. Kayano, L. Albrecht, and F. T. Costa. 2016. "MicroRNAs in the Host-Apicomplexan Parasites Interactions: A Review of Immunopathological Aspects." *Front Cell Infect Microbiol* no. 6:5. doi: 10.3389/fcimb.2016.00005.

Keeley, A., and D. Soldati. 2004. "The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa." *Trends Cell Biol* no. 14 (10):528-32. doi: 10.1016/j.tcb.2004.08.002.

Kim, H. C., L. A. Pacha, W. J. Lee, J. K. Lee, J. C. Gaydos, W. J. Sames, H. C. Lee, K. Bradley, G. G. Jeung, S. K. Tobler, and T. A. Klein. 2009. "Malaria in the Republic of Korea, 1993-2007. Variables related to re-emergence and persistence of Plasmodium vivax among Korean populations and U.S. forces in Korea." *Mil Med* no. 174 (7):762-9.

Kiszewski, A., A. Mellinger, A. Spielman, P. Malaney, S. E. Sachs, and J. Sachs. 2004. "A global index representing the stability of malaria transmission." *Am J Trop Med Hyg* no. 70 (5):486-98. doi: 10.4269/ajtmh.2004.70.486.

Kitchen, S. K. 1938. "The infection of reticulocytes by Plasmodium vivax." *Am J Trop Med Hyg* no. 18.

Kochar, D. K., A. Das, S. K. Kochar, V. Saxena, P. Sirohi, S. Garg, A. Kochar, M. P. Khatri, and V. Gupta. 2009. "Severe Plasmodium vivax malaria: a report on serial cases from Bikaner in northwestern India." *Am J Trop Med Hyg* no. 80 (2):194-8.

Kochar, D. K., V. Saxena, N. Singh, S. K. Kochar, S. V. Kumar, and A. Das. 2005. "Plasmodium vivax malaria." *Emerg Infect Dis* no. 11 (1):132-4. doi: 10.3201/eid1101.040519.

Kochar, S. K., M. Mahajan, R. P. Gupta, S. Middha, J. Acharya, A. Kochar, A. Das, and D. K. Kochar. 2009. "Acute attack of AIP (acute intermittent porphyria) with severe vivax malaria associated with convulsions: a case report." *J Vector Borne Dis* no. 46 (4):307-9. doi: http://imsear.li.mahidol.ac.th/bitstream/123456789/142704/1/jvbd2009v46n4p307.pdf.

Krishna, S., and L. Squire-Pollard. 1990. "Calcium metabolism in malaria-infected erythrocytes." *Parasitol Today* no. 6 (6):196-8.

Krotoski, W. A. 1985. "Discovery of the hypnozoite and a new theory of malarial relapse." *Trans R Soc Trop Med Hyg* no. 79 (1):1-11.

Krotoski, W. A., W. E. Collins, R. S. Bray, P. C. Garnham, F. B. Cogswell, R. W. Gwadz, R. Killick-Kendrick, R. Wolf, R. Sinden, L. C. Koontz, and P. S. Stanfill. 1982. "Demonstration of hypnozoites in sporozoite-transmitted Plasmodium vivax infection." *Am J Trop Med Hyg* no. 31 (6):1291-3.

Kugelstadt, D., D. Winter, K. Pluckhahn, W. D. Lehmann, and B. Kappes. 2007. "Raf kinase inhibitor protein affects activity of Plasmodium falciparum calcium-dependent protein kinase 1." *Mol Biochem Parasitol* no. 151 (1):111-7. doi: 10.1016/j.molbiopara.2006.10.012.

Lacerda, M. V., S. C. Fragoso, M. G. Alecrim, M. A. Alexandre, B. M. Magalhaes, A. M. Siqueira, L. C. Ferreira, J. R. Araujo, M. P. Mourao, M. Ferrer, P. Castillo, L. Martin-Jaular, C. Fernandez-Becerra, H. del Portillo, J. Ordi, P. L. Alonso and Q. Bassat (2012). "Postmortem characterization of patients

with clinical diagnosis of Plasmodium vivax malaria: to what extent does this parasite kill?" <u>Clin Infect Dis</u> **55**(8): e67-74.

Lacerda, M. V. G., A. Llanos-Cuentas, S. Krudsood, C. Lon, D. L. Saunders, R. Mohammed, D. Yilma, D. Batista Pereira, F. E. J. Espino, R. Z. Mia, R. Chuquiyauri, F. Val, M. Casapia, W. M. Monteiro, M. A. M. Brito, M. R. F. Costa, N. Buathong, H. Noedl, E. Diro, S. Getie, K. M. Wubie, A. Abdissa, A. Zeynudin, C. Abebe, M. S. Tada, F. Brand, H. P. Beck, B. Angus, S. Duparc, J. P. Kleim, L. M. Kellam, V. M. Rousell, S. W. Jones, E. Hardaker, K. Mohamed, D. D. Clover, K. Fletcher, J. J. Breton, C. O. Ugwuegbulam, J. A. Green and G. Koh (2019). "Single-Dose Tafenoquine to Prevent Relapse of Plasmodium vivax Malaria." N Engl J Med **380**(3): 215-228.

Lee, H. J., A. Georgiadou, T. D. Otto, M. Levin, L. J. Coin, D. J. Conway, and A. J. Cunnington. 2018. "Transcriptomic Studies of Malaria: a Paradigm for Investigation of Systemic Host-Pathogen Interactions." *Microbiol Mol Biol Rev* no. 82 (2). doi: 10.1128/MMBR.00071-17.

Lee, W. C., B. Malleret, Y. L. Lau, M. Mauduit, M. Y. Fong, J. S. Cho, R. Suwanarusk, R. Zhang, L. Albrecht, F. T. Costa, P. Preiser, R. McGready, L. Renia, F. Nosten, and B. Russell. 2014. "Glycophorin C (CD236R) mediates vivax malaria parasite rosetting to normocytes." *Blood* no. 123 (18):e100-9. doi: 10.1182/blood-2013-12-541698.

Leida, M. N., J. R. Mahoney, and J. W. Eaton. 1981. "Intraerythrocytic plasmodial calcium metabolism." *Biochem Biophys Res Commun* no. 103 (2):402-6.

Liu, W., Y. Li, K. S. Shaw, G. H. Learn, L. J. Plenderleith, J. A. Malenke, S. A. Sundararaman, M. A. Ramirez, P. A. Crystal, A. G. Smith, F. Bibollet-Ruche, A. Ayouba, S. Locatelli, A. Esteban, F. Mouacha, E. Guichet, C. Butel, S. Ahuka-Mundeke, B. I. Inogwabini, J. B. Ndjango, S. Speede, C. M. Sanz, D. B. Morgan, M. K. Gonder, P. J. Kranzusch, P. D. Walsh, A. V. Georgiev, M. N. Muller, A. K. Piel, F. A. Stewart, M. L. Wilson, A. E. Pusey, L. Cui, Z. Wang, A. Farnert, C. J. Sutherland, D. Nolder, J. A. Hart, T. B. Hart, P. Bertolani, A. Gillis, M. LeBreton, B. Tafon, J. Kiyang, C. F. Djoko, B. S. Schneider, N. D. Wolfe, E. Mpoudi-Ngole, E. Delaporte, R. Carter, R. L. Culleton, G. M. Shaw, J. C. Rayner, M. Peeters, B. H. Hahn, and P. M. Sharp. 2014. "African origin of the malaria parasite Plasmodium vivax." *Nat Commun* no. 5:3346. doi: 10.1038/ncomms4346.

Lopes, S. C., L. Albrecht, B. O. Carvalho, A. M. Siqueira, R. Thomson-Luque, P. A. Nogueira, C. Fernandez-Becerra, H. A. Del Portillo, B. M. Russell, L. Renia, M. V. Lacerda, and F. T. Costa. 2014. "Paucity of Plasmodium vivax mature schizonts in peripheral blood is associated with their increased cytoadhesive potential." *J Infect Dis* no. 209 (9):1403-7. doi: jiu018 [pii]10.1093/infdis/jiu018.

Lopez, F. J., M. Bernabeu, C. Fernandez-Becerra, and H. A. del Portillo. 2013. "A new computational approach redefines the subtelomeric vir superfamily of Plasmodium vivax." *BMC Genomics* no. 14:8. doi: 10.1186/1471-2164-14-8.

Lorenz, V., G. Karanis, and P. Karanis. 2014. "Malaria vaccine development and how external forces shape it: an overview." *Int J Environ Res Public Health* no. 11 (7):6791-807. doi: 10.3390/ijerph110706791.

Luo, Z., S. A. Sullivan, and J. M. Carlton. 2015. "The biology of Plasmodium vivax explored through genomics." *Ann N Y Acad Sci* no. 1342:53-61. doi: 10.1111/nyas.12708.

Luxemburger, C., W. A. Perea, G. Delmas, C. Pruja, B. Pecoul, and A. Moren. 1994. "Permethrin-impregnated bed nets for the prevention of malaria in schoolchildren on the Thai-Burmese border." *Trans R Soc Trop Med Hyg* no. 88 (2):155-9.

Machado Siqueira, A., B. M. Lopes Magalhaes, G. Cardoso Melo, M. Ferrer, P. Castillo, L. Martin-Jaular, C. Fernandez-Becerra, J. Ordi, A. Martinez, M. V. Lacerda and H. A. del Portillo (2012). "Spleen rupture in a case of untreated Plasmodium vivax infection." <u>PLoS Negl Trop Dis</u> **6**(12): e1934.

Malone, J. H., and B. Oliver. 2011. "Microarrays, deep sequencing and the true measure of the transcriptome." *BMC Biol* no. 9:34. doi: 10.1186/1741-7007-9-34.

Marsh, K., D. Forster, C. Waruiru, I. Mwangi, M. Winstanley, V. Marsh, C. Newton, P. Winstanley, P. Warn, N. Peshu, and et al. 1995. "Indicators of life-threatening malaria in African children." *N Engl J Med* no. 332 (21):1399-404. doi: 10.1056/NEJM199505253322102.

Marti, M., R. T. Good, M. Rug, E. Knuepfer, and A. F. Cowman. 2004. "Targeting malaria virulence and remodeling proteins to the host erythrocyte." *Science* no. 306 (5703):1930-3. doi: 10.1126/science.1102452.

McGready, R., B. B. Davison, K. Stepniewska, T. Cho, H. Shee, A. Brockman, R. Udomsangpetch, S. Looareesuwan, N. J. White, S. R. Meshnick, and F. Nosten. 2004. "The effects of Plasmodium falciparum and P. vivax infections on placental histopathology in an area of low malaria transmission." *Am J Trop Med Hyg* no. 70 (4):398-407.

Menard, D., E. R. Chan, C. Benedet, A. Ratsimbasoa, S. Kim, P. Chim, C. Do, B. Witkowski, R. Durand, M. Thellier, C. Severini, E. Legrand, L. Musset, B. Y. Nour, O. Mercereau-Puijalon, D. Serre, and P. A. Zimmerman. 2013. "Whole genome sequencing of field isolates reveals a common duplication of the Duffy binding protein gene in Malagasy Plasmodium vivax strains." *PLoS Negl Trop Dis* no. 7 (11):e2489. doi: 10.1371/journal.pntd.0002489.

Mendis, K., B. J. Sina, P. Marchesini, and R. Carter. 2001. "The neglected burden of Plasmodium vivax malaria." *Am J Trop Med Hyg* no. 64 (1-2 Suppl):97-106.

