Article

Distinct gastrointestinal microbial signatures predict parasite levels in controlled *Plasmodium* infections in both rhesus macaques and humans

A list of authors and their affiliations appears at the end of the paper

Functions of the gastrointestinal (GI) microbiome include maintenance of immune homeostasis and protection against infectious disease. Current assessments of the role of the GI microbiome in *Plasmodium* infection have been primarily conducted using mouse models and observational human cohorts. Here, we experimentally assessed associations between pre-infection GI microbiome composition and acute *Plasmodium* parasitemia using 16S rRNA sequencing and samples from rhesus macaques (RMs) and adult humans enrolled in a previously conducted controlled human malaria infection (CHMI) trial (NCT04072302) originally designed to test the efficacy of KAF156, a novel imidazolopiperazine class of antimalarial drugs. We identified distinct pre-infection 16S microbial signatures that were associated with increased risk for above median parasitemia in RMs infected with *P. fragile* and CHMI participants infected with *P. falciparum*. Further, we identified a *Bifidobacterium* feature set that accurately stratified parasitemia risk and could therefore serve as a foundation for a potential biomarker panel to aid prevention efforts in malaria endemic regions. Together, our findings demonstrate that pre-infection GI microbiome composition is indicative of risk for *Plasmodium* parasitemia, and our observation that the pre-infection microbiome-*P. fragile* dynamic in RMs mirrors the pre-infection microbiome-*P. falciparum* interaction in CHMI participants supports the future use of this model in pre-clinical investigations of novel microbiome-targeting approaches to reduce malaria burden.

In 2022 there were 249 million cases of malaria worldwide, resulting in over 608,000 deaths[1], demonstrating that *Plasmodium* infection, the causative agent of malaria, remains a significant global public health burden. In addition to the classic clinical features of fever and anemia, *P. falciparum* malaria causes gastrointestinal (GI) manifestations, including abdominal pain, diarrhea, and vomiting[2]. Moreover, elevated GI *P. falciparum* sequestration[3–6] causes damage to GI blood vessels and contributes to disruptions in villi formation[7], increased epithelial

permeability[7,8], and nutrient malabsorption[9–11]. These alterations may contribute to *Plasmodium*-induced microbial translocation[8,12–15] and dissemination of nontyphoidal salmonella (NTS)[16,17]. However, the precise biological mechanisms underlying *Plasmodium*-induced GI disruptions remain poorly defined.

GI microbes are essential in preserving epithelial barrier integrity and maintaining appropriate immune surveillance; suboptimal GI microbial community structure influences acquisition of enteric

✉ e-mail: jmanuzak@tulane.edu

pathogens[18]. Interestingly, GI microbial community structure may be linked with protection against severe *Plasmodium* infection. Indeed, fecal microbial community composition varied between Ugandan children that developed severe malaria anemia versus those with asymptomatic *P. falciparum* infection[19]. Additionally, an observational study of children in Mali demonstrated that time to incident *P. falciparum* blood-stage infection was significantly delayed in those with higher stool abundances of *Bifidobacterium* and *Streptococcus* species, suggesting an association between GI microbiome composition and decreased prospective risk for *Plasmodium* acquisition[20]. Further, recent work in a pediatric Malian cohort demonstrated that GI microbiome compositions enriched for inflammatory bacteria in the order Eubacteriales, which are associated with impaired gut barrier function and mucin degradation, were associated with prospective risk of febrile malaria following *P. falciparum* infection[21]. In mice, higher GI abundances of *Lactobacillus* and *Bifidobacterium* were associated with reduced parasite burden following *P. yoelii* challenge[22], while decreased abundance of *Lactobacillaceae* and increased abundances of *Enterobacteriaceae* and *Verrucomicrobiaceae* were inversely and positively correlated with *P. berghei* ANKA burden, respectively[23]. Finally, differences in GI microbial communities in mice resistant or susceptible to *P. yoelii* infection were associated with variations in the quality and specificity of anti-*Plasmodium* antibodies detected in serum, providing a potential mechanism by which the GI microbiome could influence the severity of *Plasmodium spp.* infection[24]. In sum, these findings support the theory that the pre-existing GI microbiota composition may influence prospective risk for and severity of *Plasmodium* infection.

While these prior studies have linked microbiota composition to the prospective risk for and severity of *Plasmodium* infection, the development of an animal model that replicates the complexities of intestinal immunity and mucosal barrier function in humans could accelerate assessments of the mechanistic relationships between *Plasmodium* infection and GI microbiota. In this regard, while mouse models have provided useful insights, a non-human primate (NHP) model that mimics human immune responses and *Plasmodium* infection outcomes would help to advance development of microbiome-targeting interventions, as well as reveal novel targets that could be leveraged to mitigate *Plasmodium* infection. Additionally, the use of controlled human malaria infection (CHMI) studies could aid efforts to determine the role of the microbiome in shaping *Plasmodium* infection outcome. For example, such studies could allow for highly controlled experiments to determine whether previously identified relationships between microbiome composition and risk for *Plasmodium* infection are broadly applicable across age ranges or whether microbiome-targeting therapies could be used to reduce *Plasmodium* burden. Here, we present evidence from studies conducted with adult NHPs and from a CHMI study conducted in healthy adults from the United States that (1) demonstrates that NHPs are a suitable animal model for basic and translational investigations of microbiome-*Plasmodium* interactions; and 2) identifies a significant relationship between distinct pre-infection GI microbiota signatures and the potential for above or below median parasite levels.

Results

P. fragile infection in RMs causes peripheral parasitemia and anemia

Plasmodium fragile is an NHP *Plasmodium* species with an asexual cycle of 48 h and tertian periodicity that induces parasitemia and anemia in rhesus macaques (RMs)[25]. Moreover, similar to *P. falciparum*, *P. fragile* is capable of endothelial adherence, tissue sequestration, and antigenic variation, and has been previously used as a model for cerebral malaria and for testing of anti-malarial drugs and vaccines[26–34]. Here, we intravenously (i.v.) inoculated two cohorts of male RMs (*n* = 8 each; total *n* = 16) at the Washington National Primate Research Center (WaNPRC) with *P. fragile* and assessed levels of clinical signs of

Plasmodium infection following inoculation, including parasitemia and changes in complete blood counts. We detected parasite-infected erythrocytes in peripheral blood smears of all RMs within two weeks post inoculation (p.i.) with *P. fragile* (Fig. S1A). RMs were treated with antimalarial drugs based on peripheral parasitemia thresholds. Parasitemia above 0.5% was treated with a dose of quinine sulfate once per day. Parasitemia above 1% was treated with quinine sulfate twice per day. Parasitemia above 0.5% for two consecutive weeks resulted in a switch in treatment to chloroquine. Antimalarial treatments were halted when parasitemia was observed to be below 0.5%. Anemia was characterized through complete blood counts and was defined as having a percent hematocrit (HCT) below 34.7%. Mild anemia was defined as %HCT between 27.1–34.7%, moderate anemia as 20–27%, and severe anemia as below 20%. Following *P. fragile* inoculation, anemia was sustained in RMs through Week 3 p.i. and began to resolve by Week 4 p.i., with 14 of 16 RMs returning to normal levels by Week 6 p.i. (Fig. S1B). Anemia was linked with detection of parasitemia; at Week 2 p.i., *n* = 5 RMs were mildly anemic, *n* = 5 were moderately anemic, and *n* = 5 were severely anemic (Fig. S1B and S2). Finally, we observed that *P. fragile* infection of RMs caused changes in other hematological parameters, including significant decreases in white blood cell count, red blood cell count, hemoglobin, platelet count, mean platelet volume, and absolute neutrophil count, with concurrent increases in mean corpuscular volume, mean corpuscular hemoglobin, mean corpuscular hemoglobin concentration, red blood cell distribution width, and lymphocyte and monocyte absolute count following *P. fragile* inoculation as compared to baseline (Fig. S1C-P). These findings are in line with prior work demonstrating that *Plasmodium* infection in humans leads to alterations in hematological parameters[35–38]. In sum, these data confirm that i.v. inoculation of *P. fragile* in RMs results in parasitemia, anemia, and other changes in hematological parameters indicative of malaria.

