

Gene expression of macaques infected with malaria species of zoonotic concern

Amber E. Trujillo^{1,2} and Christina M. Bergey^{3,4,*}

¹New York University, Department of Anthropology, New York, NY, USA

²New York Consortium in Evolutionary Primatology, New York, NY, USA

³Rutgers University, Department of Genetics, New Jersey, USA

⁴Rutgers University, Human Genetics Institute of New Jersey, New Jersey, USA

*To whom correspondence should be addressed: christina.bergey@rutgers.edu

ABSTRACT

A multitude of malaria species (genus *Plasmodium*) infects primates. Due to their public health importance, the human-infective species have garnered the most focus, but increased knowledge of non-human primate malaria species is warranted to improve our evolutionary understanding of host-parasite interactions. Additionally, the broad host tropism of some primate malaria parasites and their realized or theorized zoonotic potential add urgency to understanding of primate-parasite interactions. Here, we use comparative transcriptomics to understand the rhesus macaque (*Macaca mulatta*) response to two malaria parasites used as analogues to human-infective species of differing severity and which may represent emerging zoonotic threats: *P. coatneyi*, comparable to human-infective *P. falciparum*, and *P. cynomolgi*, comparable to human-infective *P. vivax*. We first validate our transcriptomics-based proxy of parasite load through comparison to gold-standard microscopy-based measures. We then find that malaria-associated host genes have functional links to immune system regulation and blood cells. Host genes with differing expression by malaria species were more likely to be involved in brain-linked functions, perhaps due to the differential central nervous system involvement of the two parasite species. Such comparative work on primate malaria species may help elucidate the essential and species-specific molecular mechanisms that underlie differing clinical presentations and zoonotic risk.

Introduction

The five species of malaria that regularly infect humans vary in their clinical presentation and severity, with the most clinically serious being *Plasmodium falciparum* and *P. vivax*. These human-infective parasites are part of lineages that separately came to infect humans thousands of years ago¹, each having distinct nearest non-human primate-infective relatives. The broad host tropism of *Plasmodium* species is underscored by the recent zoonotic exchange to humans observed for simian malaria species, including the now-widespread *P. knowlesi*^{2,3}, all of which tend to cause less severe disease than human infection specialist species of *Plasmodium*. However, despite their public health importance, the reasons for the variation in impact between the malaria parasite species remain incompletely understood. Evolutionary transcriptomic investigations of host-parasite interaction can shed light on the molecular underpinnings of this differential pathology⁴. Additionally, understanding the determinants of disease severity has critical relevance for zoonotic disease, as reservoir hosts with infection benign enough to not impact behavior pose the greatest risk of spillover⁵.

Various primate models of malaria, including these zoonotic parasites, have been used to interrogate malaria biology and better understand the host-parasite interaction differences that determine disease severity (e.g., rhesus macaques, *Macaca mulatta*^{4,6-13}; crab-eating macaques, *Macaca fascicularis*^{8,14,15}; owl monkeys, *Aotus nancymae*¹⁶; and squirrel monkeys, *Saimiri boliviensis*^{17,18}). For example, rhesus macaque monkeys infected with *P. coatneyi* are clinically comparable to human infections with the distantly-related *P. falciparum*. Experimental infections in rhesus monkeys

cause high virulence with parasitemia often rising rapidly to 10-20%. Symptoms include renal dysfunction, severe anemia, metabolic acidosis, and 30-40% mortality in the absence of effective antimalarial drug treatment^{19,20}. This severity has led to the use of the *P. coatneyi*-rhesus model for *P. falciparum*, particularly for the neurological presentations which may include coma with, importantly, sequestration of infected red blood cells in the cerebral and cerebellar brain vessels as seen in human *P. falciparum* infections¹⁹. In contrast, when rhesus macaques are instead infected with *P. cynomolgi*, most clinically comparable and also most closely related to human *P. vivax*, the monkeys exhibit a milder disease, which may include acute anemia and thrombocytopenia²¹. Like the human *P. vivax* infections it emulates, mortality risk is low except for pregnant rhesus macaques and fetuses¹⁹. The utility of the model is also augmented by a *vivax*-like rate of relapses originating from hypnozoites^{9,21-23}.

The host tropism of these parasites provides important context for their differential symptoms in various species and informs their risk for zoonotic spillover. Natural infection with *P. coatneyi* has been observed for several primate species in Southeast Asia, but crab-eating macaques (*M. fascicularis*) are thought to have co-evolved with the parasite^{20,24,25}. Although pathogenicity in crab-eating macaques is less consistent than in rhesus macaques, these monkeys still variably develop severe symptoms and clinical presentation in response to *P. coatneyi*²⁰. In contrast, *P. cynomolgi* has been found to naturally infect a range of primate hosts in Southeast Asia, including various macaque species²⁶. Alongside the established simian zoonotic *P. knowlesi*, both *P. cynomolgi*^{4,27-30} and *P. coatneyi*³⁰ parasites were recently found in blood samples of humans, though the latter appears to be far rarer, as supported by prior failed attempts to experimentally infect people³¹. Comparison of host response to these parasites may provide insight into their differing severity and risk of spillover.

Several transcriptomic studies have advanced our understanding of malaria biology by investigating non-human primate response of a macaque model to these two parasites, *P. cynomolgi* and *P. coatneyi*^{4,32}. Recent studies of *P. cynomolgi*, for example, have uncovered the causes of malaria anemia by examining bone marrow gene expression¹⁰; determined that genes related to PAMP and pro-inflammatory cytokines were upregulated in severe disease³³; modeled metabolic pathway systems to determine that patterns of purine metabolism within purine pathways for *P. cynomolgi* infection were consistent with *P. falciparum* infection in humans¹¹; and found many more genes were upregulated in acute vs. relapsing disease with enrichment for involvement in immune response¹². Transcriptomic investigations of macaque response to *P. coatneyi* are more limited, but extensive systems biology data exist from controlled experiments²². Such studies of experimentally-infected macaques have greatly improved our understanding of fundamental malaria biology^{4,32}. However, to date, no direct comparison of the gene expression response to *P. coatneyi* versus *P. cynomolgi* has been undertaken, limiting our understanding of how host response to the parasites may contribute to the differential disease severity and signs.

In this study, we compare two cohorts of rhesus macaques (*Macaca mulatta*) that have been experimentally infected with either *P. cynomolgi* and *P. coatneyi* for longitudinal analysis. We partition the macaque gene expression response to infection into components that are generalized or specific to infecting *Plasmodium* species, thereby identifying fundamental processes to malaria in this host as well as changes which may underpin the unique symptoms of each species. We also simultaneously assay *Plasmodium* gene response, determining orthologous genes with similar or divergent reactions across the two parasite species. Overall, our comparison of macaque interaction with *P. cynomolgi* or *P. coatneyi* allows insight into the transcriptomic drivers of primate-parasite interaction in general, as well as the use of these systems to model human malaria with different parasite etiology.

Results

The rhesus macaque whole blood transcriptomic datasets were generated through the efforts of the Malaria Host-Pathogen Interaction Center's (MaHPIC) to test hypotheses related to host-pathogen interactions in experimental non-human primate malaria infections. The $N = 50$ samples were from eight macaque individuals from a captive, experimental population. To identify macaque monkey genes with expression correlated to *Plasmodium* parasite levels, we aligned blood transcriptome reads to a concatenated host-parasite pseudo-reference genome (containing genomes of macaque and *Plasmodium coatneyi* or macaque and *P. cynomolgi*) and tallied reads by exon. After filtering, the median per-sample counts of reads uniquely mapping to the macaque and *P. coatneyi* genome were 30,987,028 (range: 9,945,658 to 75,084,213) and 24,177 (range: 2,899 to 26,416,731), respectively. For the macaque and *P. cynomolgi* dataset, the median uniquely mapping per-sample counts were 38,758,474 (range: 3,071,635 to 99,902,110) and 28,793 (range: 7,017 to 29,629,054), respectively.

Inferred parasitemia proxy is an accurate representation of infection

To investigate the accuracy of the parasite load proxy inferred computationally from RNA sequencing data, we tested for a correlation between these inferred parasite load proxies and previously published parasite load counts measured via microscopy as parasites per microliter, a method commonly used. We found that our parasite load proxy: *i.*) was tightly and significantly correlated with the traditional measure of parasite load (Fig. 1a; $R^2=0.8527$, $p < 2.2 \times 10^{-16}$); and *ii.*) accurately traced the severity of parasite load over the 100-day experimental period with high sensitivity (Fig. 1b). Having found this proxy to be a reliable method for inferring parasite load from RNAseq data, later analyses used this metric.