Molina-Cruz, A., G. E. Canepa, and C. Barillas-Mury. 2017. "Plasmodium P47: a key gene for malaria transmission by mosquito vectors." *Curr Opin Microbiol* no. 40:168-174. doi: 10.1016/j.mib.2017.11.029.

Mueller, I., M. R. Galinski, J. K. Baird, J. M. Carlton, D. K. Kochar, P. L. Alonso, and H. A. del Portillo. 2009. "Key gaps in the knowledge of Plasmodium vivax, a neglected human malaria parasite." *Lancet Infect Dis* no. 9 (9):555-66. doi: 10.1016/S1473-3099(09)70177-X.

Naing, C., M. A. Whittaker, V. Nyunt Wai, and J. W. Mak. 2014. "Is Plasmodium vivax Malaria a Severe Malaria?: A Systematic Review and Meta-Analysis." *PLoS Negl Trop Dis* no. 8 (8):e3071. doi: 10.1371/journal.pntd.0003071.

Neafsey, D. E., K. Galinsky, R. H. Jiang, L. Young, S. M. Sykes, S. Saif, S. Gujja, J. M. Goldberg, S. Young, Q. Zeng, S. B. Chapman, A. P. Dash, A. R. Anvikar, P. L. Sutton, B. W. Birren, A. A. Escalante, J. W. Barnwell, and J. M. Carlton. 2012. "The malaria parasite Plasmodium vivax exhibits greater genetic diversity than Plasmodium falciparum." *Nat Genet* no. 44 (9):1046-50. doi: 10.1038/ng.2373.

Noulin, F., C. Borlon, J. Van Den Abbeele, U. D'Alessandro, and A. Erhart. 2013. "1912-2012: a century of research on Plasmodium vivax in vitro culture." *Trends Parasitol* no. 29 (6):286-94. doi: 10.1016/j.pt.2013.03.012.

Orjuela-Sanchez, P., N. S. da Silva, M. da Silva-Nunes, and M. U. Ferreira. 2009. "Recurrent parasitemias and population dynamics of Plasmodium vivax polymorphisms in rural Amazonia." *Am J Trop Med Hyg* no. 81 (6):961-8. doi: 10.4269/ajtmh.2009.09-0337.

Orjuela-Sanchez, P., J. M. Sa, M. C. Brandi, P. T. Rodrigues, M. S. Bastos, C. Amaratunga, S. Duong, R. M. Fairhurst, and M. U. Ferreira. 2013. "Higher microsatellite diversity in Plasmodium vivax than in sympatric Plasmodium falciparum populations in Pursat, Western Cambodia." *Exp Parasitol* no. 134 (3):318-26. doi: 10.1016/j.exppara.2013.03.029.

Panichakul, T., J. Sattabongkot, K. Chotivanich, J. Sirichaisinthop, L. Cui, and R. Udomsangpetch. 2007. "Production of erythropoietic cells in vitro for continuous culture of Plasmodium vivax." *Int J Parasitol* no. 37 (14):1551-7. doi: 10.1016/j.ijpara.2007.05.009.

Parkyn Schneider, M., B. Liu, P. Glock, A. Suttie, E. McHugh, D. Andrew, S. Batinovic, N. Williamson, E. Hanssen, P. McMillan, M. Hliscs, L. Tilley, and M. W. A. Dixon. 2017. "Disrupting assembly of the inner membrane complex blocks Plasmodium falciparum sexual stage development." *PLoS Pathog* no. 13 (10):e1006659. doi: 10.1371/journal.ppat.1006659.

Perrin, A. J., S. J. Bartholdson, and G. J. Wright. 2015. "P-selectin is a host receptor for Plasmodium MSP7 ligands." *Malar J* no. 14:238. doi: 10.1186/s12936-015-0750-z.

Poespoprodjo, J. R., W. Fobia, E. Kenangalem, D. A. Lampah, N. Warikar, A. Seal, R. McGready, P. Sugiarto, E. Tjitra, N. M. Anstey, and R. N. Price. 2008. "Adverse pregnancy outcomes in an area where multidrug-resistant plasmodium vivax and Plasmodium falciparum infections are endemic." *Clin Infect Dis* no. 46 (9):1374-81. doi: 10.1086/586743.

Pradel, G., K. Hayton, L. Aravind, L. M. Iyer, M. S. Abrahamsen, A. Bonawitz, C. Mejia, and T. J. Templeton. 2004. "A multidomain adhesion protein family expressed in Plasmodium falciparum is essential for transmission to the mosquito." *J Exp Med* no. 199 (11):1533-44. doi: 10.1084/jem.20031274.

Price, R. N., N. M. Douglas, and N. M. Anstey. 2009. "New developments in Plasmodium vivax malaria: severe disease and the rise of chloroquine resistance." *Curr Opin Infect Dis* no. 22 (5):430-5. doi: 10.1097/QCO.0b013e32832f14c1.

Price, R. N., E. Tjitra, C. A. Guerra, S. Yeung, N. J. White, and N. M. Anstey. 2007. "Vivax malaria: neglected and not benign." *Am J Trop Med Hyg* no. 77 (6 Suppl):79-87. doi: 10.4269/ajtmh.2007.77.79.

Price, R. N., L. von Seidlein, N. Valecha, F. Nosten, J. K. Baird, and N. J. White. 2014. "Global extent of chloroquine-resistant Plasmodium vivax: a systematic review and meta-analysis." *Lancet Infect Dis* no. 14 (10):982-91. doi: 10.1016/S1473-3099(14)70855-2.

Rahimi, B. A., A. Thakkinstian, N. J. White, C. Sirivichayakul, A. M. Dondorp, and W. Chokejindachai. 2014. "Severe vivax malaria: a systematic review and meta-analysis of clinical studies since 1900." *Malar J* no. 13:481. doi: 10.1186/1475-2875-13-481.

Ramasamy, R. 2014. "Zoonotic malaria - global overview and research and policy needs." *Front Public Health* no. 2:123. doi: 10.3389/fpubh.2014.00123.

Rogerson, S. J., V. Mwapasa, and S. R. Meshnick. 2007. "Malaria in pregnancy: linking immunity and pathogenesis to prevention." *Am J Trop Med Hyg* no. 77 (6 Suppl):14-22.

Rosenberg, R. 2007. "Plasmodium vivax in Africa: hidden in plain sight?" *Trends Parasitol* no. 23 (5):193-6. doi: 10.1016/j.pt.2007.02.009.

Rosenbloom, K. R., J. Armstrong, G. P. Barber, J. Casper, H. Clawson, M. Diekhans, T. R. Dreszer, P. A. Fujita, L. Guruvadoo, M. Haeussler, R. A. Harte, S. Heitner, G. Hickey, A. S. Hinrichs, R. Hubley, D. Karolchik, K. Learned, B. T. Lee, C. H. Li, K. H. Miga, N. Nguyen, B. Paten, B. J. Raney, A. F. Smit, M. L. Speir, A. S. Zweig, D. Haussler, R. M. Kuhn, and W. J. Kent. 2015. "The UCSC Genome Browser database: 2015 update." *Nucleic Acids Res* no. 43 (Database issue):D670-81. doi: 10.1093/nar/gku1177.

Rowe, J. A., I. G. Handel, M. A. Thera, A. M. Deans, K. E. Lyke, A. Kone, D. A. Diallo, A. Raza, O. Kai, K. Marsh, C. V. Plowe, O. K. Doumbo, and J. M. Moulds. 2007. "Blood group O protects against severe Plasmodium falciparum malaria through the mechanism of reduced rosetting." *Proc Natl Acad Sci U S A* no. 104 (44):17471-6. doi: 10.1073/pnas.0705390104.

Russell, B., F. Chalfein, B. Prasetyorini, E. Kenangalem, K. Piera, R. Suwanarusk, A. Brockman, P. Prayoga, P. Sugiarto, Q. Cheng, E. Tjitra, N. M. Anstey, and R. N. Price. 2008. "Determinants of in vitro drug susceptibility testing of Plasmodium vivax." *Antimicrob Agents Chemother* no. 52 (3):1040-5. doi: 10.1128/AAC.01334-07.

Russell, B., R. Suwanarusk, C. Borlon, F. T. Costa, C. S. Chu, M. J. Rijken, K. Sriprawat, L. Warter, E. G. Koh, B. Malleret, Y. Colin, O. Bertrand, J. H. Adams, U. D'Alessandro, G. Snounou, F. Nosten, and L. Renia. 2011. "A reliable ex vivo invasion assay of human reticulocytes by Plasmodium vivax." *Blood* no. 118 (13):e74-81. doi: 10.1182/blood-2011-04-348748.

Saharan, S., U. Kohli, R. Lodha, A. Sharma, and A. Bagga. 2009. "Thrombotic microangiopathy associated with Plasmodium vivax malaria." *Pediatr Nephrol* no. 24 (3):623-4. doi: 10.1007/s00467-008-0945-4.

Severini, C., M. Menegon, M. Di Luca, I. Abdullaev, G. Majori, S. A. Razakov, and L. Gradoni. 2004. "Risk of Plasmodium vivax malaria reintroduction in Uzbekistan: genetic characterization of parasites and status of potential malaria vectors in the Surkhandarya region." *Trans R Soc Trop Med Hyg* no. 98 (10):585-92. doi: 10.1016/j.trstmh.2004.01.003.

Shanks, G. D., and N. J. White. 2013. "The activation of vivax malaria hypnozoites by infectious diseases." *Lancet Infect Dis* no. 13 (10):900-6. doi: 10.1016/S1473-3099(13)70095-1.

Sharma, J., K. Bharadawa, K. Shah, and S. Dave. 1993. "Plasmodium vivax malaria presenting as hemolytic uremic syndrome." *Indian Pediatr* no. 30 (3):369-71. doi: http://www.indianpediatrics.net/mar1993/369.pdf.

Siqueira, A. M., M. A. Alexandre, M. P. Mourao, V. S. Santos, S. K. Nagahashi-Marie, M. G. Alecrim, and M. V. Lacerda. 2010. "Severe rhabdomyolysis caused by Plasmodium vivax malaria in the Brazilian Amazon." *Am J Trop Med Hyg* no. 83 (2):271-3. doi: 10.4269/ajtmh.2010.10-0027.

Snounou, G., and N. J. White. 2004. "The co-existence of Plasmodium: sidelights from falciparum and vivax malaria in Thailand." *Trends Parasitol* no. 20 (7):333-9. doi: 10.1016/j.pt.2004.05.004.

Speir, M. L., A. S. Zweig, K. R. Rosenbloom, B. J. Raney, B. Paten, P. Nejad, B. T. Lee, K. Learned, D. Karolchik, A. S. Hinrichs, S. Heitner, R. A. Harte, M. Haeussler, L. Guruvadoo, P. A. Fujita, C. Eisenhart, M. Diekhans, H. Clawson, J. Casper, G. P. Barber, D. Haussler, R. M. Kuhn, and W. J. Kent. 2016. "The UCSC Genome Browser database: 2016 update." *Nucleic Acids Res* no. 44 (D1):D717-25. doi: 10.1093/nar/gkv1275.

Supek, F., M. Bosnjak, N. Skunca, and T. Smuc. 2011. "REVIGO summarizes and visualizes long lists of gene ontology terms." *PLoS One* no. 6 (7):e21800. doi: 10.1371/journal.pone.0021800.

Suratt, B. T., and P. E. Parsons. 2006. "Mechanisms of acute lung injury/acute respiratory distress syndrome." *Clin Chest Med* no. 27 (4):579-89; abstract viii. doi: 10.1016/j.ccm.2006.06.005.

