Total parasite biomass but not peripheral parasitaemia is associated with endothelial and haematological perturbations in *Plasmodium vivax* patients

João L Silva-Filho¹,²*, João CK Dos-Santos¹*, Carla Judice¹, Dario Beraldi², Kannan Venugopal², Diogenes Lima³, Helder Nakaya³, Erich EV Paula⁴, Stefanie CP Lopes⁶, Marcus VG Lacerda⁵,⁶, Matthias Marti²*, Fabio TM Costa¹#

¹Laboratory of Tropical Diseases – Prof. Luiz Jacintho da Silva, Department of Genetics, Evolution, Microbiology and Immunology, Institute of Biology, University of Campinas, Campinas, Brazil
²Wellcome Centre for Integrative Parasitology, Institute of Infection, Immunity & Inflammation, University of Glasgow, Glasgow, UK
³School of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil,
⁴Department of Clinical Pathology, School of Medical Sciences, University of Campinas, Campinas, Brazil,
⁵Tropical Medicine Foundation Dr. Heitor Vieira Dourado, Manaus, Brazil
⁶Institute Leônidas & Maria Deane, Fiocruz, Manaus, Brazil
*Equal contribution
#Corresponding authors coordinated this work equally.
Abstract

*Plasmodium vivax* is the major cause of human malaria in the Americas. How *P. vivax* infection can lead to poor clinical outcomes, despite low peripheral parasitaemia remains a matter of intense debate. Estimation of total *P. vivax* biomass based on circulating markers indicates existence of a predominant parasite population outside of circulation. In this study we investigate associations between both peripheral and total parasite biomass and host response in vivax malaria. We analysed parasite and host signatures in a cohort of uncomplicated vivax malaria patients from Manaus, Brazil, combining clinical and parasite parameters, multiplexed analysis of host responses and *ex vivo* assays. Patterns of clinical features, parasite burden and host signatures measured in plasma across the patient cohort were highly heterogenous. Further data deconvolution revealed two patient clusters, here termed Vivax\textsuperscript{low} and Vivax\textsuperscript{high}. These patient subgroups were defined based on differences in total parasite biomass but not peripheral parasitaemia. Overall Vivax\textsuperscript{low} patients clustered with healthy donors and Vivax\textsuperscript{high} patients showed more profound alterations in haematological parameters, endothelial cell (EC) activation and glycocalyx breakdown and levels of cytokines regulating different haematopoiesis pathways compared to Vivax\textsuperscript{low}. Vivax\textsuperscript{high} patients presented more severe thrombocytopenia and lymphopenia, along with enrichment of neutrophils in the peripheral blood and increased neutrophil-to-lymphocyte ratio (NLCR). When patients’ signatures were combined, high association of total parasite biomass with a subset of markers of EC activation, thrombocytopenia and lymphopenia severity was observed. Finally, machine learning models defined a combination of host parameters measured in the circulation that could predict the extent of parasite infection outside of circulation. Altogether, our data show that total parasite biomass is a better predictor of perturbations in host homeostasis in *P. vivax* patients than peripheral
parasitaemia. This supports the emerging paradigm of a *P. vivax* tissue reservoir, in particular in the hematopoietic niche of bone marrow and spleen.

**Keywords:** *Plasmodium vivax*, malaria parasite, total biomass, tissue infection, endothelial activation, haematopoiesis.
Introduction

Malaria remains a heavy burden across endemic regions worldwide. In 2018 *P. vivax* infection accounted for 41% of all malaria cases outside of Sub-Saharan Africa, with a total of 6.5 million cases and more than 2 billion people in 90 countries at risk. There are concerns that *P. vivax* elimination will be significantly more difficult than *P. falciparum*, as the current measures for malaria control are less effective for *P. vivax* than for *P. falciparum*, with the elimination of the former presenting a major challenge in areas that successfully reduced *P. falciparum* burden. This persistence is due to some unique biological features complicating treatment and elimination, including low peripheral parasitaemia and presence of dormant liver stages (hypnozoites) which relapse weeks or months after blood infection has been cleared.

*P. vivax* infection is associated with low peripheral parasitaemia (< 2%), as a result of a strict host cell tropism to immature reticulocytes that are exceedingly rare in peripheral blood (< 2%) but highly prevalent in the hematopoietic niche of BM and spleen. Because of limited microvascular adherence *in vivo* and endothelial cell (EC) binding *in vitro*, it was generally assumed that peripheral parasitaemia reflects the majority of *P. vivax* parasites during infection. However, discrepancy of parasite biomass based on systemic biomarkers such as *Plasmodium* lactate dehydrogenase (pLDH) compared to peripheral parasitaemia supports existence of a major *P. vivax* reservoir outside of circulation, in particular in patients with complicated outcomes. In support of this hypothesis, studies have demonstrated that late asexual blood stage *P. vivax* parasites (i.e., schizonts) are capable of cytoadhering to endothelial host receptors and present at reduced abundance compared to the other blood stages in the blood of *P. vivax* patients. In experimentally infected non-human primates (NHPs), a significant enrichment of sexual stages (gametocytes) and schizonts in bone marrow sinusoids and parenchyma has been observed, supporting previous evidence from
multiple case reports that identified *P. vivax* in bone marrow and spleen. A series of recent studies in acute and chronic human *P. vivax* infection have meanwhile provided direct evidence that bone marrow and spleen represent the major reservoir of parasite biomass in *P. vivax* infection.

*P. vivax* parasites can elicit a potent host response, including inflammation and endothelial cell (EC) activation, and cause severe and fatal manifestations at significantly lower peripheral parasitaemia than the more virulent species, *P. falciparum*. However, the pathogenic mechanisms underlying these alterations in host homeostasis and their relationship with *P. vivax* biomass are not fully understood.

Here we systematically investigated host responses in a cross-sectional cohort of uncomplicated *P. vivax* patients from Manaus, in the Brazilian Amazon region. Our analysis revealed an association between alterations in host homeostasis, including EC activation, damage and haematological disturbances, such as thrombocytopenia, lymphopenia and increased neutrophils turnover, with total parasite biomass but not peripheral parasitaemia. These findings are in line with the emerging paradigm of a clinically relevant parasite reservoir outside of circulation and merit systematic investigations of this reservoir in vivax malaria.
Results

Uncomplicated *P. vivax* patients present with haematological changes

We have conducted a cross-sectional study with uncomplicated *P. vivax* malaria patients seen at FMT-HVD in Manaus, Brazil. We included 79 adult patients (median age of 36 years) with confirmed *P. vivax* infection (smear and PCR positive) and 34 age- and sex-matched uninfected healthy donors (controls) (Table 1). All individuals within the study including controls were from the State of Amazonas, in the Amazon region of Brazil. Blood was collected at enrolment for determination of haematological parameters, peripheral parasitaemia by Giemsa staining of blood smears and PCR to determine genome copy numbers. Preparation of poor platelet plasma (PPP) was done within 15 minutes of sampling. The median peripheral parasitemia was 4,290 infected red blood cells (iRBCs)/μL of blood (25-75 interquartile range 1,860-6,620 parasites/μL) and parasite load of 26,642 copies of 18s RNA/μL (25-75 interquartile range 9,253-522,297). We also measured total parasite biomass independently of peripheral parasitaemia by quantifying levels of *P. vivax* lactate dehydrogenase (PvLDH) in plasma (Table 1).

Analysis of haematological parameters revealed significantly reduced haemoglobin levels and haematocrit across *P. vivax* patients compared to controls, with anaemia in 38% of the patients (Figure 1A). Similarly, leukocyte numbers were significantly decreased (Mean ± SD: 4.36±1.74 x10³/μL vs. 5.72±1.34 x10³/μL, p=0.0004), with 54.5% of the patients presenting with leukopenia (defined as a leukocyte count < 4000 cells/μL). In contrast, neutrophil counts were not significantly different, and only 8.3% of *P. vivax* patients were presenting with neutropenia (neutrophil counts < 1,500 cells/μL) (Figure 1B). Other myeloid cell populations however, such as monocytes, basophils and eosinophils (MXD), were significantly reduced. We also observed a significant reduction in lymphocyte and platelet counts in this cohort (Figure 1C), with 80% presenting with lymphopenia (lymphocyte...
counts < 1,000 cells/μL) and 87% with thrombocytopenia (platelet counts < 150,000 cells/μL), many of them with severely reduced levels (Figure 1C). Alterations in platelet counts were accompanied by the release of mega platelets in the peripheral circulation as a significant increase on mean platelet volume was observed (Figure 1C).

In summary, patients in our cohort presented with a wide range of parasitaemia and uncomplicated clinical signs of *P. vivax* infection at medical consultation. However, significant haematological abnormalities were present in the majority of patients during early onset of disease, in line with previous findings.8,24,28-32.

Unsupervised clustering reveals two *P. vivax* patient subgroups that differ in parasite biomass: Vivax<sup>high</sup> vs Vivax<sup>low</sup>

To determine whether the observed changes were associated with specific host signatures, in particular circulating biomarkers of haematological and endothelial changes, we applied a multiplexed microbead-based immunoassay (Luminex) in a representative subset of 31 *P. vivax* patients and 9 controls, as explained in the Methods section (Figure 1-figure supplement 1). We selected a series of circulating biomarkers associated with haematological changes, including cytokines altering thrombopoiesis (TPO and IL-11), myelopoiesis and lymphopoiesis (TNF-α, IL-1α, IL-1β, IL-6, IL-8, G-CSF)33-35. In addition, we selected markers of endothelial cell (EC) and platelet activation, coagulation (ICAM-1, VCAM-1, E-selectin, P-selectin, Angiopoietin-1 and -2, CD40L, VWF-A2, ADAMTS13, PAI-1, CXCL4, CXCL7) and EC glycocalyx breakdown (Syndecan-1).