Housing in separate facilities with different diets promotes distinct GI microbial profiles in RMs that associate with potential for elevated parasitemia

To determine if the pre-existing GI microbiome played a role in parasitemia kinetics during *P. fragile* infection in RMs, we characterized stool microbial communities two weeks prior to *P. fragile* inoculation (Week −2), the day of inoculation (Week 0), and throughout the six weeks p.i. (Week 1 through Week 6). Given our hypothesis that existing differences in GI microbial composition would impact subsequent infection outcomes, we first sought to identify any pre-existing microbial distinctions within RMs prior to *P. fragile* inoculation. Using aggregated data from the two timepoints that occurred prior to inoculation (Weeks −2 and 0) for each animal, we performed unsupervised hierarchical clustering and identified two major microbial community branches (Fig. 1A). This analysis identified a clear distinction in 16S rRNA profiles that aligned with the housing facilities used during the study (WaNPRC: Building A, teal; WaNPRC: Building B, purple; Fig. 1A). Crucially, RMs in this study were raised, purchased, and transported together to the WaNPRC before being divided into cohorts and housed in the separate buildings. While housing and feeding conditions for RMs in these separate facilities were designed to be identical, investigation of facility records identified that RMs in Building A were provided with a high fiber primate chow while RMs in Building B were provided with a high protein primate chow, a difference capable of inducing the divisions in pre-infection microbiome observed between the two groups[39,40].

Notably, the response to *P. fragile* inoculation aligned with RM housing and diet, with animals possessing the Building A/high fiber diet microbial signature experiencing significantly higher peak peripheral parasitemia, despite being challenged from the same *P. fragile* stock and at the same dose as RMs possessing the Building B/high protein diet microbial signature (Fig. 1B). Specifically, Building A/high
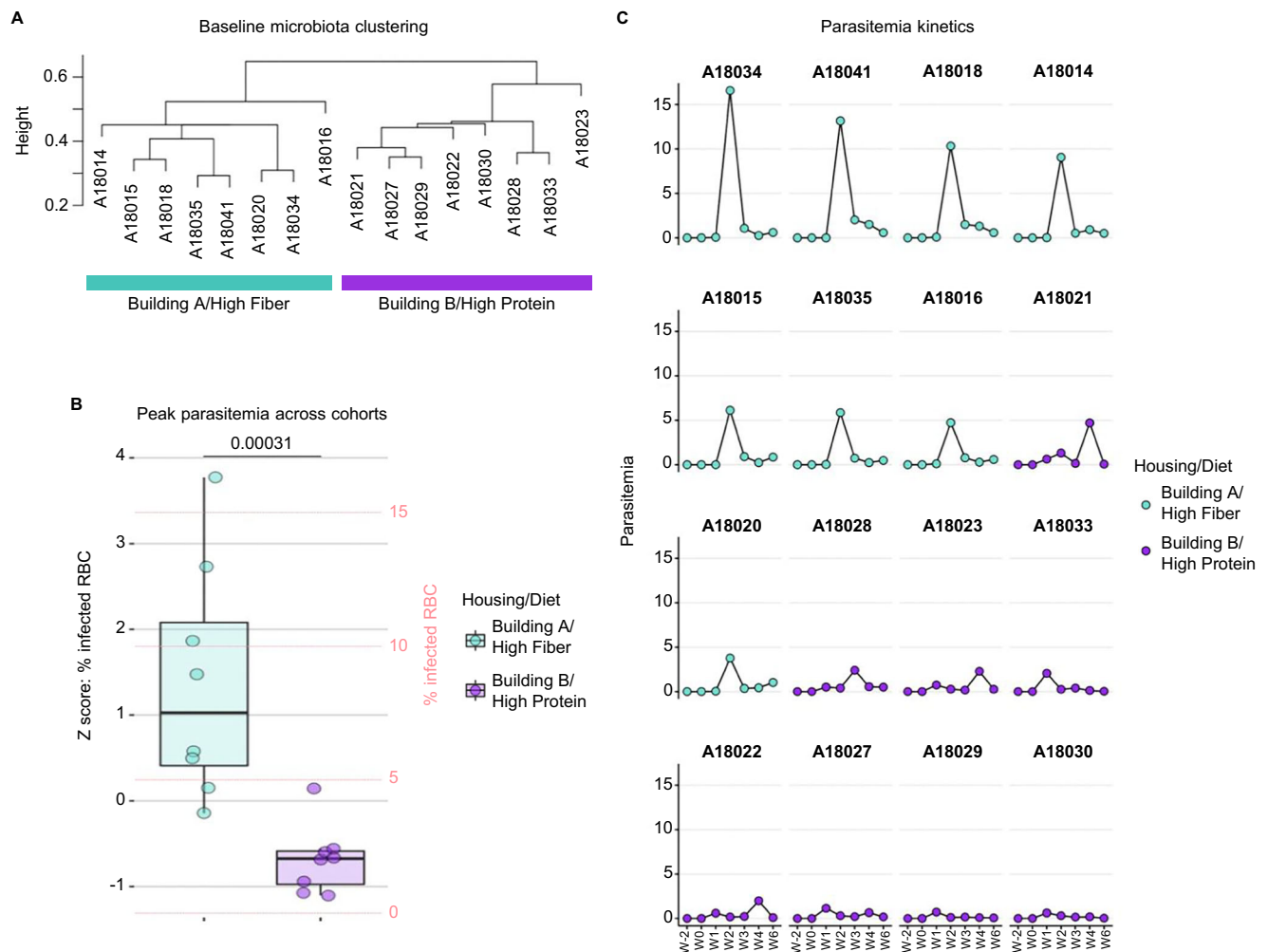
Fig. 1 | Unsupervised hierarchical clustering of 16S rRNA sequences identifies association between GI microbiota composition, housing, diet, and risk of above median *P. fragile* parasitemia in RMs. A Divisive hierarchical clustering to identify major clusters in baseline microbiota samples. RM microbiota profiles segregated according to housing location/diet using this unsupervised approach (Building A/high fiber=teal; Building B/high protein=purple; $n = 8$ RMs per group). **B** Percent infected RBC demonstrate significantly higher peak peripheral parasitemia in Building A/high fiber RMs as compared to Building B/high protein RMs ($p = 0.00031$; $n = 8$ RMs per group). Values along the left y axis indicate Z-scored % RBC, while right y axis values indicate %infected RBC. Individual data points represent Z-scored %RBC at peak parasitemia for each macaque; box limits represent the upper and lower quartiles, while box whiskers represent 1.5x the inter-quartile range. Horizontal lines within each box represent the median. Points overlaying the box and whisker bars represent data from individual RMs. Statistical significance between Building A/high fiber and Building B/high protein RM parasitemia was determined using a 2-sided Wilcoxon rank-sum test, with the *p* value shown above the horizontal line at the top of the plot. **C** %RBC data for each macaque across the entire study. Plots are ordered according to peak parasite load for each RM and colored according to cohort. W=Week.

fiber diet RMs reached peak parasite loads 3.75 times greater than Building B/high protein diet RMs ($p = 0.00031$; Fig. 1B) and were 49 times as likely to develop above median parasitemia as Building B-housed/high protein diet RMs (odds ratio = 49.00; $p = 0.010$; 95% CI = 3.83–1987). RMs with the Building A/high fiber diet microbial signature also experienced peak parasitemia sooner than RMs with the Building B/high protein diet microbial signature (Fig. 1C). The significantly different reactions to *P. fragile* infection between RM cohorts, in the context of their unique GI microbial profiles, suggests that the pre-infection microbiome composition of RMs influences subsequent parasitemia and that this risk factor may be modifiable through dietary intervention.