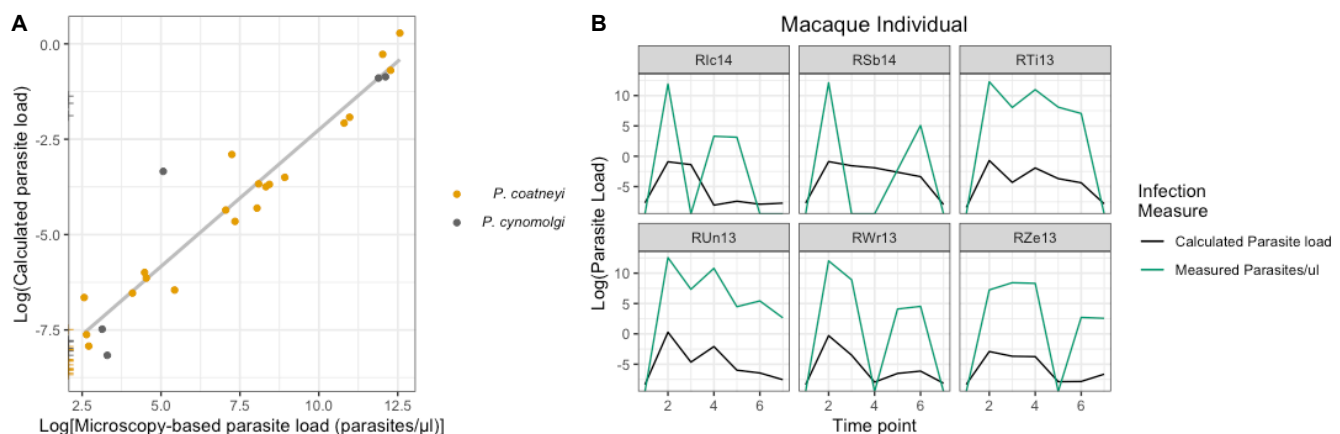


Figure 1. Correlation between previously published microscopy-based parasite load and the sequence read-based parasite load for samples infected with *P. coatneyi* (yellow) and *P. cynomolgi* (black). Samples were considered to be uninfected if there was no evidence of infection via microscopy and the proxy value was less than 0.001. Lines indicate samples with microscopy values of 0. Measures of infection were significantly and highly correlated ($R^2 = 0.853$, $p < 2.2 \times 10^{-16}$; **A**) each capturing parasite load changes throughout the 100 day experiment (**B**). Samples from individuals RCs13 and RFv13 are excluded from the plot given their lack of malaria infection and removal from the experiment, respectively.

Contrasting host response to parasitemia and infection intensity

In order to explore how the host may be responding to the malaria parasite, we analyzed our dataset at two hierarchical levels: parasitemia (infected vs. non-infected) and parasite load (inferred intensity of infection.) In analyzing $N = 50$ longitudinal samples from five individual macaques infected with *Plasmodium coatneyi* and from three individuals infected

Macaque genes with expression DE for parasitemia					
Positively correlated			Negatively correlated		
Gene	<i>logFC</i>	Adj. <i>p</i>	Gene	<i>logFC</i>	Adj. <i>p</i>
<i>HIST1H2BH</i>	1.96	2.82×10^{-5}	<i>APBB2</i>	-1.09	4.97×10^{-5}
<i>GAS6</i>	2.08	4.95×10^{-5}	<i>DHRS3</i>	-0.975	7.81×10^{-5}
<i>PTPN1</i>	0.609	4.95×10^{-5}	<i>PDE3B</i>	-0.645	2.15×10^{-4}
<i>ELL3</i>	1.40	4.95×10^{-5}	<i>NXPH3</i>	-1.68	2.47×10^{-4}
<i>SOCS1</i>	1.34	4.95×10^{-5}	<i>CUX2</i>	-1.61	2.70×10^{-4}
<i>ASAP3</i>	1.18	7.81×10^{-5}	<i>MAPT</i>	-0.834	2.96×10^{-4}
<i>AGRN</i>	1.45	0.000114	<i>PID1</i>	-1.73	2.99×10^{-4}
<i>FADS1</i>	1.25	0.000158	<i>ESRP2</i>	-0.564	2.99×10^{-4}
<i>TSTD1</i>	0.813	0.000172	<i>VANGL1</i>	-0.932	2.99×10^{-4}
<i>EGR2</i>	1.34	0.000247	<i>CA8</i>	-1.06	3.10×10^{-4}
<i>PIM2</i>	0.609	0.000247	<i>NFATC3</i>	-0.569	3.10×10^{-4}
<i>ANKRD9</i>	1.85	0.000247	<i>XPC</i>	-0.540	3.10×10^{-4}
<i>RELB</i>	0.60	0.000247	<i>DBP</i>	-0.663	3.10×10^{-4}
<i>AICDA</i>	2.66	0.000247	<i>AOX2</i>	-1.11	3.18×10^{-4}
<i>ZNF215</i>	0.779	0.000247	<i>CD1B</i>	-2.14	4.25×10^{-4}
<i>FADS2</i>	2.32	0.000247	<i>PIK3R1</i>	-0.468	4.25×10^{-4}
<i>FFAR1</i>	2.09	0.000247	<i>SAMD3</i>	-0.941	6.60×10^{-4}
<i>GPC1</i>	2.49	0.000247	<i>C1H1orf174</i>	-0.388	6.62×10^{-4}
<i>P2RX5</i>	1.31	0.000270	<i>KLRB1</i>	-1.33	7.57×10^{-4}
<i>JAK3</i>	0.675	0.000270	<i>CD96</i>	-0.731	7.64×10^{-4}

Table 1. Macaque genes with the strongest evidence of differential expression by malaria parasitemia (infected versus non-infected samples), irrespective of *Plasmodium* species. Table is limited to 20 genes, with full results present in a supplementary table (Tables S1).

with *P. cynomolgi*, we found that there were significant differences in the identified genes of interest between the two measures of infection with no overlap between the top 20 DE genes (Table 1, Table 2). Nevertheless, DE genes between the two measures were all associated with similar functions and disorders pertaining to immune system regulation and blood cells. For example, when considering parasite load (Fig. 2a), significantly DE genes include *ADAMDEC1* (LogFC=3.67, adj. $p=2.66 \times 10^{-10}$), *CYP2C76* (LogFC=2.35, adj. $p=6.45 \times 10^{-7}$), *SLCO2B1*, (LogFC=3.70, adj. $p=7.00 \times 10^{-7}$), *CA4* (LogFC=4.25, adj. $p=1.90 \times 10^{-6}$), *RABGAP1L* (LogFC=-0.604, adj. $p=1.76 \times 10^{-4}$), and *LUC7L* (LogFC=-0.517, adj. $p=0.007$). When considering parasitemia (Fig. 2b), significantly differentially expressed genes include *GAS6* (LogFC=2.08, adj. $p=4.95 \times 10^{-5}$) and *PTPN1* (LogFC=0.609, adj. $p=4.95 \times 10^{-5}$). Results for DE genes identified using the coefficients of the *Plasmodium* species and interacting variables (infection measure**Plasmodium* species) for each model are given in the supplementary tables, as are full results for parasite load- and parasitemia-DE tests (Tables S1-S2).

To further investigate the impact of the various parameters on individuals' gene expression, we conducted Principle Components Analysis (PCA; Fig. 3a and b) using the combined dataset of macaques infected with different species of *Plasmodium* and considering the two measures of infection. We found that individuals were most strongly grouping according to their parasitemia status (uninfected or infected) (PC1; Fig. 3a) with secondary grouping according to the infecting *Plasmodium* species (PC3; Fig. 3b). There was no distinct grouping according to parasite load. We additionally performed t-Distributed Stochastic Neighbor Embedding (t-SNE) analysis, but found no significant clustering.