We observed significant upregulation of multiple cytokines associated with haematological changes in the *P. vivax* patients compared to control (Table 1). In addition, patient samples exhibited a strong phenotype of increased EC activation, glycocalyx breakdown and coagulation. The high interquartile range in parasitaemia and host signatures
(Table 1) suggested a heterogenous phenotype across the patient population. In order to identify possible stratification of patients into distinct subgroups, we further analysed the clinical data (Figure 1), parasite parameters and Luminex data (Table 1) from the same 31 P. vivax patients and 9 controls as above. After z-score normalization, Principal Component Analysis (PCA) was performed for data dimensionality reduction, considering the large number of variables in our dataset. Next, we ran K-means Clustering (k) followed by bootstrapping (Figure 2A, B, Figure 2-figure supplement 1, Figure 2-figure supplement 2, Figure 2-source data 1) to identify possible subclusters of individuals. This analysis revealed consistent separation of samples into 2 clusters, one of them including all controls (Cluster 1a) and a subset of 14 patient samples (Cluster 1b) and a second one representing the remaining 17 patient samples (Cluster 2) (Figure 2A, B). In order to visualize covariables of the observed patient distribution (PCA) and clustering (K-means), we plotted the correlation (loading score) of each input variable with a principal component (Figure 2C, Figure 2-source data 1). This analysis demonstrated covariation of lymphopenia and thrombocytopenia on one hand, and markers of EC changes, platelet production, activation and parasite parameters (PvLDH and peripheral parasitemia) on the other hand as major contributors to the principal components (Figures 2C). Direct comparison of the 2 patient subgroups revealed significant higher total parasite biomass but not peripheral parasitaemia or parasite load (Figure 3A). In agreement with previous findings, z-score comparison further demonstrated that total parasite biomass was higher than and not correlated with peripheral parasitaemia levels or parasite load, in particular in patients of cluster 2 (Figure 3B, C). In addition, PvLDH was the input parasite variable with the highest loading score (correlation = 0.59) and lowest p-value (0.0000917) in the first PC dimension when compared with peripheral parasitaemia and parasite load (Figure 2C, Figure 2-source data 1). Indeed, using a best-fit classification tree model and a random
forest machine learning model defining k-means clusters as categorical outcome, PvLDH is the best parasite predictor attribute segregating patients into clusters 1b and 2 (Figure 3D, E). After both models were trained in a set of 30 individuals, randomly selected by the training algorithm set, they were tested in the 10 remaining individuals, where all cluster 1a (control) individuals and 80% of P. vivax patients were correctly classified in either cluster 1b or cluster 2. Based on these observations we designate cluster 1a as Control cluster (representing the healthy donors), cluster 1b as Vivax\textsuperscript{low} (representing patients with low P. vivax biomass) and cluster 2 as Vivax\textsuperscript{high} (representing patients with high P. vivax biomass).

Different levels of haematological alterations between Vivax\textsuperscript{high} and Vivax\textsuperscript{low} patients

The three clusters were not significantly different in patient age (median: 33; IQ 25-75: 22-57), gender (80% male; 20% female in each cluster), average days of symptoms when samples were collected, haemoglobin levels, haematocrit or RBC counts, indicating that these parameters are not confounders accounting for the differences observed between the clusters (Figure 4A). However, systematic analysis of haematological parameters between Vivax\textsuperscript{high} and Vivax\textsuperscript{low} patients revealed significant differences. Vivax\textsuperscript{high} patients showed a more intense reduction in platelet counts when compared to Vivax\textsuperscript{low} patients (Vivax\textsuperscript{high} 63,000 ± 6,413 vs Vivax\textsuperscript{low}: 100,700 ± 9,381; p = 0.002), with a higher frequency of patients with severe thrombocytopenia (Vivax\textsuperscript{high} 47% vs Vivax\textsuperscript{low} 8%) (Figure 4B). Although not significant, there was a trend in the reduction of lymphocyte counts in Vivax\textsuperscript{high} patients when compared to Vivax\textsuperscript{low}, with 88% of Vivax\textsuperscript{high} patients presenting lymphopenia versus 64% in Vivax\textsuperscript{low} patients. In addition, we observed a 4-fold increase in the frequency of patients with severe lymphopenia in the Vivax\textsuperscript{high} cluster compared to Vivax\textsuperscript{low} patients (Figure 4B). While there was no change in the number of circulating neutrophils in the different clusters of individuals, mixed cell counts (MXD), a parameter representing
monocytes, basophils and eosinophils numbers, was significantly reduced in Vivax$^{\text{high}}$ patients. As a result, there was a significant enrichment of neutrophils in the leukocyte fraction in the blood of Vivax$^{\text{high}}$ patients as well as a higher Neutrophil:Lymphocyte ratio (NLCR) (Figure 4B).

In parallel to more severe thrombocytopenia in Vivax$^{\text{high}}$ patients, plasma levels of cytokines inducing megakaryocytic differentiation in the bone marrow (BM), thrombopoietin (TPO) and IL-11, were significantly increased in this cluster (Figure 4C, Figure 4-figure supplement 1). In accordance with the pattern of immune cell fractions in the peripheral blood of P. vivax patients, the Vivax$^{\text{high}}$ cluster showed a significant increase in the levels of pro-inflammatory cytokines associated with induction of myeloid-biased haematopoietic stem cell (HSC) differentiation and inhibition of lymphopoiesis in BM (e.g., TNF-α, IL-1α, IL-1β, IL-6, IL-8; Figures 4C, Figure 4-figure supplement 1) $^{33-35}$. In addition, Vivax$^{\text{high}}$ patients had increased circulating levels of G-CSF, a major mediator of HSC-biased myelopoiesis and the neutrophil activation marker, L-Selectin (Figures 4C, Figure 4-figure supplement 1) $^{38-40}$. Together, these Luminex data support the haematological measurements, suggesting that a compensatory response is mounted in the BM to counterbalance the massive decrease of platelets in periphery. Upregulation of cytokines inducing myelopoiesis, while inhibiting lymphopoiesis $^{33-35}$, might also explain the decrease of lymphocyte counts and enrichment of activated neutrophils in the peripheral circulation of P. vivax patients.

Elevated circulating markers of EC activation and damage in Vivax$^{\text{high}}$ compared to Vivax$^{\text{low}}$ patients

Patient clustering indicated that Vivax$^{\text{high}}$ patients have increased levels of EC markers in the plasma compared to Vivax$^{\text{low}}$ patients (Figure 2C). Previous studies indicate that EC activation and damage might contribute to thrombocytopenia and inducing hematopoiesis,
resulting in HSC differentiation directed towards myelopoiesis \cite{24,28,33,35,41-43}. In our cohort, circulating levels of EC adhesion molecules (ICAM-1, VCAM-1, E-selectin and P-selectin) and other EC activation markers and procoagulant molecules (Ang-2, VWF-A2, CD40L and PAI-1) were significantly increased in the plasma of Vivax\textsuperscript{high} patients compared to Vivax\textsuperscript{low} patients and healthy controls (Figures 5A, Figure 5-figure supplement 1A, B). Likewise, Syndecan-1, a marker of EC glycocalyx breakdown (i.e., damage of EC plasma membrane) \cite{44,45}, was significantly increased Vivax\textsuperscript{high} but not in Vivax\textsuperscript{low} patients (Figures 5A, Figure 5-figure supplement 1C).

To independently test whether host factors and/or parasite products present in the plasma of the different patient groups can directly induce changes in ECs, we stimulated primary human umbilical vein endothelial cells (HUVECs) with pools of plasma from either Vivax\textsuperscript{high} patients, Vivax\textsuperscript{low} patients, or healthy controls. These experiments demonstrated that only pooled plasma from Vivax\textsuperscript{high} patients induces significant transcriptional upregulation of EC activation markers ICAM-1, IL-1\alpha, IL-8 along with downregulation of Ang-1, ADAMTS13 and NOS3 (eNOS) in HUVECs (Figure 5B, Figure 5-figure supplement 1D). In contrast, expression of Syndecan-1 and VEGF, two indicators of vascular damage, was not affected by either treatment (Figure 5B, Figure 5-figure supplement 1D).

Similarly, electric cell-substrate impedance sensing (ECIS) assays did not detect differences in functional perturbations in the endothelial cellular monolayer upon incubation with \textit{P. vivax} pooled plasma when compared to control pooled plasma (Figure 5C). In contrast, flow cytometry and immunofluorescence assays performed with stimulated HUVECs revealed increased prevalence and protein expression levels of EC activation markers ICAM-1 and VCAM-1 upon exposure with Vivax\textsuperscript{high} pooled plasma (Figure 5D, Figure 5-figure supplement 1E), in support of qRT-PCR data. These data indicate that local EC activation,
mediated by direct or indirect interactions with parasitized RBCs, can be measured systemically.

**Indirect evidence for parasite-induced changes in deep tissues**

To further investigate the interplay between host biomarkers and associated cellular responses as well as parasite parameters we constructed a network of interactions based on Pearson correlations with absolute correlation coefficient above 0.5 and \( p \)-value < 0.05 (**Figure 6A**). In addition, we also performed hierarchical clustering on matrices of Pearson correlations (\( p \)-value < 0.01) with selected modules of parasite and host parameters (**Figure 6B**). Data from Vivax\(^{\text{low}}\) and Vivax\(^{\text{high}}\) patient subgroups were combined for this analysis as they similarly contribute to the associations we found so far (**Figure 6-figure supplement 1**).