GI microbiota profiles remain distinct between RM cohorts throughout *P. fragile* infection despite variable profiles over time

Given our observation that differences in pre-infection microbiome were associated with variations in *P. fragile* parasitemia, we next

endeavored to determine what distinguishing microbiota features could subsequently promote enhanced *P. fragile* parasitemia. We first assessed all samples by principal coordinate analysis (PCoA) using Aitchison distance to confirm that the microbial distinctions observed prior to challenge represented stable differences (Fig. 2A). Importantly, the Aitchison metric allows for assessment of distance between samples while accounting for the compositional nature of sequencing data. This analysis demonstrated a durable distinction between microbial profiles, as samples from Building A/high fiber diet and Building B/high protein diet RMs rarely overlapped in the first 2 dimensions of principal coordinate space (Fig. S3A, B). This distinction was notable given that microbial profiles were relatively dynamic within each of the cohorts (Fig. 2A and Fig. S3C), with Building B/high protein diet RMs, which experienced lower parasitemia, demonstrating more variability over the course of the study (Fig. 2B). Finally, we conducted additional hierarchical clustering of 16S rRNA data by time point and found that while the microbiota profiles converged to some degree at Week 3 p.i., RM cohorts remained distinct at all timepoints
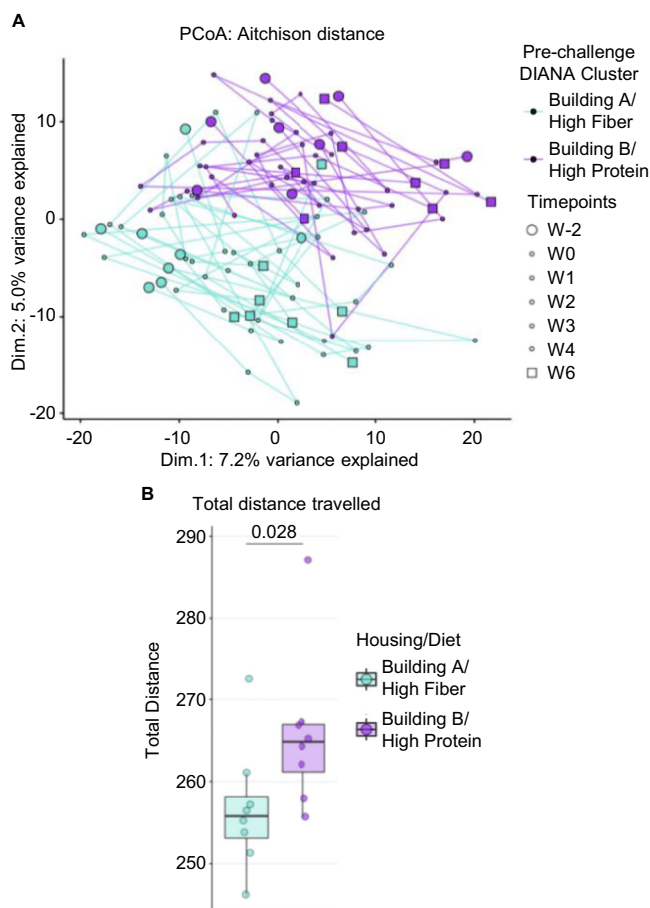
Fig. 2 | Aitchison distance demonstrates durable microbial distinctions between housing/diet cohorts. A Principal Coordinate Analysis (PCoA) of Aitchison distance for animals. Colors indicate hierarchically determined microbiota clusters for each cohort (Building A/high fiber=teal; Building B/high protein=purple; $n = 8$ RMs per group). Shapes represent unique time-points, with large circles representing the earliest sample time point (Week [W]-2), small circles representing Week 0 (W0) through Week 4 (W4), and large squares representing the final observation (Week [W]6). Lines connect timepoints for each RM. **B** Total sample-to-sample distance across time for individual animals demonstrate higher microbiota variability over time in Building B/high protein cohort as compared to Building A/high fiber cohort ($p = 0.028$; $n = 8$ RMs per group). Box limits represent the upper and lower quartiles, while box whiskers represent 1.5x the interquartile range. Horizontal lines within each box represent the median. Points overlaying the box and whisker bars represent data from individual RMs. Statistical significance between Building A/High Fiber and Building B/High Protein RM total Aitchison distance traveled was determined using a 2-sided Wilcoxon rank-sum test, with the p value shown above the horizontal line at the top of the plot.

prior to and following *P. fragile* infection, thus confirming the separation between housing/diet cohorts observed by PCoA (Fig. S4).

Microbial distinctions between cohorts are not due to differences of within-sample diversity

To determine if features of the GI community structure could explain the distinctions between RMs housed in different buildings on either a high fiber or high protein diet, we next evaluated alpha diversity metrics within and across cohorts. Microbial community structure can be assessed via 16S rRNA analysis using metrics such as Shannon diversity, which is a widely used measure of within-sample (alpha) diversity that accounts for both sample richness (the number of amplicon sequence variants (ASVs) detected) and sample evenness (the relative balance of amplicon sequence variant abundances). Our analysis of aggregated counts from the two pre-inoculation timepoints

(baseline) showed no significant differences in Shannon diversity between Building A/high fiber diet and Building B/high protein diet RMs prior to inoculation (Fig. 3A), although Building B/high protein diet RMs, who experienced lower parasitemia, trended toward higher Shannon diversity levels over time as compared to Building A/high fiber diet RMs, who experienced higher parasitemia (Fig. S5A). Notably, Shannon diversity levels were virtually equal at Week 3 p.i., when PCoA and clustering analysis suggested a convergence of microbiota communities. We also observed no differences in total observed ASVs (richness, Fig. S5B) or sequencing depth (Fig. S5C), which supports Shannon diversity estimates. Finally, we examined the relative abundance of the core microbiota (Fig. S5D). The threshold for inclusion in the core microbiome as defined by the R package "microbiome" includes all ASVs detected above 0.2% in at least 50% of samples, and by this metric, we determined that core ASVs constituted roughly the same percentage of reads across both cohorts (Fig. 3B). In sum, these data on alpha diversity levels at pre-challenge timepoints indicate that differences in overall community diversity were not a distinguishing feature between Building A/high fiber diet and Building B/high protein diet cohorts, and likely did not contribute to increased risk of parasitemia for RMs housed in Building A on a high fiber diet.

RMs that experienced elevated parasitemia were significantly enriched for *Lactobacillus* and possessed significantly lower levels of *Prevotella*

To identify organisms that were either unique to or overrepresented in each RM cohort, we utilized orthogonal approaches covering both univariate analyses (differential abundance analysis via DESeq2 and differential prevalence via Fisher's exact test) and the multivariate approach known as partial least squares discriminant analysis (PLS-DA). DESeq2 is a canonical tool in 16S rRNA studies that identifies individual ASVs with statistically significant differences in detection levels (abundances) between groups. Given our observation by PCoA that the microbiota in Building A/high fiber diet versus Building B/high protein diet RMs remained highly distinct throughout the study despite variability within animals over time, we aimed to identify significant differences in the abundances of individual taxa that were robust over time. To do so, counts were averaged within animals over time and subsequently normalized using geometric mean in DESeq2. We compared abundances at the level of individual ASVs and all higher levels of phylogenetic classification across cohorts using DESeq2. We detected 41 ASVs that significantly differed between RM cohorts (Fig. 4A; Source Data File 1). Among Building A/high fiber diet RMs, 4 of 23 ASVs that were significantly elevated could be confidently resolved at the species level, 3 of which were *Lactobacilli* including *L. coleohominis*, *L. mucosae*, and *L. salivarius* (Fig. 4A). Among Building B/high protein diet RMs, all 5 of the 18 ASVs that were significantly more abundant that could be confidently resolved at the species level were identified as *Prevotella*, including *Prevotella copri* or *Prevotella stercorea* (Fig. 4A). Importantly, these differences in *Lactobacillus* and *Prevotella* also remained significant at the genus and family levels (*Lactobacillaceae* and *Prevotellaceae*, respectively; Fig. 4A). Additionally, *Lactobacillus* remained significantly elevated in Building A/high fiber diet RMs at the levels of order (Lactobacillales) and class (Bacilli) (Fig. 4A). Other notable differences included elevated detection of *Succinivibrio*, *Moryella*, and *Parabacteroides* in Building A/high fiber diet RMs experiencing higher peak peripheral parasitemia, while Building B/high protein diet RMs with lower peak peripheral parasitemia possessed higher levels of *Aggregatibacter* and Pasteurellaceae (Fig. 4A).