Macaque genes with expression DE for parasite load					
Positively correlated			Negatively correlated		
Gene	<i>logFC</i>	Adj. <i>p</i>	Gene	<i>logFC</i>	Adj. <i>p</i>
<i>ADAMDEC1</i>	3.67	2.66×10^{-10}	<i>ASPH</i>	-1.19	7.68×10^{-6}
<i>FAM20A</i>	2.76	2.66×10^{-10}	<i>CAMK1D</i>	-0.724	1.82×10^{-5}
<i>KCNE1</i>	4.02	1.18×10^{-7}	<i>RABGAP1L</i>	-0.604	1.76×10^{-4}
<i>MARCO</i>	4.02	6.01×10^{-7}	<i>ADGRG5</i>	-1.54	4.40×10^{-4}
<i>CYP2C76</i>	2.35	6.45×10^{-7}	<i>PIK3R6</i>	-1.92	5.41×10^{-4}
<i>PLXDC1</i>	2.85	7.00×10^{-7}	<i>PDE7A</i>	-0.959	6.66×10^{-4}
<i>SLCO2B1</i>	3.71	7.00×10^{-7}	<i>MAMU.DMB</i>	-1.19	7.95×10^{-4}
<i>PODXL</i>	1.93	7.00×10^{-7}	<i>ADCY9</i>	-1.27	7.96×10^{-4}
<i>HSD3B7</i>	2.10	7.00×10^{-7}	<i>ADK</i>	-1.34	9.82×10^{-4}
<i>CDH17</i>	3.31	7.87×10^{-7}	<i>EEF2K</i>	-0.744	9.87×10^{-4}
<i>CTH</i>	4.32	7.87×10^{-7}	<i>TXNRD2</i>	-0.860	1.44×10^{-3}
<i>FCGR1A</i>	3.47	1.19×10^{-6}	<i>ABCC4</i>	-1.31	1.45×10^{-3}
<i>LIPM</i>	2.46	1.23×10^{-6}	<i>ABAT</i>	-0.981	1.49×10^{-3}
<i>TIFA</i>	2.54	1.30×10^{-6}	<i>DNAJB12</i>	-0.428	1.67×10^{-3}
<i>SYNPO2</i>	2.44	1.48×10^{-6}	<i>ULK2</i>	-0.865	1.94×10^{-3}
<i>NLRP14</i>	3.08	1.54×10^{-6}	<i>BTAF1</i>	-0.758	2.41×10^{-3}
<i>TIMP3</i>	3.21	1.54×10^{-6}	<i>SLC4A7</i>	-0.652	2.45×10^{-3}
<i>P2RY14</i>	2.97	1.54×10^{-6}	<i>MAMU.DMA</i>	-1.26	2.61×10^{-3}
<i>TUSC3</i>	3.31	1.54×10^{-6}	<i>EZH1</i>	-0.525	3.36×10^{-3}
<i>NCKAP1</i>	3.11	1.58×10^{-6}	<i>ADRB2</i>	-1.38	3.68×10^{-3}

Table 2. Macaque genes with the strongest evidence of differential expression by malaria parasite load (amount of parasite RNA detected in sample), irrespective of *Plasmodium* species. Table is limited to 20 genes, with full results present in a supplementary table (Table S2).

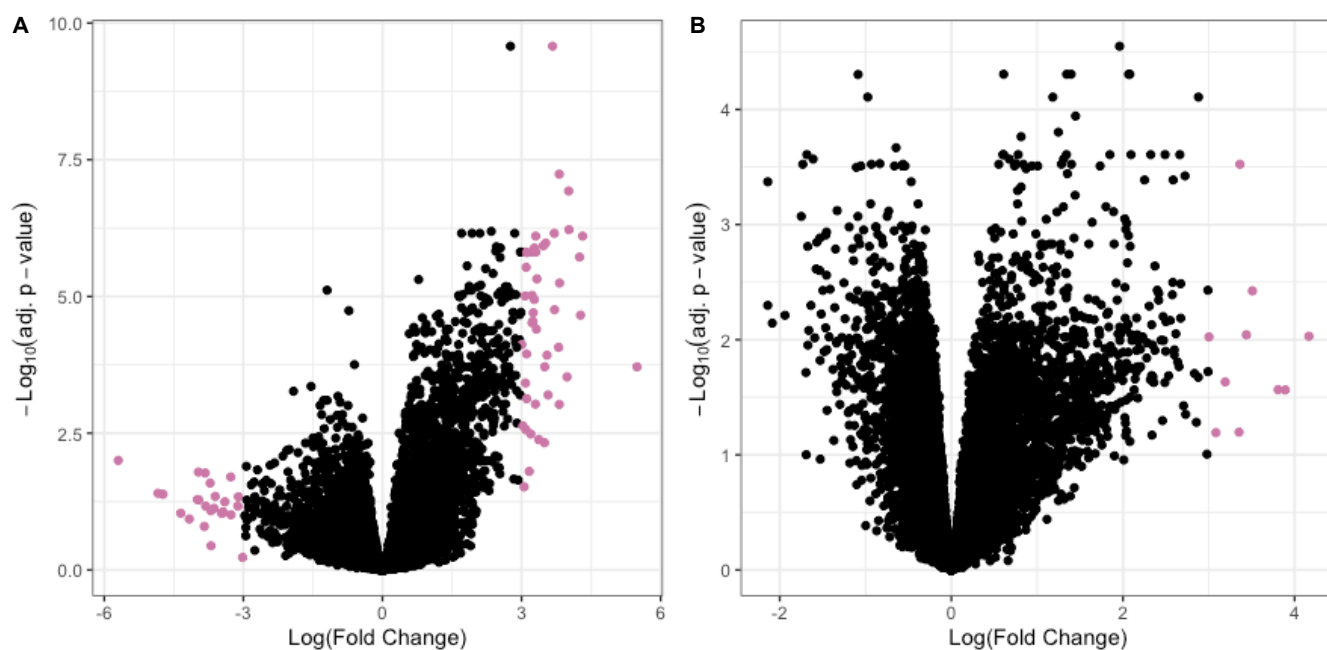


Figure 2. Summary of tests for differential expression of macaque genes by malaria infection status. Pink dots indicate differentially expressed macaque genes for two different measures of malaria infection: parasite load, or the inferred intensity of infection (**A**), and parasitemia, or samples with and without detected malaria infection (**B**). Both measures are inferred from sequencing reads. Note differing scales between two graphs.

Over-represented functions of macaque genes correlated to malaria infection

Finally, we found that up-regulated genes DE for both of the measures of parasite infection (parasite load and parasitemia) were associated with functions such as immune system process (GO:0002376; adj. $p=2.719 \times 10^{-35}$ and 1.93×10^{-8} , respectively), response to stimulus (GO:0050896; adj. $p=6.54 \times 10^{-32}$ and 1.95×10^{-15} , respectively), and response to stress (GO:0006950; adj. $p=7.09 \times 10^{-30}$ and 1.74×10^{-6} , respectively). We found that DE genes that were detected using parasite load specifically were enriched for functions such as defense response (GO:0006952; adj. $p=6.68 \times 10^{-27}$), response to other organisms (GO:0051707; adj. $p=4.41 \times 10^{-18}$), and platelet activation (KEGG:04611; adj. $p=2.803 \times 10^{-4}$), while DE genes that were detected using parasitemia were enriched for functions such as increased red cell osmotic fragility (HP:0005502; adj. $p=2.74 \times 10^{-7}$), anemia due to reduced life span of red cells (HP:0011895; adj. $p=1.88 \times 10^{-5}$), and reticulocytosis (HP:0001923; adj. $p=2.74 \times 10^{-7}$).

When exploring the DE-genes associated with the infecting *Plasmodium* species coefficient using parasitemia, we found that *P. coatneyi* (which is comparable to the *Plasmodium* species that causes cerebral malaria in humans) genes were particularly enriched for vestibular hypofunction (HP:0001756; adj. $p=1.716 \times 10^{-2}$), abnormal cochlea morphology (HP:0000375; adj. $p=3.021 \times 10^{-2}$), and subcortical cerebral atrophy (HP:0012157; adj. $p=3.021 \times 10^{-2}$; Table S3). There were no unique functional enrichment categories of interest to report when including the DE-genes associated with the infecting *Plasmodium* species coefficient using parasite load (Table S4). We found that none of our differentially expressed gene lists that were detected using the *Plasmodium* species coefficient (i.e., *P. coatneyi*- or *P. cynomolgi*-specific genes) were more than likely by chance to be overrepresented in human response genes to *Plasmodium* parasites (i.e., *P. vivax* and *P. falciparum*³⁴) with the exception of genes upregulated in *P. cynomolgi* relative to *P. coatneyi*, as these genes were more likely than expect by chance to be upregulated in *P. falciparum* response (p value=0.034).