Similar to a previous study with *P. vivax* patients and healthy donors from an endemic area in Brazil \(^{46}\), our analysis revealed a dense network of interactions with homogenous and centralized topology among the biomarkers in healthy donors (**Figure 6A, Supplementary File 1**). The network topology is drastically altered in symptomatic *P. vivax* patients, largely due to the introduction of parasite parameters in the patient graph (**Figure 6A, Supplementary File 1**). The network analysis revealed a decentralized topology, lower complexity and connectivity between the edges with data from *P. vivax* patients compared to the highly dense, homogenous and centralized network graph of healthy donors (91 edges vs 166 edges, respectively). Of note, the network pattern described in our study is similar to protein-protein associated networks described previously in *P. vivax* malaria and in other clinical contexts\(^{46,47}\). Interestingly, due to its decentralized and heterogenous patterns of interactions, the network graph of *P. vivax* patients is separated into three modules of strong interactions, with closely related biological functions. Module 1 is formed by markers of EC activation and damage, together with lymphocyte, platelet and neutrophil counts in addition
to the megakaryocyte differentiation inducing cytokines (TPO and IL-11) (Figure 6A). In support of the role of EC activation and damage in the haematological changes observed in this cohort, hierarchical clustering revealed a positive correlation between adhesion molecules VCAM-1 and E-selectin and EC glycocalyx breakdown (Syndecan-1) (Figure 6B). In addition, VCAM-1, E-selectin, Ang-2 and VWF-A2, and Syndecan-1 are negatively correlated with platelet and lymphocyte counts, while ICAM-1 is positively correlated with neutrophil counts (Figures 6A, B). Module 2 is formed by proinflammatory cytokines with myelopoiesis inducing effects and molecules associated with platelet activation and coagulation cascades (Figures 6A, B). Interestingly, EC activation markers and Syndecan-1 (EC damage) from module 1 also display positive correlations with myelopoiesis-inducing cytokines from module 2 (Figure 6B). Finally, module 3 is formed by Ang-2 and the proinflammatory cytokine IL-1β negatively associated with haemoglobin, haematocrit and RBC numbers (anemia markers) (Figures 6A, B). Most notably, PvLDH connects the two main functional modules 1 and 2 (Figures 6A, B). Accordingly with Figures 2 and 6A,B, the biological significance of total parasite biomass, but not peripheral parasitaemia or parasite load, in affecting host response is also corroborated by the high significant and positive associations of PvLDH with multiple host parameters, including Syndecan-1 (EC damage), VCAM-1, VWF (EC activation and platelet pooling), and IL-6, IL-8, TNF-α (inflammation and myelopoiesis inducing cytokines) (Figure 6C). Meanwhile, platelet, lymphocyte and neutrophil counts are negatively correlated with high significance (p-value < 0.0001) with total parasite biomass, but not with peripheral parasitaemia or parasite load (Figure 6C). The association between endothelial activation, Syndecan-1 and parasite biomass (PvLDH) indicates a positive feedback loop between glycocalyx breakdown, activation of endothelial receptors such as ICAM-1 and VCAM-1 and parasite accumulation in deep tissues⁹,12. Similar to Figure 2E, application of a best-fit classification tree model demonstrated that Syndecan-1,
IL-6 and platelet counts are the most dominant predictor attributes capable to classify *P. vivax* patients based on total parasite biomass levels (Figure 6D). Using this model all *P. vivax* patients were correctly classified in either low (Vivax\textsubscript{low}) or high (Vivax\textsubscript{high}) total parasite biomass (PvLDH). In turn, PvLDH is a relevant predictor attribute (high information gain) in predicting thrombocytopenia severity and it is associated with increased severity of thrombocytopenia and lymphopenia in our cohort (Figure 6-figure supplement 2, Figure 6-figure supplement 3). Together, these data further support the hypothesis that a parasite population outside of circulation, as represented by total parasite biomass, is driving the host response including EC activation and damage as well as haematological disturbances (i.e., lymphopenia, thrombocytopenia and anaemia) in *P. vivax* patients (Figure 6-figure supplement 2, Figure 6-figure supplement 3).

**Discussion**

In this study, we performed a comprehensive analysis of host and parasite signatures detected in the plasma of a cross-sectional cohort of uncomplicated *P. vivax* malaria. Initial analysis of a series of circulating host biomarkers revealed significant levels of thrombocytopenia, lymphopenia and anaemia, as well as EC activation and damage across *P. vivax* patients compared to healthy controls. Deconvolution of heterogeneity across patients revealed two patient subgroups (Vivax\textsuperscript{high} and Vivax\textsuperscript{low}) characterized by differences in total parasite biomass (based on circulating PvLDH levels) but not peripheral parasitaemia (based on blood smears). We observed a significant correlation between total parasite biomass (but not peripheral parasitaemia) and systemic levels of markers of EC activation and damage and hematopoietic perturbations. In addition, by applying a supervised machine learning tree-structured model, we were able to associate EC damage and thrombocytopenia with parasite biomass. In agreement with a previous study\textsuperscript{8,37} our observations further suggest that total biomass.
parasite biomass as measured by PvLDH is a better predictor of *P. vivax* host responses and pathogenesis than peripheral parasitaemia. Furthermore, these findings support the emerging paradigm of a major *P. vivax* parasite reservoir outside of circulation, in particular in the haematopoietic niche of bone marrow and spleen. 

The existence of a significant *P. vivax* reservoir outside of circulation was first predicted by disproportionately high PvLDH levels in peripheral circulation compared to parasitemia by blood smear (in particular in patients with complicated outcomes), and by modelling using experimental *P. cynomolgy* infections in non-human primates. Recent studies provide direct evidence that bone marrow and spleen represent the major reservoir of parasite biomass in *P. vivax* infection. PvLDH is produced by viable or recently killed parasites and hence considered a proxy for ongoing rather than past infection. PvLDH antigen capture ELISA established a direct relationship between pLDH levels and *P. vivax* parasitemia in *ex vivo* experiments, demonstrating that pLDH reflects total *P. vivax* parasite biomass. Our study further explores the relevance of PvLDH as a prognostic marker of host perturbations and disease severity, with a particular focus on markers of changes in the hematopoietic niches of bone marrow and spleen. A major observation in the network graph of *P. vivax* patients is the central position of the total parasite biomass marker PvLDH, due to its equally strong interactions with the two main functional modules 1 and 2. Given that the hematopoietic niches of the BM and the spleen are the major reservoir of parasite biomass, interactions of PvLDH with these two main modules indicate an interplay between parasite infection in these niches and endothelial activation/damage as well as the proinflammatory response that results in myeloid-biased differentiation, thrombocytopenia and lymphopenia. Furthermore, the highly significant and positive associations between endothelial activation, Syndecan-1 and parasite biomass (PvLDH) indicates a positive feedback loop between glycocalyx breakdown, activation of endothelial receptors such as...
ICAM-1 and VCAM-1 and parasite accumulation in deep tissues. Vivax<sup>High</sup> patients show higher plasma levels of all these markers. Consistent with previous reports, we propose that elevated EC activation and glycocalyx damage increases the exposure of adhesion molecules, which in turn favours endothelial cytoadherence of *P. vivax*-infected RBCs, in particular in the splenic red pulp cords and in the BM. Accordingly, application of a best-fit classification tree model identifies Syndecan-1 as a putative host biomarker (EC glycocalyx breakdown marker) predicting total parasite biomass in *P. vivax* patients. We hypothesise that elevated endothelial activation and damage in Vivax<sup>High</sup> patients results in increased cytoadherence of *P. vivax* iRBCs and hence accumulation and growth in deep tissues, thus reducing the fraction of the parasite biomass in circulation.

In contrast to *P. falciparum*-infected individuals, a wide range of complicated clinical syndromes occurs in *P. vivax* patients even at low or subpatent parasitemia – thus indicating that peripheral parasitemia is a poor predictor of clinical outcomes. Two lines of evidence support our conclusion that severity of infection is dependent on parasite biomass instead. First, the discrepancy between PvLDH levels and peripheral parasitaemia determined by blood smears is more evident in *P. vivax*-infected patients with complicated outcomes: the ratio of plasma pLDH to peripheral parasitaemia is 6-fold higher than in non-severe patients. The same comparison between severe and non-severe *P. falciparum* patients reveals only a 1.4-fold difference. Second, although thrombocytopenia and lymphopenia are not included in the World Health Organization (WHO) criteria for defining severe malaria, it has been associated with severe manifestations and the need for blood and platelet transfusions in severe vivax malaria. This points out their clinical relevance in malaria diagnosis and treatment, suggesting that these haematological complications could be explored as markers of severity for this species. Both severe thrombocytopenia and lymphopenia were more frequent in patients in the Cluster 2 (Vivax<sup>high</sup>) in our study. By integrating these
clinical perturbations with host biomarker measurements and parasite parameters, we demonstrated the high attribute value of total parasite biomass in predicting the severity of thrombocytopenia and lymphopenia and highly significant correlations with endothelial activation, glycocalyx breakdown and other markers of inflammation.

Thrombocytopenia, lymphopenia and anaemia are the most frequent \textit{P. vivax} and \textit{P. falciparum} associated haematological complications\textsuperscript{24-27}. In our cohort, 34\%, 85\% and 87\% of patients exhibited anaemia, lymphopenia and thrombocytopenia, respectively. Various mechanisms have been proposed to explain the damage or excessive removal of platelets in \textit{P. vivax} infection, including oxidative stress, platelet phagocytosis, IgG binding to platelet-bound malaria antigens, spleen pooling, or increased circulating nucleic acids levels\textsuperscript{24,25,55,57}. EC activation and damage also play a role in intravascular platelet agglutination and increased platelet clearance from the circulation\textsuperscript{31,32}. Our data also demonstrate that thrombocytopenia is associated with an increase in IL-1, IL-6, IL-8, IL-10 and TNF-\(\alpha\). We also observed elevated levels of cytokines inducing megakaryocyte differentiation, TPO and IL-11, suggesting that a compensatory response is mounted in the BM to counterbalance the massive decrease of platelets in the periphery. In contrast, the relatively large drop in peripheral lymphocyte numbers we observed in the \textit{P. vivax} patients is likely non-specific effect, e.g. pooling in the enlarged spleen rather than a response by \textit{Plasmodium}-specific lymphocytes\textsuperscript{58}. Corroborating the potential role of total parasite biomass, rather than peripheral parasitaemia, in haematological disturbances (i.e., lymphopenia, thrombocytopenia and anaemia), Figures S7 and S8 show that total parasite biomass increases accordingly with thrombocytopenia and lymphopenia severity. Patients with severe thrombocytopenia also show more severe leukopenia, lymphopenia and mega platelets (higher MPV). In addition, plasma levels of cytokines, such as TNF-\(\alpha\), IL1-\(\beta\), IL-8, IL-10; EC activation/damage markers, VCAM-1, E-selectin, VWF-A2, Ang-2, Ang-2:Ang1 ratio; Syndecan-1;
thrombopoiesis-inducing cytokines, TPO and IL-11; platelet activation marker, CD40L and neutrophil activation marker, L-selectin, follow the increase in thrombocytopenia severity (Figure 6-figure supplement 2). A similar pattern is observed when stratifying patients based on lymphopenia severity (Figure 6-figure supplement 3). Interestingly, a tree-structured model demonstrated that PvLDH, along with VCAM-1 and Syndecan-1, is a relevant predictor attribute (high information gain) in predicting thrombocytopenia severity in our cohort (Figure 6-figure supplement 2).