While the DESeq2 approach to differential abundance is rigorous, it was originally designed to measure gene transcript, a data type where sparsity, or frequent zeros, is not common. In microbiome data, however, bacteria that are detectable in one cohort may be frequently missing in another. While DESeq2 can provide information about
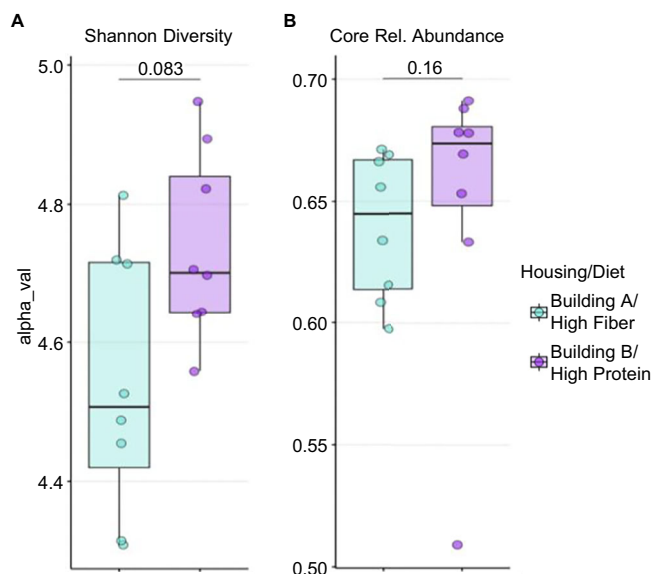
Fig. 3 | Shannon diversity and relative abundance of core microbiota do not differ across housing/diet cohorts. A Shannon diversity assessment of aggregated baseline samples (Weeks -2 and 0) shows a trend toward increased baseline diversity in the Building B/high protein cohort (purple) as compared to Building A/ high fiber cohort (teal; $p = 0.083$; $n = 8$ RMs per group). **B** The relative abundance of the core microbiota does not significantly differ between Building A/high fiber and Building B/high protein RMs ($p = 0.150$; $n = 8$ RMs per group). In both A and B, box limits represent the upper and lower quartiles, while box whiskers represent 1.5x the interquartile range. Horizontal lines within each box represent the median. Points overlaying the box and whisker bars represent data from individual RMs. Statistical significance between Building A/high fiber and Building B/high protein RM Shannon diversity and core relative abundance was determined using a 2-sided Wilcoxon rank-sum test, with the p value shown above the horizontal lines at the top of each plot.

whether an organism is more abundant in one cohort over another, it does not identify these differences in prevalence. Therefore, we performed Fisher's exact tests to evaluate ASV prevalence, allowing us to see if differences in abundance were due to the absence of specific microbes in either housing cohort. Indeed, when we assessed all differentially abundant ASVs, we detected significant differences in the prevalence of core microbiota members as well as several less common ASVs, which were often detected in only one RM cohort (Fig. 4B; Source Data File 2). Building A/high fiber diet RMs that experienced higher peripheral parasitemia possessed *L. coleohominis*, *L. mucosae*, and *L. salivarius* significantly more often and possessed some *P. copri* strains less frequently or not at all, as compared to Building B/high protein diet RMs that experienced lower peripheral parasitemia (Fig. 4B). In addition to considering these strains independently, we also evaluated the *Lactobacillus* to *Prevotella* ratio (Fig. 5A), a metric that has been associated with altered immune cell recruitment and survival in the GI tract[41]. Building A/high fiber diet RMs that experienced higher peripheral parasitemia consistently possessed a far greater *Lactobacillus* to *Prevotella* ratio than Building B/high protein diet RMs that experienced relatively low peripheral parasitemia (Fig. 5A). Together, these findings suggest a potential protective role for *Prevotella* species or strains during *P. fragile* infection. Interestingly, lower levels of *Prevotella* may foster higher levels of *Lactobacillus*, a genus associated with animals experiencing high peak peripheral parasitemia.

To move beyond case-by-case evaluation of individual ASVs, we next performed multivariate partial least squares discriminant analysis (PLS-DA), which considers all features in the data set simultaneously to identify ASVs that best distinguish between groups (Fig. 5B; Fig. S6A, B). This analysis supported several key findings from earlier

assessments, including the defining nature of the imbalance between *Lactobacillus* and *Prevotella* ASVs between Building A/high fiber diet and Building B/high protein diet RMs. PLS-DA also highlighted bacterial families frequently associated with both cohorts by way of ASVs, including *Coriobacteriaceae*, *Lachnospiraceae*, and *Ruminococcaceae* (Fig. S6B). Taken together, these results identify *Lactobacillus* and *Prevotella* as central features of RM microbiomes associated with *P. fragile* parasitemia and highlight the need for more comprehensive sequencing to resolve strain-level differences and discern more finely the importance of differential abundance and prevalence.

Pre-infection human microbiota composition associates with parasitemia during CHMI

Our investigation into associations between the microbiome and peripheral parasitemia in RMs suggested a potential role for the GI microbiota in determining the response to *Plasmodium* challenge. To increase the relevance to humans, we also assessed the stool microbiota in healthy U.S. adults that participated in a controlled human malaria infection (CHMI) study[42]. As previously reported, study participants were inoculated via infected mosquito bite ($n = 35$). Of note, only participants that progressed to detectable parasitemia by qRT-PCR ($n = 32$) were included in the sequencing analysis performed here[42]. All participants were treated with malarone (atovaquone/proguanil), which was initiated according to predetermined criteria of (1) two consecutive positive qRT-PCR test results, with at least one measurement greater than 250 parasites/ml in an afebrile participant, (2) one positive qRT-PCR result in a febrile participant, or (3) one positive thick blood smear in a febrile or symptomatic participant[42]. Stool samples were collected (1) prior to challenge, (2) at peak parasitemia (prior to antimalarial treatment), (3) post-malarone treatment, and (4) during follow-up.