Suwanarusk, R., M. Chavchich, B. Russell, A. Jaidee, F. Chalfein, M. Barends, B. Prasetyorini, E. Kenangalem, K. A. Piera, U. Lek-Uthai, N. M. Anstey, E. Tjitra, F. Nosten, Q. Cheng, and R. N. Price. 2008. "Amplification of pvmdr1 associated with multidrug-resistant Plasmodium vivax." *J Infect Dis* no. 198 (10):1558-64. doi: 10.1086/592451.

Suwanarusk, R., B. M. Cooke, A. M. Dondorp, K. Silamut, J. Sattabongkot, N. J. White, and R. Udomsangpetch. 2004. "The deformability of red blood cells parasitized by Plasmodium falciparum and P. vivax." *J Infect Dis* no. 189 (2):190-4. doi: 10.1086/380468.

Suwanarusk, R., B. Russell, M. Chavchich, F. Chalfein, E. Kenangalem, V. Kosaisavee, B. Prasetyorini, K. A. Piera, M. Barends, A. Brockman, U. Lek-Uthai, N. M. Anstey, E. Tjitra, F. Nosten, Q. Cheng, and R. N. Price. 2007. "Chloroquine resistant Plasmodium vivax: in vitro characterisation and association with molecular polymorphisms." *PLoS One* no. 2 (10):e1089. doi: 10.1371/journal.pone.0001089.

Ta, T. H., S. Hisam, M. Lanza, A. I. Jiram, N. Ismail, and J. M. Rubio. 2014. "First case of a naturally acquired human infection with Plasmodium cynomolgi." *Malar J* no. 13:68. doi: 10.1186/1475-2875-13-68.

Tan, L. K., S. Yacoub, S. Scott, S. Bhagani, and M. Jacobs. 2008. "Acute lung injury and other serious complications of Plasmodium vivax malaria." *Lancet Infect Dis* no. 8 (7):449-54. doi: 10.1016/S1473-3099(08)70153-1.

Tewari, R., U. Straschil, A. Bateman, U. Bohme, I. Cherevach, P. Gong, A. Pain, and O. Billker. 2010. "The systematic functional analysis of Plasmodium protein kinases identifies essential regulators of mosquito transmission." *Cell Host Microbe* no. 8 (4):377-87. doi: 10.1016/j.chom.2010.09.006.

Tjitra, E., N. M. Anstey, P. Sugiarto, N. Warikar, E. Kenangalem, M. Karyana, D. A. Lampah, and R. N. Price. 2008. "Multidrug-resistant Plasmodium vivax associated with severe and fatal malaria: a prospective study in Papua, Indonesia." *PLoS Med* no. 5 (6):e128. doi: 10.1371/journal.pmed.0050128.

Tremp, A. Z., V. Sharma, V. Carter, E. Lasonder, and J. T. Dessens. 2017. "LCCL protein complex formation in Plasmodium is critically dependent on LAP1." *Mol Biochem Parasitol* no. 214:87-90. doi: 10.1016/j.molbiopara.2017.04.005.

Udomsangpetch R, S. R., Williams J, Sattabongkot J. 2001. "Modified techniques to establish a continuous culture of Plasmodium vivax [abstract]. ." *American Society of Tropical Medicine and Hygiene 51st Annual Meeting program; 2002 Nov 10-14; Denver, CO. .*

Udomsanpetch, R., K. Thanikkul, S. Pukrittayakamee, and N. J. White. 1995. "Rosette formation by Plasmodium vivax." *Trans R Soc Trop Med Hyg* no. 89 (6):635-7.

van Dijk, M. R., C. J. Janse, J. Thompson, A. P. Waters, J. A. Braks, H. J. Dodemont, H. G. Stunnenberg, G. J. van Gemert, R. W. Sauerwein, and W. Eling. 2001. "A central role for P48/45 in malaria parasite male gamete fertility." *Cell* no. 104 (1):153-64.

Westenberger, S. J., C. M. McClean, R. Chattopadhyay, N. V. Dharia, J. M. Carlton, J. W. Barnwell, W. E. Collins, S. L. Hoffman, Y. Zhou, J. M. Vinetz, and E. A. Winzeler. 2010. "A systems-based analysis of Plasmodium vivax lifecycle transcription from human to mosquito." *PLoS Negl Trop Dis* no. 4 (4):e653. doi: 10.1371/journal.pntd.0000653.

Westermann, A. J., K. U. Forstner, F. Amman, L. Barquist, Y. Chao, L. N. Schulte, L. Muller, R. Reinhardt, P. F. Stadler, and J. Vogel. 2016. "Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions." *Nature* no. 529 (7587):496-501. doi: 10.1038/nature16547.

Westermann, A. J., S. A. Gorski, and J. Vogel. 2012. "Dual RNA-seq of pathogen and host." *Nat Rev Microbiol* no. 10 (9):618-30. doi: 10.1038/nrmicro2852.

White, M. T., S. Karl, K. E. Battle, S. I. Hay, I. Mueller, and A. C. Ghani. 2014. "Modelling the contribution of the hypnozoite reservoir to Plasmodium vivax transmission." *Elife* no. 3. doi: 10.7554/eLife.04692.

White, M. T., G. Shirreff, S. Karl, A. C. Ghani, and I. Mueller. 2016. "Variation in relapse frequency and the transmission potential of Plasmodium vivax malaria." *Proc Biol Sci* no. 283 (1827):9. doi: 10.1098/rspb.2016.0048.

White, N. J. 2011. "Determinants o   f relapse periodicity in Plasmodium vivax malaria." *Malar J* no. 10:297. doi: 10.1186/1475-2875-10-297.

WHO. 2010. "Guidelines for the Treatment of Malaria. 2nd edition." *World Health Ornanization* no. 2:1-194.

WHO. 2018. World Malaria Report 2018. In *World Malar. Rep.* Geneva.

Wickramasinghe, S. N., S. Looareesuwan, B. Nagachinta, and N. J. White. 1989. "Dyserythropoiesis and ineffective erythropoiesis in Plasmodium vivax malaria." *Br J Haematol* no. 72 (1):91-9.

Williams, T. N., K. Maitland, L. Phelps, S. Bennett, T. E. Peto, J. Viji, R. Timothy, J. B. Clegg, D. J. Weatherall, and D. K. Bowden. 1997. "Plasmodium vivax: a cause of malnutrition in young children." *QJM* no. 90 (12):751-7.

Yamagishi, J., A. Natori, M. E. Tolba, A. E. Mongan, C. Sugimoto, T. Katayama, S. Kawashima, W. Makalowski, R. Maeda, Y. Eshita, J. Tuda, and Y. Suzuki. 2014. "Interactive transcriptome analysis

of malaria patients and infecting Plasmodium falciparum." *Genome Res* no. 24 (9):1433-44. doi: 10.1101/gr.158980.113.

Zhang, R., W. C. Lee, Y. L. Lau, L. Albrecht, S. C. Lopes, F. T. Costa, R. Suwanarusk, F. Nosten, B. M. Cooke, L. Renia, and B. Russell. 2016. "Rheopathologic Consequence of Plasmodium vivax Rosette Formation." *PLoS Negl Trop Dis* no. 10 (8):e0004912. doi: 10.1371/journal.pntd.0004912.

Zhu, L., S. Mok, M. Imwong, A. Jaidee, B. Russell, F. Nosten, N. P. Day, N. J. White, P. R. Preiser, and Z. Bozdech. 2016. "New insights into the Plasmodium vivax transcriptome using RNA-Seq." *Sci Rep* no. 6:20498. doi: 10.1038/srep20498.

ANNEX 1

"Genetic sequence characterization and naturally acquired immune response to *Plasmodium vivax* Rhoptry Neck Protein 2 (PvRON2)"

Overview

The genetic diversity of malaria antigens often results in allele variant-specific immunity, imposing a great challenge to vaccine development. Rhoptry Neck Protein 2 (PvRON2) is a blood-stage antigen that plays a key role during the erythrocyte invasion of *Plasmodium vivax*. This study investigates the genetic diversity of PvRON2 and the naturally acquired immune response to *P. vivax* isolates.

Here, the genetic diversity of PvRON2$_{1828-2080}$ and the naturally acquired humoral immune response against PvRON2$_{1828-2080}$ in infected and non-infected individuals from a vivax malaria endemic area in Brazil was reported. The diversity analysis of PvRON2$_{1828-2080}$ revealed that the protein is conserved in isolates in Brazil and worldwide. A total of 18 (19%) patients had IgG antibodies to PvRON2$_{1828-2080}$. Additionally, the analysis of the antibody response in individuals who were not acutely infected with malaria, but had been infected with malaria in the past indicated that 32 patients (33%) exhibited an IgG immune response against PvRON2.

PvRON2 was conserved among the studied isolates. The presence of naturally acquired antibodies to this protein in the absence of the disease suggests that PvRON2 induces a long-term antibody response. These results indicate that PvRON2 is a potential malaria vaccine candidate.

"Efforts into

Plasmodium vivax Drug Discovery"

Introduction

Currently, while it has been recognized that *P. vivax* infection can cause severe clinical complications [1], the availability of novel therapies (*e.g.* antimalarials) remains limited. Most importantly, resistance has been identified against all available antiplasmodial drugs, from the long-used chloroquine to the more recently instituted artemisinin and antifolate agents, such as pyrimethamine. Presently, effective control policies for decision-making are hampered and the recent increase of malaria in some endemic areas of the world is of great concern, seriously threating the prospects for malaria elimination in the near future [2].

Methods to combat malaria using small molecules have been impeded by complex and quick antigen variant switch and development of resistance. According to the WHO guidelines [2-4], the first line of *P. vivax* chemotherapy is chloroquine (CQ) plus primaquine (PQ), the only approved drug targeting the latent parasite form (hypnozoite) [5]. Recent efforts have resulted in the global approval of Tafenoquine (TAF), a drug analog of PQ developed by GlaxoSmithKline. Today, TAF is a better alternative antimalarial than PQ, especially because while it targets *P. vivax* hypnozoites and asexual stages like PQ, it has a considerably shorter course of treatment and lower dosage. It has been hypothesized that TAF may exert its effect by inhibiting hematin polymerization and inducing mitochondrial dysfunction leading to the apoptotic-like death of the organism. However, the mechanism of action of both TAF and PQ drugs are not well characterized at this time. Further, the restrictions concerning immunosuppressed and glucose-6-phosphate dehydrogenase (G6PD) deficient patients apply for both PQ and TAF [6], leaving no current therapies available for this patient population.

In high transmission areas presenting cases of drug resistance, an artemisinin based combination therapy (ACT) is recommended [3, 7]. The constant increase and spread of anti-malarial drug resistance of *P. vivax* remains of great concern [8-15], pointing to the need for augmented efforts in new drug discovery. Given the fact that in many parts of the world, after *P. falciparum* transmission rates have decreased abruptly, *P. vivax* becomes the predominant species, causing persistent infections with high morbidity. Therefore, there is a great need for novel experimental approaches targeting this parasite.

For thousands of years, natural products from autochthonous flora have been used to treat and alleviate malaria symptoms, not only in traditional medicine settings, but later as lead compound sources for new drug investigation. Some examples of these are artemisinins, quinine atovaquone, clindamycin, erythromycin, azithromycin, tetracyclines, doxycycline, and others [16]. The majority of natural products were based on several existing small molecules or pharmacophores. Nevertheless, several natural compounds have neither their active components nor their origin identified, and their efficacy has not yet been systematically studied. As a result, this source of potential leads discovery is almost completely neglected.

