Immunogenetic response to a malaria-like parasite in a wild primate

Amber E. Trujillo[1,2] and **Christina M. Bergey**[3,*]

[1] Department of Anthropology, New York University, 25 Waverly Place, New York, NY, 10003.
[2] New York Consortium in Evolutionary Primatology.
[3] Department of Genetics, Rutgers University, 604 Allison Road, Piscataway, NJ, 08840.

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

Abstract

Malaria is infamous for the massive toll it exacts on human life and health. In the face of this intense selection, many human populations have evolved mechanisms that confer some resistance to the disease, such as sickle-cell hemoglobin or the Duffy null blood group. Less understood are adaptations in other vertebrate hosts, many of which have a longer history of co-evolution with malaria parasites. By comparing malaria resistance adaptations across host species, we can gain fundamental insight into host-parasite co-evolution. In particular, understanding the mechanisms by which non-human primate immune systems combat malaria may be fruitful in uncovering transferable therapeutic targets for humans. However, most research on primate response to malaria has focused on a single or few loci, typically in experimentally-infected captive primates. Here, we report the first transcriptomic study of a wild primate response to a malaria-like parasite, investigating gene expression response of red colobus monkeys (*Piliocolobus tephroscles*) to natural infection with the malaria-like parasite, *Hepatocystis*. We identified colobus genes with expression strongly correlated with parasitemia, including many implicated in human malaria and suggestive of common genetic architecture of disease response. For instance, the expression of *ACKR1* (alias *DARC*) gene, previously linked to resistance in humans, was found to be positively correlated with parasitemia. Other similarities to human parasite response include induction of changes in immune cell type composition and, potentially, increased extramedullary hematopoiesis and altered biosynthesis of neutral lipids. Our results illustrate the utility of comparative immunogenetic investigation of malaria response in primates. Such inter-specific comparisons of transcriptional response to pathogens afford a unique opportunity to compare and contrast the adaptive genetic architecture of disease resistance, which may lead to the identification of novel intervention targets to improve human health.

Author Summary

The co-evolutionary arms race between humans and malaria parasites has been ongoing for millennia. Fully understanding the evolved human response to malaria is impossible without comparative study of parasites in our non-human primate relatives. Though laboratory primates are fruitful models, the complexity of wild primates infected in a natural transmission system may be a more suitable comparison for contextualizing malaria infections in human patients. Here, we investigate the genetic mechanisms

underlying the immune response to *Hepatocystis*, a close relative of human-infective malaria, in a population of wild Ugandan red colobus monkeys. We find that the genes involved have considerable overlap with those active in human malaria patients. Like *Plasmodium*, *Hepatocystis* induces changes in blood cell type and may cause the host to produce blood components outside of the bone marrow or alter metabolism related to the production of lipids. Our work helps to identify the genetic mechanisms underlying the arms race between primates and malaria parasites, providing fundamental evolutionary insight. Such comparative work on the interaction between wild non-human primates and malaria parasites can identify ways in which primates have evolved resistance to malaria parasites, and further investigation of such implicated genes may lead to novel potential therapeutic and vaccine targets.

Introduction

Though malaria parasites (order Haemosporida) infect a broad variety of vertebrates via a broad variety of dipteran insect vectors (Galen et al., 2018; Perkins, 2014), the five species that have caused incalculable human deaths have understandably garnered the most research attention. Evidence of this lengthy battle is evident from scars in the human genome, as malaria has exerted one of the strongest selection pressures on the human immune system (Kwiatkowski, 2005; Kariuki and Williams, 2020). Recent investigations simultaneously assaying human and malaria gene expression in the blood have illuminated the genetic architecture of this host-parasite interaction, permitting fundamental insight into the immunology of this disease (Yamagishi et al., 2014; Otto et al., 2018; Lee et al., 2018; Mukherjee and Heitlinger, 2019). However, a full understanding of the adaptive human immunogenetic response to malaria requires comparative data from closely related primates.

Few investigations of malaria and malaria-like parasite response in non-human primates have been undertaken, and fewer still have investigated gene expression response in free-ranging individuals in a natural pathogen transmission system. Most studies in non-human primates have instead made use of experimentally-infected laboratory primates (*e.g.*, rhesus macaques, *Macaca mulatta*; Ylostalo et al., 2005; Tang et al., 2017, 2018; Galinski, 2019; Tang et al., 2019). However, tightly controlled laboratory infections likely do not capture the complexity in parasite behavior and host response characterizing natural infections, as has been found for human patients (Daily et al., 2007). In one of the few studies of malaria in wild primates, variation in the *ACKR1* (alias *DARC*) gene was found to be associated with phenotypic diversity in susceptibility to *Hepatocystis* (a malaria-like parasite) in yellow baboons (*Papio cynocephalus*; Tung et al., 2009). High throughput transcriptomic profiling of host-parasite interactions in wild primates living in natural transmission systems is needed to fully understand the evolution of malaria response in primates.

Contained within the polyphyletic *Plasmodium* clade (Galen et al., 2018), members of genus *Hepatocystis* commonly parasitize bats and non-human primates. *Hepatocystis* is often highly prevalent in primates (20% to 90% across host species; Thurber et al., 2013; Ayouba et al., 2012), and parasitemia may persist for over a year without clinical symptoms (Vickerman, 2005) though reports exist of anemia and liver scarring due to monocytes infiltration (Garnham, 1966). At least one species is able to host-switch between primate species (Thurber et al., 2013), but no human infections have been reported. Though nested within the *Plasmodium* genus as a sister taxon to bat- and rodent-infecting *Plasmodium* in particular (Galen et al., 2018), *Hepatocystis* parasites differ in their apparently lack of asexual replication in red blood cells and their use of midges as vectors (Aunin et al., 2020; Garnham, 1966).

Here, we report the first study of genome-wide gene expression response to a

malaria-like parasite in wild non-human primates. We investigate the genetic architecture of response to *Hepatocystis* in a population of wild red colobus monkeys (*Piliocolobus tephrosceles*) living in Kibale National Park, Uganda. We identify colobus monkey and *Hepatocystis* genes and genetic pathways that show differential expression in response to *Hepatocystis* infection. We find considerable overlap between the genes involved in the colobus response to *Hepatocystis* and human response to *Plasmodium* malaria. Overall, our study illustrates the importance of comparative host-parasite genomic investigations for understanding the genetic architecture of human disease response.

Results and Discussion

The Ugandan red colobus monkey whole blood transcriptomic dataset was collected by Simons and colleagues (2019) as part of the effort to generate the red colobus monkey reference genome. The sampled individuals ($N = 29$) are from a habituated population in Kibale National Park, Uganda, where red colobus have been the subjects of many disease-related studies as a part of the the Kibale EcoHealth Project (Chapman et al., 2005, 2013; Goldberg et al., 2008; Lauck et al., 2011; Simons et al., 2019; Aunin et al., 2020). To identify colobus monkey genes with expression correlated to *Hepatocystis* parasite levels, we aligned blood transcriptome reads to a concatenated host-parasite pseudo-reference genome (containing genomes of red colobus and *Hepatocystis* sp.) and tallied reads by gene. After filtering, the median per-sample count of reads uniquely mapping to the colobus genome was 9,531,580 (range: 4,149,243 to 12,907,207). We next computed a proxy of *Hepatocystis* parasitemia as the ratio of reads mapped uniquely to the primate versus parasite genome (Fig. 1; Table S1). For *Hepatocystis* reads, we used published read counts for these samples (Aunin et al., 2020) excluding genes with expression in the top and bottom quartile to minimize the impact of genes with outlier expression.

Gene	$\log_2$FC	p-value	adj. p-value
UBE2K	1.84	4.37×10^{-6}	0.0176
PP2D1	3.88	1.23×10^{-5}	0.0300
TMEM167A	2.53	1.48×10^{-5}	0.0302
LSM14A	2.52	2.99×10^{-5}	0.0449
UBE2B	3.31	3.49×10^{-5}	0.0449
TTLL12	3.14	3.67×10^{-5}	0.0449
AGFG1	2.61	5.52×10^{-5}	0.0454
APOBEC2	6.36	5.53×10^{-5}	0.0454
NAGA	-4.63	2.18×10^{-6}	0.0133
TBC1D9B	-1.33	4.36×10^{-5}	0.0454
ZNF682	-5.46	5.39×10^{-5}	0.0454

Table 1. Red colobus monkey genes with expression most strongly correlated with parasitemia. The $\log_2$ fold change in response to inferred parasitemia ("$\log_2$FC") is computed per unit change in inferred parasitemia proxy. All genes lacking orthology information (*i.e.*, with identifiers beginning with "LOC") were excluded from this table. Genes with Benjamini-Hochberg FDR-adjusted p-value < 0.05 show. Full results available in Table S6.

***Hepatocystis* parasitemia correlates with proportions of monocytes and neutrophils.** To investigate the cellular composition of the red colobus immune
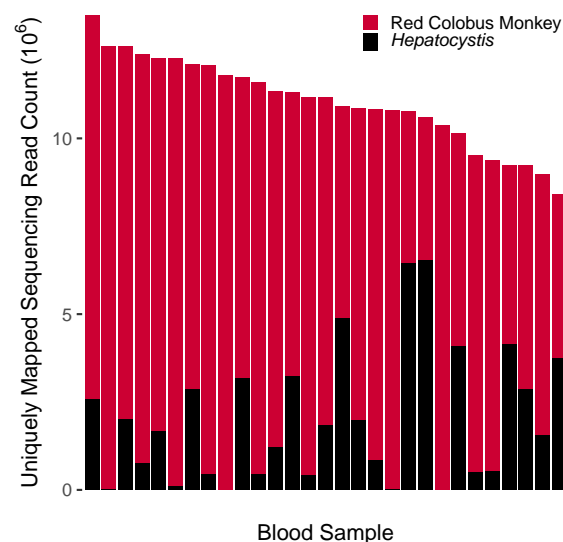
Figure 1. Uniquely mapped read counts from red colobus monkey (red) and *Hepatocystis* parasite (black) in monkey blood samples. *Hepatocystis* counts based on those of Aunin et al., 2020, excluding genes with expression in the highest and lowest quartiles. The ratio of *Hepatocystis* to colobus reads for each individual was used as a proxy for parasitemia.

response, we first tested if the relative abundance of red colobus monkey immune cell type subsets (Simons et al., 2019) correlated with inferred *Hepatocystis* parasitemia (Table S2; Fig. 2, S1). Though *Hepatocystis* differs from some *Plasmodium* malaria species in causing long-lasting, chronic infections (Garnham, 1966), our findings suggest *Hepatocystis* parasite load causes similar changes in blood cell counts. In *Plasmodium* malaria, such changes in hematologic parameters have been hypothesized to be associated with changes in the host immunity or ability to mount an effective immune response (Ingersoll et al., 2010).