As in our RM study, we hypothesized that the magnitude of acute *Plasmodium falciparum* peripheral parasitemia would vary according to participants' pre-challenge GI microbiota composition. To test this hypothesis, we clustered study participants according to their baseline microbiota profiles by performing unsupervised, divisive hierarchical clustering of CHMI participant 16S rRNA gene sequencing data. Baseline microbiota compositions could be segregated into three distinct clusters (Fig. 6A). When peak peripheral parasitemia was compared across clusters, we found that participants whose baseline microbiota was categorized as human microbiome cluster 1 (HC1; gold; median parasitemia=9491 parasites/ml) developed significantly higher levels of parasitemia as compared to HC3 participants (dark cyan; $p = 0.043$; median parasitemia=1873 parasites/ml; Fig. 6B). HC1 parasitemia was also elevated compared to HC2 participants (gray; median parasitemia=5,293 parasites/ml), although this difference did not reach statistical significance ($p = 0.17$; Fig. 6B). Differences in parasitemia between HC2 and HC3 were also non-significant ($p = 0.23$; Fig. 6B). To quantify the risk of elevated peripheral parasitemia associated with distinct microbiota profiles, we stratified participants according to their peak peripheral parasitemia, with "high" or "low" parasitemia defined as above or below the median parasite level, respectively. Participants in HC1 were eight times more likely to be identified as having above median parasitemia than HC3 participants (OR = 8.17; $p = 0.047$; 95% CI = 1.17–83.38; Fig. 6C). HC2 participants were more than four times as likely to be identified as having above median parasitemia than HC3 participants, but these results did not reach statistical significance (OR = 4.08; $p = 0.149$; 95% CI = 0.67–35.22; Fig. 6C). Importantly, neither participant age or BMI were significantly different across cluster (Fig. S7A, B). Moreover, while post-challenge treatment with KAF156 would not impact pre-challenge microbiome[42], we also evaluated the distribution of treatment modalities across these cohorts and found the differences to be insignificant by Fisher's exact test ($p = 0.195$; Fig. S7C). Taken together, GI microbiota composition prior to *P. falciparum* challenge was significantly associated with peak peripheral parasitemia in CHMI participants.
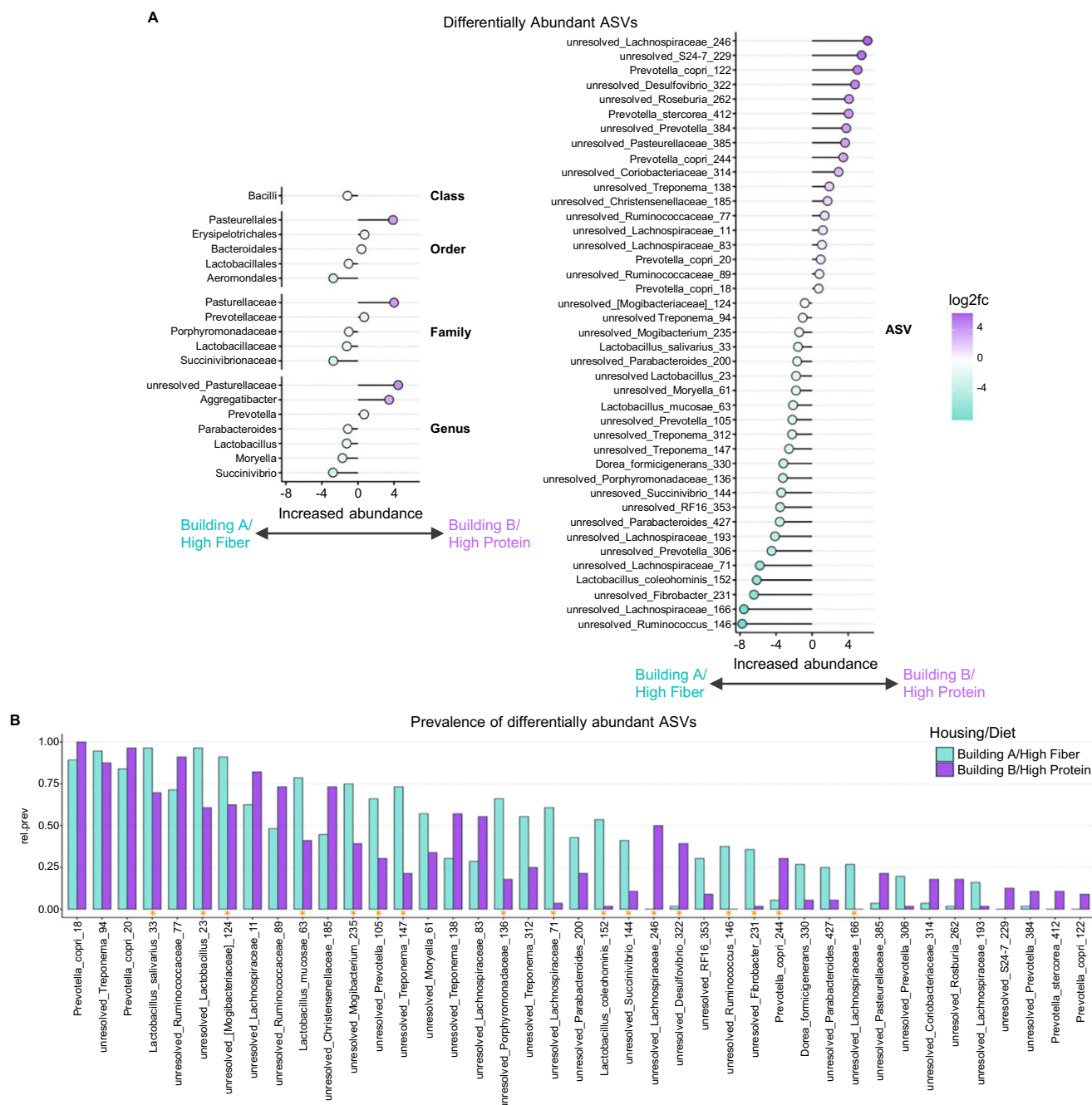
Fig. 4 | Univariate and multivariate assessments identify ASVs and higher taxonomic levels that differ and distinguish microbiota cohorts. A Differential abundance of amplicon sequence variants (ASVs) across cohorts (Building A/high fiber=teal; Building B/high protein=purple; n = 8 RMs per group). Each lolliplot represents the log2fc for ASVs or higher taxonomic levels that differ significantly, as determined by DESeq2 analysis, with points to the left being more prevalent in Building A/high fiber and those to the right more prevalent in Building B/high protein. Taxonomic levels include the class, order, family, and genus levels. The color intensity indicates the degree of log2fc. **B** Bar graph depicting the prevalence levels of differentially abundant ASVs in Building A/high fiber and Building B/high protein RMs (n = 8 RMs per group). The height of each bar indicates the proportion of samples for each cohort that were positive for detection of a particular ASV. Significant differences in prevalence (p < 0.05) as determined using a 2-sided Wilcoxon rank-sum test with Benjamini-Hochberg correction are indicated by orange stars (*Prevotella_copri_18*, p = 0.318; *Lactobacillus_salivarius_33*, p = 0.0008; *unresolved_Ruminococcaceae_77*, p = 0193; *unresolved_Lactobacillus_23*, p = 1.85e-05; *unresolved_[Mogibacteriaceae]_124*, p = 0.0015; *unresolved_Lachnospiraceae_11*,