The complex life cycle of *P. vivax* parasite, the impossibility of cultivating them *in vitro,* and their vast genetic diversity pose great obstacles to efficiently discovering antimalarial drug candidates. To aid the malaria research community in this context, genome-wide drug sensitivity screens of yeast mutants (chemogenomic profiling [17-19]) have been successfully used to identify the key molecular targets of several frontline antimalarial drugs, such as quinine [20] and artemisinin [21]. *Saccharomyces cerevisiae* is a model organism used for its genetic malleability, high degree of genetic and cellular process conservation with human cells, and forgiving growth conditions requirements. *S. cerevisiae* can be engineered to express heterologous proteins, providing a well-characterized platform for automated screens, due to its fast grow rate at low costs. Chemogenomics screening also provides evidence of possible off-target effects [22], allowing the prediction of potential side effects (*e.g.*, cytotoxicity).

Heterologous expression of *in silico* triaged molecular targets [23] in yeast offers expedient experimental access for evaluation of those parasite proteins, otherwise practically impossible in an *in vitro* functional assay, performed for all parasite-stage forms found in the human host. Yeast strains expressing heterologous parasite (*e.g.* plasmodial) drug transporters have been constructed [24], providing the opportunity to specifically test for species-specific uptake or efflux propensity for particular drug candidates. Thus, yeast has great potential as an *in vitro* system and experimental platform for drug candidate screening in *P. vivax* [25, 26], combining high versatility by virtue of a very large molecular genetics toolbox with ease of cultivation and amenability to robotized high-throughput, high-density, technologies [27].

Bilsland *et al.* 2013 [26] developed a novel drug-screening method based on bioengineering of the model organism yeast *S. cerevisiae*, which enable us the identification of compounds that selectively inhibits parasite molecular target and not their human counterpart. This system can be robotized and used in automated high-throughput screens to identify anti-parasitic compounds, while eliminating the ones having potential side effects (*e.g.* cytotoxicity) [27]. Bilsland and colleagues successfully engineered *S. cerevisiae* strains deleted for one essential gene (as examples, *dhfr, nmt* and *pgk* genes) and functionally complemented by the heterologous expression of the orthologous gene from parasite and human origin [25]. In this system, the major drug-export pump in *S. cerevisiae*, codified by the *pdr5* gene, was also deleted to sensitize the yeast to a larger range of chemicals. According to this method, the addition of chemical compounds to the culture medium can allow the discovery of potential drugs by monitoring over-time the growth of those strains. If a specific compound inhibits the parasite target protein function, the genetically engineered yeast functionally complemented by the parasite orthologous protein will not be able to grow. Furthermore, the growth of the yeast functionally complemented by the human orthologue would reveal if that compound has little or no effect on the human protein, thus ruling out cytotoxicity. The growth of the initial yeast strain expressing its own essential gene, could help us to discriminate for such compounds if they are cytotoxic for the yeast system. The researchers proved the viability of this high-throughput assays by identifying compounds from The Malaria Box assortment inhibiting each one of the selected targets tested but failing to inhibit the corresponding human counterpart. The "hit compounds" found against yeast strains encoding *P. falciparum*, *P. vivax*, *Schistosoma mansoni*, *Tripanosoma brucei*, *T. cruzi* or *Leishmania major* proteins were then validated by demonstrating their effectiveness *in vivo* [25]. More recently, this methodological approach was also applied to identify novel compounds active against *Brugia malayi* [28]. We grounded our work on findings from Bilsland *et al.* and continued to extend her efforts in order to create a platform for drug discovery against the *P. vivax* human malaria parasite.

Aim

Worldwide, there is the urgent need for novel drugs against malaria to defeat this devastating disease with its severe social and economic burdens and impacts on individual welfare. As we push to discover new *P. vivax* molecular targets to develop novel drugs, this prospective collaborative project was initiated with the purpose of pathways and specific intracellular drug targets identification, using yeast-based drug screen methods [25-27] (Fig. 1). This yeast-based robust anti-parasitic drug-screening method allow high-throughput drug screens for the selection of compounds specifically targeting parasites and not their host counterpart ("hits") from *in silico* prediction pool of known pre-selected *P. vivax* molecular targets. Prospective compounds discovered are then evaluated *in vitro* (*P. falciparum*) and *in vivo* (*P. berghei*) for their bioactivity. Since the most likely molecular targets are known, we expect to reveal the compounds biological mode of action. Furthermore, interesting compounds might be further re-design and synthesize for improved bioactive selective properties (Fig. 1.).
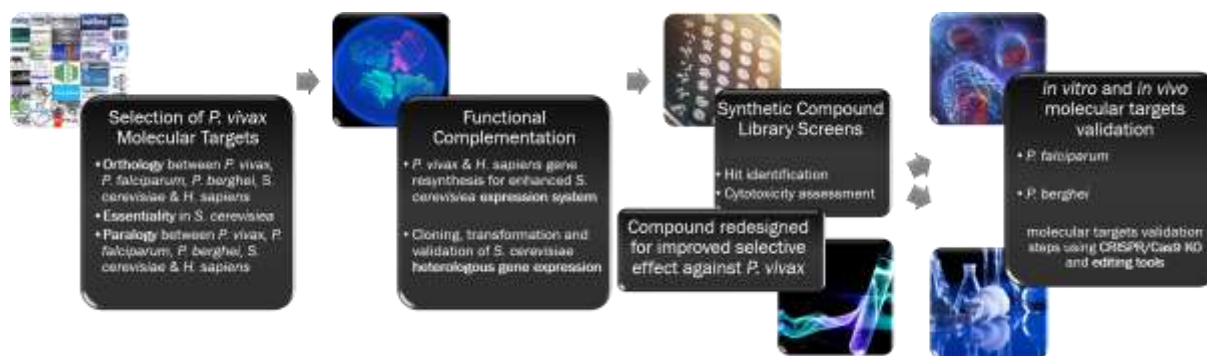
Figure 1. Scheme depicting the methodological overview of the collaborative project on a chemical-genomic profiling approach for drug target discovery in *P. vivax*.

Methodology overview and Accomplished Aims

To initiate this challenge, I have performed the following experimental steps:

(i) In silico selection of 13 *P. vivax* molecular targets, fulfilling three main criteria:

- gene orthology between *P. vivax*, *S. cerevisiae*, (acting as a surrogate for expressing antiparasitic targets), *H. sapiens* (for drug selectivity determination), *P.*

falciparum and *P. berghei* (for downstream *in vitro* and *in vivo* validation);

- gene essentiality in *S. cerevisiae*. Gives efficient selection for complementation of the yeast loss-of-function mutation by the heterologous *P. vivax* and *H. sapiens* genes;

- no paralogs within all 3 *Plasmodium* spp. (*P. vivax*, *P. falciparum* and *P. berghei*), *S. cerevisiae* and human to avoid functional redundancy.

*(ii)*Functional expression of *P. vivax* molecular targets and their human orthologs in *S. cerevisiae* surrogate system (Fig. 2):

- all coding sequences from *P. vivax* and *H. sapiens* gene targets were resynthesized by GeneArt® Optomizer for suitable and efficient expression in yeast model organism;

- the synthetic constructs were sub-cloned into a yeast expression plasmid (pCM188);

- *S. cerevisiae* BY4743 (diploid background) strains containing a heterozygous deletion on the candidate gene targets (BY4743 *gene△/*GENE) were heterologous deleted for *pdr5* gene, which codifies a major pleiotropic drug pump;

- plasmid constructs pCM*Hs/Pv*GENE were transformed into the BY4743 *gene△/*GENE::*pdr5△/*PDR5 strains, lacking the major pleiotropic efflux drug pump PDR5;

- *S. cerevisiae* BY4743 *gene△/*GENE::*pdr5△/*PDR5 pCM*Hs/Pv*GENE strains were sporulated, tetrads dissected for haploids selection with the desired heterologous gene expression for drug screens;

- *S. cerevisiae* BY4741 (haploid mate-type a background) strains containing the *pdr5* gene deletion (BY4741 *pdr5△*) were transformed with plasmid constructs pCM*Hs/Pv*GENE;

- *S. cerevisiae* BY4741 *pdr5△* pCM*Hs/Pv*GENE strains were deleted for the essential gene target (*gene△*) of interest.
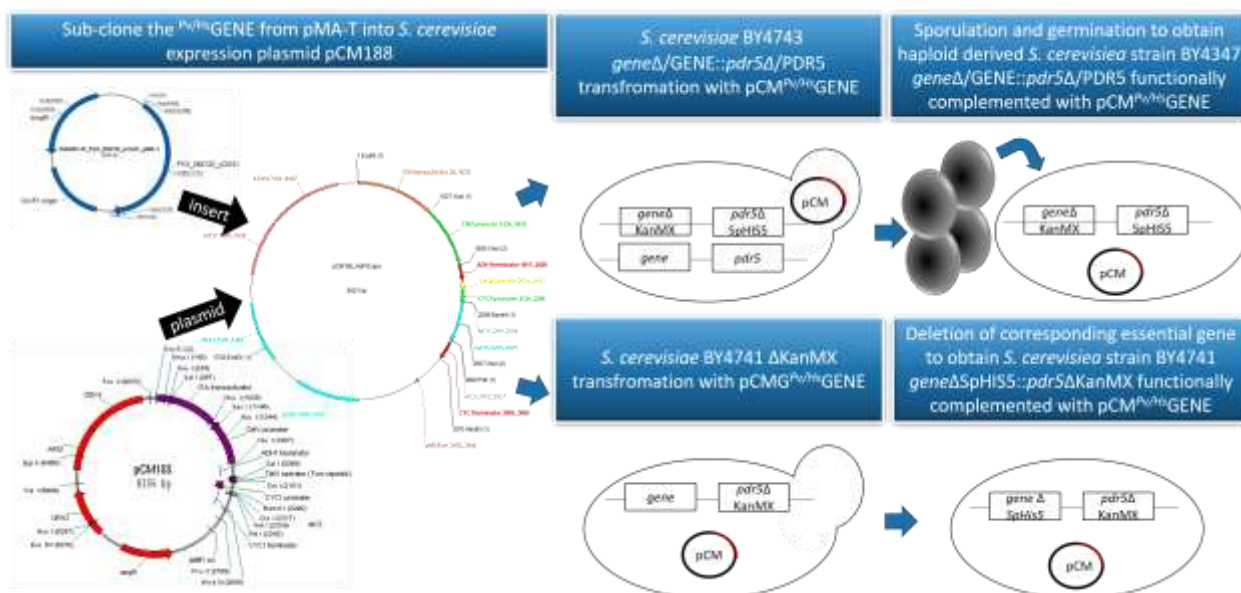
Figure 2. Methodological overview of functional expression of *P. vivax* molecular targets and their human orthologs in *S. cerevisiae* surrogate system.

(iii) <u>Synthetic Compound Library Screens</u>:

• Using protein structure modeling and virtual ligand screening, prospective compounds against 3 different *Plasmodium* spp. targets have been identified from synthetic compound libraries;

• Compounds selectively inhibiting the *P. vivax* orthologue over the human one were/are being identified from available libraries, using our target-based yeast system.

Important note: Since this is a prospective project, all researchers involved decided to not disclose any information relative to the identification of the *P. vivax* molecular targets and/or compound hits IDs or codes.