The relative proportion of monocytes was significantly negatively associated with parasitemia ($p = 0.0134$, Pearson's product moment correlation coefficient $r^2 = -0.454$; Fig. 2). Unsurprisingly, given their involvement in protection from diverse pathogens, monocytes have central roles in the innate immune response to malaria, with involvement in diverse processes including phagocytosis, cytokine production and antigen presentation (Ortega-Pajares and Rogerson, 2018). However, products of *Plasmodium*, including hemozoin and extracellular vesicles, may impair monocyte function. Investigations of the relationship between malaria parasitemia and monocyte proportion in humans have yielded contrasting results with some studies reporting low monocyte counts in relation to high parasitemia (Kotepui et al., 2015) similar to our findings, while others report the opposite (Berens-Riha et al., 2014). This apparent contradiction may be due to differences in the hematological profile of individuals from different locales (Gansane et al., 2013).

In contrast to monocytes, the proportion of neutrophils was significantly positively correlated with parasitemia ($p = 0.0469$, $r^2 = 0.372$; Fig. 2). Differences in observed response to malaria parasitemia have also been observed in neutrophil counts for cohorts from different geographic areas, possibly attributable to immune status or age (Aitken et al., 2018). Our findings are similar to those of a comparative study of uncomplicated malaria in pediatric patients across sub-Saharan Africa, in which neutrophil counts were
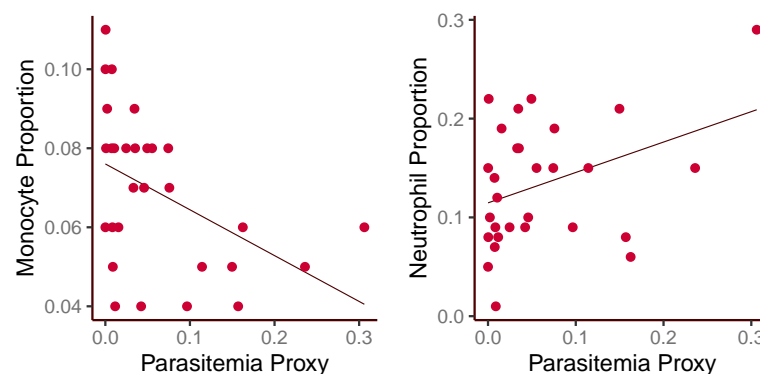
Figure 2. **Monocyte and neutrophil proportions significantly associated with parasitemia..** Relationship between proportions of (left) monocytes and (right) neutrophils with inferred *Hepatocystis* parasitemia. The *p*-values listed summarize *t*-test for significant association (*e.g.*, Pearson's correlation coefficient not equal to zero). Proportions of monocytes and neutrophils were significantly negatively and positively correlated with parasite load, respectively.

found to be significantly positively correlated with parasitemia (Olliaro et al., 2011). 95

Similarly, we next tested whether the proportion of *Hepatocystis* parasites in 96
particular stages of the parasite life cycle was correlated with parasite load. Using the 97
recently generated single cell transcriptomic data for *Plasmodium berghei* malaria 98
(Howick et al., 2019), we deconvoluted the bulk *Hepatocystis* gene expression data in 99
order to infer parasite life stage proportions for each sample (after Aunin et al., 2020; 100
Table S3; Fig. S2). We found no significant correlation between inferred parasitemia 101
and parasite life stage proportion (Table S4; Fig. S3). This is consistent with 102
expectations for chronic *Hepatocystis* infection, in which parasite life stages persist 103
throughout the infection, potentially due to constant generation of new parasites rather 104
than re-infection (Aunin et al., 2020). 105

Finally, we tested whether relative abundances of *Hepatocystis* developmental stages 106
and colobus immune cell types were correlated. The proportion of memory B cells was 107
significantly negatively correlated with *Hepatocystis* female gametocytes (r^2 = -0.416, p 108
= 0.0246; Table S5). Our results are consistent with findings that memory B cells 109
dominate early response after a malaria rechallenge in mice, while sexual stages of the 110
lifecycle do not appear until late infection (Krishnamurty et al., 2016). 111

Numerous genes are implicated in immunogenetic response to *Hepatocystis*. 112
Of the 12,223 red colobus genes considered after filtration, 324 and 235 genes were 113
significantly positively or negatively correlated with inferred parasitemia, respectively, 114
allowing a False Discovery Rate (FDR) of 20% (Fig. 3). The red colobus genes with 115
expression most strongly positively correlated with parasitemia (Table 1, S6; Fig. 4) 116
included *UBE2K*, *PP2D1*, *TMEM167A*, *LSM14A*, *UBE2B*, *TTLL12*, *AGFG1*, and 117
APOBEC2. 118

These genes with the strongest evidence of being up-regulated in response to 119
Hepatocystis parasitemia were often associated with red blood cell phenotypes or 120
malaria response. For instance, *LSM14A* has been linked to red blood cell (RBC) 121
distribution width (*i.e.*, red blood cell volume variation; Kichaev et al., 2019) and 122
antiviral cytokine induction (Liu et al., 2016), and likely contributes to stress granule 123
formation in mice infected with *Plasmodium* (Hanson and Mair, 2014). *APOBEC2* is a 124
part of the APOBEC family of (deoxy)cytidine deaminases which play an integral role 125
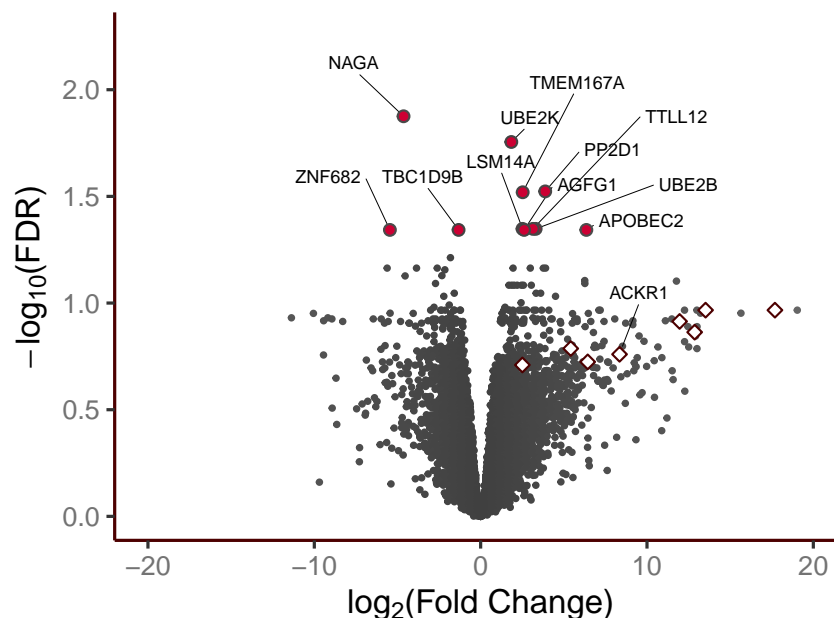
Figure 3. Volcano plot showing magnitude and significance of colobus gene expression changes correlated with inferred *Hepatocystis* parasitemia. Labelled red circles indicate colobus genes significantly up- or down-regulated in response to parasitemia ($p < 0.05$ after Benjamini-Hochberg adjustment for FDR). The $\log_2$ fold change in response to inferred parasitemia (x-axis) is computed per unit change in inferred parasitemia proxy. Diamond points indicate genes associated with erythrocyte-associated disorders such as stomatocytosis that were enriched among up-regulated genes. All genes lacking orthology information (*i.e.*, with identifiers beginning with "LOC") were excluded.
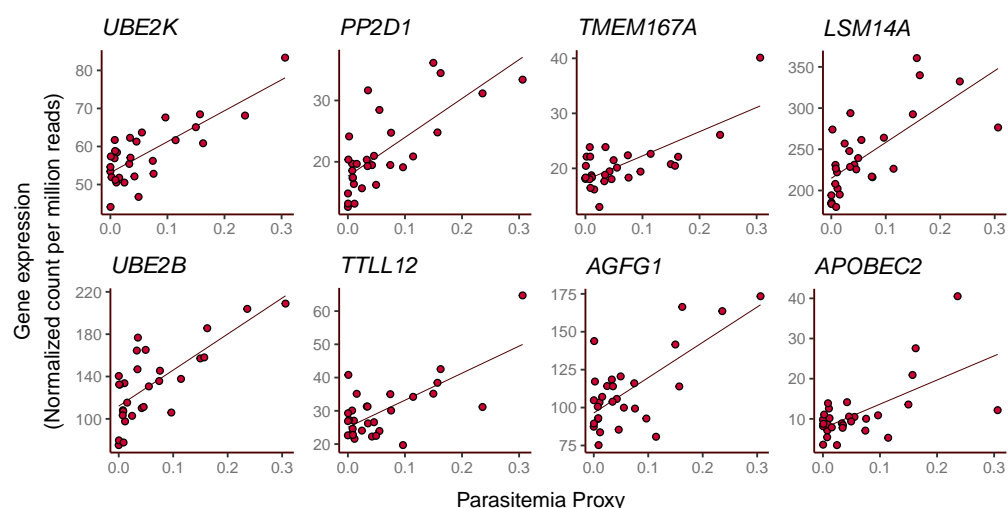

Figure 4. Expression of significantly up-regulated colobus genes positively correlated with *Hepatocystis* parasitemia. Expression summarized as normalized read count per million reads sequenced.

Figure 5. **Expression of significantly down-regulated colobus genes negatively correlated with *Hepatocystis* parasitemia.** Expression summarized as normalized read count per million reads sequenced.
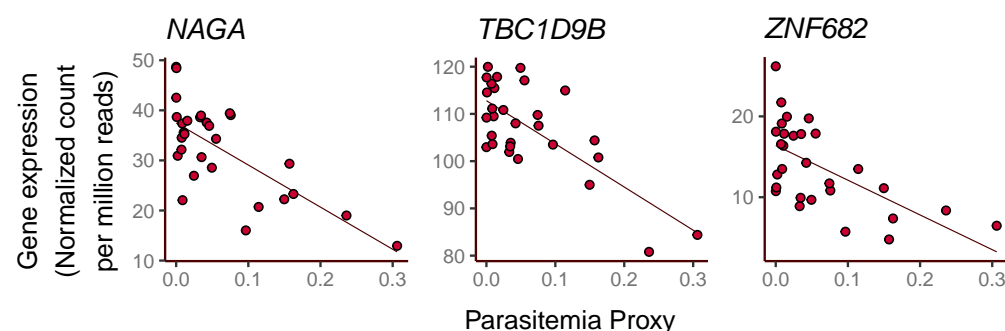
in innate immunity (Jha et al., 2012) and were found to be upregulated in early liver-stage *Plasmodium* infection in mice (Albuquerque et al., 2009). 126 127

The red colobus genes with expression most strongly negatively correlated with parasitemia (Table 1, S6; Fig. 5) included *NAGA*, *TBC1D9B*, and *ZNF682*. 128 129