Our data support previous studies suggesting a role for endothelial cell (EC) activation and damage in increased leukocyte adhesion, intravascular platelet agglutination with increased platelet clearance from the circulation and skewing of haematopoiesis toward the myeloid lineage (likely at the expense of lymphopoiesis) in the BM. P. vivax elicits a stronger inflammatory response and more pronounced endothelial activation when compared with other Plasmodium infections with similar or higher peripheral parasitaemia, however the role of EC activation in P. vivax pathogenesis is not yet understood. Damage of the EC plasma membrane, as represented by glycocalyx breakdown, has been associated with poor prognostic outcome in P. falciparum, but there is no data available for P. vivax. In our cohort soluble EC activation biomarkers (e.g., ICAM-1, VCAM-1, E-selectin, Ang-2, CD40L, vWF-A2) and the EC damage product, Syndecan-1, are positively correlated with thrombocytopenia, lymphopenia, anaemia and neutrophil enrichment in the peripheral blood. In addition, these biomarkers are positively correlated with increased circulating levels of cytokines inducing megakaryocyte-differentiation (e.g., IL-11 and TPO) and with cytokines inducing myeloid-biased HSC differentiation (e.g., TNF-α, IL1-α, IL6, IL-8 and G-CSF), suggesting both direct and indirect links between EC activation and damage and haematological perturbations. Total parasite biomass inducing EC activation might act synergistically with inflammatory changes potentially leading to splenic
platelet pooling and platelet clumping in the vasculature without DIC\textsuperscript{19,45,59}. Likewise, increased activation-induced cell death (AICD) in T cells, splenic T-cell accumulation\textsuperscript{58} or decreased lymphopoiesis due to myeloid-biased HSC differentiation induced by inflammatory cytokines and EC activation in the BM\textsuperscript{33,34,37} might explain the severe lymphopenia and neutrophilia in vivax patients. Together such mechanisms could explain the link between parasite biomass and EC activation/damage with haematological changes observed in vivax patients that might contribute to pathogenesis and disease severity.

In a second series of experiments, we performed \textit{ex vivo} stimulation of HUVECs with the plasma of the \textit{P. vivax} cohort demonstrating that the mixture of parasite and host factors can directly induce EC activation in absence of parasitized RBCs. Of note, functional differences between HUVECs and adult vascular endothelium, including lack of ABO blood group antigen expression, have been reported\textsuperscript{60,61}. Hence EC stimulation with patient plasma may be further evaluated using primary vascular ECs.

ECs are capable of responding to pathogens by sensing pathogen-associated molecular patterns (PAMPs) through pattern-recognition receptors (PRRs), which might play a key role in inducing EC activation when detecting \textit{P. vivax} molecules enriched in the tissues where the parasite accumulates. ECs also express specific cytokine/chemokine receptors to detect proinflammatory signals released systemically or locally by activated immune cells in response to infection\textsuperscript{62,63}. As result, EC activation induces exocytosis of secretory granules known as Weibel–Palade bodies that leads to the release of Ang-2 and VWF, as well as transcriptional programs that activate expression of adhesion molecules such as ICAM-1, VCAM-1, E-selectin, and secreted cytokines and chemokines\textsuperscript{29,62,63}. However, EC pathophysiology is complex, and changes represent a heterogenous spectrum ranging from simple perturbation to activation and even endothelial cell damage\textsuperscript{28}. Our Luminex data clearly confirm such heterogeneity in the spectrum of EC changes due to \textit{P. vivax} infection,
with systemic increase of markers of EC activation and damage only detected in Vivax$^{\text{high}}$
patients. The ex vivo data show that increased systemic host proinflammatory factors and/or
parasite products can alter EC properties, including activation of adhesion molecules and
proinflammatory cytokines and downregulation of ADAMTS13. In contrast, vascular
integrity was not affected. These data indicate that systemic inflammatory responses in P.
vivax patients can lead to local EC activation but not vascular damage, central events in
malaria pathogenesis. It is likely that other circulating factors that we have not directly
measured in our study are also contributing to EC activation and vascular permeability. In
particular, extracellular vesicles (EV) originating from ECs, platelets, and RBCs are present
during malaria infection and are known to modulate the host immune response to the
parasite$^{52,64,65}$. In *P. falciparum*, infected RBCs release EVs containing immunogenic
parasite antigens, that activate macrophages, induce neutrophil migration and alter
endothelial barrier function$^{64,65}$. In *P. vivax*, plasma-derived EVs from iRBCs are taken up by
human spleen fibroblasts (hSFs). This event signals NF-kB translocation and upregulation of
ICAM-1 expression, facilitating cytoadherence of *P. vivax*-infected reticulocytes$^{52}$.

Although our study lacks longitudinal information, the findings might have clinical
implications during and after treatment of vivax malaria. Several case reports demonstrate
progressive clinical deterioration after commencement of treatment in *P. vivax* patients,
associated with parasite killing that result in haemolysis of iRBCs and intravascular
inflammation and oedema in response to the products released from these cells$^{66-69}$. Patients
presenting with a strong host response during acute infection might therefore be at increased
risk of deteriorating and developing severe symptoms after commencement of treatment
(Figure 5-figure supplement 2). Thus, identification of unique biological signatures in *P.
vivax* patients might help to build rational approaches to the diagnosis, prognosis and
individualized treatment to modulate the host response to vivax malaria.
Altogether, our data indicate that changes in clinical parameters and biomarkers detected in the plasma of *P. vivax* patients are the result of both, systemic host responses and local infection in tissue reservoirs such as BM and spleen. Our analysis shows that measuring a combination of host parameters (e.g., Syndecan-1, IL-6, platelet levels) and total parasite biomass (PvLDH) could predict the extent of parasite infection outside of circulation. Our data also instigate future investigations of systemic signatures with parallel analysis focused on tissue responses, in particular in reservoirs such as the haematopoietic niche of BM and spleen, which has great potential to advance better diagnosis and treatment of *P. vivax*.

**Materials and Methods**

**Key Resources Table**

<table>
<thead>
<tr>
<th>Reagent type (species) or resource</th>
<th>Designation</th>
<th>Source or reference</th>
<th>Identifiers</th>
<th>Additional information</th>
</tr>
</thead>
<tbody>
<tr>
<td>biological sample (<em>Homo sapiens</em>)</td>
<td>Human umbilical vein endothelial cells (HUVEC)</td>
<td>Hematology Center - University of Campinas, Campinas, SP, Brazil</td>
<td></td>
<td>Primary cells isolated from the umbilical vein.</td>
</tr>
<tr>
<td>antibody</td>
<td>ICAM-1 (mouse monoclonal)</td>
<td>Abcam</td>
<td>Cat. #ab2213; RRID: AB_302892</td>
<td>Antibody clone MEM-111 IF (1:100)</td>
</tr>
<tr>
<td>antibody</td>
<td>VCAM-1 (mouse monoclonal)</td>
<td>Abcam</td>
<td>Cat. #ab212937; RRID: AB_2892824</td>
<td>Antibody clone 1.4C3 IF (1:500) FC (1:100)</td>
</tr>
<tr>
<td>antibody</td>
<td>IgG1 isotype control (mouse monoclonal)</td>
<td>Dako</td>
<td>Cat. #X0931; RRID: AB_2892825</td>
<td>IF (1:10)</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------------------------</td>
<td>---------------------------</td>
<td>--------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>antibody</td>
<td>ICAM-1 (mouse monoclonal)</td>
<td>Biolegend</td>
<td>Cat. #322714; RRID: AB_535986</td>
<td>Antibody clone HCD54 FC (1:100)</td>
</tr>
<tr>
<td>antibody</td>
<td>pLDH Vivax-specific (mouse monoclonal)</td>
<td>Vista Diagnostics International LLC, WA, USA</td>
<td>Cat. #3h8; RRID: AB_2892826</td>
<td>Antibody clone 3h8 ELISA (1μg/mL)</td>
</tr>
<tr>
<td>antibody</td>
<td>pLDH detection antibody (goat monoclonal)</td>
<td>Vista Diagnostics International LLC, WA, USA</td>
<td>Cat. #6c9; RRID: AB_2892827</td>
<td>Antibody clone 6c9 ELISA (1:4000)</td>
</tr>
<tr>
<td>sequence-based reagent</td>
<td>qRT-PCR Oligonucleotides</td>
<td>This study</td>
<td>See Supplementary file 2</td>
<td></td>
</tr>
<tr>
<td>commercial assay or kit</td>
<td>Customized multiplex suspension detection system</td>
<td>R&amp;D Systems</td>
<td></td>
<td></td>
</tr>
<tr>
<td>commercial assay or kit</td>
<td>Accutase® Cell Detachment Solution</td>
<td>Biolegend</td>
<td>Cat. #423201</td>
<td></td>
</tr>
<tr>
<td>chemical compound, drug</td>
<td>Fixable Viability Dye eFluor™ 506</td>
<td>ThermoFisher</td>
<td>Cat. # 65-0866-14</td>
<td></td>
</tr>
<tr>
<td>software, algorithm</td>
<td>FlowJo software (v10)</td>
<td>Ashland, OR, USA</td>
<td><a href="https://www.flowjo.com">https://www.flowjo.com</a></td>
<td></td>
</tr>
</tbody>
</table>
Peripheral blood and plasma samples was collected from 79 patients infected with *P. vivax*, as diagnosed by light microscopy, seen at FMT-HVD and 34 healthy donors (controls). Patients and healthy donors were age and sex-matched, with a frequency of 30% female and 70% male individuals in both groups. All individuals within the study were from a local *vivax* malaria epidemic area in the Amazon region of Brazil. All patients included were outpatients that did not meet World Health Organization (WHO) criteria for severe malaria. Diagnosis was further confirmed by quantitative PCR (qPCR) for both *P. vivax* and *P. falciparum*, using previously published nucleotide sequences. Excluding other coinfections could have been of interest. However, the differential diagnosis for an acute febrile illness is very broad and it would be impractical to track all other possible diseases. In addition, the patients included in the present work had mild disease, and therefore were discharged from hospital after a positive malaria diagnosis. No further investigation on other infections was done. The main coinfection to be considered for an acute febrile illness with no localizing signs in our context is Dengue Fever. Although
Dengue coinfection in our cohort is possible, the incidence at the Hospital is only 2.8% \( (P. \text{ vivax}/\text{Dengue coinfection}) \). Thus, it is unlikely that such a coinfection would have a major impact on our findings. Exclusion criteria were i) under 18 years of age, ii) pregnancy, iii) use of antimalarials, iv) chronic disease, v) medication known to interfere with platelet count/function and vi) smoking.