Results for the functional enrichment analyses of macaque genes DE for the interaction between infection measure

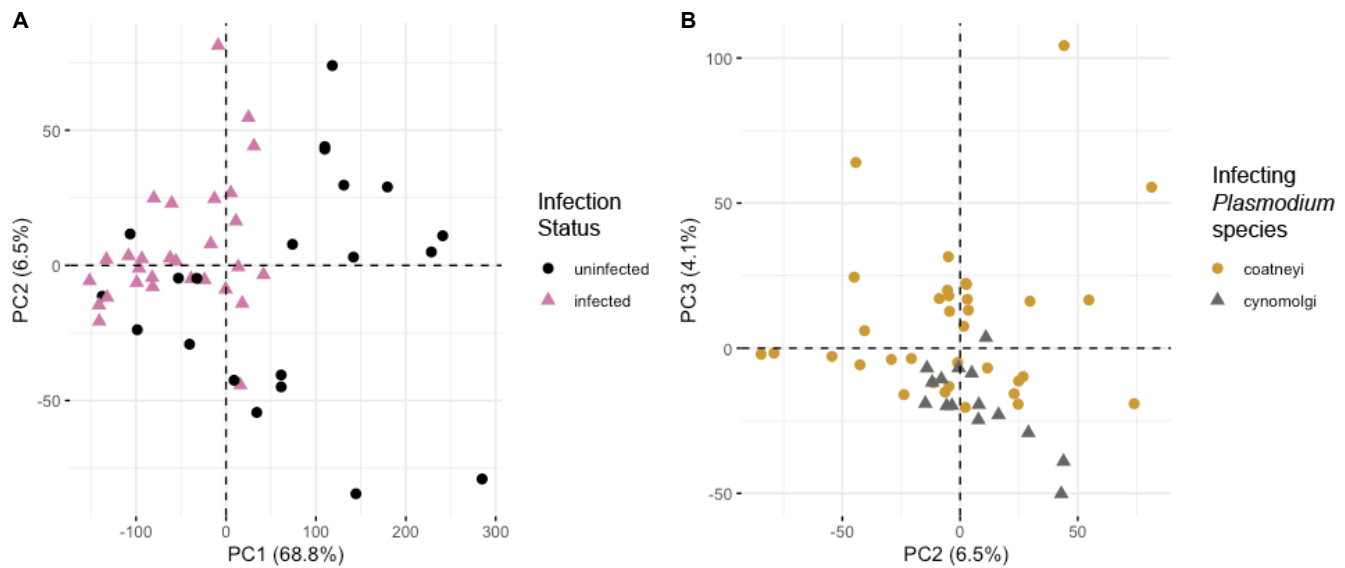


Figure 3. Visualization of Principal Components Analysis (PCA) of macaque gene expression. **(A)** The first principal component, explaining 68.8% of the variation, corresponded approximately to parasitemia, distinguishing infected (pink triangles) versus non-infected (black circles) samples. **(B)** The third principal component, explaining 4.1% of the variation, partially distinguished samples infected with differing *Plasmodium* species: *P. coatneyi* (yellow circles) versus *P. cynomolgi* (gray triangles).

and *Plasmodium* species are included in the supplementary tables (Tables S3-S4).

Malaria parasite response across infection

To explore further the relationship between the host and pathogen during infection, we identified orthologous genes of the two *Plasmodium* species (Tables S5-S6) and then tested for those with expression correlated to parasite load, the infecting *Plasmodium* species, or the interaction between parasite load and *Plasmodium* species (Table S7). Here, we report the *P. coatneyi* gene accession for the orthologous pairs. For the parasite load variable we found no significantly up-regulated genes but 151 down-regulated *Plasmodium* genes including PCOAH_00027310 (LogFC= -7.224; adj. $p=1.36 \times 10^{-5}$), which encodes an ATPase and has been shown to be essential for invasion, as *Plasmodium* uses ATP to determine which host cell to invade (Gunalan et al., 2013) and PCOAH_00031020 (LogFC=-6.947; adj. $p=3.66 \times 10^{-5}$), which encodes a highly conserved protease that is involved in the release of merozoite from erythrocyte and cell invasion (Withers-Martinez et al., 2012). For the *Plasmodium* species predictor, we found 351 *P. cynomolgi*-associated genes (e.g., PCOAH_00025970: LogFC=4.722; adj. $p=3.28 \times 10^{-13}$) and 310 *P. coatneyi*-associated genes (e.g., PCOAH_00034280: LogFC=-5.034; adj. $p=1.76 \times 10^{-14}$). Finally, for the interaction variable (parasite load**Plasmodium* species) we found no significantly up-regulated genes and only 20 down-regulated genes (e.g., PCOAH_00021590: LogFC=-13.480; adj. $p=6.133 \times 10^{-4}$); and PCOAH_00041960: LogFC=-17.955; adj. $p=6.05 \times 10^{-5}$).

In order to further explore the relationships between gene expression and biological factors, we conducted PCA and t-SNE analyses. The PCA clustered samples according to the degree of parasite load (PC1) and infecting *Plasmodium* species (PC2) (Fig. 4a). The t-SNE analysis suggested that *P. coatneyi* expression was most similar to that of *P. cynomolgi* when parasite load was high (Fig. 4b).

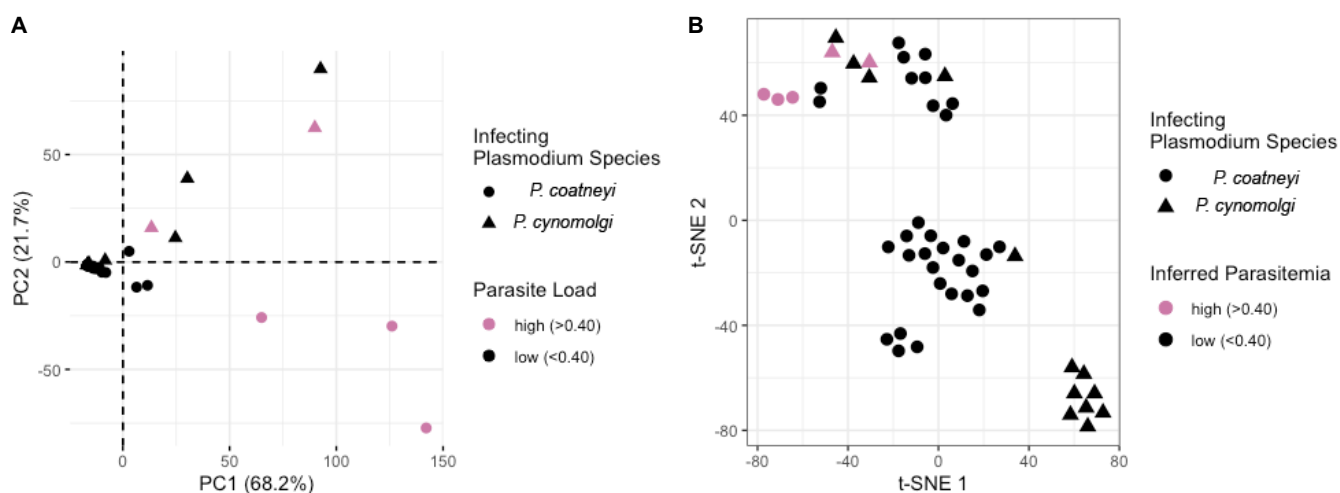


Figure 4. (A) Visualization of Principal Components Analysis (PCA) of *Plasmodium* gene expression. The first principal component, explaining 68.2% of the variation, captured parasite load, distinguishing high (pink) versus low (black) parasite load samples. The second principal component, explaining 21.7% of the variation, distinguished samples infected with differing *Plasmodium* species: *P. coatneyi* (circles) versus *P. cynomolgi* (triangles). (B) Visualization of results of t-Distributed Stochastic Neighbor Embedding (t-SNE) analyses on macaque gene expression with perplexity of nine. Samples clustered largely by infecting *Plasmodium* species: *P. coatneyi* (circles) versus *P. cynomolgi* (triangles). Of note, *P. coatneyi*-infected macaque samples with the highest parasite load (pink circles) clustered with *P. coatneyi*.

Over-represented functions of malaria genes linked to parasite load

We conducted functional enrichment analyses using annotations from PlasmoDB (Bahl et al. 2003) for the orthologous gene pairs in *P. coatneyi* and *P. cynomolgi* that were significantly differentially expressed for parasite load, *Plasmodium* species, or their interaction (Table S8). Parasite load-DE genes were enriched for reproduction (GO:0000003; p-value=0.003) and microtubule-based movement (GO:0007018; p-value= 5.51×10^{-5}). For the *Plasmodium* species-DE genes, we found that *P. cynomolgi*-associated genes were enriched for multiple binding molecular functions including ion binding (GO:0043167; p-value= 6.96×10^{-6}) and microtubule binding (GO:0008017; p-value=0.009), and *P. coatneyi*-associated genes were enriched for transferase activities including those involved in one-carbon (GO:0016741; p-value=0.030) and nitrogenous groups (GO:0016769; p-value=0.0466). Finally, for genes DE for the interaction of *Plasmodium* species and parasite load, *i.e.*, those that respond differentially to high parasite load between species, these DE *Plasmodium* genes were enriched for integral (GO:0016021; p-value=0.017) and intrinsic (GO:0031224; p-value=0.017) components of the membrane.

Discussion

Primates (including humans) and malaria have been co-evolving for thousands of years, leading to the evolution of both shared and species-specific coping strategies in both host and parasite. These primate-parasite relationships have created an opportunity to study model non-human primates in order to understand the biological mechanisms that are involved in human infections that differ in severity and symptomatology. For example, macaques infected with *Plasmodium coatneyi* serve as a useful model for human *P. falciparum* infections, particularly for the study of cerebral malaria^{35–37}, and for malaria during pregnancy²⁰. In contrast, macaques infected with *P. cynomolgi* have been used to model human *P. vivax* infections, particularly relapsing malaria in which both *Plasmodium* species can remain dormant in the liver as hypnozoites⁴. The differential symptoms are also important factors in determining the risk of zoonotic spillover to humans, as infections that are sufficiently benign to not impact behavior may be hypothesized to pose a

greater spillover risk. This is in accordance with observations for these macaque malarias, as spillover appears rare for the more dangerous *P. coatneyi* but increasingly observed for the lower severity *P. cynomolgi*. This differential severity and spillover risk motivated our comparison of macaque host response to these two parasites.