Results

In silico selection of *P. vivax* molecular targets

Using bioinformatic tools and databases [23] (Table 1), I performed an *in silico* screen under a restricted set of criteria, from an initial ~5631 *P. vivax* genes to the final 30 candidate gene targets, from which 13 candidate genes were selected to start this project (Fig. 2). The three most key points in this *in silico* gene selection were:

(i) gene orthology between *P. vivax*, our organism of interest; *P. falciparum* and *P. berghei* for downstream *in vitro* and *in vivo* validation steps using CRISPR/Cas9 knockout and editing tools. *S. cerevisiae* that acts as a surrogate for expressing anti-parasitic targets, and *H. sapiens* for drug selectivity determination;

(ii)gene essentiality in *S. cerevisiae*, so that expressing genes from *P. vivax* and human orthologs complement yeast loss-of function mutations;

(iii) gene paralogy identification within the *P. vivax*, *P. falciparum*, *P. berghei*, *S. cerevisiae* and *H. sapiens* genes to avoid functional redundancy.

Table 1. Principal databases and omics resources for the *in silico* selection of eleven *P. vivax* molecular targets.

Database	Database Search Goal	Webpage
Inparanoid	Orthology search between *P. vivax*, *S. cerevisiae*, *P. berghei*, *P. falciparum* and *H. sapiens*	http://inparanoid.sbc.su.se/cgi-bin/index.cgi
Yeast Mine	Search and retrieve *S. cerevisiae* data with YeastMine, populated by SGD and powered by InterMine	http://yeastmine.yeastgenome.org /yeastmine/begin.do
DIOPT - DRSC Integrative Ortholog Prediction Tool	Confirmation of best orthology between *S. cerevisiae* and *H. sapiens* genes (and vice-versa)	http://www.flyrnai.org /cgi-bin/DRSC_orthologs.pl
Saccharomyces Genome Database **(SGD)**	Comprehensive integrated biological information for the budding yeast *Saccharomyces cerevisiae*	http://www.yeastgenome.org/
Yeast-Human Functional Complementation Data (SGD)	Search for human orthologs proven to functionally complement *S. cerevisiae* genes	http://www.yeastgenome.org/yeast-human-functional-complementation-data-now-in-sgd
Yeast Genome Database (YGD)	Search and download protein sequence FASTA files from *S. cerevisiae*	http://downloads.yeastgenome.org /sequence/S288C_reference/orf_protein/
PlasmoDB	*Plasmodium* spp. database: searches of orthology, paralogy and data (DNA, RNA and protein) download	http://plasmodb.org/plasmo/
Ensemble	Search for P. vivax – H. sapiens and S. cerevisiae – H. sapiens paralogy using emceed BLAST tool from NCBI	http://www.ensembl.org /Multi/Tools/Blast?db=core

Duplicated Genes Database	Search for *H. sapiens* paralogs	http://dgd.genouest.org/
UNIPROT	Download *H. sapiens* protein sequences from Ensemble Protein Ref IDs	http://www.uniprot.org/
NCBI BLAST	Protein blast too used to blast between *P. vivax* and *P. falciparum* proteins	https://blast.ncbi.nlm.nih.gov /Blast.cgi
Gene Ontology (GO) Consortium	GO terms for biological processes, molecular function, cellular compartments, pathways and protein classes in *P. falciparum*, *S. cerevisiae*	http://geneontology.org
Open Targets	Search on drugs developed or in development for the correspondent human orthologs	https://www.opentargets.org/
The TDR Targets Database	A chemo-genomics resource for neglected tropical diseases	http://tdrtargets.org/
eMolecules	Compound database	https://www.emolecules.com/

Functional expression of *P. vivax* molecular targets and their human orthologs in *S. cerevisiae* surrogate system

The set of 13 *P. vivax* molecular targets and their human orthologues are being expressed in *S. cerevisiae* to assess the degree of functional complementation of the yeast orthologues [25]. In this study, we selected two yeast strains with the same genetic background based on their genotype: BY4743 and BY4741, a diploid and a haploid strain, respectively. These strains come from an extensive library denominated Yeast Genome Deletion Project (http://www-sequence.stanford.edu/group/yeast_deletion_project/deletions3.html), and have been engineered to minimize the homology to the marker genes in commonly used vectors and to reduce plasmid integration events [29]. Besides, heterozygous deletion mutants for each of the selected essential genes are available for the diploid yeast strain. Both BY4743 and BY4741 were used to generate the final haploid strains functionally complemented with the *P. vivax* and human orthologs. The heterologous expression of the ortholog genes is done through a yeast expression plasmid (pCM188, ATCC®
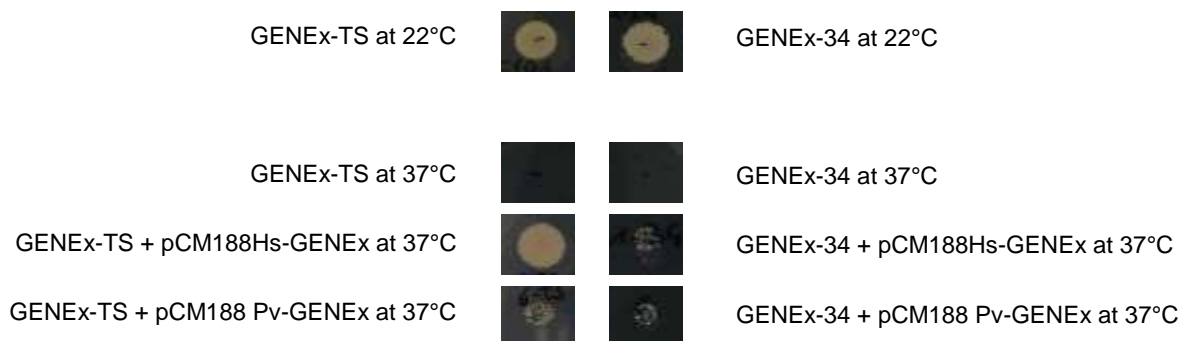
87660™) under a tetracycline-controlled transcription activator (TetOFF promoter under doxycycline (DOX) presence)[30], to control its expression levels inside the functionally complemented strains.

All the 17 coding sequences from *P. vivax* and *H. sapiens* were designed and resynthesized by GeneArt® Optimizer (https://www.thermofisher.com/se/en/home/life-science/cloning/gene-synthesis/geneart-gene-synthesis/geneoptimizer.html) for suitable subsequent sub-cloning and efficient expression in *S. cerevisiae*. From the synthetic constructs, delivered in pMA-T*Pv/Hs*GENE (Amp^R marker) plasmids, 15 were successfully sub-cloned into a yeast expression plasmid (pCM188 ordered from Euroscarf - Europen *Saccharomyces cerevisiae* Archive for Functional Analysis, www.euroscarf.de) obtaining pCM*Hs/Pv*GENE constructs.

S. cerevisiae BY4743 (diploid background) strains containing a heterozygous deletion on the candidate gene targets (BY4743 *gene*Δ/GENE) were heterologously deleted for *pdr5* gene, which codifies a major pleiotropic drug pump. The 15 plasmid constructs pCM*Hs/Pv*GENE are being transformed into the BY4743 *gene*Δ/GENE::*pdr5*Δ/PDR5 strains, lacking the major pleiotropic efflux drug pump PDR5. In alternative, *S. cerevisiae* BY4741 (haploid mate-type A background) strains containing the *pdr5* gene deletion (BY4741 *pdr5*Δ) are being transformed with the plasmid constructs pCM*Hs/Pv*GENE. At this stage, sporulation of heterozygous diploid strains and tetrad dissection, and/or yeast orthologous gene deletion for haploid selection resulted in <u>heterologous expression of 6 chosen *P. vivax* and human molecular targets</u>.

Furthermore, we used *S. cerevisiae* Thermosensitive (TS) strains from Charlie Boone's collection to assay the functional complementation ability of our *P. vivax* and human ortholog genes by transformation with plasmid constructs pCM*Hs/Pv*GENE followed by temperature tests. Thermosensitive mutants are sensible to high temperature, which leads to a truncated protein unable to accomplish its functions. This loss of function leads to a loss of growth at specific temperature points. The expression of a heterologous protein functionally complementing the truncated mutant protein will allow the thermosensitive yeast mutant to grow beyond its thermosensitive temperature. The use of this technique, when TS strains are available, has revealed important for the exclusion of genes (*P. vivax* or human) that cannot functionally complement on our yeast surrogate

system. For one of our chosen *P. vivax* molecular targets, this has proven to be the case, and will not be further perused, but other for other molecular targets, TS strains allowed the confirmation of functional complementation (Fig. 2).

GENEx-TS at 22°C   GENEx-34 at 22°C

GENEx-TS at 37°C   GENEx-34 at 37°C

GENEx-TS + pCM188Hs-GENEx at 37°C   GENEx-34 + pCM188Hs-GENEx at 37°C

GENEx-TS + pCM188 Pv-GENEx at 37°C   GENEx-34 + pCM188 Pv-GENEx at 37°C

Figure 2. Functional complementation assay in thermosensitive yeast mutants. Thermosensitive strains GENEx-TS and GENEx-34 lose their ability to grow at 37°C. The thermosensitive mutants expressing *H. sapiens* or *P. vivax* GENEx protein can grow at 37°C. *H. sapiens* and *P. vivax* GENEx protein can functionally complement the yeast GENEx ortholog.

In silico P. vivax and their human ortholog proteins structural modelling and database virtual screen against compound libraries

In order to narrow down the available choices of compounds to use on our drug screen, our collaborators under supervision of Prof. Dr. Leif Eriksson modeled the protein structure of both parasite (*P. vivax*, *P. falciparum* and *P. berghei*) and human molecular targets using ClustalW as a measure of sequence similarity (conserved and polymorphic regions) through multiple protein sequence alignment and the YASARA-Homology modeler (http://www.yasara.org/homologymodeling.htm) and made predictions for active and/or biding site identification using MOE (http://cbm.msoe.edu/scienceOlympiad/designEnvironment/practice.html). This protein models were then used to screen against the compounds from three big libraries of interest, The Pathogen Box (https://www.pathogenbox.org/), FDA-approved (https://www.accessdata.fda.gov/scripts/cder/daf/), Maybridge (https://www.maybridge.com/portal/alias__Rainbow/lang__en/tabID__146/DesktopDefault.aspx) and ZincDB (http://zinc.docking.org/), in order to find compounds predicted to dock and bind to our molecular targets with enough strength to decide the preferred

compound choices to initiate our drug screens (evaluation performed individually to each species through docking scores and for compounds binding more than one parasite species and/or human target by MASC scores). Both ClustalW and Schrödinger/Glide computational protein modelers were used. Finally, inverse docking was performed as to measure the potential selectivity of each compound towards each four protein targets (*P. vivax*, *P. falciparum*, *P. berghei* and *H. sapiens*).

With this data, I further selected the compounds towards each molecular target in three different groups concerning their overlap in docking/binding capacity. One group had compounds predicted to dock/bind to at least two parasite proteins, the second group had compounds docking/binding to one or more parasite proteins and human ortholog and the last group of compounds was selected from those having the highest docking scores against the *P. vivax* protein. We have already performed drug screens with compounds from The Pathogen Box library towards yeast strains expressing 3 molecular targets. Given the availability of constructed strains for another 3 targets, we are planning to proceed with the screens towards those.