Anaemia is defined as haemoglobin < 12.5 g/dL; haematocrit < 37%; RBCs counts < 4.45x10^6/\( \mu \)L. Thrombocytopenia is defined as a decrease in platelet counts to below 150,000/\( \mu \)L. Based on platelet levels, patients were grouped into (i) non-thrombocytopenia (NT: platelet counts > 150,000/\( \mu \)L), (ii) mild thrombocytopenia (MT: platelet counts 100,000-150,000/\( \mu \)L), (iii) moderate thrombocytopenia (MDT: platelet counts 50,000-100,000/\( \mu \)L), and (iv) severe thrombocytopenia (ST: platelet counts < 50,000/\( \mu \)L). Lymphopenia was defined as a lymphocyte count of less than 1,000 cells/\( \mu \)L. Neutropenia was defined as a neutrophil count of less than 1,500 cells/\( \mu \)L and neutrophilia as a neutrophil count of more than 7,000 cells/\( \mu \)L.

Preparation of poor platelet plasma

After signing the informed consent, 20 mL of venous blood were drawn by venepuncture in a syringe with 15% acid citrate dextrose as anticoagulant to minimize \textit{in vitro} platelet activation. Complete blood counts were done within 15 minutes of blood sampling with a Sysmex KX21N counter. Whole blood was centrifuged at 180 g for 18 minutes at room temperature, without brake for gradient formation, to obtain the platelet rich plasma (PRP). PRP was centrifuged at 100 g for 10 minutes for removal of residual leukocytes, and subsequently centrifuged at 800 g for 20 minutes to obtain the platelet pellet.
E1 at 300 nM was used to minimize platelet aggregation. The supernatant was centrifuged at 1000 g for 10 minutes to obtain platelet poor plasma (PPP).

**Multiplex bead array assay**

The biomarkers were analysed in thawed plasma with a customized multiplex suspension detection system (R&D Systems) for quantification of the following biomarkers:

a) pro-inflammatory and myelopoiesis-inducing cytokines: TNF-α, IL-1α, IL-1β, IL-6, IL-8, G-CSF

b) endothelial cell (EC) activation and coagulation markers: ICAM-1, VCAM-1, E-selectin, P-selectin, Angiopoietin-1 (Ang-1), Angiopoietin-2 (Ang-2), von Willebrand Factor (vWF-A2), CD40L, PAI-1, ADAMTS13

c) glycocalyx breakdown and EC damage marker: Syndecan-1

d) platelet activation markers: CXCL4 and CXCL7

e) megakaryocyte differentiation-inducing cytokines: thrombopoietin (TPO) and IL-11; and other proteins such as IL-10, L-selectin and SCF.

A representative set of 31 *P. vivax* patients were selected for the multiplex assay (Figure 1-figure supplement 1). These patients were selected to encompass the wide range of peripheral parasitemia present in the cohort (260 to 25,150 infected RBCs/μL) and to match age, gender and other haematological parameters to those that were not selected. Nine healthy donors matched for age and sex were also selected.

**Plasmodium vivax lactate dehydrogenase (PvLDH) ELISA**

To measure PvLDH in patient plasma samples, ELISA was performed using a matching pair of capture and detection antibodies (Vista Diagnostics International LLC, Greenbank, WA,
USA). Briefly, 96-well microtiter plate was coated with monoclonal anti-pLDH Vivax-
specific (clone 3H8 - Vista Diagnostics International LLC, WA, USA; RRID: AB_2892826)
at a concentration of 1μg/mL in PBS (pH 7.4) and incubated overnight at 4°C. The plate was
washed and incubated with blocking buffer (reagent diluent) at room temperature for 1h.
After washing, samples were added and incubated for 2h. Next, plates were washed and
biotinylated anti-PvLDH detection antibody (clone 6c9 - Vista Diagnostics International
LLC, WA, USA; RRID: AB_2892827), diluted 1:4000 in blocking buffer, was incubated for
2h at room temperature, followed by streptavidin-HRP incubation for 20 minutes at room
temperature. Plates were washed and incubated for 20 min with substrate solution. Optical
density was determined at 450 nm. Cut-off of positivity was defined by correcting
absorbance values generated in the plasma samples from healthy donors (controls) by blank
values (plate controls), with both values being in the same range. Absorbance values higher
than controls were considered positive. In parallel, we used schizont extracts to perform
standard curves and lower absorbance values were in the range of O.D = 0.03-0.04. All
positive patient samples gave O.D. values equal or higher than 0.05.

Principal component analysis and K-means hierarchical clustering

Haematological parameters (haemoglobin levels, haematocrit, differential blood cell counts),
parasite parameters (peripheral parasitaemia by blood smear, parasite load by qPCR and
parasite biomass PvLDH ELISA) and Luminex data (24 biomarkers) from the selected 9
healthy donors and 31 P. vivax patients were normalized to avoid variable-specific bias and
z-score values were determined. Since the host response is complex and multidimensional
(one dimension per Luminex biomarker), we applied dimension reduction and clustering for
ease of downstream analysis. For this, all variables were used as input for principal
component analysis (PCA) to reduce the dimensionality of data using the PCA function in the
FactoMineR package in R. For visualization of PCA results ggplot2, factoextra and corrplot packages were used. For each principal component (PC), we determined which variables are better represented and the contribution (correlation or loading score) of each variable for each (PC). Investigation of eigenvalues and the percentage of explained variances retained by the PCs demonstrated that the first 10 PCs accounted for the variance of the data (Figure 2-figure supplement 1). However, variables were well represented by the first 2 PCs (Dim 1 and Dim 2), which were therefore retained for further analysis. In parallel, we performed K-means Clustering (k) followed by bootstrapping, which produced the most stable clusters with k = 2 (Cluster 1 = 21 individuals; Cluster 2 = 18 individuals), which seemed to be the most consistent with the data (Figure 2A, Figure 2-figure supplement 1, Figure 2-figure supplement 2, Figure 2-figure supplement 2-source data1). Figure 2-figure supplement 2-source data1 contains the numerical data representing cluster stability via bootstrapping. The metrics of interest is jaccard_index which measures the cluster similarity across bootstrap samples. Similar to the above, k=2 gives stable clusters for all configurations (jaccard_index 0.9 and 0.86). Using Monte Carlo Reference-based Consensus Clustering (M3C) analysis (M3C function in the M3C package in R) indicated that k=2 is the optimal number of clusters when using K-means clustering (Figure 2-figure supplement 2C, D), but when determining spectral clusters, different from elliptical k-means clusters, k=3 gives the best number of clusters (Figure 2-figure supplement 2E-G).

Correlation plots and heatmap visualization
Heatmaps were created to visualize variable values using R function Complex Heatmap. They represent z-scores using row scaling obtained by centring represented variables with the scale function, followed by column clustering using average cluster method and Euclidean distance metric in R. The same software was used to determine pairwise Pearson’s
correlation coefficients between variables by running the function `cor` in the `ggcorrplot` package and visualized as a correlogram using R function `corrplot` in the `Hmisc` package displaying positive correlations in red and negative correlations in blue using $p \leq 0.01$ as a cut-off.

Recursive partitioning decision-tree classification and machine learning models

We used recursive partitioning decision-tree classification models to evaluate dominant signatures (attributes) predicting a specific outcome. For decision-tree construction we applied the C4.5 algorithm, using the `RWeka`, `caret` (Classification and Regression Training) and `e1071` packages or the `rpart` package in R. First, the library `caret` is used to create a 10-fold training set to train the model. Then, the algorithm implements decision trees (using the J48 method, which is an open-source Java implementation of the C4.5 algorithm) starting with all instances in the same group, then repeatedly divides the data based on attributes until each item is classified. The attribute on which to divide is selected by information gain, a statistical technique for determining which attribute split will most cleanly divide the data. To avoid overfitting, sometimes the tree is pruned back. In parallel, the algorithm performs k-fold cross validation to measure the performance of a given predictive model and indicates which one has the higher accuracy. Here, we used $k=10$ to yield test error rate estimates that suffer neither from excessively high bias nor from very high variance. In parallel, features with mean decrease accuracy larger than 6 were used for random forest. In the random forest analysis, a thousand trees were built using R package `randomForest` (version 4.6.14). The normalized additive predicting probability was computed as the final predicting probability. Those selected important features were used for the random forest analysis on the test cohort for model validation.
Stimulation of HUVEC with patients’ plasma pools

After standardization procedures, primary human umbilical vein endothelial cells (HUVEC) were stimulated or not (mock control) in culture media for 6h – to evaluate mRNA expression – or for 18h – to evaluate protein expression - with complete EGM-2 medium (Lonza) containing 30% (v/v) plasma pools generated from the three subgroups - Healthy control, Vivax\textsuperscript{low} and Vivax\textsuperscript{high} - and 3U/L heparin.

Real-Time Quantitative RT-PCR

After 6h stimulation, total RNA was isolated from the cell lysate using the miRVana miRNA Extraction kit (Ambion) according to the instructions of the manufacturer. cDNA was synthesized with TaqMan Reverse Transcriptase (Applied Biosystems, Foster City, CA) and mRNA expression of genes were determined by qRT-PCR. Real-time qRT-PCR was performed on an ABI-Prism 7000 PCR cycler (Applied Biosystems) or on the CFX96 Real-Time PCR Detection System (Bio-Rad). Cycling parameters were 95°C for 1min and then 35 cycles of 95°C (15s) and 60°C (1min), followed by a melting curve analysis. The median cycle threshold (C\textsubscript{T}) value and \(2^{-\Delta\Delta C_T}\) method were used for relative quantification analysis, and all C\textsubscript{T} values were normalized to the GAPDH mRNA expression level. Results expressed as means and SEM of biologic replicates are shown. The mock sample (HUVECs incubated with culture media only) was used as reference. The oligonucleotides used are described in the Supplementary File 2.