We were particularly interested in the relationship between *Hepatocystis* parasitemia and expression of Atypical Chemokine Receptor 1 (*ACKR1*, alias *DARC*) which defines the Duffy blood group. *ACKR1* is the only gene associated with *Hepatocystis* protection in a wild primate population, with polymorphisms in its promoter region associated with variation in susceptibility to *Hepatocystis kochi* in yellow baboons (*Papio cynocephalus*; Tung et al., 2009). In the red colobus monkey population, we found that the *ACKR1* (*DARC*) gene expression was positively correlated with parasitemia ($p = 0.00351$, adjusted $p = 0.159$; Fig. S4). The Duffy antigen is an important receptor by which *Plasmodium vivax* and *P. knowlesi* infect human and non-human primate erythrocytes, and humans with the Duffy negative blood group—meaning cells that do not express this receptor—are semi-protected from infection by these malaria species (Miller et al., 1975; Mason et al., 1977; Michon et al., 2001). Underscoring the selective advantage and evolutionary importance of this gene in humans, the allele that confers protection against malaria (a variant at a single polymorphic site in the promoter region; Tournamille et al., 1995) has independently arisen in humans in malaria endemic regions, particularly Papua New Guinea and parts of sub-Saharan Africa (Miller et al., 1975; Zimmerman et al., 1999), and has rapidly reached high frequency following movement of people into areas with high *P. vivax* prevalence (Pierron et al., 2018; Hodgson et al., 2014). Our results add to the prior evidence (Tung et al., 2009) that the alleles at the *ACKR1* gene may confer a selective advantage to non-human primates mediated by *Hepatocystis* susceptibility, similar as has been established in humans for malaria-mediated selection. 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151

Strategies for host organism response to infection and mitigation of pathology include alteration of immune cell type proportions. To investigate whether the differential expression we observed in response to parasitemia was wholly or in part due to changes in cell-type composition, we added all relative cell-type proportions alongside the parasitemia proxy in the model of read counts. We found the results of the test for differentially expressed genes were broadly similar to the simple model (Table S7; Fig. S5). Very few genes that had been associated with parasitemia in the simple model were no longer significant when relative immune cell-type proportions were included as parameters, which would indicate their response to parasitemia may be indirect and associated with changes in cell-type composition. Of the top 12 named genes associated 152 153 154 155 156 157 158 159 160 161
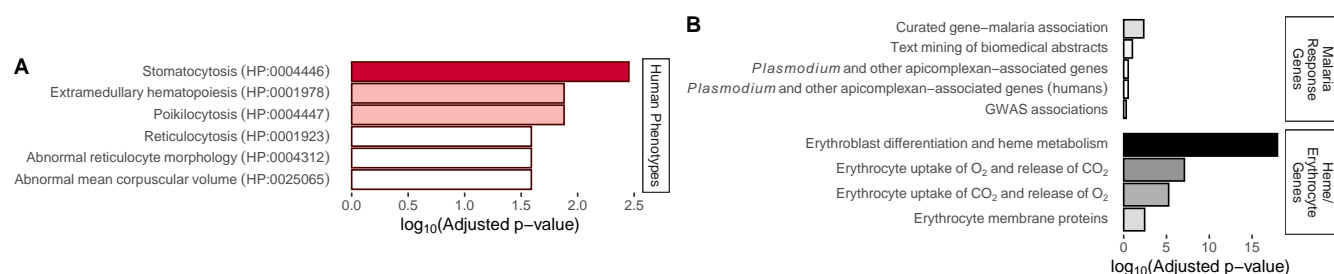
Figure 6. Overlap of colobus *Hepatocystis* response genes with human phenotype-associated genes and genes implicated in malaria response. Shading indicates strength of enrichment of Hepatocystis response genes with each functional category. **(A)** Over-represented human phenotypes, typically red blood cell disorders, were the result of a blind test for functional enrichment across all human phenotype annotation categories. **(B)** Malaria response gene sets were included *a priori* to compare *Hepatocystis* response to prior findings on *Plasmodium* response, while selection of heme and erythrocyte gene sets was informed by blind functional enrichment results.

with the parasitemia proxy with the highest confidence (uncorrected $p < 0.0001$), only two (*LSM14A* and *NAGA*) remained associated at that strict significance cutoff, though under a more lenient cutoff of $\alpha = 0.01$ to allow for the decrease in statistical power associated with the more complex model, 10 of the 12 original genes were significant (*TTLL12* and *ZNF682* were not). Our results suggest that genes such as *LSM14A* and *NAGA* are cell composition-independent in their response to parasitemia, but other genes such as *TTLL12* and *ZNF682* likely act at least in part via changes in immune cell composition.

Immunogenetic response to parasite involves erythrocyte-related genes.
We next tested whether red colobus genes with expression significantly correlated to parasitemia were united in having certain biological characteristics more likely than expected by chance. Using known orthologous human-colobus genes, we tested whether genes up-regulated in response to parasitemia were enriched for involvement in disease phenotypes (using Human Phenotype Ontology data; Köhler et al., 2019). The phenotypes with the strongest evidence of enrichment among our up-regulated genes included many red blood cell (RBC) disorders (Fig. 6; Table S8). These included stomatocytosis, or the production of RBCs with an abnormally shaped zone of central pallor (HP:0004446; adj. $p = 0.00352$); poikilocytosis (HP:0004447; adj. $p = 0.0132$), or the production of abnormally shaped RBCs; and reticulocytosis, or the increased production of immature erythrocytes which is disrupted during malaria anaemia in humans (HP:0001923; adj. $p = 0.0257$).

Red colobus genes involved in *Hepatocystis* response were often implicated in all three of the enriched RBC disorders (Fig. S6). Such genes with significant association with *Hepatocystis* parasitemia (FDR adjusted $p < 0.2$) included *RHAG* (adj. $p = 0.108$), *SPTA1* (adj. $p = 0.108$), *SLC4A1* (adj. $p = 0.137$), *ABCB6* (adj. $p = 0.163$), and *SLC2A1* (adj. $p = 0.174$). Genes associated with a subset of the RBC phenotypes include *KLF1*, associated with reticulocyte disorders (adj. $p = 0.122$), as well as *STEAP3* (adj. $p = 0.189$) and *JAK2* (adj. $p = 0.195$), both linked to poikilocytosis. Many of these RBC disorder genes have been associated with malaria response in mice or humans. For instance, in mice vaccinated with erythrocyte ghosts isolated from *P. chaubaudi*-infected erythrocytes, *RHAG*, *SPTA1*, *KLF1*, and *SLC4A1* were among the erythropoiesis-involved genes that were significantly upregulated in response to infection with blood stage *Plasmodium chabaudi* (Delic et al., 2020). Furthermore, variation in *SPTA1*, which encodes a component of the RBC plasma membrane, was recently found

to impact *P. falciparum* growth in cultured human RBCs (Ebel et al., 2020). Lastly, [196]
heterozygous deletion in *SLC4A1* causes South Asian hereditary ovalocytosis, which [197]
confers protection in humans against many *Plasmodium* species (Jarolim et al., 1991; [198]
Paquette et al., 2015), and this gene has been found to be under differential selection [199]
pressure in other primate species (Steiper et al., 2012). [200]

In addition to the above blood-related disorders, the monkey genes with expression [201]
most strongly positively correlated to *Hepatocystis* parasitemia were also more likely [202]
than expected to be related to extramedullary hematopoiesis (HP:0001978; adj. $p =$ [203]
0.0132), or the production of cellular blood components outside of bone marrow. [204]
Extramedullary or stress hematopoiesis is typically an indication of pathology in adults [205]
(Kim, 2010), and an increase in splenic erythropoiesis has been observed in mice with [206]
malarial anemia (reviewed in Lamikanra et al., 2007) and suggested as a mechanism by [207]
which humans compensate for anemia during *P. vivax* infection (del Portillo et al., [208]
2004). [209]

No RBC-related disorders were enriched among genes negatively correlated to [210]
parasitemia. No Gene Ontology (Ashburner et al., 2000) functional annotations were [211]
over-represented among either significantly up- or down-regulated genes in response to [212]
inferred *Hepatocystis* parasitemia (Table S8). [213]

Genetic architecture of *Hepatocystis* response overlaps that of *Plasmodium*. [214]
As the blood stage of *Hepatocystis* differs from that of *Plasmodium* in being limited to [215]
the gametocyte sexual stages, the *Hepatocystis* erythrocytic phase is predicted to not [216]
induce a cytokine response in the host (Ejotre et al., 2020). Despite this key difference, [217]
we found significant overlap between the genetic architecture of the colobus *Hepatocystis* [218]
response and the human response to malaria (Fig. 6; Table S9). Genes with significant [219]
increased gene expression response of colobus to *Hepatocystis* were significantly more [220]
likely than expected by chance to be associated with human response to malaria [221]
through curated gene-disease association (Davis et al., 2009; $p = 0.00455$) or text [222]
mining of biomedical abstracts, albeit not significantly (Pletscher-Frankild et al., 2015; [223]
$p = 0.0963$). However, *Hepatocystis* response genes did not significantly overlap malaria [224]
response genes defined via GWAS associations (Becker et al., 2004; $p = 0.509$), possibly [225]
highlighting the difficulty in generalizing GWAS results across species or populations. [226]
Down-regulated *Hepatocystis* response genes were enriched for only malaria response [227]
genes defined via text mining of biomedical abstracts (Pletscher-Frankild et al., 2015; p [228]
$= 0.0128$). [229]

When intersected with genes implicated in mammalian response to *Plasmodium* and [230]
other apicomplexan parasites (Ebel et al., 2017; Fig. 6; Table S9), we found the overlap [231]
with upregulated colobus *Hepatocystis*-response genes to be no larger than expected by [232]
chance ($p = 0.295$). The concordance was no higher when the former was restricted to [233]
genes with evidence from human studies ($p = 0.300$), in contrast to our above results. [234]
The same was true for significantly down-regulated genes ($p = 0.395$ and $p = 0.240$ [235]
respectively.) [236]

Prompted by the preponderance of RBC disorders among the enriched phenotypes, [237]
we next tested whether upregulated host genes were significantly enriched for association [238]
with selected erythrocyte functions (Fig. 6; Table S9). Significant enrichments included [239]
genes involved in erythroblast differentiation and metabolism of heme [240]
($p = 1.09 \times 10^{-18}$), genes involved in the take up of oxygen and release of carbon dioxide [241]
($p = 8.11 \times 10^{-8}$) and the reverse ($p = 5.59 \times 10^{-6}$), and genes encoding erythrocyte [242]
membrane protein genes ($p = 3.57 \times 10^{-3}$). Colobus genes that were down-regulated in [243]
response to *Hepatocystis* were not significantly enriched for these erythrocyte-related [244]
functional annotations ($p = 1$ for all). Although genes involved in interactions with red [245]
blood cells have been lost in *Hepatocystis* (Aunin et al., 2020), *Hepatocystis* has been [246]

reported to induce anemia in primates (Garnham, 1966; Chang et al., 2011), which may mediate the link between parasitemia and these monkey RBC-related genes.