<u>Synthetic compound library screens</u>

To create the best conditions for high sensitivity drug screening experiments in yeast strains expressing the *P. vivax* and human ortholog molecular targets, we use doxycycline (DOX, analogue of tetracycline) to control the expression of the genes under the control of TetO2 promoter of the pCM188 constructed plasmids. To determine the appropriate DOX concentration to control each gene expression, we performed calibration plate and BioScreen assays (more accurate and sensitive method that will be after used for the actual drug screen) with a gradient of concentrations ranging from 0.01 to 30 mg/L (0.01, 0.05, 0.5, 1, 5, 10, 20 and 30 mg/L of DOX). These results gave us the optimal DOX concentration to perform drug screening experiments for each gene target expression, i.e. the optimal DOX concentration at which there is enough molecular target expression to allow yeast strain survival (given the fact that all molecular targets chosen are essential genes), but at such lower lever that could enable us to see the slightest effect that a drug might have, even when the compound concentrations used are restricted.

Upon choosing the best conditions for each strain (medium and DOX concentration), we proceeded towards the drug screen experiments using the BioScreen platform (http://www.growthcurvesusa.com/description.html). All strains were pre-cultured for 3 days in 5mL of media (YPD or YNB-ura) at 30ºC and at 200 rpms. For each yeast culture the O.D. was measured and diluted to a final O.D. of 2.0. Honey comb plates were filled in with 300 to 400µL of appropriate media (YPD or YNB-ura or YNB + DOX) and 10µL of the different yeast strain cultures. For the initial screens, a 10µM compound concentration was used. The experiments were performed in triplicate for the yeast strains heterologous expressing the *P. vivax* and *H. sapiens* genes, as well as a BY4741 strain containing the empty pCM188 vector as a positive control. Since the compounds used were dissolved in DMSO and to exclude and normalize data towards eventual DMSO toxic effects over cell growth rates, we run growth curves of this strains growing solely in media (YPD or YNB-ura or YNB) with the appropriate DOX concentration for each strain and 10µM DMSO. Negative technical controls (wells with the experiment set of medias used) were performed in all plates to exclude contaminations. The runs at the BioScreen O.D. plate readers were executed for a total of 3 days at 30ºC, continuously shaking on low amplitude and slow speed, stopping at 20 minute intervals for O. D. read after 10 minutes using the 420-580nm wave length filter. By the end of each experiment, data was saved in csv format for further analysis.

For analyzing yeast growth data, we use the software PRECOG (http://precog.lundberg.gu.se/). PRECOG automatically estimates the three main growth parameters, lag time, growth rate and yield, and was designed to take care of dataset problems. For example, BioScreen errant readings often caused by gas bubbles in the wells, misread O.D. values, and extreme datapoints. PRECOG gives growth curves a smoother and more realistic appearance, by using algorithms to compensate underestimations of cell density at higher O.D. reads and adjust the growth curves so that they better represent the cell density over time. All statistics are calculated based on the average of our triplicate data.

Chemical screens for novel antiplasmodial agents are currently being performed on the set of 6 *P. vivax* molecular targets and their human orthologues, expressed in *S. cerevisiae* by functional complementation.
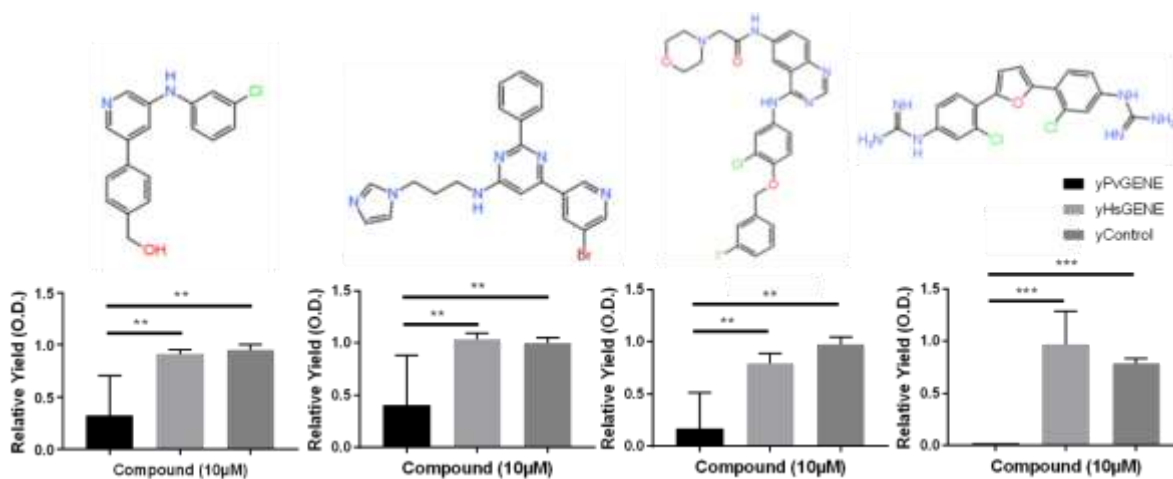
Figure 2. Graphs showing the results for the 4 compound hits found against one *P. vivax* molecular target. ** *p-value* < 0.005; *** *p-value* < 0.0001.

From the Pathogen Box library, we have already found 4 promising compounds, selectively inhibiting the yeast strain expressing one of the *P. vivax* molecular targets (Fig. 2). Importantly, all 4 compounds show no relevant cytotoxic effects towards the yeast strain heterologous expressing the human gene, they also have no significant effects on yeast control strains (Fig. 2).

Discussion

We believe this high-throughput target-based approach is a powerful method to find high efficacy anti-parasitic lead molecules and has already given proofs with the discovery of 4 compound hits against one molecular target showing very strong selectivity and low human cytotoxicity (Fig. 2). These and further compound hits are being further tested for accurate IC50 determination on our yeast system.

Some drawbacks must be overcome related to our target-based yeast screen method. From all experimental tasks yet performed, sporulation of *S. cerevisiae* BY4743 *gene∆/*GENE::*pdr5∆/*PDR5 pCM*Hs/Pv*GENE strains are, as expected, very inefficient. Tetrads dissection for haploids selection with the desired heterologous gene expression for the drug screens is time consuming and has a very low success rate. Further

optimization and other sporulation protocols have been attempted. To overcome this problem, we immediately started a similar approach, but this time using the *S. cerevisiae* BY4741 *pdr5Δ* pCM*Hs/Pv*GENE haploid strain, in which we deleted the essential gene target (*geneΔ*) of interest. From this alternative only, we were able to obtain half of the 6 different strains, needed for drug screen experimentation.

In parallel, we acquired recently a new yeast background yeast strains (Y13363 and Y13118), which might make the construction of functional complemented strains faster and easier and enhance the screen capacity of this platform (Y13363: MATalpha his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 pdr1::natMX pdr3::Klura3 snq2::KI.LEU2 and Y13118 (with SGA reporters): MATalpha his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 pdr1::natMX pdr3::Klura3 snq2::KI.LEU2 can::STEpr-Sp_his5 lyp1Δ). These strains were genetically engineered with SGA reporters for haploid selection on sporulation protocols when mated with an *S. cerevisiae* BY4743 *geneΔ*/GENE and the 3 major efflux pump genes or precursors (*pdr1*, *pdr3* and *snq2*) deleted, sensitizing further the strain to drug uptake and ameliorating the efficiency of the system for drug effects identification. Another future option includes the enhancement of all surrogate yeast systems using new CRISPR/Cas9 techniques to both create a more sensitized yeast strain, where multiple export pumps can be deleted, and the deletion of the yeast gene of interest would be easily done in a one-step switch. This could be done together with insertion and expression of the *Plasmodium* spp. or human counterpart genes in the *S. cerevisiae* genome under TetOFF regulable yeast expression system.

On the other hand, TS strains, when available, were a valuable tool that aided us to understand the capacity of each *P. vivax* and human genes to functionally complement the *S. cerevisiae* ortholog, and thus, the possibility of generating the correspondent work strains.

The use of protein structure modeling and virtual ligand screening allowed the identification of prospective compounds against 3 different *Plasmodium* spp. targets from synthetic compound libraries. This bioinformatic screening step narrows down the list of compounds to be tested on our experimental screens, by selecting those based on the strength of binding (docking scores) towards the structural protein features of each *Plasmodium* spp. and human molecular targets. Following this *in silico* approach,

compounds predicted to selectively inhibit the *P. vivax* orthologue (over the human counterpart) are being identified from extensive libraries of synthetic molecules.

Importantly, *in vitro*, *in vivo* and *ex vivo* assays for compound efficacy validation are ongoing, especially the test of candidate bioactive molecules against (drug-resistant) *P. falciparum, P. berghei* and *P. vivax* strains.

Concluding Remarks and Future Prospects

The yeast model under development will surely provide a target-based drug screen platform able to identify molecules selective for the yeast heterologous expressing *P. vivax* molecular target and enhance parasite selectivity. Most importantly, it will give us further information on the molecular mechanisms of drug targeting to avoid the so dangerous acquisition of parasite resistance. Also, the use of chemical genomics technology in yeast can help us to reveal the biological mode-of-action of natural antiplasmodial compounds. Molecular target-compound *in vitro* assays will be done to evaluate the interaction between the *P. vivax* target protein and the candidate bioactive molecules through thermal shift and enzymatic activity assays. Within this collaborative project, we wish to do compound redesign and synthesis, for those hit compounds showing a selective effect against *P. vivax*. Modelling and structure–activity relationship (SAR) studies to explore the relationship between the chemical structure of a molecule and its biological activity will also be applicable in these cases. Finally, the best hits should be assayed, and our *P. vivax* molecular targets validated using the powerful and new CRISPR/Cas9 knockout and editing tools, both through *in vitro* in *P. falciparum* and/or *in vivo* in *P. berghei* assays. Automated high-throughput screens can be performed to identify new anti-parasitic compounds and allow a cheaper and faster drug development against Malaria, which also could be extended to other neglected tropical diseases [27]. This endeavor will help to develop novel antiplasmodial agents targeting the proteins expressed in *P. vivax* life cycle stages which are otherwise difficult to access.

Research collaborations

This work enforces the already existent international collaboration between Brazilian, Swedish and Cambodia research groups managed by the Principal Investigators Prof. Dr. Per Sunnerhagen, Prof. Dr. Elizabeth Bilsland, Prof. Dr. Fabio T. M. Costa, Prof. Dr. Morten Grotli and Prof. Dr. Benoit Witkowski, within the ongoing research grants entitled "Superior bioactive molecules against *Plasmodium vivax* and *Plasmodium falciparum* through genetic screening in yeast" and "Combatting antimicrobial resistance by novel antimalarial molecules against *Plasmodium vivax* and *P. falciparum* from South America and South-East Asia". Other collaboration opportunities are currently being explored, principally those that can bring us further support for high throughput robotic systems for the screen, libraries of compounds, CRISPR/Cas9 technology learning and/or establishment and, most importantly, scientific result cooperation for *P. vivax* drug discovery breakthroughs.