Endothelial cell Flow Cytometry (FC) and Immunofluorescence Analysis (IFA)

For immunofluorescence analysis, cells were grown in 8-well chambered coverslips (IBIDI) until confluence. After 18h stimulation with plasma pools, cells were washed with PBS and fixed/permeabilized with ice cold 100% methanol for 5 minutes at -20°C. In brief,
cells were incubated with 10% goat serum (ThermoFisher) to avoid secondary antibody
nonspecific binding for 1h at room temperature and then incubated with specific primary
antibodies to human ICAM-1 (mouse monoclonal clone MEM-111; Abcam; cat. number
ab2213; RRID: AB_302892) (used at a dilution of 1:100 in 10% goat serum); VCAM-1
(mouse monoclonal clone 1.4C3; Abcam; cat. number ab212937; RRID: AB_2892824) (used
at a dilution of 1:500 in 10% goat serum) and mouse IgG1 isotype control (Dako; cat.
Number X0931; RRID: AB_2892825) (used at a dilution of 1:10 in 10% goat serum)
overnight at 4°C. After washing, wells were overlaid for 1h with AF488-conjugated
secondary antibody (used at a dilution of 1:500 in 10% goat serum) and Hoechst (diluted at
1:2000) at room temperature. For controls, primary antibodies were omitted from the staining
procedure and were negative for any reactivity. The chambers were placed at 4°C until use
for immunofluorescence assay (IFA). Percentage of positive cells and expression profiles for
ICAM-1 and VCAM-1 were then determined using ImageJ software (NIH, Bethesda, MD,
USA).

In flow cytometry, after 18h stimulation with 30% plasma pools, cells were washed
2x with DPBS and treated with Accutase® Cell Detachment Solution (Biolegend) at room
temperature for up to 3 minutes, or until the cells are detached. Cell count and viability with
trypan blue dye were determined and cells were resuspended in ice cold DPBS without
calcium/magnesium, 0.5mM EDTA, 10% foetal bovine serum (FBS) (GIBCO). Cells were
incubated with FcBlock (BD Biosciences, San Jose, CA, USA), followed by incubation with
unconjugated anti-VCAM (mouse monoclonal clone 1.4C3; Abcam; cat. number ab212937)
or AF488-conjugated anti-ICAM-1 (mouse monoclonal clone HCD54; Biolegend; cat.
number 322714; RRID:AB_535986) or unconjugated mouse IgG1 isotype control (Dako; cat.
Number X0931) for 1h at 4°C. Cells were then washed and incubated for 1h at 4°C with
secondary antibody AF488-conjugated anti-mouse IgG (ThermoFisher). Finally, cells were
incubated with Fixable Viability Dye eFluor™ 506 (ThermoFisher) in DPBS without calcium/magnesium, 0.5mM EDTA for 30 min at 4°C. Cells were washed and resuspended in buffer and acquired using a BD FACS Celesta cytometer (100,000 events/sample). Percentage of positive cells and expression profiles for ICAM-1 and VCAM-1 were then determined by the mean fluorescence intensity using FlowJo software (v10; Ashland, OR, USA).

Ex vivo Evaluation of Endothelial cell Monolayer Function

Endothelial cell monolayer function was measured using ECIS, an electric cell-substrate impedance sensing system (ECIS Z0, Applied BioPhysics, Troy, NY), as previously described. The system then applies weak alternating currents through the electrode array and continuously measures the ability of the cell monolayer to impede the movement of electrons between adjacent endothelial cells (resistance). Briefly, cells were seeded at 2.5 × 10^5 cells/well on fibronectin-coated (10 µg/ml) eight-well arrays (8WE10, Applied BioPhysics, Troy, NY) containing interdigitated gold electrodes. Endothelial cells were seeded 48 h before experiments and the resistance started to be recorded after 48 h. Only wells with resistance > 1,500 ohms and stable impedance/resistance readings were used. Before stimulation, resistance was continuously monitored for 2 h, to confirm monolayer stability represented by a plateau in the resistance curve. Stimuli (20% v/v pooled plasma in complete medium) was then added to wells under continuous impedance/resistance monitoring for 12 h. A baseline resistance value was recorded immediately prior to the addition of each stimuli and results are expressed as a ratio from baseline resistance (normalized resistance).

Network analysis

The values of each circulating factor measured in the plasma samples, as well as haematological parameters and parasite biomass in healthy donors and P. vivax malaria
patients were input in the RStudio software (Version 1.4.1106, 2021) to determine pair-wise Pearson’s correlation coefficients to generate correlation networks and the \( p \)-value to test for non-correlation was evaluated using \( p \leq 0.05 \) as a cut-off. In order to analyse the structure of the networks, edges list was generated in R using the functions `melt` (reshape2 package), `graph_from_edgelist` (igraph package). Graphs were customized in the Cytoscape software (v 3.8.1) using the force-directed layout, which simulates a system of forces, determined by the correlation strength. In the in the equilibrium state edges tend to have uniform length, and nodes that are not connected by an edge tend to be drawn further apart. **Network topology and module analysis were performed using the NetworkAnalyzer, jActiveModules and MCODE plugins in Cytoscape**. **Supplementary File1 shows the results for all parameters quantified in the comparative network topology analysis between the graphs for healthy donors and *P. vivax* patients.**

**Statistical Analysis**

Fisher’s exact test was used for categorical data. Data normality was checked by the Shapiro-Wilk test. Student’s t-test was used to compare means between groups with normally distributed data, and data sets with non-normal distributions were compared using the Mann–Whitney test. All tests were performed two-sided, using a nominal significance threshold of \( p < 0.05 \) unless otherwise specified. When appropriate to adjust for multiple hypothesis testing, Tukey’s or Bonferroni corrected multiple comparisons test significance at the \( p \)-value < 0.05 threshold was performed unless otherwise specified. Data are presented as scatter plots with median and 25%-75% interquartile range, box plots showing minimum to maximum range or means and SEM, unless otherwise stated. Analyses were performed and the graphs generated in GraphPad Prism 9 (Version 9.1.1 (223), 2021) and RStudio software (Version 1.4.1106; 2021). To ensure that differences observed between *P. vivax* - infected patients and controls,
as well as between the clusters, were due to disease status and not confounded by age or sex, the clinical parameters were fitted as response variables in a linear model with sex and/or age fitted as explanatory variables. Age and sex were included in the model if their coefficients were estimated as different from zero with p-value < 0.05. The residuals from the linear model were then used as age and/or sex corrected parameters in subsequent analyses.

Study approval

All subjects enrolled in the study were adults. Written informed consent was obtained from all participants and the study was conducted according to the Declaration of Helsinki principles. The study was approved by the local Research Ethics Committee at Fundação de Medicina Tropical Dr. Heitor Vieira Dourado (FMT-HVD, Manaus, Brazil), under #CAAE: 54234216.1.0000.0005.

Ethics Approval and Consent to Participate:

All subjects enrolled in the study were adults. Written informed consent was obtained from all participants and the study was conducted according to the Declaration of Helsinki principles. The study was approved by the local Research Ethics Committee at Fundação de Medicina Tropical Dr. Heitor Vieira Dourado (FMT-HVD, Manaus, Brazil), under #CAAE: 54234216.1.0000.0005.

Acknowledgements

We would like to thank all patients enrolled in this research and the support of the malaria diagnosis and field team at field team in the Fundação de Medicina Tropical Dr. Heitor Vieira Dourado (FMT-HVD) in Manaus, Brazil. The authors also gratefully acknowledge the help and assistance provided by the Central Laboratory of High-Performance Technologies.
(LaCTAD, University of Campinas) and the Institute of Infection, Immunity and Inflammation Flow Core Facility in the generation of some of the data reported in this manuscript. M.M. is supported by a Wolfson Merit Award from the Royal Society and Wellcome Trust Center award (number 104111). J.L.S.F. was supported by the Sao Paulo Research Foundation (FAPESP grant 2019/01578-2 and 2016/12855-9), and F.T.M.C is supported by the Sao Paulo Research Foundation (FAPESP grant 2017/18611-7). M.V.G.L. and F.T.M.C. are CNPq research fellows.

Conflict of interest

The authors have declared that no conflict of interest exists.

Data Availability Statement

All data generated or analysed during this study are included in the manuscript and supporting files. Numerical tables and source data files have been provided. Table 1, Figure 2-source data 1 and Figure 2-figure supplement 2-source data 1 contain the numerical data used to generate the figures.
References


Figure legends

Figure 1: Clinical data of *P. vivax* patients (Pv) and healthy donors (HDs). (A) Red blood cell parameters. Shown are red blood cell counts, hemoglobin levels and hematocrit. (B) Other blood cell parameters. Shown are numbers of leukocytes, neutrophils, and monocytes, basophils and eosinophils (MXD). (C) Number of lymphocytes, platelets and mean platelet volume (MPV). Parameters are depicted as box plots showing each individual value and the median with maximum and minimum values. Dashed lines in black mark the minimum threshold for normal reference values, while lines in red mark threshold for severe lymphopenia and thrombocytopenia, respectively. Two-tailed student’s t-test was used to compare variables with normally distributed data, and Mann-Whitney test was used to compare variables with non-normal distributions; *p*-value is indicated above the graph when *p* < 0.05. HDs = healthy donors (controls, n = 34); Pv = *P. vivax*-infected patients (n = 79).

Figure 1-figure supplement 1: Demographic and clinical features of all *P. vivax*-infected patients compared with selected 31 patients for multiplex-bead based assay and downstream analysis. (A) Gender, age and hematological parameters compared between all 79 *P. vivax*-infected patients (All) and those 31 selected for downstream molecular analysis. (B) Comparison of gender, age and hematological parameters between 31 selected (S) *P. vivax*-infected patients with the remaining 48 non-selected (NS) patients. Parameters are depicted as box plots showing each individual value and the median with maximum and minimum values.

Figure 2: Characterization of heterogeneity in symptomatic *P. vivax* patients defines clusters of individuals. (A, B) Clustering of patients and healthy controls. After z-score normalization, Principal Component Analysis (PCA) was performed for data dimensionality.
reduction. K-means Clustering using k=2 followed by bootstrapping (1,000 times) in a PCA plot was performed and produced the most stable clusters regardless of the starting point (ln 1,000/1,000): Cluster 1 = 23 individuals comprising of 9 healthy donors and 14 *P. vivax* patients. The jaccard_index measures cluster similarity across bootstrap samples (jaccard_index ranges from 0 to 1, an index < 0.6 hints at a weak, unreliable cluster while > 0.85 means generally reliable). As indicated in the PCA plot, k=2 gives stable clusters for all configurations (jaccard_index 0.9 and 0.86) and withinss (wss) = 1,122. Open ovals represent 95% confidence interval ellipses around group mean points. Principal Component Analysis (PCA) was performed for data dimensionality reduction, in parallel with K-means Clustering (k) followed by bootstrapping (1,000 times). Open ovals represent 95% confidence interval ellipses around group mean points. (B) The resulting clusters represent healthy controls (1a) and patients (1b, 2). (C) Contribution of variables to clustering. In the circular plot the correlation between each input variable and principal components is used as coordinates (loading score). Plots shows how covariables determine patient distribution in the PCA plot.