Hepatocystis response genes are selected in primates. Apicomplexan parasites such as malaria have exerted arguably the strongest selective force on humans (Kwiatkowski, 2005; Kariuki and Williams, 2020) and other primates (Tung et al., 2009; Oliveira et al., 2012; Ebel et al., 2017). The evolutionary arms race between host and pathogen or parasite may result in rapid evolution of gene sequences, and scans for signatures of natural selection in human and non-human genomes have repeatedly indicated immune genes to be the targets of adaptive evolution (Deschamps et al., 2016; Sironi et al., 2015; Fumagalli et al., 2011; Enard et al., 2016; Voight et al., 2006; Nielsen et al., 2005). In light of this, we tested whether red colobus genes involved in *Hepatocystis* response had accelerated evolution in primates (van der Lee et al., 2017). Colobus genes with expression that was positively correlated with *Hepatocystis* parasitemia were significantly more likely than expected by chance to be under positive selection in primates ($p = 0.00847$), but the same was not true for down-regulated genes ($p = 1$). Our findings are suggestive of potential *Hepatocystis*-mediated selection or generalized selection pressure from apicomplexan or haemosporidian parasites in primates.

Few parasite genes have expression that varies with parasitemia. In parallel to our tests for colobus host genes that were associated with parasitemia, we also tested whether any *Hepatocystis* gene expression was similarly correlated with parasite load (Fig. S7; Table S10). The three genes with the strongest evidence had expression inversely proportional to parasitemia. Two were proteins with unknown function (HEP_00496100: $p = 5.00 \times 10^{-5}$, adj. $p = 0.222$; and HEP_00472500: $p = 1.40 \times 10^{-4}$, adj. $p = 0.275$), while the third (HEP_00452800: $p = 1.86 \times 10^{-4}$; adj. $p = 0.275$) is a putative circumsporozoite- and thrombospondin-related anonymous protein (TRAP)-related protein (CTRP), based on it exhibiting a thrombospondin type 1 domain (Pfam:PF00090.19). CTRPs have been previously implicated in parasite motility and host cell invasion, making them promising drug targets (Morahan et al., 2009). Our finding that this CTRP is most active when parasitemia is low may be linked to findings that TRAP is essential for early (liver) stage infection but not necessary for later (blood) stage infection in *Plasmodium* (Sultan et al., 1997).

Parasitemia-associated co-expressed genes associated with lipid biosynthesis. Finally, to explore the interaction between colobus monkeys and *Hepatocystis* parasites from a systems biology perspective, we identified co-expressed colobus gene modules, and tested whether the expression of the modules was correlated with inferred *Hepatocystis* parasitemia or the expression of specific *Hepatocystis* genes. We identified two colobus gene modules for which the eigengene expression (expression profile summary) values were significantly correlated with inferred parasitemia ($p < 0.01$), one positively and one negatively (Table S11).

We found one cluster of 143 co-expressed colobus genes to be positively correlated with inferred parasitemia ($p = 0.00277$, $r^2 = 0.535$), and these genes were enriched (Table S12) for being in the extracellular region (GO:0005576, $p = 3.68 \times 10^{-5}$) and for involvement in antimicrobial humoral immune response mediated by antimicrobial peptides (GO:0061844, $p = 5.41 \times 10^{-3}$; Table S12). No *Hepatocystis* gene had expression that was significantly correlated with the eigengene expression of this colobus module.

Another cluster of 215 genes had expression negatively correlated with inferred parasitemia ($p = 0.00528$, $r^2 = -0.504$), and these genes were enriched for being

involved in various processes, including biosynthesis of neutral lipids (GO:0046460, $p = 5.20 \times 10^{-3}$) specifically acylglycerol (GO:0046463, $p = 5.20 \times 10^{-3}$; Table S12). Acylglycerols (triacylglycerol and, secondarily, diacylglycerol) have been found to accumulate in erythrocytes infected with *P. falciparum* during the late trophozoite and schizont stages (Nawabi et al., 2003). Our findings are consistent with knowledge that *Plasmodium* parasites often hijack host lipid receptors, to invade other tissues, increasing infection severity (reviewed in Alsultan et al., 2020; Orikiiriza et al., 2017; O'Neal et al., 2020) and a recent study of malaria in Rwandan children which found a metabolic shift to increased production of long-chain unsaturated fatty acid associated triacylglycerols in response to infection with *P. falciparum* (Orikiiriza et al., 2017). The genes in this module with expression that most strongly correlated to parasitemia included *NAGA* (logFC = -4.63, $p = 2.18 \times 10^{-6}$, adj. $p = 0.0133$) and *TBC1D9B* (logFC = -1.33, $p = 4.36 \times 10^{-5}$; adj. $p = 0.0454$) which have been linked in humans to platelet count or count (via association of enhancers rs972578 or rs8137128; Astle et al., 2016) and blood protein measurement (via association of enhancer rs73351608; Sun et al., 2018), respectively. The expression of a large number of *Hepatocystis* genes (N=512) was significantly ($p < 0.01$) correlated with the eigengene expression of this module (Table S13), and these co-expressed *Hepatocystis* genes were enriched for protein channel activity (GO:0015252; $p = 0.0476$; Table S14).

Conclusions and future directions. Overall, our study reveals considerable overlap between the colobus response to *Hepatocystis* and human response to *Plasmodium*. By studying the gene expression response to these parasites in non-human hosts, particularly our close non-human primate relatives, similar future studies will improve our understanding of broad patterns of host-parasite co-evolution. Such comparative context is critical for understanding the evolution of this important disease, and mechanistic insight into disease resistance may inform prioritization of possible intervention targets for human malaria patients. Understanding the genetic architecture of malaria response in a comparative context has implications for the determination of host specificity. This has additional practical importance, as multiple instances of non-human primate to human (*i.e.*, zoonotic) transmission have occurred. These zoonoses include, most pressingly, *Plasmodium knowlesi*, a monkey-origin malaria species which crossed over to humans in the late 20th century. We predict that similar future studies in diverse primate species that have co-evolved with these important parasites will further improve our understanding of host-parasite co-evolution.

Materials and Methods

Dataset: We downloaded RNA sequence (mRNA-seq) data previously generated (by Simons et al., 2019) from whole blood samples of 29 red colobus monkey (*Piliocolobus tephrosceles*) individuals (study PRJNA413051) from the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA). Methods for sample collection, RNA extraction, library preparation, and transcriptome sequencing are described in Simons et al., 2019. Briefly, Simons and colleagues (2019) prepared mRNA-seq libraries with the KAPA Biosystems Stranded mRNA-seq kit and sequenced these on the Illumina HiSeq 4000 platform with 150 paired-end cycles.

Gene expression inference: 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 Trimmomatic (Bolger et al., 2014). We removed reads that were less than 36 bp after

trimming. In order to determine if sequence reads were from the colobus (GCF_002776525.3_ASM277652v3; Simons et al., 2019) or *Hepatocystis* (GCA_902459845.2_HEP1; Aunin et al., 2020), we performed a competitive mapping analysis as follows. Using the STAR alignment program (Dobin et al., 2013), we mapped trimmed, quality-filtered RNA-seq sequencing reads to concatenated reference genomes that include colobus and *Hepatocystis* genomes, allowing each read to map multiple locations with up to eight mismatches per 100 bp paired-end read (after Lee et al., 2018). 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 colobus. Libraries which had been split and sequenced in multiple lanes were then combined by red colobus monkey individual (N=29).

We used the Rsubread R package (Liao et al., 2019) to create a colobus monkey read-count matrix for each sample, using the "uniquely" mapped host reads and allowing for overlapping reads. Reads were tallied at the level of exons which were then summed to create a per-gene count for the colobus monkey, using gene annotations (GCF_002776525.3_ASM277652v3). For *Hepatocystis*, we used published read counts per gene to be concordant with prior work (Aunin et al., 2020). Read counts per gene and the results of downstream analyses were largely consistent between analyses using the published *Hepatocystis* read count dataset and sums of read counts per gene from our mapping analysis (supplemental text).

Inference of parasite load: We inferred a proxy of parasitemia by calculating the ratio of the number of pathogen reads (data from Aunin et al., 2020) to the number of host colobus reads recovered (after Yamagishi et al., 2014; Table S1; Fig. 1). For this measure, we used the *Hepatocystis* read counts for genes, excluding the upper and lower quartiles of the distribution to avoid noise from genes with extremely high or low expression. However, using the full set of genes yielded very similar results (supplemental text).

Tests for cell type or stage correlates to inferred parasitemia: Using our inferred *Hepatocystis* parasitemia and previously published relative immune cell composition abundance data (Simons et al., 2019), we tested whether parasitemia and immune cell proportions were correlated by testing the hypothesis that Pearson's correlation coefficient is not equal to zero using a t-test. We applied the same approach to test for association between inferred parasitemia and fraction of *Hepatocystis* cells in each developmental stage, which we estimated via bulk RNAseq deconvolution as follows.

In order to infer the ratio of different *Hepatocystis* cell types (*i.e.*, life stages) in our bulk data and determine how these may covary with parasitemia and the expression of differentially expressed genes, we deconvoluted the bulk *Hepatocystis* RNAseq read count data. To do so, we used published stage-specific "pseudo-bulk" samples (Aunin et al., 2020) constructed from data gathered from the Malaria Cell Atlas (Howick et al., 2019), which was based on single cell sequencing of cells in known life stages to deconvolute previously published *Hepatocystis* read count data. For the analysis, we used relationships between *Hepatocystis* genes and their one-to-one orthologous *Plasmodium berghei* genes, which were determined by Aunin et al., 2020 using OrthoMCL (Li et al., 2003) and IQ-TREE (Nguyen et al., 2014). We excluded those *Hepatocystis* genes without one-to-one orthology with *P. berghei* genes. All analyses for estimating cell type proportions in our *Hepatocystis* read count data were implemented in BSeq-sc (Baron et al., 2016) and CIBERSORTx (Newman et al., 2019) which we ran with 1,000 permutations (Table S3). For downstream analyses, we combined all

non-erythrocytic life stages into an "Other" category to remain consistent with previously published data.

Finally, we tested for an association between these relative abundances of *Hepatocystis* developmental stages inferred via bulk RNAseq deconvolution and colobus immune cell types (Simons et al., 2019) using a *t*-test to establish if Pearson's correlation coefficient differed significantly from zero.

Identification of differentially expressed genes by parasitemia: To identify host colobus monkey and *Hepatocystis* genes with expression correlated to parasitemia, we analyzed the two datasets independently with the Bioconductor package edgeR (Robinson et al., 2010) as follows. We first excluded low coverage genes, defined as any gene with expression below a counts per million (CPM) threshold corresponding to a read count of 10 in at least two individuals for *Hepatocystis* genes or corresponding to a read count of 5 in at least four individuals for colobus genes. 12,223 colobus genes and 4,432 *Hepatocystis* genes remained after filtering. We then normalized the read count matrices to control for library size and compositional bias.

From the read-count matrices for the host and parasite separately, we identified significantly differentially expressed genes with expression correlated with parasitemia ("parasitemia-DE genes") using the edgeR package (Robinson et al., 2010) as follows. After estimating dispersion using the Cox-Reid profile-adjusted likelihood (CR) method and fitting a negative binomial model, we identified significantly correlated genes using the generalized linear model (GLM) likelihood ratio test. To do so, we used the model:

$$\text{read count} \sim \text{inferred parasitemia} + \text{correction factors}$$

where correction factors include those adjusting for differences in number of reads and library complexity. We corrected for multiple tests via false discovery rate (FDR) estimation. We considered a gene to be a parasitemia-DE gene if the adjusted *p*-value was less than 0.2 for colobus genes, or if the unadjusted *p*-value was less than 0.01 for *Hepatocystis*, using a lower threshold due to reduced statistical power in analyses of parasite gene expression.