As has been previously established for human patients infected with *P. falciparum* (e.g.,³⁸), we confirmed that the proportion of whole blood RNA sequencing reads that map to the *Plasmodium* genome can be successfully used as a proxy for both parasitemia (presence or absence of parasite) and parasite load (proportion of infected red blood cells in macaques infected with *P. coatneyi* and *P. cynomolgi*). Owing to the release of microscopy data by the MaHPIC project, we were able to validate this measure in macaques to demonstrate that our bioinformatic proxy: *i.*) was tightly correlated with the traditionally-used microscopic measure of parasitemia/parasite load, and *ii.*) mirrored parasite load levels over the infection period. These findings support the use of this measure in future studies of wild populations for which microscopic measures of parasite load are infeasible. As sequencing data generation becomes increasingly inexpensive, bioinformatic inference of parasitemia or parasite load may make this computational method preferable to labor intensive microscopy.

Across the two *Plasmodium* species, we identified many significantly differentially expressed genes in the macaque host correlated with parasite load and parasitemia associated with immune function, red blood cell function, and morphology. Many of these genes were similarly linked to malaria infection in humans, particularly those we identified as correlated with the magnitude of parasite load or parasitemia. Such up-regulated genes include *ADAMDEC1* which is involved in inflammatory response during placental *P. falciparum* malaria³⁹, *SLCO2B1* which has been associated with parasitemia clearance rate in human *P. vivax* infection⁴⁰, and *CA4* which is negatively correlated with hemoglobin concentrations in children with *P. falciparum* malaria⁴¹. Similarly down-regulated genes were often implicated in human malaria response. These include *RABGAP1L* which has roles in hematopoiesis⁴² and has been found to have been under selection⁴³; *LUC7L* which contains SNPs associated with malaria prevalence in a Chinese population and with the blood disorder α -thalassemia that is protective against malaria⁴⁴; and *ADRB2* which contains SNPs significantly associated with malaria in a Southern Indian population⁴⁵. Such overlap between rhesus macaque genes responsive to malaria and human genes linked to malaria susceptibility emphasize the similarity across these two primates in response to diverse malaria parasites and supports the suitability of these non-human primate models for research aiming to improve our fundamental understanding of malaria.

The parallel infection of rhesus macaques with *Plasmodium coatneyi* and *P. cynomolgi* allowed us comparative insight into malaria species-level differences that may mirror those seen between human *P. falciparum* and *P. vivax* infections. Although both macaque malarias are more closely related to *P. vivax*, *P. coatneyi* infection of macaques results in *falciparum*-like cerebral malaria and associated high mortality. Consistent with this severity, genes involved in the macaque response to *P. coatneyi* specifically (in contrast to *P. cynomolgi*) were significantly enriched for relationships to brain morphology and function such as subcortical cerebral atrophy. No phenotypes were enriched for *P. cynomolgi* species-specific response genes, in contrast. This association between *P. coatneyi* and brain-associated functions indicates that this *Plasmodium* species may be inducing transcriptomic response that is similar to the comparable *P. falciparum* infection in humans, supporting use of this model in studies of cerebral malaria.

There are several noteworthy limitations to our study, which was possible only through the generous data sharing of the original MaHPIC project. The most important caveat is that *Plasmodium* species and experimental trial were conflated in our analysis. Although some of the impact of technical variation can be mitigated during bioinformatic library normalization, it is likely that some differentially expressed genes which we attribute to malaria species-specific response are instead due to differences in experimental setup, macaque genetics, or biases introduced during RNA extraction, cDNA synthesis, library preparation, or sequencing. An improved experimental design to identify malaria species-specific response could be structured to minimize these confounding factors in the future. However, our prior

hypothesis of *falciparum*-like nervous system involvement in *P. coatneyi* infections lends support to our finding of a broad enrichment of brain-associated functions, which is unlikely attributable to technical artifacts.

As treatment regimens varied between the studies, were subcurative, and were timed to coincide with high parasite load, several genes and pathways we identified as being differentially expressed in response to malaria may instead be linked to anti-malaria drug response. For instance, we found *CYP2C76* to be upregulated in response to parasite load which has previously been attributed to metabolism of the anti-malarial mefloquine in crab eating macaques (*M. fascicularis*)⁴⁶. Additionally, genes that were differentially responsive to *P. coatneyi* parasitemia were enriched for vestibular hypofunction and abnormal cochlear morphology, which may reflect the acute ototoxicity commonly associated with certain antimalarials^{47,48}. Here again, malaria species and antimalarial drug are confounded: The *P. coatneyi*-infected macaques were given subcurative and curative treatments of artemether alone while the *P. cynomolgi*-infected macaques were given subcurative blood-stage treatment of artemether and curative treatment of chloroquine and primaquine. Quinoline derivatives such as chloroquine and primaquine are known to demonstrate ototoxicity, but the risk for artemisinin derivatives like artemether remains controversial⁴⁷. Future study of macaque transcriptomic response to antimalarials may support our findings and could lead to insight into the pathways involved in ototoxicity and identification of biomarkers for this important side-effect.

Overall, our investigation into the gene expression response of a primate model to two malaria parasites of emerging concern identified many similarities with the comparatively well-studied human malaria response. This common architecture of *Plasmodium* transcriptomic response supports the use of experimentally-infected macaques to better understand fundamental primate malaria pathophysiology to improve human treatment. Similarly our comparative study revealed brain-associated differences between the macaque response to malaria species *P. coatneyi* and *P. cynomolgi*, supporting the use of the former in models of *P. falciparum* cerebral malaria. Future comparative studies of non-human primate response to malaria and malaria-like parasites can benefit from the sequencing read-based proxy, which we newly validated in these species. Such “dual-transcriptomic” research that simultaneously assays host and parasite response will allow insight into primate-parasite coevolution and may identify new therapeutic targets for human malaria. Finally, there is pressing practical motivations to study these parasites in particular: Although *P. coatneyi* and *cynomolgi* are traditionally deemed simian malaria species, their recent identification in human blood samples heightens the necessity of understanding their interaction with their primate hosts, both macaque and human.

Methods

Experimental set-up and dataset

We analyzed RNA sequence (mRNA-seq) data previously generated by the Malaria Host-Pathogen Interaction Center (MaHPIC)^{6,22}. The datasets used were generated from 50 whole blood samples of 5 malaria-naïve rhesus macaque (*Macaca mulatta*) individuals that had been experimentally inoculated with *Plasmodium coatneyi* Hackeri strain (study PRJNA400695) and 5 malaria-naïve rhesus macaque individuals that had been experimentally inoculated with *Plasmodium cynomolgi* strain B (study PRJNA388645). Methods for sample collection, pathogen inoculation, RNA extraction, library preparation, and transcriptome sequencing are fully described in the Gene Expression Omnibus page (GSE103259 and GSE99486, respectively) for these experiments.

Briefly, for the first experiment, researchers in the MaHPIC consortium⁶ infected 5 adult malaria-naïve male macaques with *P. coatneyi* sporozoites from the salivary glands of *Anopheles dirus*, *An. gambiae*, and *An. stephensi* mosquitoes, which were produced and isolated at the Centers for Disease Control and Prevention. These macaques were then monitored for 101-days. Subcurative treatment of artemether was administered to all subjects at the initial peak of infection and complete blood-treatment with artemether was given at the end of the experiment. Only one individual

(RCs13) did not develop infection and was not given additional subcurative treatments for subsequent recrudescence peaks. Whole blood samples were taken at 7 time-points over the 101-day period (35 samples total). Parasite load for each time point was measured as follows: 1) pre-infection or uninfected control; 2) acute infection or the peak infection as determined by symptomatic state and parasitological assessment; 3) early post-subcurative treatment or the earliest observation after peak infection and treatment; 4) late post-subcurative treatment; 5) early chronic or observation of low levels of persistent parasitemia; 6) late chronic; and 7) final measurement at end of experimental period.