Figure 2-figure supplement 1: Principal components analysis metrics. (A, B) Analysis of eigenvalues (measure of the amount of variation retained by each principal component) and the percentage of explained variances retained by the PCs accounted for the variance of the data. (C) However, most of the variables were highly represented in the first 2 PCs (Dim 1 and Dim 2), which were therefore retained for further analysis.

Figure 2-figure supplement 2: Methods determining the number of clusters best representing the data. (A, B) PCA plots indicating different K-means cluster configurations, using k= 3 and k=4 clusters, respectively after performing bootstrapping. With k=3 different starting points give different clusters. The two most common clusters (top row) are very similar and they are obtained in 241 and 179 starts out of 1,000, respectively. However, the clustering that best represents the data when k = 3 is the third one found in 168/1000 starting points as its withinss (wss) metric is lower (highlighted in red). Indeed, this configuration is more equivalent to those clustering configurations when k=2. (B) Clusters seem more stable when k=4. Accordingly, the best clustering appears to be the ones represented in the bottom row, which contains two main groups and two small groups with just 2 patients. (C, D) The second method used was the Monte Carlo Reference-based Consensus Clustering (M3C), which also indicated that k=2 is the optimal number of clusters, as indicated in (C) the flat line in the CDF plot and (D) in the highest Relative Cluster Stability Index (RCSI) plot. (E, F, G). Using spectral clusters, instead of elliptical K-means clusters, M3C analysis indicates that k=3 gives the optimal number of clusters, as indicated in the (E) CDF plot, (F) RCSI plot and (G) the NXN consensus matrix, where each element represents the fraction of times two samples clustered together.

Figure 2-source data 1: Correlation (loading score) of variables to principal components.

Figure 2-figure supplement 2-source data 1: Measurements of K-means cluster stability, using k=2, k=3 and k=4 clusters, via bootstrapping.

Figure 3: Unsupervised clustering analysis reveals two *P. vivax* patient subgroups that differ in parasite biomass. (A) Parasite parameters vs patient clusters. Comparison of the
two patient clusters (Cluster 1b and Cluster 2) across parasite parameters reveals significant
differences with total parasite biomass (PvLDH) but not peripheral parasitemia or parasite
load (copies of 18s rRNA/μL of blood). (B) **Parasite biomass vs parasitemia across**
clusters. Heatmap represents z-scores of PvLDH with peripheral parasitemia or parasite load,
respectively. Black boxes highlight patients with relatively lower peripheral parasitemia
compared to PvLDH levels, indicating the underestimation of total parasite biomass based on
peripheral parasitemia values. (C) **Correlation between parasite biomass and parasitemia.**
Scatter plot showing lack of correlation between PvLDH and peripheral parasitemia or
parasite load, respectively. Regression line in red, with 95% confidence interval shown in
shaded gray. (D, E) **Predicting parasite clusters.** (D) Top parameters prioritized by random
forest analysis ranked by the mean decrease in accuracy. (E) Best-fit decision trees and
random forest machine learning models corroborate PvLDH value as the most important
parasite signature in segregating patients into clusters 1b and 2. Cut-off values of the attribute
that best divided groups were placed in the root of the tree according to the parameter value.
The total of classified registers for each class and the percentage of observations used at that
node are given in each terminal node.

**Figure 4: More severe haematological alterations in Vivax\textsuperscript{high} compared to Vivax\textsuperscript{low} patients.** (A) **Patient data and hematological parameters.** Comparison of patient age,
average days of symptoms when samples were collected, haemoglobin levels, haematocrit or
RBC counts across patient clusters (Control: n=9; Vivax\textsuperscript{low}: n=14; Vivax\textsuperscript{high}: n=17). Data are
depicted as plots showing individual values and the median (black lines) and the interquartile
range. (B) **Blood cell counts.** Comparison of differential haematological counts across
clusters. Shown are, numbers of platelets, lymphocytes, neutrophils and monocytes, basophils
and eosinophils (MXD), neutrophil to total leucocyte ratio and neutrophil to lymphocyte ratio
(NLCR). Top dashed lines mark the minimal threshold for normal reference values, while
bottom dashed lines mark the threshold for severe lymphopenia and thrombocytopenia,
respectively. Parameters are depicted as plots showing individual values and the median
(black lines) and the interquartile range. One-way analyses of variance with Bonferroni-
corrected multiple comparisons test were performed. *p*-value is indicated above the graph
when reached significance of *p* < 0.05. (C) **Cytokine response and neutrophil activation**
across clusters. Heatmap represents z-scores obtained by centering values of Luminex data.
Shown are thrombopoiesis-inducing cytokines, myelopoiesis-inducing cytokines and
neutrophil activation markers. Biomarker concentrations were normalized (scale function in
R) and the average scaled value is showed in blue and yellow scales. Blue shading represents
the highest average scaled value; yellow shading, the lowest average scaled value. Each
column (i.e., individual) in the heatmap is matched with color-coded cluster assignment:
Cluster Control – green bar; Cluster Vivax\textsuperscript{low} – blue bar and Cluster Vivax\textsuperscript{high} – red bar.

**Figure 4-figure supplement 1: Increase of thrombopoiesis- and myelopoiesis inducing
cytokines in the plasma of Vivax\textsuperscript{high} patients.** (A) Levels of myelopoiesis-inducing
cytokines, (B) thrombopoiesis-inducing cytokines thrombopoietin (TPO) and IL-11 and (C)
neutrophil activation markers in the acute-phase patients’ plasma samples of our cross-
sectional cohort in Manaus, Brazil, were determined by multiplex-bead based assay
(Luminex): Control (healthy donors, n = 9), Vivax\textsuperscript{low} patients (n=14) and Vivax\textsuperscript{high} patients
(n=17). Biomarkers’ concentration is depicted as scatter plots showing individual data points
and the median (black lines) and the interquartile range. One-way analyses of variance with
Bonferroni-corrected multiple comparisons test were performed. *p*-value is indicated above
the graph when reached *p* < 0.05.
Figure 5: Elevated circulating markers of EC activation and damage in Vivax\textsuperscript{high} compared to Vivax\textsuperscript{low} patients. (A) Endothelial changes across clusters: Luminex. Heatmap represents z-scores obtained by centering values of Luminex data. Shown are markers of endothelial cell (EC) activation, procoagulation and glycocalyx damage. Each column (each individual) in the heatmap is matched with color-coded cluster assignment: Cluster Control – green bar; Cluster Vivax\textsuperscript{low} – blue bar and Cluster Vivax\textsuperscript{high} – red bar. (B) Endothelial changes across clusters: qRT-PCR. Transcriptional response of HUVECs incubated for 6h with 30% v/v pooled plasma from different clusters. Heatmap reflects relative mRNA expression intensity (average scaled value) after results were normalized to GAPDH housekeeping gene expression and untreated condition (mock). Data shown represent the mean of three independent experiments. For each experiment, 2 technical replicates were performed for each condition. (C) Endothelial changes across clusters: impedance changes. Endothelial monolayer integrity was measured during 20% v/v of pooled plasma incubation. Each line represents the mean ± SD of normalized resistance of HUVECs measured by ECIS at 4,000 Hz. Data shown are representative of three independent experiments. For each experiment, 2 technical replicates were performed for each condition. (D) Endothelial changes across clusters: imaging and flow cytometry. HUVECs were incubated for 18h with 30% v/v of pooled plasma of individuals in the different clusters or left untreated (mock). Percentage of cells expressing EC activation markers (adhesion molecules) ICAM and VCAM as well as quantification of protein expression was determined by flow cytometry and immunofluorescence analysis (scale bar = 33μM). Isotype antibodies were used as control to define positive populations. Significance was calculated for comparisons between conditions at the corresponding time point using. One-way analyses of variance statistical test with Tukey’s corrected multiple comparisons test. $p$-value is indicated above the graph when $p < 0.05$. Data shown are representative mean ± SEM of three independent experiments.

Figure 5-figure supplement 1: Increase of markers of endothelial cell activation, damage (glycocalyx breakdown) and procoagulation in the plasma of Vivax\textsuperscript{high} patients. (A) Levels of EC activation markers, (B) procoagulant phenotype and (C) EC damage (glycocalyx breakdown) in the acute-phase patients’ plasma samples of our cross-sectional cohort in Manaus, Brazil, were determined by multiplex-bead based assay (Luminex): Control (healthy donors, $n = 9$), Vivax\textsuperscript{low} patients ($n=14$) and Vivax\textsuperscript{high} patients ($n=17$), as indicated in the legend (top right corner). Biomarkers’ concentration is depicted as scatter plots showing individual data points and the median (black lines) and the interquartile range. One-way analyses of variance with Bonferroni-corrected multiple comparisons test were performed. $p$-value is indicated above the graph when reached significance of $p < 0.05$. (D) Quantitative mRNA expression was determined by qRT-PCR in RNA extracted from HUVECs incubated for 6h with 30% v/v pooled plasma of individuals in the different clusters, as indicated in the legend (top right corner). Graphs depict relative expression after results were normalized to GAPDH housekeeping gene expression. The data shown are mean ± SEM representative of three independent experiments. Significance was calculated for comparisons between conditions at the corresponding time point using One-way analyses of variance with Tukey’s corrected multiple comparisons test. $p$-value is indicated above the graph when reached significance of $p < 0.05$. (E) Schematics of HUVECs gating strategy used for flow cytometry analysis. Endothelial cells gate (ECs) was defined based on the cells’ forward scatter (FSC) and side scatter (SSC). Further, gated single cells on FSC-A vs FSC-H scatter plot and selected live cells based on Fixable Viability Dye eFluor™ 506 staining.
**Figure 5-figure supplement 2: Haemolysis potentiates Vivax\textsuperscript{high} -induced EC activation.**

In order to mimic the environment associated with commencement of anti-malarial treatment, such as content released from haemolysis of RBCs and dead parasites biproducts, HUVECS were stimulated with either *P. vivax*-infected (schizont enriched) or uninfected RBCs lysates in 30% v/v pooled-plasma of individuals in the different clusters. For the parasite lysates, batch pellets of *P. vivax* iRBCs enriched of schizonts, isolated from blood of *P. vivax*-infected patients using Percoll gradient centrifugation, and healthy donor RBCs were stored at −80 °C without any cryopreservative agent. Pellets were twice freeze–thawed for use as whole lysates. Total RNA was extracted 6h after treatment and relative mRNA expression determined by real-time quantitative PCR. Graphs depict relative expression after results were normalized to GAPDH housekeeping gene expression. The data shown are mean ± SEM representative of three independent experiments. Significance was calculated for comparisons between conditions at the corresponding time point using One-way analyses of variance with Tukey’s corrected multiple comparisons test. p-value is indicated above the graph when reached significance of p < 0.05. Haemolysis of either *P-vivax*-infected or uninfected RBCs potentiates the effect of Vivax\textsuperscript{high} pooled plasma in inducing transcriptional upregulation of EC activation markers. Different from the stimulation only with plasma, in the presence of haemolysis, we also observed upregulation of Ang-2 and VEGF, and downregulation of NOS3, markers associated with perturbation of the vascular integrity and function.