To test for an influence of immune cell composition on differential expression by parasitemia, we repeated the above analysis adding in parameters for relative abundance of memory B cells, CD8 naïve CD4+ T cells, memory resting CD4+ T Cells, memory activated CD4+ T cells, natural killer cells, monocytes, macrophages, and neutrophils. All fractions of colobus immune cells were taken from Simons et al., 2019 and are based on bulk RNAseq deconvolution analysis. The filtering and testing for differential expression by parasitemia took place as in the simpler model without cell type composition parameters. Results are presented joined with those of the simpler differential expression analysis (Table S7).

Functional enrichment analyses: We next tested whether the genes with expression that was significantly positively correlated with parasitemia (upregulated parasitemia-DE genes) or significantly negatively correlated with parasitemia (downregulated parasitemia-DE genes) were enriched for particular biological functions. Specifically, we tested whether these genes were more likely than expected by chance to share Gene Ontology molecular functions, biological processes, or cell component annotations (Ashburner et al., 2000; The Gene Ontology Consortium, 2019). For the colobus dataset, we also tested for functional enrichment of Human Phenotype Ontology (Köhler et al., 2019) annotations, using annotations for human orthologs matched to our colobus genes using the Ensembl gene mappings (accessed via g:Orth; Raudvere et al., 2019). To test for enrichment, we used a hypergeometric test as implemented in

g:Profiler (Raudvere et al., 2019) and used all annotated genes after filtration for low read count as a comparative background list for differential expression determination.

To compare colobus response to *Hepatocystis* with parasite response in humans and other mammals, we further analyzed the colobus dataset as follows. We tested whether there was significantly more overlap than expected by change between our up- or downregulated parasitemia-DE genes and particular sets of malaria- or blood-related genes. As these genes sets were based on human gene identifiers, we used all human genes that had orthology with colobus genes sets for our significantly up- and down-regulated genes, as well as with a comparative background set all colobus genes with read count data after filtering. The sets of *a priori* genes related to malaria included genes from curated gene-malaria association (N=19; Davis et al., 2009), genes associated with malaria through text mining of biomedical abstracts (N=701; Pletscher-Frankild et al., 2015), and malaria response genes as defined via GWAS associations (N=51; Becker et al., 2004). We also tested sets of genes implicated in mammalian response to *Plasmodium* and related parasites (N=490; Ebel et al., 2017), and the same dataset reduced to only include those with evidence of association in humans (N=301). Our final gene sets were selected due to their connections with erythrocyte functions. Gene sets were sourced from MSigDB (Subramanian et al., 2005; Liberzon et al., 2015) and included genes involved in erythroblast differentiation and metabolism of heme (N=200; MSigDB ID M5945), genes involved in the uptake of oxygen and release of carbon dioxide (N=11; MSigDB ID M29536; Jassal et al., 2020) and the uptake of carbon dioxide and release of oxygen (N=16; MSigDB ID M26928; Jassal et al., 2020), and genes encoding erythrocyte membrane protein genes (N=14; MSigDB ID M2442; Steiner et al., 2009).

Evolutionary analyses: Finally we tested whether the colobus parasitemia-linked genes were enriched for genes previously inferred to be under positive natural selection in primates. As above, we tested for greater overlap than expected by chance between our up- and downregulated colobus parasitemia-DE genes and a list of 331 protein-coding genes previously found to be under strong positive selection (*i.e.*, having a significantly better fit for a selection model than a neutral model in a maximum likelihood dN/dS analysis and containing at least one significant positively selected amino acid residue) in a comparative analysis of nine whole-genome sequenced primates (van der Lee et al., 2017).

Identification of co-expressed gene modules associated with parasitemia: To identify key sets of genes involved in host-parasite interactions between colobus and *Hepatocystis*, we inferred gene co-expression networks using Weighted Gene Co-expression Network Analysis of the colobus genes (WGCNA; Langfelder and Horvath, 2008). We used the same read count matrix for the identification of parasitemia-DE genes. For clustering, we selected a soft-thresholding power (β) via analysis of network topology, visually examining scale-free topology fit index as a function of soft-thresholding power and selecting $\beta = 9$ as the value for which the scale free topology model fit R^2 is approximately 0.9 ($R^2 = 0.894$).

We performed automatic network construction and module detection in a block-wise manner using the WGCNA function blockwiseModules. Signed network options were selected for the adjacency function (networkType = "unsigned") and topological overlap (TOMType = "signed"). For tree cutting, the minimum module size was set to 30, and for module merging, the merge cut height was set to 0.15. The gene reassignment threshold was set to 0.

Using the same parasitemia proxy as the differential expression analysis, we next tested for co-expression modules with expression correlated to parasite load by

calculating the Pearson correlation coefficient and computing a p-value via the   492
Student's t-test with asymptotic null distribution. For each significant module, we   493
identified significantly over-represented functional annotations using a hypergeometric   494
test as implemented in g:Profiler (Raudvere et al., 2019), using the same parameters as   495
the previous function enrichment analysis of parasite-DE colobus genes. Similarly, we   496
also used all annotated genes after filtration for low read count as a comparative   497
background list.   498

Finally, we tested if any significantly parasitemia-associated modules co-expressed   499
with *Hepatocystis* gene expression using the same read count data as the differential   500
expression analysis above. To do so, we tested for a correlation between the colobus   501
module's eigen expression and *Hepatocystis* gene expression using Pearson's product   502
moment correlation coefficient and testing for significance using a t-test. For each   503
module, we next tested whether *Hepatocystis* genes with correlated expression were   504
enriched for biological annotations. To associate genes with functional annotations, we   505
replaced *Hepatocystis* gene identifiers to their one-to-one *Plasmodium berghei*   506
orthologues, as inferred by Aunin et al. (2020) using OrthoMCL (Li et al., 2003). We   507
identified significantly over-represented functional annotations using a hypergeometric   508
test as implemented in g:Profiler (Raudvere et al., 2019), using all annotated genes as a   509
comparative background list.   510

Acknowledgements   518

Supplemental Figure and Table Captions   526

Supplemental Table Captions   527

Table S1. Inferred parasitemia by sample. Competitively mapped, unique   528
colobus read counts and previously published *Hepatocystis* read counts from each   529
colobus sample.   530

Table S2. Immune cell type proportion correlation with parasitemia.   531
Pearson coefficient for correlation between colobus immune cell type proportions and   532
inferred *Hepatocystis* parasitemia, along with t-test for significant association (deviation   533
from zero). Proportions of immune cell types are from Simons et al., 2019 and include   534
memory B cells ("MemB"), CD8 Naïve CD4+ T Cells ("naiveCD4plus"), Memory   535

Resting CD4+ T Cells ("MemrestCD4plus"), Memory Activated CD4+ T Cells   536
("MemactCD4plus"), natural killer cells ("restNK"), monocytes, macrophages, and   537
neutrophils.   538

Table S3. Inferred *Hepatocystis* life stage proportions for each sample.   539
Inferred *Hepatocystis* life stage proportions for each colobus sample after deconvolution   540
analysis using a cell atlas based on *Plasmodium berghei* single-cell data.   541

Table S4. *Hepatocystis* life stage proportion correlation with parasitemia.   542
Pearson coefficient for correlation between *Hepatocystis* developmental stage   543
proportions and inferred *Hepatocystis* parasitemia, along with *t*-test for significant   544
association (deviation from zero).   545

Table S5. Results of test for correlation between *Hepatocystis* life stages   546
and immune cell composition.   Pearson coefficient for correlation between   547
Hepatocystis developmental stage proportions and colobus immune cell type proportions,   548
along with *t*-test for significant association (deviation from zero). Proportions of   549
immune cell types are from Simons et al., 2019 and include memory B cells ("MemB"),   550
CD8 Naïve CD4+ T Cells ("naiveCD4plus"), Memory Resting CD4+ T Cells   551
("MemrestCD4plus"), Memory Activated CD4+ T Cells ("MemactCD4plus"), natural   552
killer cells ("restNK"), monocytes, macrophages, and neutrophils.   553

Table S6. Colobus genes differentially expressed by parasitemia.   Full   554
results of test for differential expression of colobus genes in response to *Hepatocystis*   555
parasitemia. The $\log_2$ fold change in response to inferred parasitemia ("logFC") is   556
computed per unit change in inferred parasitemia proxy.   557

Table S7. Colobus genes differentially expressed by parasitemia controlling   558
for colobus immune cell type proportions.   559

Table S8. Functional enrichment of colobus genes DE by parasitemia.   560

Table S9. Comparison of colobus *Hepatocystis*-response genes with malaria   561
response-associated gene sets.   Significantly up-regulated *Hepatocystis* response   562
genes are overrepresented in some sets of genes connected to *Plasmodium* malaria as   563
well as those associated with erythrocytes.   564

Table S10. *Hepatocystis* genes differentially expressed by parasitemia.   Full   565
results of test for differential expression of *Hepatocystis* genes in response to   566
Hepatocystis parasitemia. The $\log_2$ fold change in response to inferred parasitemia   567
("logFC") is computed per unit change in inferred parasitemia proxy.   568

Table S11. Co-expression modules correlated with parasitemia.   Clusters of   569
genes with similar co-expression, inferred via weighted gene co-expression network   570
analysis (WGCNA). Correlation with parasitemia summarized by *p*-value indicating   571
result of *t*-test that Pearson's correlation coefficient is not equal to zero.   572

Table S12. Co-expression module functional enrichment.   Functional   573
categories (Gene Ontology and Human Phenotype) overrepresented among colobus   574
genes in the two co-expression modules with the strongest evidence for correlation with   575
parasitemia.   576

Table S13. *Hepatocystis* genes with expression correlated to that of colobus 577
co-expression modules.   Correlation with parasitemia summarized by p-value 578
indicating result of t-test that Pearson's correlation coefficient is not equal to zero. 579

Table S14. Functional enrichment of *Hepatocystis* genes with expression 580
correlated to that of colobus co-expression modules.   Functional categories 581
(Gene Ontology) overrepresented among *Hepatocystis* genes that contrived with the 582
colobus co-expression modules with the strongest evidence for correlation with 583
parasitemia. 584

Supplemental Figure Captions 585

Figure S1. Relationship between relative abundance of immune cell types 586
and inferred *Hepatocystis* parasitemia.   Proportions of immune cell types are 587
from Simons et al., 2019 and include memory B cells ("MemB"), CD8 naïve CD4+ T 588
cells ("naiveCD4plus"), memory resting CD4+ T cells ("MemrestCD4plus"), memory 589
activated CD4+ T cells ("MemactCD4plus"), natural killer cells ("restNK"), monocytes, 590
macrophages, and neutrophils. 591

Figure S2. Proportion of *Hepatocystis* reads inferred to be in each parasite 592
developmental stage via deconvolution.   Deconvolution based on previously 593
published pseudobulk single-cell *Plasmodium berghei* data constructed from the Malaria 594
Cell Atlas. Results are grouped by colobus individual blood sample. Proportions are 595
consistent with previously published data. 596