For the second experiment, researchers in the MaHPIC consortium²² infected 5 adult malaria-naïve male macaques with *P. cynomolgi* sporozoites from the salivary glands of *Anopheles dirus*, *An. gambiae*, and *An. stephensi* mosquitoes, which were produced and isolated at the Centers for Disease Control and Prevention. These macaques were also monitored for 101-days and subcurative treatment was administered to all subjects at the initial peak of infection. One individual (RFv13) was euthanized during the experiment due to clinical complications of malaria and only two blood samples were taken. Whole blood samples were taken at 7 time-points over the 101-day period (30 samples total; 29 from the surviving subjects.) RNAseq data was not collected for RSb14 at time point 5, and we exclude a blood sample taken from the transfusion donor that was used for individual RFv13 to treat severe clinical complications. For the present study, individuals that experienced phenotypically severe malaria (RMe14 and RFa14) were also excluded as they were given alternative treatments to cope with clinical complications. Blood was sampled and parasite load was measured across the following time points: 1) pre-infection control; 2) at acute infection or the peak infection as determined by symptomatic state and parasitological assessment; 3) 7 days post-peak infection; 4) during relapses; 5) during inter-relapse intervals; and 6) at the end of the experimental period.

Gene expression inference

For both RNA-seq datasets we removed Illumina-specific adapter regions (using parameters ILLUMINACLIP:TruSeq3-PE.fa:2:30:10) and low quality bases (quality score < 3) at the beginning and end of each raw sequencing read, and trimmed the read when the average quality of bases in a sliding 4-base wide window fell below 15 using Trimmomatic⁴⁹. All statistical analyses were performed in R⁵⁰ unless otherwise stated. We removed reads that were less than 36 basepairs after trimming. In order to determine if sequence reads were from the host, we performed a competitive mapping analysis as follows. Using the STAR alignment program⁵¹, we mapped trimmed, quality-filtered reads to concatenated reference genomes that include rhesus macaque (GCF_003339765.1.Mmul_10) and parasite (*P. coatneyi*: GCF_001680005.1_ASM168000v1⁵²; or *P. cynomolgi*: GCF_000321355.1_PcynB⁵³), allowing each read to map multiple locations with up to eight mismatches per 100 basepair paired-end read (after⁵⁴). The best matching, primary alignments were extracted and then merged into pathogen and host “unique” datasets (*i.e.*, primary reads that have only mapped to the host or parasite), while excluding unplaced scaffolds for macaque.

We used Rsubread⁵⁵ to create a macaque and *Plasmodium* read-count matrix for each host individual, using the “uniquely” mapped host and pathogen reads, respectively, and allowing for overlapping reads. Reads for each analysis were tallied at the level of exons which were then summed to create a per-gene count for the sample, using gene annotations (GCF_003339765.1.Mmul_10 and GCF_001680005.1_ASM168000v1 for the first dataset, and GCF_003339765.1.Mmul_10 and GCF_000321355.1_PcynB for the second dataset).

To compare *Plasmodium* gene expression between the two experiments (the first using *P. coatneyi* and the second, *P. cynomolgi*) from the two macaque cohorts, we combined the above *Plasmodium* read count matrices matched by orthologous genes. We identified orthologous genes using orthologue groups from PlasmoDB⁵⁶ and excluded any genes for which orthologous genes between the two *Plasmodium* species did not exist. This combined read count matrix was then used to explore *Plasmodium* differentially expressed genes responding to parasite load.

Correlation of sequencing read- and microscopy-based parasite load inference

We inferred a proxy of parasite load (“sequencing read-based parasite load”) for all samples by calculating the ratio of the number of *Plasmodium* pathogen reads to the number of host macaque reads recovered (after^{38,57}). To validate this metric, we tested for a correlation between our parasite load proxy and previously published microscopy-based parasite load calculated from thick and thin blood smears for these samples. Microscopy-based parasite load was defined as parasites count per microliter of whole blood taken on the same day as RNA data was collected. We tested whether our inferred parasitemia proxy and the original studies’ microscopy-based estimates were significantly correlated using a linear regression model comparing our calculated parasite load against the microscopy-based estimate (parasites per microliter) and by computing Spearman’s correlation coefficient.

Identification of differentially expressed host genes across two measures of pathogen infection and infecting malaria species

We next identified rhesus macaque genes with expression correlated to either parasite load (relative amount of parasite in blood) or parasitemia (presence of parasite in blood), as well as the species of infecting *Plasmodium*. To do so, we estimated the sequencing read-based parasite load (hereafter “parasite load”) as the ratio of pathogen to host reads, and sequencing read-based parasitemia (hereafter “parasitemia”), which we defined as positive if the proxy were greater than 0.001 with confirmation from microscopy. Using the combined read count matrix (with data from PRJNA400695 and PRJNA388645), we first calculated normalization factors and filtered for lowly expressed genes, defined as any gene with expression below a counts per million (CPM) threshold of 1 in at least 1 individual for macaque genes using the Bioconductor package edgeR⁵⁸. With these calculated normalization factors, we then used the limma-voom package^{59,60} to include biological replicates as a random effect (i.e., blocking variable) and each parasitemia measure (parasite load and parasitemia) as a fixed effect in the following models:

read count \sim parasite load + (1 | individual) + *Plasmodium* species + parasite load**Plasmodium* species + correction factors

read count \sim parasitemia + (1 | individual) + *Plasmodium* species + parasitemia**Plasmodium* species + correction factors

where correction factors include those adjusting for differences in number of reads and library complexity. We extracted the coefficients and adjusted p -values for all predictors for each gene from this model, correcting for multiple tests via FDR. We considered a gene to be a parasitemia-DE gene if the adjusted p -value was less than 0.05.

To explore potential biological relationships (e.g., between parasite load or parasitemia and infecting *Plasmodium* species) using normalized and filtered read counts, we performed dimensional reduction analyses using PCA and t-SNE.

We next tested whether macaque DE genes with expression that was significantly positively correlated (up-regulated DE genes) or significantly negatively correlated (down-regulated DE genes) with the measure of infection (i.e., parasite load or parasitemia), infecting *Plasmodium* species, and the interaction between the measure of infection and the infecting *Plasmodium* species were enriched for particular biological functions. Specifically, we tested whether these genes were more likely than expected by chance to share molecular functions, biological processes, or cell component annotations from the Gene Ontology database^{61,62}. For each dataset, we also tested for functional enrichment of Human Phenotype Ontology⁶³ annotations, using annotations for human orthologs matched to our macaque genes using the Ensembl gene mappings (accessed via g:Orth⁶⁴). To test for enrichment, we used a hypergeometric test as implemented in g:Profiler⁶⁴ and used all annotated genes after filtration for low read count as a comparative background list for differential expression determination. We finally tested whether these genes were more likely than expected by

chance to also be differentially expressed in human response to *Plasmodium* parasites (*i.e.*, *P. vivax* and *P. falciparum*³⁴). We did so using a hypergeometric test after converting our parasite species-DE genes to their one-to-one human orthologs in g:Profiler⁶⁴.

Identification of differentially expressed *Plasmodium* genes between malaria species

To identify *Plasmodium* genes with expression correlated to parasite load, we analyzed the combined datasets (PR-JNA400695 and PRJNA388645) after collapsing *Plasmodium* genes into orthologous groups based on one-to-one orthologs between *P. coatneyi* and *P. cynomolgi* (Tables S5-S6), which were gathered from PlasmoDB⁵⁶. Using the combined read count matrix, we first calculated the normalization factors and, because there were very few *Plasmodium* reads, filtered for lowly expressed genes using the filterByExpr function of edgeR which considers datasets relative to their design matrix⁵⁸. Using these calculated normalization factors, we then used the limma-voom package^{59,60} to include biological replicates as a random effect (*i.e.*, blocking variable) and parasite load as a fixed effect in the following model:

read count \sim parasite load + (1 | individual) + *Plasmodium* species + parasite load**Plasmodium* species + correction factors

where, as above, correction factors include those adjusting for differences in number of reads and library complexity. We corrected for multiple tests via FDR and again considered a gene to be a parasitemia-DE gene if the adjusted *p*-value was less than 0.05. We then used Empirical Bayes *t*-statistics to identify significantly up- and down-regulated parasitemia-response genes.

Finally, to investigate potential biological relationships between parasite load and *Plasmodium* species identity using normalized and filtered read counts, we performed dimensional reduction analyses including PCA and t-SNE analyses.

Similar to the method used for macaque genes, we next tested whether *P. coatneyi*-*P. cynomolgi* orthologous gene pairs with expression that was significantly positively correlated (up-regulated) or significantly negatively correlated (down-regulated) with the measure of infection (*i.e.*, parasite load or parasitemia), *Plasmodium* species, or the interaction between the measure of infection and the *Plasmodium* species were enriched for particular functional annotations using Gene Ontology analysis as implemented in PlasmoDB⁵⁶.

Code used in analyses is available at https://github.com/ambertrujillo/Comparative_macaque.