References

1. Anstey, N.M., et al., *The pathophysiology of vivax malaria.* Trends Parasitol, 2009. **25**(5): p. 220-7.
2. WHO, *World Malaria Report 2018.*, in *World Malar. Rep.* 2018: Geneva.
3. Reyburn, H., *New WHO guidelines for the treatment of malaria.* BMJ, 2010. **340**: p. c2637.
4. WHO, *Guidelines for the Treatment of Malaria. 2nd edition.* World Health Ornanization, 2010. **2**: p. 1-194.
5. Beutler, E., S. Duparc, and G.P.D.W. Group, *Glucose-6-phosphate dehydrogenase deficiency and antimalarial drug development.* Am J Trop Med Hyg, 2007. **77**(4): p. 779-89.
6. Frampton, J.E., *Tafenoquine: First Global Approval.* Drugs, 2018. **78**(14): p. 1517-1523.
7. Gogtay, N., et al., *Artemisinin-based combination therapy for treating uncomplicated Plasmodium vivax malaria.* Cochrane Database Syst Rev, 2013(10): p. CD008492.
8. Tjitra, E., et al., *Multidrug-resistant Plasmodium vivax associated with severe and fatal malaria: a prospective study in Papua, Indonesia.* PLoS Med, 2008. **5**(6): p. e128.
9. Poespoprodjo, J.R., et al., *Adverse pregnancy outcomes in an area where multidrug-resistant plasmodium vivax and Plasmodium falciparum infections are endemic.* Clin Infect Dis, 2008. **46**(9): p. 1374-81.
10. Baird, J.K., *Chloroquine resistance in Plasmodium vivax.* Antimicrob Agents Chemother, 2004. **48**(11): p. 4075-83.
11. de Santana Filho, F.S., et al., *Chloroquine-resistant Plasmodium vivax, Brazilian Amazon.* Emerg Infect Dis, 2007. **13**(7): p. 1125-6.
12. Suwanarusk, R., et al., *Chloroquine resistant Plasmodium vivax: in vitro characterisation and association with molecular polymorphisms.* PLoS One, 2007. **2**(10): p. e1089.
13. Russell, B., et al., *Determinants of in vitro drug susceptibility testing of Plasmodium vivax.* Antimicrob Agents Chemother, 2008. **52**(3): p. 1040-5.
14. Price, R.N. and N.M. Douglas, *Artemisinin combination therapy for malaria: beyond good efficacy.* Clin Infect Dis, 2009. **49**(11): p. 1638-40.
15. Price, R.N., et al., *Global extent of chloroquine-resistant Plasmodium vivax: a systematic review and meta-analysis.* Lancet Infect Dis, 2014. **14**(10): p. 982-91.

16. Ginsburg, H. and E. Deharo, *A call for using natural compounds in the development of new antimalarial treatments - an introduction.* Malar J, 2011. **10 Suppl 1**: p. S1.
17. Giaever, G., et al., *Functional profiling of the Saccharomyces cerevisiae genome.* Nature, 2002. **418**(6896): p. 387-91.
18. Hillenmeyer, M.E., et al., *The chemical genomic portrait of yeast: uncovering a phenotype for all genes.* Science, 2008. **320**(5874): p. 362-5.
19. Ho, C.H., et al., *A molecular barcoded yeast ORF library enables mode-of-action analysis of bioactive compounds.* Nat Biotechnol, 2009. **27**(4): p. 369-77.
20. Khozoie, C., R.J. Pleass, and S.V. Avery, *The antimalarial drug quinine disrupts Tat2p-mediated tryptophan transport and causes tryptophan starvation.* J Biol Chem, 2009. **284**(27): p. 17968-74.
21. Li, W., et al., *Yeast model uncovers dual roles of mitochondria in action of artemisinin.* PLoS Genet, 2005. **1**(3): p. e36.
22. Ericson, E., et al., *Off-target effects of psychoactive drugs revealed by genome-wide assays in yeast.* PLoS Genet, 2008. **4**(8): p. e1000151.
23. Ludin, P., et al., *In silico prediction of antimalarial drug target candidates.* Int J Parasitol Drugs Drug Resist, 2012. **2**: p. 191-9.
24. Augagneur, Y., et al., *Identification and functional analysis of the primary pantothenate transporter, PfPAT, of the human malaria parasite Plasmodium falciparum.* J Biol Chem, 2013. **288**(28): p. 20558-67.
25. Bilsland, E., et al., *Functional expression of parasite drug targets and their human orthologs in yeast.* PLoS Negl Trop Dis, 2011. **5**(10): p. e1320.
26. Bilsland, E., et al., *Yeast-based automated high-throughput screens to identify anti-parasitic lead compounds.* Open Biol, 2013. **3**(2): p. 120158.
27. Williams, K., et al., *Cheaper faster drug development validated by the repositioning of drugs against neglected tropical diseases.* J R Soc Interface, 2015. **12**(104): p. 20141289.
28. Bilsland, E., et al., *Yeast-Based High-Throughput Screens to Identify Novel Compounds Active against Brugia malayi.* PLoS Negl Trop Dis, 2016. **10**(1): p. e0004401.
29. Brachmann, R.K. and J.D. Boeke, *Tag games in yeast: the two-hybrid system and beyond.* Curr Opin Biotechnol, 1997. **8**(5): p. 561-8.
30. Gari, E., et al., *A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in Saccharomyces cerevisiae.* Yeast, 1997. **13**(9): p. 837-48.

ANNEX 3

Ethical Approval

CAAE-0044.0.114.000-11

CAAE- 54234216.0000.0005

SisGen nº A5493E1

PARECER Nº 2082-CEP/FMT-HVD

Registro CEP: Protocolo Nº0013/2011- CEP/FMT-HVD

CAAE – 0044.0.114.000-11.

Projeto de Pesquisa: "Estudo da biologia de *Plasmodium vivax*: imunopatogênese, citoadesão e invasão".
Pesquisadora Responsável: Marcus Vinícius Guimarães de Lacerda.
Instituição Proponente: Fundação de Medicina Tropical "Doutor Heitor Vieira Dourado"
Instituição Co-Participante: Centro de Pesquisa Leônidas e Maria Deane.
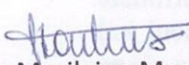**Área Temática Especial:** Não se aplica.
Patrocinador: .Fundação de Amparo à Pesquisa do Estado de São Paulo - Fapesp.
Registro para armazenamento de material Biológico humano: Não.

Ao se proceder a análise do projeto de pesquisa em questão, o Comitê de Ética em Pesquisa em Seres Humanos da Fundação de Medicina Tropical "Doutor Heitor Vieira Dourado" - CEP/FMT-HVD, em sessão ordinária do dia 16 de dezembro de 2011 e de acordo com as atribuições definidas na Resolução CNS 196/96 e suas complementares, manifesta-se favorável à **aprovação** do protocolo. A partir desta data o Pesquisador fica autorizado a dar inicio à coleta de dados da pesquisa e pela **entrega do relatório final deste estudo, neste CEP, em janeiro de 2014,** conforme Resolução do Conselho Nacional de Saúde nº 196, de 10.10.1996, item IX. 2 letra "c".

Situação do Protocolo: APROVADO

Manaus, 16 de dezembro de 2011.

Dra Marilaine Martins
Coordenadora do CEP/FMT-HVD, em exercício

PARECER CONSUBSTANCIADO

Identificação do Projeto:
   Estudo da biologia de *Plasmodium vivax*: imunopatogênese, citoadesão e invasão.

Pesquisadores responsáveis:
   Marcus Vinicius Guimarães de Lacerda

Instituição onde se realizará a pesquisa:
   Gerência de Malária da Fundação Medicina Tropical – Heitor Vieira Dourado

Data da reapresentação ao CEP:
   29 de novembro de 2011 – Protocolo n° 0013/2011
   Entrega para parecer – 02 de dezembro de 2011

Objetivo da Reapresentação:
   Correção das pendências apresentadas no parecer do dia 16/10/2011

Parecer:
   O pesquisador principal fez todas as alterações solicitadas no projeto:
   1- Mencionar como os pacientes selecionados serão tratados e acompanhados:
   Foi incluído um novo item na metodologia com todos os dados com acompanhamento dos pacientes = **Pendência atendida**
   2- Mencionar o tamanho da amostra de pacientes controles, como serão obtidos e tratamento em relação TCLE.
   Foi incluído os detalhes dos pacientes do grupo controle que serão 20 pacientes com diagnóstico microscópico negativo para malária e que concordarem e assinarem o TCLE, incluído também os critérios de inclusão e exclusão do grupo controle = **Pendência atendida**
   3- Considerando mulheres grávidas serão excluídas como será feito o processo de identificação da gravidez.
   Foi incluído no questionário (anexo 2) questionamento sobre gravidez vigente, só sendo considerado para inclusão no estudo aquelas que responderam convictamente Não = **Pendência atendida**

1

Todos os documentos exigidos estão anexados. O Termo de Consentimento Livre e esclarecido está anexado. **Achamos que o projeto deva ser aprovado**.

Salvo melhor juízo

Este é o meu parecer

Dra. Mariiaine Martins
Gerente de Parasitologia
FMTAM

Data: Manaus, 09/12/2011

2

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Patogênese de Plasmodium vivax
Pesquisador: Stefanie Costa Pinto Lopes
Área Temática:
Versão: 1
CAAE: 54234216.1.0000.0005
Instituição Proponente: Fundação de Medicina Tropical do Amazonas - FMT/IMT/AM
Patrocinador Principal: FUNDACAO DE AMPARO A PESQUISA DO ESTADO DE SAO PAULO

DADOS DO PARECER

Número do Parecer: 1.457.458

Apresentação do Projeto:

O Plasmodium vivax é responsável por mais de 80 milhões de casos de malária por ano no mundo, apresentando forte impacto social fora do continente africano, principalmente na Ásia e nas Américas. O Brasil responde por 50-60% do total de casos de malária notificados nas Américas, sendo 85% dessas infecções causadas por P. vivax e com transmissão restrita quase exclusivamente à região Amazônica (99,8%). Achados

anatomopatológicos semelhantes aos que são observados nos casos de P. falciparum foram recentemente constatados em infecções por P. vivax, as quais podem também evoluir para formas graves da doença. Essas observações desafiam a visão pré-estabelecida de que P. vivax é um parasita "benigno". No entanto, a realização de ensaios funcionais para o estudo da patogênese e infectividade de P. vivax permanecem restritos a centros

hospitalares de referência regional em áreas endêmicas. Este fato se deve, principalmente, a impossibilidade de sistema de cultivo de longa duração in vitro que seja confiável e reprodutível. Sendo assim, em parceria com centros hospitalares de referência em áreas endêmicas para malária, fomos capazes de desenvolver ensaio de invasão ex vivo, e mostramos que eritrócitos infectados de P. vivax (Ei-Pv) coletados de pacientes infectados são capazes de aderir ex vivo ao endotélio pulmonar, cerebral e na placenta. Esses achados sugerem a participação desta capacidade adesiva nos

processos patológicos de P. vivax nesses órgãos. Baseado na capacidade de campo instalada e no conhecimento já adquirido pretendemos: (i) ampliar a compreensão dos mecanismos de patogênese relacionados à citoadesão e à formação de rosetas de Ei-Pv (ensaios funcionais), (ii) identificar os potenciais ligante(s) parasitários envolvidos (por meio de análises moleculares e geração de anticorpos monoclonais) e (iii) avaliar a

participação das plaquetas nesse processo adesivo e seu efeito na geração e amplificação da ativação endotelial na malária por P. vivax.

Objetivo da Pesquisa:

Objetivo Primário:

O objetivo geral deste subprojeto é investigar detalhadamente o papel de ligantes parasitários e receptores endoteliais envolvidos no processo de sequestro de P. vivax, via estudos de formação de rosetas e de citoadesão no endotélio, bem como a análise das potenciais consequências destes fenômenos no hospedeiro. Pretende-se ainda, avaliar a participação das plaquetas e mediadores solúveis noplasma nesse processo adesivo e seu efeito na geração e amplificação da ativação endotelial na malária por P. vivax.