Figure S3. Relationship between inferred parasitemia and parasite life 597
stage proportion.   Proportions shown for schizonts, merozoites, ring-form 598
trophozoites, mature trophozoites, female and male gametocytes, and other. 599

Figure S4. Expression of colobus *ACKR1* gene plotted against inferred 600
***Hepatocystis* parasitemia.**   Expression summarized as normalized read count per 601
million reads sequenced. One outlier data point was removed from the plot. 602

Figure S5. Comparison of results of test for DE genes using model with 603
and without immune cell composition parameters.   Points shown indicate 604
unadjusted p-value for each test, plotted on a negative logarithmic scale. Red points 605
were significant in the simple model without immune cell composition parameters 606
(FDR-adjusted $p < 0.05$). Only *ANP32E* and *SLC18A2* (labeled) were significant at 607
this strict threshold (FDR-adjusted $p < 0.05$) in the more complex model that included 608
cell composition parameters. 609

Figure S6. Overlap of parasitemia-DE genes associated with RBC 610
disorders.   Various red blood cell-related disorders had the strongest evidence of 611
enrichment among parasitemia-DE genes. Heatmap shows overlap among their 612
associated up-regulated genes. The degree of log-fold change is indicated by the color 613
gradient, while grey squares indicate no association between gene and phenotype. 614

Figure S7. Volcano plot showing magnitude and significance of 615
Hepatocystis* gene expression changes correlated with inferred *Hepatocystis 616
parasitemia.   Labelled red circles indicate *Hepatocystis* genes significantly up- or 617
down-regulated in response to parasitemia (uncorrected $p < 0.05$). The $\log_2$ fold change 618

in response to inferred parasitemia (x-axis) is computed per unit change in inferred parasitemia proxy.

References

Aitken, E. H., Alemu, A., and Rogerson, S. J. (2018). Neutrophils and malaria. *Frontiers in Immunology*, 9:3005. doi: `10.3389/fimmu.2018.03005`.

Albuquerque, S. S., Carret, C., Grosso, A., Tarun, A. S., Peng, X., Kappe, S. H., Prudêncio, M., and Mota, M. M. (2009). Host cell transcriptional profiling during malaria liver stage infection reveals a coordinated and sequential set of biological events. *BMC Genomics*, 10(1):1–13. doi: `10.1186/1471-2164-10-270`.

Alsultan, M., Morriss, J., Contaifer, D., Kumar, N. G., and Wijesinghe, D. S. (2020). Host lipid response in tropical diseases. *Current Treatment Options in Infectious Diseases*, 12(3):243–257. doi: `10.1007/s40506-020-00222-9`.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., and Sherlock, G. (2000). Gene Ontology: Tool for the unification of biology. *Nature Genetics*, 25(1):25–29. doi: `10.1038/75556`.

Astle, W. J., Elding, H., Jiang, T., Allen, D., Ruklisa, D., Mann, A. L., Mead, D., Bouman, H., Riveros-Mckay, F., Kostadima, M. A., Lambourne, J. J., Sivapalaratnam, S., Downes, K., Kundu, K., Bomba, L., Berentsen, K., Bradley, J. R., Daugherty, L. C., Delaneau, O., Freson, K., Garner, S. F., Grassi, L., Guerrero, J., Haimel, M., Janssen-Megens, E. M., Kaan, A., Kamat, M., Kim, B., Mandoli, A., Marchini, J., Martens, J. H., Meacham, S., Megy, K., O'Connell, J., Petersen, R., Sharifi, N., Sheard, S. M., Staley, J. R., Tuna, S., van der Ent, M., Walter, K., Wang, S.-Y., Wheeler, E., Wilder, S. P., Iotchkova, V., Moore, C., Sambrook, J., Stunnenberg, H. G., Di Angelantonio, E., Kaptoge, S., Kuijpers, T. W., Carrillo-de-Santa-Pau, E., Juan, D., Rico, D., Valencia, A., Chen, L., Ge, B., Vasquez, L., Kwan, T., Garrido-Martín, D., Watt, S., Yang, Y., Guigo, R., Beck, S., Paul, D. S., Pastinen, T., Bujold, D., Bourque, G., Frontini, M., Danesh, J., Roberts, D. J., Ouwehand, W. H., Butterworth, A. S., and Soranzo, N. (2016). The allelic landscape of human blood cell trait variation and links to common complex disease. *Cell*, 167(5):1415–1429.e19. doi: `10.1016/j.cell.2016.10.042`.

Aunin, E., Böhme, U., Sanderson, T., Simons, N. D., Goldberg, T. L., Ting, N., Chapman, C. A., Newbold, C. I., Berriman, M., and Reid, A. J. (2020). Genomic and transcriptomic evidence for descent from *Plasmodium* and loss of blood schizogony in *Hepatocystis* parasites from naturally infected red colobus monkeys. *PLOS Pathogens*, 16(8):e1008717. doi: `10.1371/journal.ppat.1008717`.

Ayouba, A., Mouacha, F., Learn, G. H., Mpoudi-Ngole, E., Rayner, J. C., Sharp, P. M., Hahn, B. H., Delaporte, E., and Peeters, M. (2012). Ubiquitous *Hepatocystis* infections, but no evidence of *Plasmodium Falciparum*-like malaria parasites in wild greater spot-nosed monkeys (*Cercopithecus Nictitans*). *International Journal for Parasitology*, 42(8):709–713. doi: `10.1016/j.ijpara.2012.05.004`.

Baron, M., Veres, A., Wolock, S. L., Faust, A. L., Gaujoux, R., Vetere, A., Ryu, J. H., Wagner, B. K., Shen-Orr, S. S., Klein, A. M., Melton, D. A., and Yanai, I. (2016). A Single-cell transcriptomic map of the human and mouse pancreas reveals inter- and

intra-cell population structure. *Cell Systems*, 3(4):346–360.e4. doi: 10.1016/j.cels.2016.08.011.

Becker, K. G., Barnes, K. C., Bright, T. J., and Wang, S. A. (2004). The Genetic Association Database. *Nature Genetics*, 36(5):431–432. doi: 10.1038/ng0504-431.

Berens-Riha, N., Kroidl, I., Schunk, M., Alberer, M., Beissner, M., Pritsch, M., Kroidl, A., Fröschl, G., Hanus, I., Bretzel, G., von Sonnenburg, F., Nothdurft, H. D., Löscher, T., and Herbinger, K.-H. (2014). Evidence for significant influence of host immunity on changes in differential blood count during malaria. *Malaria Journal*, 13(1):155. doi: 10.1186/1475-2875-13-155.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120. doi: 10.1093/bioinformatics/btu170.

Chang, Sun, Wang, Yi, Song, Peng, Lu, Zhou, and Jiang (2011). Identification of *Hepatocystis* species in a macaque monkey in northern Myanmar. *Research and Reports in Tropical Medicine*, 2:141–146. doi: 10.2147/RRTM.S27182.

Chapman, C. A., Ghai, R., Jacob, A., Koojo, S. M., Reyna-Hurtado, R., Rothman, J. M., Twinomugisha, D., Wasserman, M. D., and Goldberg, T. L. (2013). Going, going, gone: A 15-year history of the decline of primates in forest fragments near Kibale National Park, Uganda. In Marsh, L. K. and Chapman, C. A., editors, *Primates in Fragments*, pages 89–100. Springer New York, New York, NY. doi: 10.1007/978-1-4614-8839-2_7.

Chapman, C. A., Struhsaker, T. T., and Lambert, J. E. (2005). Thirty years of research in Kibale National Park, Uganda, reveals a complex picture for conservation. *International Journal of Primatology*, 26(3):539–555. doi: 10.1007/s10764-005-4365-z.

Daily, J. P., Scanfeld, D., Pochet, N., Roch, K. L., Plouffe, D., Kamal, M., Sarr, O., Mboup, S., Ndir, O., Wypij, D., Levasseur, K., Thomas, E., Tamayo, P., Dong, C., Zhou, Y., Lander, E. S., Ndiaye, D., Wirth, D., Winzeler, E. A., Mesirov, J. P., and Regev, A. (2007). Distinct physiological states of *Plasmodium falciparum* in malaria-infected patients. *Nature*, 450:7. doi: 10.1038/nature06311.

Davis, A. P., Murphy, C. G., Saraceni-Richards, C. A., Rosenstein, M. C., Wiegers, T. C., and Mattingly, C. J. (2009). Comparative Toxicogenomics Database: A knowledgebase and discovery tool for chemical-gene-disease networks. *Nucleic Acids Research*, 37(Database):D786–D792. doi: 10.1093/nar/gkn580.

del Portillo, H. A., Lanzer, M., Rodriguez-Malaga, S., Zavala, F., and Fernandez-Becerra, C. (2004). Variant genes and the spleen in *Plasmodium vivax* malaria. *International Journal for Parasitology*, 34(13-14):1547–1554. doi: 10.1016/j.ijpara.2004.10.012.

Delic, D., Wunderlich, F., Al-Quraishy, S., Abdel-Baki, A.-A. S., Dkhil, M. A., and Araúzo-Bravo, M. J. (2020). Vaccination accelerates hepatic erythroblastosis induced by blood-stage malaria. *Malaria Journal*, 19(1):49. doi: 10.1186/s12936-020-3130-2.

Deschamps, M., Laval, G., Fagny, M., Itan, Y., Abel, L., Casanova, J.-L., Patin, E., and Quintana-Murci, L. (2016). Genomic signatures of selective pressures and introgression from archaic hominins at human innate immunity genes. *The American Journal of Human Genetics*, 98(1):5–21. doi: 10.1016/j.ajhg.2015.11.014.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2013). STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1):15–21. doi: `10.1093/bioinformatics/bts635`.

Ebel, E. R., Kuypers, F., Lin, C., Petrov, D. A., and Egan, E. S. (2020). Common host variation drives malaria parasite fitness in healthy human red cells. Preprint. doi: `10.1101/2020.10.08.332494`.

Ebel, E. R., Telis, N., Venkataram, S., Petrov, D. A., and Enard, D. (2017). High rate of adaptation of mammalian proteins that interact with *Plasmodium* and related parasites. *PLOS Genetics*, 13(9):e1007023. doi: `10.1371/journal.pgen.1007023`.

Ejotre, I., Reeder, D. M., Matuschewski, K., and Schaer, J. (2020). *Hepatocystis. Trends in Parasitology*, page S1471492220301987. doi: `10.1016/j.pt.2020.07.015`.

Enard, D., Cai, L., Gwennap, C., and Petrov, D. A. (2016). Viruses are a dominant driver of protein adaptation in mammals. *eLife*, 5:e12469. doi: `10.7554/eLife.12469`.

Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admettla, A., Pattini, L., and Nielsen, R. (2011). Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLOS Genetics*, 7(11):e1002355. doi: `10.1371/journal.pgen.1002355`.

Galen, S. C., Borner, J., Martinsen, E. S., Schaer, J., Austin, C. C., West, C. J., and Perkins, S. L. (2018). The polyphyly of *Plasmodium* : Comprehensive phylogenetic analyses of the malaria parasites (order Haemosporida) reveal widespread taxonomic conflict. *Royal Society Open Science*, 5(5):171780. doi: `10.1098/rsos.171780`.