p = 0.0374; *unresolved_Ruminococcaceae_89*, p = 0.0168; *Lactobacillus_mucosae_63*, p = 0.0004; *unresolved_Christensenellaceae_185*, p = 0.0077; *unresolved_Mogibacterium_235*, p = 0.0008; *unresolved_Prevotella_105*, p = 0.0008; *unresolved_Treponema_147*, p = 6.16e-07; *unresolved_Moryella_61*, p = 0.0286; *unresolved_Treponema_138*, p = 0.0124; *unresolved_Lachnospiraceae_83*, p = 0.0124; *unresolved_Porphyromonadaceae_136*, p = 2.49e-06; *unresolved_Treponema_312*, p = 0.0043; *unresolved_Lachnospiraceae_71*, p = 9.83e-10; *unresolved_Parabacteroides_200*, p = 0.0315; *Lactobacillus_coleohominis_152*, p = 2.49e-09; *unresolved_Succinivibrio_144*, p = 0.0012; *unresolved_Lachnospiraceae_246*, p = 1.60e-09; *unresolved_Desulfovibrio_322*, p = 3.12e-06; *unresolved_RF16_353*, p = 0.0124; *unresolved_Ruminococcus_146*, p = 7.73e-07; *unresolved_Fibrobacter_231*, p = 1.63e-05; *Prevotella_copri_244*, p = 0.0023; *Dorea_formicigenerans_330*, p = 0.0077; *unresolved_Parabacteroides_427*, p = 0.0124; *unresolved_Lachnospiraceae_166*, p = 8.49e-05; *unresolved_Pasteurellaceae_385*, p = 0.0124; *unresolved_Prevotella_306*, p = 0.0079; *unresolved_Coriobacteriaceae_314*, p = 0.0329; *unresolved_Rosburia_262*, p = 0.0124; *unresolved_Lachnospiraceae_193*, p = 0.0215; *unresolved_S24-7_229*, p = 0.0181; *Prevotella_stercorea_412*, p = 0.0318).
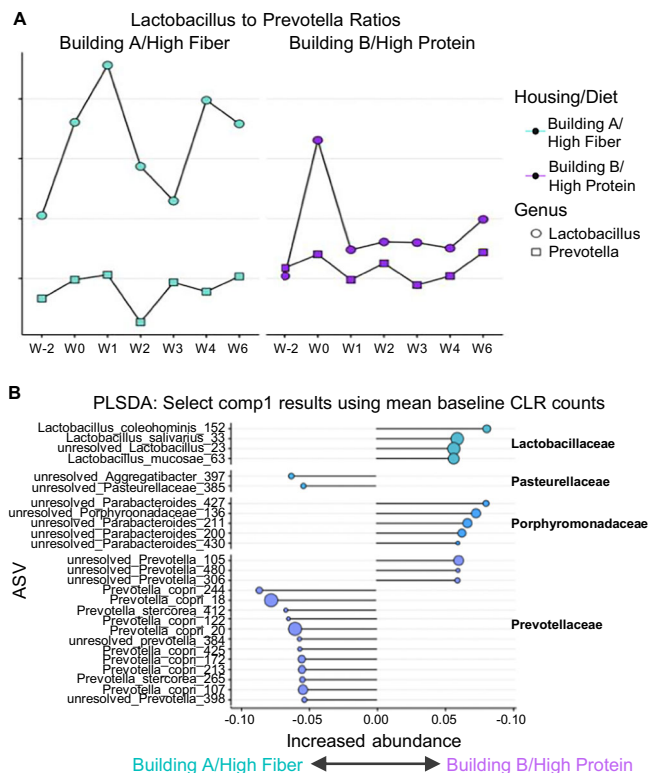
Fig. 5 | Enrichment in *Lactobacillus* and lower levels of *Prevotella* observed in RMs with elevated *P. fragile* parasitemia. A Line graph with line height along the y axis illustrating the CLR-transformed levels of *Lactobacillus* (circles) and *Prevotella* (squares) and time represented linearly on the x axis for RM cohorts (Building A/high fiber=teal; Building B/high protein=purple; *n* = 8 RMs per group). Data points are faceted by cohort. Week=W. **B** Partial Least Squares Discriminant Analysis (PLSDA) showing the separation of microbial communities between the two cohorts based on mean aggregated CLR-transformed counts of ASVs. The x axis represents the first discriminant component (comp1) which captures the maximum variance between the two cohorts. Points represent individual ASVs, with the color indicating the family and the size reflecting the relative prevalence of each ASV. The direction of the axis indicates increased abundance in Building A/high fiber to the right and Building B/high protein to the left.

HC1 CHMI participants experiencing high peripheral parasitemia exhibited significantly greater levels of microbial diversity and intra-cluster heterogeneity

Given the statistical differences in outcome that associated with baseline microbiota, we next sought to identify distinctions between microbiota clusters observed among CHMI participants. As with our RM study, we visualized the Aitchison distance between all collected samples using PCoA (Fig. 7A). While we could infer from hierarchical clustering of baseline samples that HC1 would be peripheral to HC2 and HC3 (Fig. 7A), PCoA indicated that two additional traits distinguished participants in HC1. First, samples collected at subsequent timepoints from HC1 participants remained outside of the space of specimens collected from HC2 and HC3 (Fig. 7A; Fig. S8A, B). Second, participants in HC1 experienced more shifts in GI microbiota between timepoints (Fig. 7B), with the distance between baseline and follow-up being significantly greater for participants in HC1 as compared to HC2 and HC3 (*p* = 0.0015 and 0.0076, respectively; Fig. 7C). Together, these data demonstrate that the microbiota composition from CHMI participants who developed elevated acute peripheral parasitemia was statistically different from CHMI participants who developed below median parasitemia.

We next sought to identify community characteristics that could distinguish the HC1 microbiota from HC2 and HC3. Prior to *P. falciparum* challenge, baseline samples for participants in HC1 possessed

significantly more diversity than samples classified as HC2 (*p* = 0.0011) and HC3 (*p* = 0.0021) and remained more diverse when comparing aggregated post-challenge timepoints (*p* = 0.03 and 0.022, respectively; Fig. 7D). In light of these findings, we also assessed the cumulative number of unique ASVs detected at any time within each microbiota cluster. The number of unique ASVs at baseline in HC1 was significantly greater than in HC2 (*p* = 0.01) and HC3 (*p* = 0.0055; Fig. S8C), indicating a greater range of taxa and explaining the wide dispersion of HC1 samples in PCoA space; importantly these results were robust to sequencing depth (Fig. S8D). The significant differences in diversity prompted us to evaluate the relative abundance of the "core" microbiome across microbiota clusters, which demonstrated 25 core ASVs. Importantly, this core community of microbes was a significantly smaller component of samples within HC1 compared to HC2 and HC3 at baseline and when comparing aggregated post-challenge timepoints (Fig. 7E). Further, decreasing core abundance was significantly associated with increased peak peripheral parasitemia when core microbiome relative abundance was regressed against peak peripheral parasitemia (*p* = 0.026; Fig. 7F). Together, these findings suggest that a relative reduction in the core microbiome is associated with significant increases in both microbiota diversity and risk for elevated peripheral parasitemia.

Genus-level distinctions in intestinal microbial composition are linked with parasitemia in CHMI participants

To identify ASVs with the potential to distinguish HC1 from HC2 and HC3, we again conducted univariate differential analyses using the DESeq2 pipeline to make comparisons at all taxonomic levels, including the ASV rank (Fig. 8A). Significant differences existed between HC1 and participants from HC2/HC3 even at the taxonomic rank of Phylum, with participants in HC1 possessing significantly lower levels of Actinobacteria and Firmicutes and significantly higher levels of Proteobacteria (Fig. 8A; Source Data File 3). The specific taxa driving these differences can be observed by descending the taxonomic ranks. Specifically, participants in HC1 had lower levels of several genera from the core microbiome, including *Lactococcus*, *Aldercreutzia*, *Streptococcus*, *Bifidobacterium*, and *Blautia* (Fig. 8A). Lower levels of these organisms were balanced by higher levels of *Escherichia*, *Finegoldia*, *Peptoniphilus*, and the S24-7 group (Fig. 8A). Several ASVs not included in these genera were also differentially abundant between participants in HC1 and those in HC2/HC3. Specifically, HC1 microbiotas were enriched for unresolved *Catenibacterium*, *Coriobacteriaceae*, and *Phascolarctobacterium*, while HC2 and HC3 participants possessed higher levels of several *Roseburia* and *Bacteroides* ASVs (Fig. 8A).