**Figure 6: Network analysis and clustering of parasite and host signatures indicate parasite-induced changes in deep tissues. (A) Network analysis.** Networks of the parasite biomass (PvLDH) and host signatures in healthy donors (left graph) and in *P. vivax*-infected patients (right graph), using a force-directed layout. The symbols of the nodes represent biological functions: triangle represents markers of platelet activation and thrombopoiesis-inducing cytokines; V shape represents haematological parameters (neutrophil, lymphocyte and platelet counts); circles represent endothelial cell activation markers; squares represent myelopoiesis-inducing cytokines and neutrophil activation markers. The colors in the nodes represent the fold change in relation to control levels. Because healthy donors do not have parasitemia, PvLDH node is represented in black. Each connecting line (edge) represents a significant interaction detected by the network analysis using R. Correlation strength is represented by edge color transparency and width. Positive correlations are represented by red edges and negatives correlations are represented by blue edges. (B,C) Correlation matrix and heatmap. (B) Representative image of Pearson’s correlation matrix calculated for all *P. vivax* patients. Only correlations with *p*-value < 0.05 are shown and hierarchical clustering was applied. Red circles highlight positive correlations in the functional modules depicted in (A), and blue circles highlight negative correlations in the functional modules also depicted in (A). (C) Heatmap showing *p*-values of the correlations between different parasite parameters, parasite biomass (PvLDH) and peripheral parasitemia and host signatures (haematological and Luminex parameters). (D) Decision tree model. Best-fit classification tree model generated with the C4.5 algorithm showing Syndecan-1, IL-6 and platelet counts are the dominant variables capable to predict total parasite biomass in *P. vivax* patients. Cut-off values of the attribute that best divided groups were placed in the root of the tree according to the parameter value (pg/mL for soluble markers or number of cells x 1,000/μL of blood for platelet counts). The total of classified registers for each class are
given in parentheses for each terminal node with the k-fold cross-validation (k-fold CV) accuracy indicated.

**Figure 6-figure supplement 1:** Representative images of Pearson’s correlation matrix calculated separately for each *P. vivax* patient cluster. (A) Vivax<sub>low</sub> patients (n=14). (B) Vivax<sub>high</sub> patients (n=17). A reduced complexity model was established by focusing on informative interactions between *P. vivax* and host signatures determined by Pearson’s correlation coefficients. Only correlations with associated *p*-value <0.01 are shown and hierarchical clustering was applied.

**Figure 6-figure supplement 2:** Validation of patients’ clusters and correlations when segregating patients based on thrombocytopenia severity. Box plots showing (A) parasite parameters, clinical parameters and biomarkers measured on plasma samples were generated segregating patients based on levels of thrombocytopenia severity (normal, mild, moderate and severe) color-coded. (B) Recursive partitioning classification tree model generated with the rpart function in R showing the high value of VCAM-1, PvLDH and Syndecan-1 to predict thrombocytopenia severity in *P. vivax* patients. Cut-off values of the attribute that best divided groups were placed in the root of the tree according to the parameter value (pg/mL for soluble markers or O.D. for PvLDH).

**Figure 6-figure supplement 3:** Validation of patients’ clusters and correlations when segregating patients based on lymphopenia severity. Box plots showing parasite parameters, clinical parameters and biomarkers measured on plasma samples were generated segregating patients based on levels of lymphopenia severity (normal, moderate and severe) color-coded.

**Supplementary Files**

**Supplementary File 1:** Topological analysis of the network graphs of healthy donors and *P. vivax* patients.

**Supplementary File 2:** Oligonucleotides sequences used in the qRT-PCRs
Table 1: Demographic, parasite and multiplexed microbead-based immunoassay (Luminex) data obtained from the plasma of a representative subset of 31 *P. vivax* patients and 9 healthy donors (controls).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Healthy donors (n=36)</th>
<th>Symptomatic Pv patients (n=79)</th>
<th>P-value (Pv vs control)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Median [IQ 25-75]</td>
<td>Median [IQ 25-75]</td>
<td>0.06</td>
</tr>
<tr>
<td>Age</td>
<td>32 [23-49]</td>
<td>36 [28-45]</td>
<td></td>
</tr>
<tr>
<td>Parasitemia (10⁶/mL)</td>
<td>-</td>
<td>4.29 [1.86-6.62]</td>
<td></td>
</tr>
<tr>
<td>Parasitemia (%)</td>
<td>-</td>
<td>0.76 [0.57-1.25]</td>
<td></td>
</tr>
<tr>
<td>Parasite load (copies 18s RNA/mL)</td>
<td>-</td>
<td>26,642 [9,253-522,297]</td>
<td></td>
</tr>
<tr>
<td>PvLDH (O.D.)</td>
<td>-</td>
<td>0.18 [0.005-0.34]</td>
<td></td>
</tr>
<tr>
<td>Plasma biomarkers</td>
<td>Healthy donors (n=9)</td>
<td>Symptomatic Pv patients (n=31)</td>
<td>P-value (Pv vs control)</td>
</tr>
<tr>
<td>TNF-α (pg/mL)</td>
<td>17.2 [11.0-22.3]</td>
<td>38.4 [30.0-69.6]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>IL-1α (pg/mL)</td>
<td>11.9 [10.0-19.5]</td>
<td>25.4 [19.8-33.5]</td>
<td>0.0004</td>
</tr>
<tr>
<td>IL-1β (pg/mL)</td>
<td>12.0 [8.0-12.8]</td>
<td>21.4 [14.5-27.6]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Value</td>
<td>Median</td>
<td>P-Value</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>IL-6 (pg/mL)</td>
<td>3.0 [2.5-3.7]</td>
<td>33.4 [7.6-133.1]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>IL-8 (pg/mL)</td>
<td>2.2 [0.6-2.4]</td>
<td>6.4 [2.7 – 19.9]</td>
<td>0.0005</td>
</tr>
<tr>
<td>IL-10 (pg/mL)</td>
<td>--*</td>
<td>314 [169-562]</td>
<td>--</td>
</tr>
<tr>
<td>G-CSF (pg/mL)</td>
<td>9.485 [9.485-9.485]</td>
<td>101.5 [33.49-239.6]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>L-Selectin (ng/mL)</td>
<td>326 [287-391]</td>
<td>481 [386-579]</td>
<td>0.0019</td>
</tr>
<tr>
<td>ICAM-1 (ng/mL)</td>
<td>323 [260-464]</td>
<td>634 [456-849]</td>
<td>0.0026</td>
</tr>
<tr>
<td>VCAM-1 (ng/mL)</td>
<td>819 [623-959]</td>
<td>2875 [1753-5108]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>E-Selectin (ng/mL)</td>
<td>26.4 [22.5-33.7]</td>
<td>56.7 [41.5-74.1]</td>
<td>0.0001</td>
</tr>
<tr>
<td>P-selectin (ng/mL)</td>
<td>17.0 [15.4-20.6]</td>
<td>22.2 [17.6-25.7]</td>
<td>0.0621</td>
</tr>
<tr>
<td>Angiopoietin-1 (ng/mL)</td>
<td>0.4 [0.3-0.6]</td>
<td>0.5 [0.2-0.9]</td>
<td>0.8874</td>
</tr>
<tr>
<td>Angiopoietin-2 (ng/mL)</td>
<td>1.8 [1.5-2.1]</td>
<td>4.3 [2.7-5.3]</td>
<td>0.0003</td>
</tr>
<tr>
<td>Ang-2:Ang-1 ratio</td>
<td>4.2 [2.7-5.6]</td>
<td>12.14 [2.7-40.2]</td>
<td>0.03</td>
</tr>
<tr>
<td>VWF-A2 (pg/mL)</td>
<td>126 [120-150]</td>
<td>218 [199-277]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>ADAMTS13 (ng/mL)</td>
<td>1110 [483-1740]</td>
<td>776 [572-1328]</td>
<td>0.5485</td>
</tr>
<tr>
<td>PAI-1 (pg/mL)</td>
<td>78.9 [62.4-96.4]</td>
<td>112 [69.3-242]</td>
<td>0.1541</td>
</tr>
<tr>
<td>CD40L (ng/mL)</td>
<td>0.5 [0.4-0.7]</td>
<td>1.0 [0.7-1.3]</td>
<td>0.0001</td>
</tr>
<tr>
<td>Syndecan-1 (ng/mL)</td>
<td>1.8 [1.6-2.4]</td>
<td>3.7 [2.9-6.0]</td>
<td>0.0003</td>
</tr>
<tr>
<td>IL-11 (ng/mL)</td>
<td>3.5 [2.9-4.3]</td>
<td>5.7 [4.7-6.4]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TPO (ng/mL)</td>
<td>2.0 [1.7-2.2]</td>
<td>3.0 [2.6-3.4]</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>CXCL4 (ng/mL)</td>
<td>0.8 [0.6-1.2]</td>
<td>1.4 [0.7-2.8]</td>
<td>0.1236</td>
</tr>
<tr>
<td>CXCL7 (ng/mL)</td>
<td>0.4 [0.4-0.5]</td>
<td>0.73 [0.4-1.7]</td>
<td>0.1958</td>
</tr>
<tr>
<td>SCF (pg/mL)</td>
<td>47.61 [37.34 - 89.34]</td>
<td>45.68 [36.22 - 61.39]</td>
<td>0.1594</td>
</tr>
</tbody>
</table>