References

1. Nosten, F., Richard-Lenoble, D. & Danis, M. A brief history of malaria. *Press. Med.* **51**, 104130, DOI: [10.1016/j.lpm.2022.104130](https://doi.org/10.1016/j.lpm.2022.104130) (2022).
2. Chin, W., Contacos, P. G., Robert Coatney, G. & Kimball, H. R. A naturally acquired quotidian-type malaria in man transferable to monkeys, DOI: [10.1126/science.149.3686.865.a](https://doi.org/10.1126/science.149.3686.865.a) (1965).
3. Singh, B. *et al.* A large focus of naturally acquired *Plasmodium knowlesi* infections in human beings. *Lancet* **363**, 1017–1024, DOI: [10.1016/S0140-6736\(04\)15836-4](https://doi.org/10.1016/S0140-6736(04)15836-4) (2004).
4. Galinski, M. R. Functional genomics of simian malaria parasites and host–parasite interactions. *Brief. Funct. Genomics* **18**, 270–280, DOI: [10.1093/bfpg/elz013](https://doi.org/10.1093/bfpg/elz013) (2019).
5. van de Straat, B. *et al.* Zoonotic malaria transmission and land use change in Southeast Asia: what is known about the vectors. *Malar. J.* **21**, 109, DOI: [10.1186/s12936-022-04129-2](https://doi.org/10.1186/s12936-022-04129-2) (2022).
6. Cordy, R. J. *et al.* Distinct amino acid and lipid perturbations characterize acute versus chronic malaria. *JCI Insight* **4**, DOI: [10.1172/jci.insight.125156](https://doi.org/10.1172/jci.insight.125156) (2019).

7. Fonseca, L. L. *et al.* Analysis of erythrocyte dynamics in rhesus macaque monkeys during infection with *Plasmodium cynomolgi*, DOI: [10.1186/s12936-018-2560-6](https://doi.org/10.1186/s12936-018-2560-6) (2018).
8. Gupta, A., Styczynski, M. P., Galinski, M. R., Voit, E. O. & Fonseca, L. L. Dramatic transcriptomic differences in *Macaca mulatta* and *Macaca fascicularis* with *Plasmodium knowlesi* infections. *Sci. Rep.* **11**, 19519, DOI: [10.1038/s41598-021-98024-6](https://doi.org/10.1038/s41598-021-98024-6) (2021).
9. Joyner, C. J. *et al.* Humoral immunity prevents clinical malaria during *Plasmodium* relapses without eliminating gametocytes. *PLoS Pathog.* **15**, e1007974, DOI: [10.1371/journal.ppat.1007974](https://doi.org/10.1371/journal.ppat.1007974) (2019).
10. Tang, Y. *et al.* Integrative analysis associates monocytes with insufficient erythropoiesis during acute *Plasmodium cynomolgi* malaria in rhesus macaques. *Malar. J.* **16**, 384, DOI: [10.1186/s12936-017-2029-z](https://doi.org/10.1186/s12936-017-2029-z) (2017).
11. Tang, Y. *et al.* Metabolic modeling helps interpret transcriptomic changes during malaria. *Biochim. Biophys. Acta Mol. Basis Dis.* **1864**, 2329–2340, DOI: [10.1016/j.bbadis.2017.10.023](https://doi.org/10.1016/j.bbadis.2017.10.023) (2018).
12. Tang, Y. *et al.* Multi-omics integrative analysis of acute and relapsing malaria in a non-human primate model of *P. vivax* infection (2019). DOI: [10.1101/564195](https://doi.org/10.1101/564195).
13. Ylostalo, J. *et al.* Transcriptome profiles of host gene expression in a monkey model of human malaria. *J. Infect. Dis.* **191**, 400–409, DOI: [10.1086/426868](https://doi.org/10.1086/426868) (2005).
14. Gupta, A., Galinski, M. R. & Voit, E. O. Dynamic control balancing cell proliferation and inflammation is crucial for an effective immune response to malaria. *Front Mol Biosci* **8**, 800721, DOI: [10.3389/fmolb.2021.800721](https://doi.org/10.3389/fmolb.2021.800721) (2022).
15. Peterson, M. S. *et al.* Clinical recovery of *Macaca fascicularis* infected with *Plasmodium knowlesi*. *Malar. J.* **20**, 486, DOI: [10.1186/s12936-021-03925-6](https://doi.org/10.1186/s12936-021-03925-6) (2021).
16. Taghavian, O. *et al.* Antibody profiling by proteome microarray with multiplex isotype detection reveals overlap between human and *Aotus nancymae* controlled malaria infections. *Proteomics* **18**, e1870115, DOI: [10.1002/pmic.201870115](https://doi.org/10.1002/pmic.201870115) (2018).
17. Anderson, D. C., Lapp, S. A., Barnwell, J. W. & Galinski, M. R. A large scale *Plasmodium vivax*–*Saimiri boliviensis* trophozoite-schizont transition proteome. *PLoS One* **12**, e0182561, DOI: [10.1371/journal.pone.0182561](https://doi.org/10.1371/journal.pone.0182561) (2017).
18. Peterson, M. S. *et al.* *Plasmodium vivax* parasite load is associated with histopathology in *Saimiri boliviensis* with findings comparable to *P. vivax* pathogenesis in humans. *Open Forum Infect Dis* **6**, ofz021, DOI: [10.1093/ofid/ofz021](https://doi.org/10.1093/ofid/ofz021) (2019).
19. Galinski, M. R. & Barnwell, J. W. Nonhuman primate models for human malaria research, DOI: [10.1016/b978-0-12-381366-4.00005-5](https://doi.org/10.1016/b978-0-12-381366-4.00005-5) (2012).
20. Lombardini, E. D., Gettayacamin, M., Turner, G. D. H. & Brown, A. E. A review of *Plasmodium coatneyi*–macaque models of severe malaria. *Vet. Pathol.* **52**, 998–1011, DOI: [10.1177/0300985815583098](https://doi.org/10.1177/0300985815583098) (2015).
21. Joyner, C. *et al.* *Plasmodium cynomolgi* infections in rhesus macaques display clinical and parasitological features pertinent to modelling *vivax* malaria pathology and relapse infections. *Malar. J.* **15**, 451, DOI: [10.1186/s12936-016-1480-6](https://doi.org/10.1186/s12936-016-1480-6) (2016).
22. DeBarry, J. D. *et al.* MaHPIC malaria systems biology data from *Plasmodium cynomolgi* sporozoite longitudinal infections in macaques. *Sci Data* **9**, 722, DOI: [10.1038/s41597-022-01755-y](https://doi.org/10.1038/s41597-022-01755-y) (2022).
23. Voorberg-van der Wel, A., Kocken, C. H. M. & Zeeman, A.-M. Modeling relapsing malaria: Emerging technologies to study parasite-host interactions in the liver. *Front. Cell. Infect. Microbiol.* **10**, DOI: [10.3389/fcimb.2020.606033](https://doi.org/10.3389/fcimb.2020.606033) (2021).
24. Kotepui, M., Masangkay, F. R., Kotepui, K. U. & De Jesus Milanez, G. Preliminary review on the prevalence, proportion, geographical distribution, and characteristics of naturally acquired *Plasmodium cynomolgi* infection in mosquitoes, macaques, and humans: A systematic review and meta-analysis, DOI: [10.1186/s12879-021-05941-y](https://doi.org/10.1186/s12879-021-05941-y) (2021).
25. Zhang, X. *et al.* Distribution and prevalence of malaria parasites among long-tailed macaques (*Macaca fascicularis*) in regional populations across Southeast Asia. *Malar. J.* **15**, 1–8, DOI: [10.1186/s12936-016-1494-0](https://doi.org/10.1186/s12936-016-1494-0) (2016).
26. Kojom Foko, L. P., Kumar, A., Hawadak, J. & Singh, V. *Plasmodium cynomolgi* in humans: Current knowledge and future directions of an emerging zoonotic malaria parasite. *Infection* 1–18, DOI: [10.1007/s15010-022-01952-2](https://doi.org/10.1007/s15010-022-01952-2) (2022).