To provide additional context for these results, we assessed the prevalence of differentially abundant ASVs and found significant differences in prevalence across multiple genera (Source Data File 4). Specifically, the differentially abundant ASV *Bifidobacterium adolescentis* was frequently missing from HC1 participants who experienced higher peripheral parasitemia (Fig. 8B; Fig. S9A). Interestingly, viewing these results in the context of differentially abundant genera revealed that several *Bifidobacterium* ASVs were also frequently missing from these participants, including *B. longum*, *B. pseudolongum*, and *B. bifidum*, a finding that agrees with surveillance-based studies that linked increased *Bifidobacterium* abundance with increased time to first malaria episode in humans[20] and reduced malaria severity in mice[22] (Fig. 8B). We also detected similar trends within the *Adlercreutzia* and *Streptococcus* genera (Fig. 8B). For participants with higher peripheral parasitemia, the frequent absences of these taxa were filled by the presence of ASVs from *Escherichia*, *Finegoldia*, and *Peptoniphilus* (Fig. 8B). These significant differences in ASV prevalence, together with several genera-wide trends, provides expanded evidence for the association between participants' GI microbiota composition and odds of elevated parasitemia following *P. falciparum* challenge.
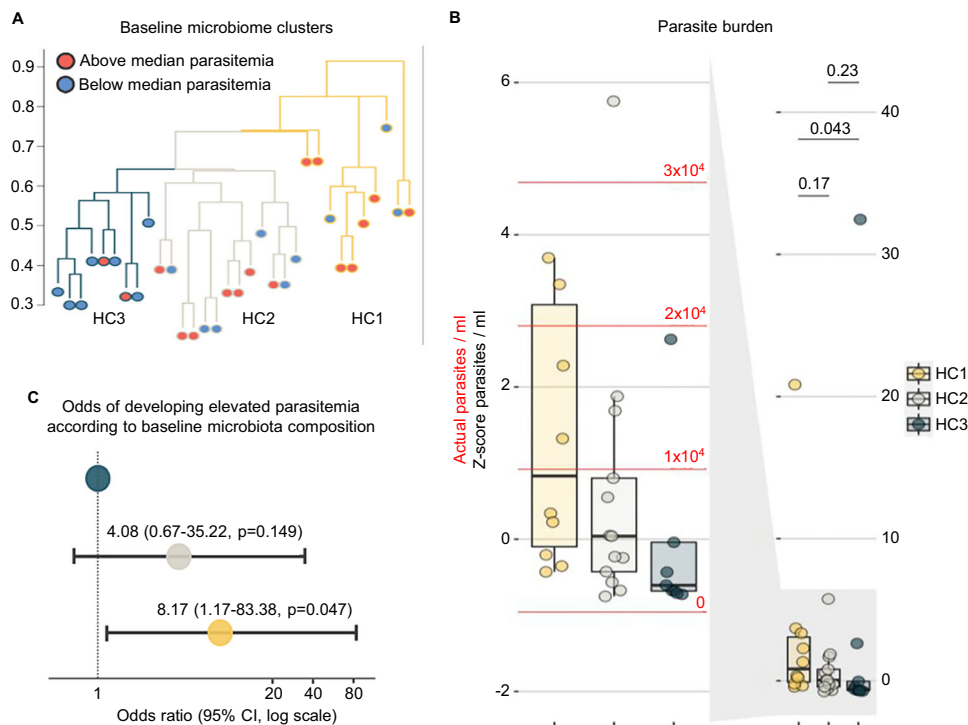
Fig. 6 | Unsupervised hierarchical clustering of 16S rRNA sequences identifies association between GI microbiota composition and risk of elevated _P. falciparum_ parasitemia. A Divisive hierarchical clustering identifies 3 major clusters in baseline microbiota samples. Major clusters were defined as human cluster 1 (HC1; gold; $n = 10$ participants), human cluster 2 (HC2; gray; $n = 13$ participants) and human cluster 3 (HC3; dark cyan; $n = 9$ participants). Branch tips indicate the parasitemia classification of each participant, which was based on detection of peak parasite levels (above median parasitemia=red or below median parasitemia=blue) via qRT-PCR. **B** Log2-transformed parasitemia values (qRT-PCR) for participants in each microbiota cluster show significantly higher peak peripheral parasitemia in HC1 participants (HC1 [$n = 10$] vs HC3 [$n = 9$], $p = 0.043$; HC1 vs HC2 [$n = 13$], $p = 0.17$;

HC2 vs HC3, $p = 0.23$). Box limits represent the upper and lower quartiles, while box whiskers represent 1.5x the interquartile range. Horizontal lines within each box represent the median. Points overlaying the box and whisker bars represent data from individual CHMI participants. The left side of the figure magnifies the plot on right, which was compressed by outliers. Significant differences in parasitemia between the three groups were assessed using a 2-sided Wilcoxon rank-sum test. **C** Odds ratio (OR) assessment to determine likelihood of participants within each microbiota cluster progressing to above median parasitemia (HC1 [$n = 10$] vs HC3 [$n = 9$]: OR = 8.17, 95% confidence interval [CI] = 1.17-83.38, $p = 0.047$; HC2 [$n = 13$] vs HC3: OR = 4.08, 95% CI = 0.67–34.22, $p = 0.149$).

We conducted a similar differential abundance analysis with participants segregated into either above or below median peak peripheral parasitemia, but this approach only yielded one ASV, an unresolved _Catenibacterium_, that associated with increased peripheral parasitemia (Fig. S9B). To overcome statistical limitations associated with univariate approaches, we again utilized the multivariate approach PLS-DA to identify the ASVs that most distinguished participants with elevated parasitemia from those with low peak parasite levels (Fig. S9C). This supervised approach ranks ASVs according to their association with a particular group for each component of the plot through a value called the loading score. In our analysis, samples from participants with above median peak peripheral parasitemia could be clearly distinguished along the first component by their positive loading scores. To maximize our ability to identify discriminatory ASVs that were associated with elevated peripheral parasitemia, we performed an identical analysis that leveraged the clustering assignments based on baseline microbiota profiles, with participants from HC1 again clustering along the x axis (Fig. S9D). We then combined the loadings from both data sets to identify (1) taxa that were associated with both elevated peripheral parasitemia and HC1; and (2) taxa associated with both lower peripheral parasitemia and HC2/HC3 (Fig. 9A; Fig. S10). This analysis demonstrated a pattern wherein ASVs that associated with HC1 and elevated peripheral parasitemia consistently mapped to taxa that were identified as the same species or resolved with a similar level of certainty. Specifically, these ASVs repeatedly mapped to _Prevotella copri_ and several unresolved _Coriobacteriaceae_, _Catenibacterium_, and

Phascolarctobacterium (Fig. 9A; Fig. S10). A large number of these ASVs were only detected in HC1, a finding that points back to earlier measurements of increased alpha diversity in this group. When considering ASVs that most associated with both HC2/HC3 and lower peripheral parasitemia, we observed similar taxonomic grouping, including _Bifidobacterium adolescentis_ and _Bifidobacterium longum_ grouping with an unresolved Bifidobacteriaceae, and several ASVs that were unresolved at the species level, including multiple _Bacteroides_ and _Erysipelotrichaceae_ (Fig. 9A; Fig. S10). Interestingly, many of the ASVs that distinguished participants from HC2 and HC3 who experienced low peak peripheral parasitemia were previously identified as part of the core microbiome in this study. Taken together, these data further indicate that the reduced detection of core microbiota in HC1 participants was associated with many low prevalence ASVs, and that this phenotype is either a direct risk factor for elevated malaria burden or a biomarker for increased risk of elevated acute parasitemia.