Galinski, M. R. (2019). Functional genomics of simian malaria parasites and host–parasite interactions. *Briefings in Functional Genomics*, 18(5):270–280. doi: `10.1093/bfgp/elz013`.

Gansane, A., Ouedraogo, I. N., Henry, N. B., Soulama, I., Ouedraogo, E., Yaro, J.-B., Diarra, A., Benjamin, S., Konate, A. T., Tiono, A., and Sirima, S. B. (2013). Variation in haematological parameters in children less than five years of age with asymptomatic *Plasmodium* infection: Implication for malaria field studies. *Memórias do Instituto Oswaldo Cruz*, 108(5):644–650. doi: `10.1590/0074-0276108052013017`.

Garnham, P. C. C. (1966). *Malaria Parasites and Other Haemosporidia.* Blackwell Scientific Publications Ltd., Oxford.

Goldberg, T. L., Gillespie, T. R., and Rwego, I. B. (2008). Health and disease in the people, primates, and domestic animals of Kibale National Park: Implications for conservation. In Wrangham, R. and Ross, E., editors, *Science and Conservation in African Forests*, pages 75–87. Cambridge University Press, Cambridge. doi: `10.1017/CBO9780511754920.010`.

Hanson, K. K. and Mair, G. R. (2014). Stress granules and *Plasmodium* liver stage infection. *Biology Open*, 3(1):103–107. doi: `10.1242/bio.20136833`.

Hodgson, J. A., Pickrell, J. K., Pearson, L. N., Quillen, E. E., Prista, A., Rocha, J., Soodyall, H., Shriver, M. D., and Perry, G. H. (2014). Natural selection for the Duffy-null allele in the recently admixed people of Madagascar. *Proceedings of the Royal Society B: Biological Sciences*, 281(1789):20140930. doi: `10.1098/rspb.2014.0930`.

Howick, V. M., Russell, A. J. C., Andrews, T., Heaton, H., Reid, A. J., Natarajan, K., Butungi, H., Metcalf, T., Verzier, L. H., Rayner, J. C., Berriman, M., Herren, J. K., Billker, O., Hemberg, M., Talman, A. M., and Lawniczak, M. K. N. (2019). The Malaria Cell Atlas: Single parasite transcriptomes across the complete *Plasmodium* life cycle. *Science*, 365(6455):eaaw2619. doi: `10.1126/science.aaw2619`.

Ingersoll, M. A., Spanbroek, R., Lottaz, C., Gautier, E. L., Frankenberger, M., Hoffmann, R., Lang, R., Haniffa, M., Collin, M., Tacke, F., Habenicht, A. J. R., Ziegler-Heitbrock, L., and Randolph, G. J. (2010). Comparison of gene expression profiles between human and mouse monocyte subsets. *Blood, The Journal of the American Society of Hematology*, 115(3):e10–e19. doi: `10.1182/blood-2009-07-235028`.

Jarolim, P., Palek, J., Amato, D., Hassan, K., Sapak, P., Nurse, G. T., Rubin, H. L., Zhai, S., Sahr, K. E., and Liu, S. C. (1991). Deletion in erythrocyte band 3 gene in malaria-resistant Southeast Asian ovalocytosis. *Proceedings of the National Academy of Sciences*, 88(24):11022–11026. doi: `10.1073/pnas.88.24.11022`.

Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., Loney, F., May, B., Milacic, M., Rothfels, K., Sevilla, C., Shamovsky, V., Shorser, S., Varusai, T., Weiser, J., Wu, G., Stein, L., Hermjakob, H., and D'Eustachio, P. (2020). The Reactome Pathway Knowledgebase. *Nucleic Acids*, page 6. doi: `10.1093/nar/gkz1031`.

Jha, P., Sinha, S., Kanchan, K., Qidwai, T., Narang, A., Singh, P. K., Pati, S. S., Mohanty, S., Mishra, S. K., Sharma, S. K., Awasthi, S., Venkatesh, V., Jain, S., Basu, A., Xu, S., Mukerji, M., and Habib, S. (2012). Deletion of the *APOBEC3B* gene strongly impacts susceptibility to *falciparum* malaria. *Infection, Genetics and Evolution*, 12(1):142–148. doi: `10.1016/j.meegid.2011.11.001`.

Kariuki, S. N. and Williams, T. N. (2020). Human genetics and malaria resistance. *Human Genetics*, 139(6-7):801–811. doi: `10.1007/s00439-020-02142-6`.

Kichaev, G., Bhatia, G., Loh, P.-R., Gazal, S., Burch, K., Freund, M. K., Schoech, A., Pasaniuc, B., and Price, A. L. (2019). Leveraging polygenic functional enrichment to improve GWAS power. *The American Journal of Human Genetics*, 104(1):65–75. doi: `10.1016/j.ajhg.2018.11.008`.

Kim, C. (2010). Homeostatic and pathogenic extramedullary hematopoiesis. *Journal of Blood Medicine*, page 13. doi: `10.2147/JBM.S7224`.

Köhler, S., Carmody, L., Vasilevsky, N., Jacobsen, J. O. B., Danis, D., Gourdine, J.-P., Gargano, M., Harris, N. L., Matentzoglu, N., McMurry, J. A., Osumi-Sutherland, D., Cipriani, V., Balhoff, J. P., Conlin, T., Blau, H., Baynam, G., Palmer, R., Gratian, D., Dawkins, H., Segal, M., Jansen, A. C., Muaz, A., Chang, W. H., Bergerson, J., Laulederkind, S. J. F., Yüksel, Z., Beltran, S., Freeman, A. F., Sergouniotis, P. I., Durkin, D., Storm, A. L., Hanauer, M., Brudno, M., Bello, S. M., Sincan, M., Rageth, K., Wheeler, M. T., Oegema, R., Lourghi, H., Della Rocca, M. G., Thompson, R., Castellanos, F., Priest, J., Cunningham-Rundles, C., Hegde, A., Lovering, R. C., Hajek, C., Olry, A., Notarangelo, L., Similuk, M., Zhang, X. A., Gómez-Andrés, D., Lochmüller, H., Dollfus, H., Rosenzweig, S., Marwaha, S., Rath, A., Sullivan, K., Smith, C., Milner, J. D., Leroux, D., Boerkoel, C. F., Klion, A., Carter, M. C., Groza, T., Smedley, D., Haendel, M. A., Mungall, C., and Robinson, P. N. (2019). Expansion of the Human Phenotype Ontology (HPO) knowledge base and resources. *Nucleic Acids Research*, 47(D1):D1018–D1027. doi: `10.1093/nar/gky1105`.

Kotepui, M., Piwkham, D., PhunPhuech, B., Phiwklam, N., Chupeerach, C., and Duangmano, S. (2015). Effects of malaria parasite density on blood cell parameters. *PLOS ONE*, 10(3):e0121057. doi: 10.1371/journal.pone.0121057.

Krishnamurty, A. T., Thouvenel, C. D., Portugal, S., Keitany, G. J., Kim, K. S., Holder, A., Crompton, P. D., Rawlings, D. J., and Pepper, M. (2016). Somatically hypermutated *Plasmodium*-specific IgM+ Memory B cells are rapid, plastic, early responders upon malaria rechallenge. *Immunity*, 45(2):402–414. doi: 10.1016/j.immuni.2016.06.014.

Kwiatkowski, D. P. (2005). How malaria has affected the human genome and what human genetics can teach us about malaria. *The American Journal of Human Genetics*, 77(2):171–192. doi: 10.1086/432519.

Lamikanra, A. A., Brown, D., Potocnik, A., Casals-Pascual, C., Langhorne, J., and Roberts, D. J. (2007). Malarial anemia: Of mice and men. *Blood*, 110(1):18–28. doi: 10.1182/blood-2006-09-018069.

Langfelder, P. and Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1):559. doi: 10.1186/1471-2105-9-559.

Lauck, M., Hyeroba, D., Tumukunde, A., Weny, G., Lank, S. M., Chapman, C. A., O'Connor, D. H., Friedrich, T. C., and Goldberg, T. L. (2011). Novel, divergent simian hemorrhagic fever viruses in a wild Ugandan red colobus monkey discovered using direct pyrosequencing. *PLOS ONE*, 6(4):e19056. doi: 10.1371/journal.pone.0019056.

Lee, H. J., Georgiadou, A., Walther, M., Nwakanma, D., Stewart, L. B., Levin, M., Otto, T. D., Conway, D. J., Coin, L. J., and Cunnington, A. J. (2018). Integrated pathogen load and dual transcriptome analysis of systemic host-pathogen interactions in severe malaria. *Science Translational Medicine*, page 16. doi: 10.1126/scitranslmed.aar3619.

Li, L., Stoeckert, C., and Roos, D. (2003). OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Research*, 13(9):2178–2189. doi: 10.1101/gr.1224503.

Liao, Y., Smyth, G. K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Research*, 47(8):e47–e47. doi: 10.1093/nar/gkz114.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems*, 1(6):417–425. doi: 10.1016/j.cels.2015.12.004.

Liu, T.-T., Yang, Q., Li, M., Zhong, B., Ran, Y., Liu, L.-L., Yang, Y., Wang, Y.-Y., and Shu, H.-B. (2016). *LSm14A* plays a critical role in antiviral immune responses by regulating MITA level in a cell-specific manner. *The Journal of Immunology*, 196(12):5101–5111. doi: 10.4049/jimmunol.1600212.

Mason, S. J., Miller, L. H., Shiroishi, T., Dvorak, J. A., and Mcginniss, M. H. (1977). The Duffy blood group determinants: Their role in the susceptibility of human and animal erythrocytes to *Plasmodium knowlesi* malaria. *British Journal of Haematology*, 36(3):327–335. doi: 10.1111/j.1365-2141.1977.tb00656.x.

Michon, P., Woolley, I., Wood, E. M., Kastens, W., Zimmerman, P. A., and Adams, J. H. (2001). Duffy-null promoter heterozygosity reduces *DARC* expression and abrogates adhesion of the *P. vivax* ligand required for blood-stage infection. *FEBS Letters*, 495(1-2):111–114. doi: 10.1016/S0014-5793(01)02370-5.

Miller, L., Mason, S., Dvorak, J., McGinniss, M., and Rothman, I. (1975). Erythrocyte receptors for (*Plasmodium knowlesi*) malaria: Duffy blood group determinants. *Science*, 189(4202):561–563. doi: 10.1126/science.1145213.

Morahan, B. J., Wang, L., and Coppel, R. L. (2009). No TRAP, no invasion. *Trends in Parasitology*, 25(2):77–84. doi: 10.1016/j.pt.2008.11.004.

Mukherjee, P. and Heitlinger, E. (2019). Dual RNA-Seq meta-analysis in *Plasmodium* infection. Preprint. doi: 10.1101/576116.

Nawabi, P., Lykidis, A., Ji, D., and Haldar, K. (2003). Neutral-lipid analysis reveals elevation of acylglycerols and lack of cholesterol esters in *Plasmodium falciparum*-infected erythrocytes. *Eukaryotic Cell*, 2(5):1128–1131. doi: 10.1128/EC.2.5.1128-1131.2003.

Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., Khodadoust, M. S., Esfahani, M. S., Luca, B. A., Steiner, D., Diehn, M., and Alizadeh, A. A. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nature Biotechnology*, 37(7):773–782. doi: 10.1038/s41587-019-0114-2.

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., and Minh, B. (2014). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32(1):7. doi: 10.1093/molbev/msu300.

Nielsen, R., Bustamante, C., Clark, A. G., Glanowski, S., Sackton, T. B., Hubisz, M. J., Fledel-Alon, A., Tanenbaum, D. M., Civello, D., White, T. J., J. Sninsky, J., Adams, M. D., and Cargill, M. (2005). A scan for positively selected genes in the genomes of humans and chimpanzees. *PLOS Biology*, 3(6):e170. doi: 10.1371/journal.pbio.0030170.

Oliveira, T. Y. K., Harris, E. E., Meyer, D., Jue, C. K., and Silva, W. A. (2012). Molecular evolution of a malaria resistance gene (*DARC*) in primates. *Immunogenetics*, 64(7):497–505. doi: 10.1007/s00251-012-0608-2.

Olliaro, P., Mårtensson, A., Karema, C., Vaillant, M., Dorsey, G., and Djimdé, A. (2011). Hematologic parameters in pediatric uncomplicated *Plasmodium falciparum* malaria in Sub-Saharan Africa. *The American Journal of Tropical Medicine and Hygiene*, 85(4):619–625. doi: 10.4269/ajtmh.2011.11-0154.

O'Neal, A. J., Butler, L. R., Rolandelli, A., Gilk, S. D., and Pedra, J. H. (2020). Lipid hijacking: A unifying theme in vector-borne diseases. *eLife*, 9:e61675. doi: 10.7554/eLife.61675.

Orikiiriza, J., Surowiec, I., Lindquist, E., Bonde, M., Magambo, J., Muhinda, C., Bergström, S., Trygg, J., and Normark, J. (2017). Lipid response patterns in acute phase paediatric *Plasmodium falciparum* malaria. *Metabolomics*, 13(4):41. doi: 10.1007/s11306-017-1174-2.

Ortega-Pajares, A. and Rogerson, S. J. (2018). The rough guide to monocytes in malaria infection. *Frontiers in Immunology*, 9:2888. doi: 10.3389/fimmu.2018.02888.

Otto, T. D., Gilabert, A., Crellen, T., Böhme, U., Arnathau, C., Sanders, M., Oyola, S. O., Okouga, A. P., Boundenga, L., Willaume, E., Ngoubangoye, B., Moukodoum, N. D., Paupy, C., Durand, P., Rougeron, V., Ollomo, B., Renaud, F., Newbold, C., Berriman, M., and Prugnolle, F. (2018). Genomes of all known members of a *Plasmodium* subgenus reveal paths to virulent human malaria. *Nature Microbiology*, 3(6):687–697. doi: 10.1038/s41564-018-0162-2.

Paquette, A., Harahap, A., Laosombat, V., Patnode, J., Satyagraha, A., Sudoyo, H., Thompson, M., Yusoff, N., and Wilder, J. (2015). The evolutionary origins of Southeast Asian ovalocytosis. *Infection, Genetics and Evolution*, 34:153–159. doi: 10.1016/j.meegid.2015.06.002.

Perkins, S. L. (2014). Malaria's many mates: Past, present, and future of the systematics of the Order Haemosporida. *Journal of Parasitology*, 100(1):11–25. doi: 10.1645/13-362.1.

Pierron, D., Heiske, M., Razafindrazaka, H., Pereda-loth, V., Sanchez, J., Alva, O., Arachiche, A., Boland, A., Olaso, R., Deleuze, J.-F., Ricaut, F.-X., Rakotoarisoa, J.-A., Radimilahy, C., Stoneking, M., and Letellier, T. (2018). Strong selection during the last millennium for African ancestry in the admixed population of Madagascar. *Nature Communications*, 9(1):932. doi: 10.1038/s41467-018-03342-5.

Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J. X., and Jensen, L. J. (2015). DISEASES: Text mining and data integration of disease–gene associations. *Methods*, 74:83–89. doi: 10.1016/j.ymeth.2014.11.020.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). G:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, 47(W1):W191–W198. doi: 10.1093/nar/gkz369.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140. doi: 10.1093/bioinformatics/btp616.

Simons, N. D., Eick, G. N., Ruiz-Lopez, M. J., Hyeroba, D., Omeja, P. A., Weny, G., Zheng, H., Shankar, A., Frost, S. D. W., Jones, J. H., Chapman, C. A., Switzer, W. M., Goldberg, T. L., Sterner, K. N., and Ting, N. (2019). Genome-wide patterns of gene expression in a wild primate indicate species-specific mechanisms associated with tolerance to natural simian immunodeficiency virus infection. *Genome Biology and Evolution*, 11(6):1630–1643. doi: 10.1093/gbe/evz099.

Sironi, M., Cagliani, R., Forni, D., and Clerici, M. (2015). Evolutionary insights into host–pathogen interactions from mammalian sequence data. *Nature Reviews Genetics*, 16(4):224–236. doi: 10.1038/nrg3905.

Steiner, L. A., Maksimova, Y., Schulz, V., Wong, C., Raha, D., Mahajan, M. C., Weissman, S. M., and Gallagher, P. G. (2009). Chromatin architecture and transcription factor binding regulate expression of erythrocyte membrane protein genes. *Molecular and Cellular Biology*, 29(20):5399–5412. doi: 10.1128/MCB.00777-09.

Steiper, M. E., Walsh, F., and Zichello, J. M. (2012). The *SLC4A1* gene is under differential selective pressure in primates infected by *Plasmodium falciparum* and related parasites. *Infection, Genetics and Evolution*, 12(5):1037–1045. doi: 10.1016/j.meegid.2012.02.019.

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., and Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550. doi: 10.1073/pnas.0506580102.

Sultan, A. A., Thathy, V., Frevert, U., Robson, K. J., Crisanti, A., Nussenzweig, V., Nussenzweig, R. S., and Ménard, R. (1997). TRAP is necessary for gliding motility and infectivity of *Plasmodium* sporozoites. *Cell*, 90(3):511–522. doi: 10.1016/S0092-8674(00)80511-5.

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., Burgess, S., Jiang, T., Paige, E., Surendran, P., Oliver-Williams, C., Kamat, M. A., Prins, B. P., Wilcox, S. K., Zimmerman, E. S., Chi, A., Bansal, N., Spain, S. L., Wood, A. M., Morrell, N. W., Bradley, J. R., Janjic, N., Roberts, D. J., Ouwehand, W. H., Todd, J. A., Soranzo, N., Suhre, K., Paul, D. S., Fox, C. S., Plenge, R. M., Danesh, J., Runz, H., and Butterworth, A. S. (2018). Genomic atlas of the human plasma proteome. *Nature*, 558(7708):73–79. doi: 10.1038/s41586-018-0175-2.

Tang, Y., Gupta, A., Garimalla, S., Galinski, M. R., Styczynski, M. P., Fonseca, L. L., and Voit, E. O. (2018). Metabolic modeling helps interpret transcriptomic changes during malaria. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1864(6):2329–2340. doi: 10.1016/j.bbadis.2017.10.023.

Tang, Y., Joyner, C. J., Cabrera-Mora, M., Saney, C. L., Lapp, S. A., Nural, M. V., Pakala, S. B., DeBarry, J. D., Soderberg, S., Kissinger, J. C., Lamb, T. J., Galinski, M. R., and Styczynski, M. P. (2017). Integrative analysis associates monocytes with insufficient erythropoiesis during acute *Plasmodium cynomolgi* malaria in rhesus macaques. *Malaria Journal*, 16(1):384. doi: 10.1186/s12936-017-2029-z.

Tang, Y., Joyner, C. J., Cordy, R. J., Malaria Host-Pathogen Interaction Center (MaHPIC), Galinski, M. R., Lamb, T. J., and Styczynski, M. P. (2019). Multi-omics integrative analysis of acute and relapsing malaria in a non-human primate model of *P. vivax* infection. Preprint. doi: 10.1101/564195.

The Gene Ontology Consortium (2019). The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Research*, 47(D1):D330–D338. doi: 10.1093/nar/gky1055.

Thurber, M. I., Ghai, R. R., Hyeroba, D., Weny, G., Tumukunde, A., Chapman, C. A., Wiseman, R. W., Dinis, J., Steeil, J., Greiner, E. C., Friedrich, T. C., O'Connor, D. H., and Goldberg, T. L. (2013). Co-infection and cross-species transmission of divergent *Hepatocystis* lineages in a wild African primate community. *International Journal for Parasitology*, 43(8):613–619. doi: 10.1016/j.ijpara.2013.03.002.

Tournamille, C., Colin, Y., and Cartron, J. P. (1995). Disruption of a GATA motif in the Duffy gene promoter abolishes erythroid gene expression in Duffy-negative individuals. *Nature Genetics*, 10:5. doi: 10.1038/ng0695-224.

Tung, J., Primus, A., Bouley, A. J., Severson, T. F., Alberts, S. C., and Wray, G. A. (2009). Evolution of a malaria resistance gene in wild primates. *Nature*, 460(7253):388–391. doi: 10.1038/nature08149.

van der Lee, R., Wiel, L., van Dam, T. J., and Huynen, M. A. (2017). Genome-scale detection of positive selection in nine primates predicts human-virus evolutionary conflicts. *Nucleic Acids Research*, 45(18):10634–10648. doi: 10.1093/nar/gkx704.

Vickerman, K. (2005). The lure of life cycles: Cyril garnham and the malaria parasites of primates. *Protist*, 156(4):433–449. doi: `10.1016/j.protis.2005.09.003`.

Voight, B. F., Kudaravalli, S., Wen, X., and Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLOS Biology*, 4(3):e72. doi: `10.1371/journal.pbio.0040072`.

Yamagishi, J., Natori, A., Tolba, M. E., Mongan, A. E., Sugimoto, C., Katayama, T., Kawashima, S., Makalowski, W., Maeda, R., Eshita, Y., Tuda, J., and Suzuki, Y. (2014). Interactive transcriptome analysis of malaria patients and infecting *Plasmodium falciparum*. *Genome Research*, 24(9):1433–1444. doi: `10.1101/gr.158980.113`.

Ylostalo, J., Randall, A. C., Myers, T. A., Metzger, M., Krogstad, D. J., and Cogswell, F. B. (2005). Transcriptome profiles of host gene expression in a monkey model of human malaria. *The Journal of Infectious Diseases*, 191(3):400–409. doi: `10.1086/426868`.

Zimmerman, P. A., Woolley, I., Masinde, G. L., Miller, S. M., McNamara, D. T., Hazlett, F., Mgone, C. S., Alpers, M. P., Genton, B., Boatin, B. A., and Kazura, J. W. (1999). Emergence of FY^*A^{null} in a *Plasmodium vivax*-endemic region of Papua New Guinea. *Proceedings of the National Academy of Sciences*, 96(24):13973–13977. doi: `10.1073/pnas.96.24.13973`.