27. Imwong, M. *et al.* Asymptomatic natural human infections with the simian malaria parasites *Plasmodium cynomolgi* and *Plasmodium knowlesi*. *J. Infect. Dis.* **219**, 695–702, DOI: [10.1093/infdis/jiy519](https://doi.org/10.1093/infdis/jiy519) (2019).
28. Singh, B. *et al.* Naturally acquired human infections with the simian malaria parasite, *Plasmodium cynomolgi*, in Sarawak, Malaysian Borneo. *Int. J. Infect. Dis.* **73**, 68, DOI: [10.1016/j.ijid.2018.04.3581](https://doi.org/10.1016/j.ijid.2018.04.3581) (2018).
29. Ta, T. H. *et al.* First case of a naturally acquired human infection with *Plasmodium cynomolgi*, DOI: [10.1186/1475-2875-13-68](https://doi.org/10.1186/1475-2875-13-68) (2014).
30. Yap, N. J. *et al.* Natural human infections with *Plasmodium cynomolgi*, *P. inui*, and 4 other simian malaria parasites, Malaysia. *Emerg. Infect. Dis.* **27**, 2187–2191, DOI: [10.3201/eid2708.204502](https://doi.org/10.3201/eid2708.204502) (2021).
31. Coatney, G. R. *The Primate Malaria* (U.S. National Institute of Allergy and Infectious Diseases, 1971).
32. Galinski, M. R. Systems biology of malaria explored with nonhuman primates. *Malar. J.* **21**, 177, DOI: [10.1186/s12936-022-04199-2](https://doi.org/10.1186/s12936-022-04199-2) (2022).
33. Yan, Y. H., Moncada, D. M., Trippe, E. D. & Gutierrez, J. B. Correlates of severity of disease in *Macaca mulatta* infected with *Plasmodium cynomolgi*. (2017). [1706.08836](https://doi.org/10.1101/1706.08836).
34. Bobowik, K. *et al.* Transcriptomic profiles of *Plasmodium falciparum* and *Plasmodium vivax*-infected individuals in Indonesia (2021). DOI: [10.1101/2021.01.07.425684](https://doi.org/10.1101/2021.01.07.425684).
35. Aikawa, M. *et al.* A primate model for human cerebral malaria: *Plasmodium coatneyi*-infected rhesus monkeys. *Am. J. Trop. Med. Hyg.* **46**, 391–397, DOI: [10.4269/ajtmh.1992.46.391](https://doi.org/10.4269/ajtmh.1992.46.391) (1992).
36. Kawai, S., Aikawa, M., Kano, S. & Suzuki, M. A primate model for severe human malaria with cerebral involvement: *Plasmodium coatneyi*-infected *Macaca fuscata*, DOI: [10.4269/ajtmh.1993.48.630](https://doi.org/10.4269/ajtmh.1993.48.630) (1993).
37. Sugiyama, M. *et al.* Cerebral metabolic reduction in severe malaria: fluorodeoxyglucose-positron emission tomography imaging in a primate model of severe human malaria with cerebral involvement. *Am. J. Trop. Med. Hyg.* **71**, 542–545 (2004).
38. Yamagishi, J. *et al.* Interactive transcriptome analysis of malaria patients and infecting *Plasmodium falciparum*. *Genome Res.* **24**, 1433–1444, DOI: [10.1101/gr.158980.113](https://doi.org/10.1101/gr.158980.113) (2014).
39. Muehlenbachs, A., Mutabingwa, T. K., Edmonds, S., Fried, M. & Duffy, P. E. Hypertension and maternal–fetal conflict during placental malaria. *PLoS Med.* **3**, e446, DOI: [10.1371/journal.pmed.0030446](https://doi.org/10.1371/journal.pmed.0030446) (2006).
40. Sortica, V. A. *et al.* *SLCO1A2*, *SLCO1B1* and *SLCO2B1* polymorphisms influences chloroquine and primaquine treatment in *Plasmodium vivax* malaria. *Pharmacogenomics* **18**, 1393–1400, DOI: [10.2217/pgs-2017-0077](https://doi.org/10.2217/pgs-2017-0077) (2017).
41. Boldt, A. B. W. *et al.* The blood transcriptome of childhood malaria, DOI: [10.1016/j.ebiom.2018.12.055](https://doi.org/10.1016/j.ebiom.2018.12.055) (2019).
42. Roberti, M. C. *et al.* *RABGAP1L* gene rearrangement resulting from a der(Y)t(Y;1)(q12;q25) in acute myeloid leukemia arising in a child with Klinefelter syndrome. *Virchows Arch.* **454**, 311–316, DOI: [10.1007/s00428-009-0732-z](https://doi.org/10.1007/s00428-009-0732-z) (2009).
43. Tennesen, J. A. & Duraisingh, M. T. Three signatures of adaptive polymorphism exemplified by malaria-associated genes, DOI: [10.1093/molbev/msaa294](https://doi.org/10.1093/molbev/msaa294) (2021).
44. Ren, N. *et al.* High incidence of malaria along the Sino-Burmese border is associated with polymorphisms of *CR1*, *IL-1A*, *IL-4R*, *IL-4*, *NOS*, and *TNF*, but not with G6PD deficiency. *Medicine* **94**, e1681, DOI: [10.1097/MD.0000000000001681](https://doi.org/10.1097/MD.0000000000001681) (2015).
45. Saadi, A. V. *et al.* Single nucleotide polymorphisms of *ADRB2* gene and their association with susceptibility for *Plasmodium falciparum* malaria and asthma in an Indian population, DOI: [10.1016/j.meegid.2013.08.026](https://doi.org/10.1016/j.meegid.2013.08.026) (2013).
46. Uno, Y., Fujino, H., Kito, G., Kamataki, T. & Nagata, R. CYP2C76, a novel cytochrome P450 in cynomolgus monkey, is a major CYP2C in liver, metabolizing tolbutamide and testosterone. *Mol. Pharmacol.* **70**, 477–486, DOI: [10.1124/mol.106.022673](https://doi.org/10.1124/mol.106.022673) (2006).
47. Jozefowicz-Korczynska, M., Pajor, A. & Grzelczyk, W. L. The ototoxicity of antimalarial Drugs—A state of the art review, DOI: [10.3389/fneur.2021.661740](https://doi.org/10.3389/fneur.2021.661740) (2021).
48. Yu, D., Ding, D., Jiang, H., Stolzberg, D. & Salvi, R. Mefloquine damage vestibular hair cells in organotypic cultures. *Neurotox. Res.* **20**, 51–58, DOI: [10.1007/s12640-010-9221-z](https://doi.org/10.1007/s12640-010-9221-z) (2011).

49. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120, DOI: [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170) (2014).
50. R Core Team. R: A language and environment for statistical computing (2021).
51. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21, DOI: [10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635) (2013).
52. Chien, J.-T. *et al.* High-quality genome assembly and annotation for *Plasmodium coatneyi*, generated using Single-Molecule Real-Time PacBio technology. *Genome Announc.* **4**, DOI: [10.1128/genomeA.00883-16](https://doi.org/10.1128/genomeA.00883-16) (2016).
53. Tachibana, S.-I. *et al.* *Plasmodium cynomolgi* genome sequences provide insight into *Plasmodium vivax* and the monkey malaria clade. *Nat. Genet.* **44**, 1051–1055, DOI: [10.1038/ng.2375](https://doi.org/10.1038/ng.2375) (2012).
54. Lee, H. J. *et al.* Integrated pathogen load and dual transcriptome analysis of systemic host-pathogen interactions in severe malaria. *Sci. Transl. Med.* **10**, DOI: [10.1126/scitranslmed.aar3619](https://doi.org/10.1126/scitranslmed.aar3619) (2018).
55. Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, DOI: [10.1093/nar/gkz114](https://doi.org/10.1093/nar/gkz114) (2019).
56. Bahl, A. *et al.* PlasmoDB: The *Plasmodium* genome resource. a database integrating experimental and computational data. *Nucleic Acids Res.* **31**, 212–215, DOI: [10.1093/nar/gkg081](https://doi.org/10.1093/nar/gkg081) (2003).
57. Mukherjee, P., Burgio, G. & Heitlinger, E. Dual RNA sequencing meta-analysis in *Plasmodium* infection identifies host-parasite interactions. *mSystems* **6**, DOI: [10.1128/mSystems.00182-21](https://doi.org/10.1128/mSystems.00182-21) (2021).
58. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140, DOI: [10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616) (2010).
59. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29, DOI: [10.1186/gb-2014-15-2-r29](https://doi.org/10.1186/gb-2014-15-2-r29) (2014).
60. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47, DOI: [10.1093/nar/gkv007](https://doi.org/10.1093/nar/gkv007) (2015).
61. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29, DOI: [10.1038/75556](https://doi.org/10.1038/75556) (2000).
62. The Gene Ontology Consortium. The Gene Ontology resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338, DOI: [10.1093/nar/gky1055](https://doi.org/10.1093/nar/gky1055) (2019).
63. Köhler, S. *et al.* Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Res.* **47**, D1018–D1027, DOI: [10.1093/nar/gky1105](https://doi.org/10.1093/nar/gky1105) (2019).
64. Raudvere, U. *et al.* g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198, DOI: [10.1093/nar/gkz369](https://doi.org/10.1093/nar/gkz369) (2019).

Acknowledgements

We thank the MaHPIC researchers for making their data available. Funding provided by the Ford Foundation and N.S.F. BCS-2118108 to A.E.T.

Author contributions statement

A.E.T. conceived the study, analyzed the results, and wrote the manuscript. C.M.B. conceived the study, interpreted the results, and wrote the manuscript. All authors reviewed the manuscript.

Data availability statement

Data used in the present study are available from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) under accessions PRJNA400695 and PRJNA388645.

Additional information

The authors declare no competing interests.