Objetivo Secundário:

i. Verificar a existência de desproporção entre diferentes estágios parasitários no sangue periférico de pacientes infectados com P. vivax por meio da análise cuidadosa de esfregaços sanguíneos e citometria de fluxo; ii. Determinar a participação no processo citoadesivo e na formação de rosetas de diferentes estágios parasitários (trofozoítas e esquizontes) de Ei-Pv, bem como o grau de envolvimento de células e/ou moléculas (e.g. plaquetas, micropartículas, citocinas, imunoglobulinas, sistema complemento, ADAMTS13 e VWF) presentes no plasma de indivíduos infectados; iii. Verificar a natureza bioquímica (proteica) de antígenos da superfície de Ei-Pv responsáveis pela formação de rosetas por meio de tratamento prévio com enzimas (e.g. condroitinase, hialuronidase, neuroaminidase, tripsina, quimiotripsina, heparinase, etc.), bem como a participação de diferentes grupos

sanguíneos (ABO); iv. Realizar ensaios de adesão (estático e fluxo) com Ei-Pv isolados a partir do sangue periférico de pacientes em células endoteliais após diferentes tratamentos enzimáticos (e.g. condroitinase, hialuronidase, etc.) e na presença de CSA e heparan sulfato; v. Avaliar o papel do endotélio, por meio de sua ativação ou não, com moléculas imunoestimulatórias (e.g. TNF-, IFN- e LT-), determinando a expressão de

potenciais receptores e do grau adesivo de Ei-Pv; vi. Analisar a natureza genética, por meio do sequenciamento do transcrito em plataforma Illumina, de diferentes estágios de Ei-Pv envolvidos,

Endereço: Av. Pedro Teixeira, 25
Bairro: D. Pedro I   CEP: 69.040-000
UF: AM   Município: MANAUS
Telefone: (92)2127-3572   Fax: (92)2127-3572   E-mail: cep@fmt.am.gov.br

Página 02 de 05

Continuação do Parecer: 1.457.458

ou não, na adesão e na formação de rosetas. vii. Caracterização da diversidade genética de diferences isolados de P. vivax coletados em área endêmica da Amazônia, quer por detecção de microssatélites, quer por análise da distribuição de

polimorfismos (SNPs). viii. Expressar variantes destas proteínas para avaliação da sua imunogenicidade utilizando um painel de soros de pacientes infectados com P. vivax na Amazônia Brasileira.

Avaliação dos Riscos e Benefícios:

Riscos:

Os únicos riscos aos participantes deste pesquisa são os riscos associados com a coleta de sangue, que incluem: dor, hematoma, ou outro desconforto no local da coleta. Raramente desmaio ou infecções no local de punção podem ocorrer. Cuidados devem ser tomados para minimizar esses riscos.

Benefícios:

O participante não será beneficiado diretamente pela participação neste estudo. Trata-se de um estudo sobre a biologia do Plasmodium vivax e imunopatogênese da malária, que gerará resultados a longo prazo. Porém, os resultados obtidos com este estudo poderão contribuir e impactar futuramente o tratamento e desenvolvimento de vacinas e drogas para malária vivax.

Comentários e Considerações sobre a Pesquisa:

A análise do protocolo revelou todos os itens requeridos pela Resolução CNS nº 466 de 2012. Diante do exposto na justificativa apresentada, considera-se o tema de extrema relevância para a saúde pública. Trata-se de pesquisa com cooperação de pesquisadores da UNICAMP, FIOCRUZ e da FMT-HVD, cujo protocolo permitiu comprovar a participação de cada membro da equipe, de adequada capacidade científica. Além disso, observa-se que os

benefícios advindos do processo de investigação e dos resultados da pesquisa suplantam seu ônus e potenciais riscos.

Considerações sobre os Termos de apresentação obrigatória:

A análise do protocolo revelou todos os itens requeridos pela Resolução CNS nº 466 de 2012: 1. PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO; 2.Folha de rosto adequadamente preenchida; 3. Projeto detalhado; 4. Carta de anuência da Gerência de Malária da FMT-HVD; 5.TCLE adequadamente redigido; 6. Instrumento de coleta de dados.

Recomendações:

Endereço: Av. Pedro Teixeira, 25	
Bairro: D. Pedro I	**CEP:** 69.040-000
UF: AM **Municipio:** MANAUS	
Telefone: (92)2127-3572 **Fax:** (92)2127-3572	**E-mail:** cep@fmt.am.gov.br

Página 03 de 05

Continuação do Parecer: 1.457.458

Conclusões ou Pendências e Lista de Inadequações:

Diante da análise do protocolo, este relator não observou nenhuma pendência. Voto pela aprovação do mesmo na sua versão atual.

Considerações Finais a critério do CEP:

O presente projeto está APROVADO e os interessados ficam informados de apresentar a este CEP os relatórios parciais e final do estudo, conforme prevê a Resolução CNS nº 466/2012, utilizando o formulário de Roteiro para Relatório Parcial/Final de estudos clínicos Unicêntricos e Multicêntricos, proposto pela CONEP em nossa home page.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_P ROJETO_558680.pdf	10/03/2016 17:58:29		Aceito
Outros	ANUENCIA.pdf	10/03/2016 17:49:17	João Conrado Khouri dos Santos	Aceito
Folha de Rosto	FOLHADEROSTO.pdf	10/03/2016 17:46:37	João Conrado Khouri dos Santos	Aceito
Projeto Detalhado / Brochura Investigador	COMITEDEETICA.docx	09/03/2016 14:47:05	Stefanie Costa Pinto Lopes	Aceito
Outros	Questionario.pdf	03/02/2016 13:21:21	Catarina Baeta da Luz Bourgard	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.pdf	03/02/2016 13:19:47	Catarina Baeta da Luz Bourgard	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Endereço: Av. Pedro Teixeira, 25
Bairro: D. Pedro I   **CEP:** 69.040-000
UF: AM   **Município:** MANAUS
Telefone: (92)2127-3572   **Fax:** (92)2127-3572   **E-mail:** cep@fmt.am.gov.br

Continuação do Parecer: 1.457.458

MANAUS, 18 de Março de 2016

_____

Assinado por:
Marilaine Martins
(Coordenador)

Comprovante de Cadastro de Acesso

Cadastro nº A5493E1

A atividade de acesso ao Patrimônio Genético, nos termos abaixo resumida, foi cadastrada no SisGen, em atendimento ao previsto na Lei nº 13.123/2015 e seus regulamentos.

Número do cadastro:	**A5493E1**
Usuário:	**UNICAMP**
CPF/CNPJ:	**46.068.425/0001-33**
Objeto do Acesso:	**Patrimônio Genético**
Finalidade do Acesso:	**Pesquisa**

Espécie

Plasmodium vivax

Plasmosium vivax

Plasmodium vivax

Plasmodium vivax

Título da Atividade:	**DESENVOLVIMENTO DE NOVAS FERRAMENTAS PARA BUSCA E VALIDAÇÃO DE ALVOS MOLECULARES PARA TERAPIA E PROFILAXIA CONTRA Plasmodium vivax**

Equipe

Fabio Trindade Maranhão Costa	**UNICAMP**
Carla Cristina Judice Maria	**Universidade Estadual de Campinas - UNICAMP**
Najara Carneiro Bittencourt	**Universidade Estadual de Campinas - UNICAMP**
João Conrado Khouri dos Santos	**Universidade Estadual de Campinas - UNICAMP**
Juliana Almeida Leite	**Universidade Estadual de Campinas - UNICAMP**

Tatyana Almeida Tavella	**Universidade Estadual de Campinas - UNICAMP**
Ana Carolina Andrade Vitor kayano	**Universidade Estadual de Campinas - UNICAMP**
Catarina Baeta da Luz Bourgard	**Universidade Estadual de Campinas - UNICAMP**
Kaira Cristina Peralis Tomaz	**Universidade Estadual de Campinas - UNICAMP**
João Luiz da Silva Filho	**Universidade Estadual de Campinas - UNICAMP**
Letícia Cristina Scarapicchia Monteiro	**Universidade Estadual de Campinas - UNICAMP**
Letusa Albrecht	**Instituto Carlos Chagas - FIOCRUZ Paraná**
Pedro Cravo	**Instituo de Higiene e Medicina Tropical (IHMT) - Universi**
Marcus Vinicius Lacerda	**Fundação de Medicina Tropical do Amazonas**
Daniel Youssef Bargieri	**Universidade de São Paulo**
Irene da Silva Soares	**Universidade de São Paulo**
Laurent Claude Stéphane Renia	**Singapore Immunology Network (SIgN) :: Agency for Sci**
Carolina Horta Andrade	**Universidade Federal de Goiania**
Stefanie Costa Pinto Lopes	**Instituto Leônidas e Maria Deane - FIOCRUZ Amazônia**
Elizabeth Bilsland	**Universidade Estadual de Campinas - UNICAMP**
Wuelton M. Monteiro	**Fundação de Medicina Tropical do Amazonas**
Rogerio Amino	**Institut Pasteur Paris**
Matthias Marti	**University of Glasgow**
Samuel Wassmer	**London School of Hygiene & Tropical Medicine**
Per Johan Sunnerhagen	**University of Gothenburg**
Renato Beilner Machado	**Instituição de Patologia Tropical e Saúde Pública da Univ**
Ana Beatriz Silva	**Instituto Carlos Chagas - FIOCRUZ Paraná**
Tamirys S Pimenta	**Instituto Evandro Chagas**
Helder I Nakaya	**Universidade de São Paulo**
Ana Maria Revorêdo da Silva Ventura	**Instituto Evandro Chagas**
Ricardo LD Machado	**Instituto Evandro Chagas**
Matthew William Alfred Dixon	**University of Melbourne**
Leticia Tiburcio	**Universidade Estadual de Campinas - UNICAMP**
Gustavo Capatti Cassiano	**Universidade Estadual de Campinas - UNICAMP**

Parceiras Nacionais

04.534.053/0001-43 / Fundacao de Medicina Tropical Doutor Heitor Vieira Dourado

33.781.055/0001-35 / Centro de Pesquisas Leônidas e Maria Deane - Fiocruz

Parceiras no Exterior

University of Gothenburg

University of Glasgow

Singapore Immunology Network, Agency for Science, Te

Institute Pasteur - Paris

London School of Hygiene & Tropical Medicine

Data do Cadastro: **11/10/2018 11:19:53**

Situação do Cadastro: **Concluído**

Conselho de Gestão do Patrimônio Genético

Situação cadastral conforme consulta ao SisGen em **11:49** de **11/10/2018**.

SISTEMA NACIONAL DE GESTÃO
DO PATRIMÔNIO GENÉTICO
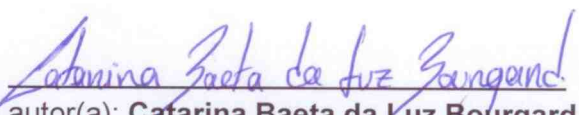E DO CONHECIMENTO TRADICIONAL
ASSOCIADO - **SISGEN**

ANNEX 4

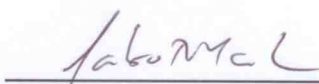**Authorship rights declaration**

Declaração

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **Analyses of the immunopathological and molecular mechanisms involved in cytoadherence of Plasmodium vivax**, não infringem os dispositivos da Lei n.° 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 11 de março de 2019

Assinatura : *Catarina Baeta da Luz Bourgard*

Nome do(a) autor(a): **Catarina Baeta da Luz Bourgard**

RG n.° V999127-A

Assinatura : *Fabio MC*

Nome do(a) orientador(a): **Fabio Trindade Maranhão Costa**

RG n.° 1263942