Because 16S rRNA sequencing analysis is not yet scalable for population-wide risk assessment, we next sought to define a reduced feature set that could serve as a potential biomarker panel to predict risk of elevated parasite burdens in individuals living in malaria endemic regions. Our first attempt sought to utilize any ASV within genera that were differentially abundant between HC1 and HC2/HC3 (Fig. 8A), creating a panel of 194 ASVs. Given the low number of participants in our data set ($n = 32$), we established a model training setup using Leave-One-Out Cross Validation (LOOCV) to ensure robust evaluation, with a focus on optimizing the Receiver
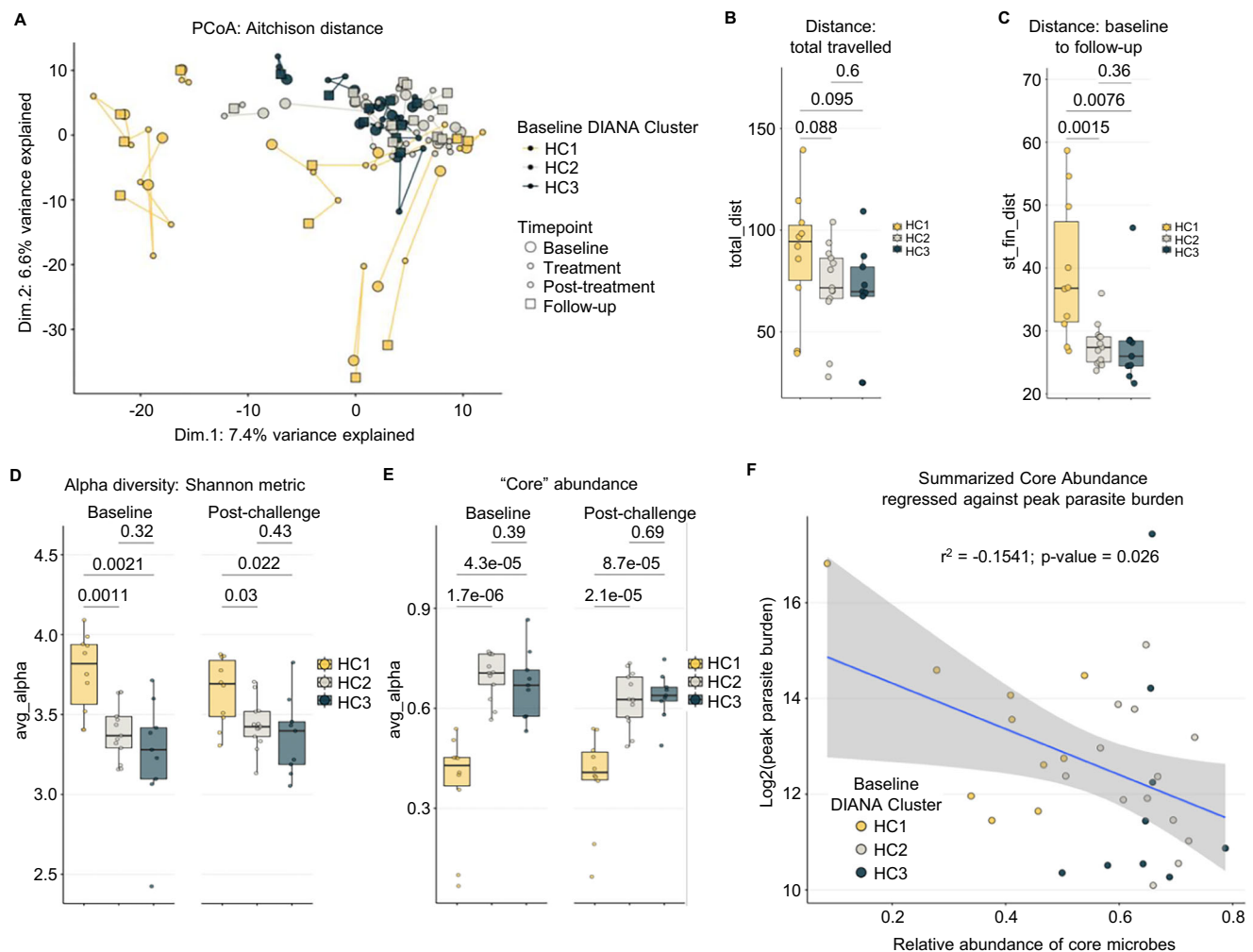
Fig. 7 | Higher GI microbiota diversity and lower relative abundance of core microbes is linked to likelihood of developing elevated peripheral parasitemia. **A** Principal Coordinate Analysis (PCoA) of Aitchison distance for CHMI participants. Colors indicate hierarchically determined microbiota clusters (HC1=gold, n = 10 participants; HC=gray, n = 13 participants; HC3=dark cyan, n = 9 participants). Shapes represent unique time-points, with large circles representing baseline samples, small circles representing pre-treatment and post-treatment timepoints, and large squares representing follow-up; lines connect timepoints for each participant. **B** Boxplots showing the total (Aitchison) distance travelled for participants within each microbiota cluster (HC1 [n = 10] vs HC3 [n = 9], p = 0.095; HC1 vs HC2 [n = 13], p = 0.088; HC2 vs HC3, p = 0.600). **C** Boxplots showing that the distance between baseline and follow-up samples within each microbiota cluster (HC1 [n = 10] vs HC3 [n = 9], p = 0.0076; HC1 vs HC2 [n = 13], p = 0.0015; HC2 vs HC3, p = 0.36). **D** Boxplots showing Shannon diversity levels at baseline and at aggregated points post-challenge (baseline: HC1 [n = 10] vs HC3 [n = 9], p = 0.0021; HC1 vs HC2 [n = 13], p = 0.0011; HC2 vs HC3, p = 0.32; post-challenge: HC1 [n = 10] vs HC3 [n = 9], p = 0.022; HC1 vs HC2 [n = 13], p = 0.03; HC2 vs HC3, p = 0.43). **E** Boxplots showing the relative abundance of the core microbiota at baseline and at aggregated points post-challenge (baseline: HC1 [n = 10] vs HC3 [n = 9], p = 4.3e-05; HC1 vs HC2 [n = 13], p = 1.7e-06; HC2 vs HC3, p = 0.39; post challenge: HC1 [n = 10] vs HC3 [n = 9], p = 8.7e-05; HC1 vs HC2 [n = 13], p = 2.1e-05; HC2 vs HC3, p = 0.69). For **B**–**E** box limits represent the upper and lower quartiles, while box whiskers represent 1.5x the interquartile range. Horizontal lines within each box represent the median. Points overlaying the box and whisker bars represent data from individual CHMI participants. Statistical significance between CHMI participant groups was determined using a 2-sided Wilcoxon rank-sum test, with the p value shown above the horizontal line at the top of each plot. **F** Scatterplot visualizing the relationship between the relative abundance of core microbes and the log2-transformed peak peripheral parasitemia. Each point represents data from individual CHMI participants (HC1, n = 10 participants; HC2, n = 13 participants; HC3, n = 9 participants). The blue line represents the linear model fit (least-squares regression line) and the shaded error bands represent the 95% confidence interval of the estimated mean response. The line indicates a negative association between peripheral parasitemia and core abundance (r² = 0.01541, p = 0.026).

Operating Characteristic (ROC) area. This model achieved an Area Under the ROC (AUROC) of 0.957, which in the setting of high feature numbers and low sample numbers indicates overfitting. Looking to avoid overfitting and to produce a panel that could be more easily deployed in a clinical setting, we evaluated strategies to reduce the feature set. Recent papers exploring machine learning in the microbiome context have identified limitations associated with large, granular feature sets, finding that Family and Genus-level feature sets outperform ASVs by AUROC metrics and are more generalizable, due in large part to sparsity of 16S rRNA data sets[43,44]. We thus chose to focus on the family Bifidobacteriaceae for 3 reasons: (1)

Bifidobacteriaceae was detected as differentially abundant between microbiome clusters (Fig. 8A) and contained several ASVs prioritized by PLS-DA analysis; (2) several ASVs from *Bifidobacteriaceae* were part of the "core" microbiome and likely more generalizable to a wider population; and (3) previous observational studies have demonstrated a relationship between *Bifidobacterium* and risk of contracting malaria in endemic regions. Our new feature set included 9 ASVs: 2 ASVs each for *Bifidobacterium adolescentis*, *Bifidobacterium longum*, and unresolved *Bifidobacterium*; and 1 ASV each for *Bifidobacterium pseudolongum*, *Bifidobacterium bifidum*, and an unresolved *Bifidobacteriaceae*. When the model was trained and evaluated as
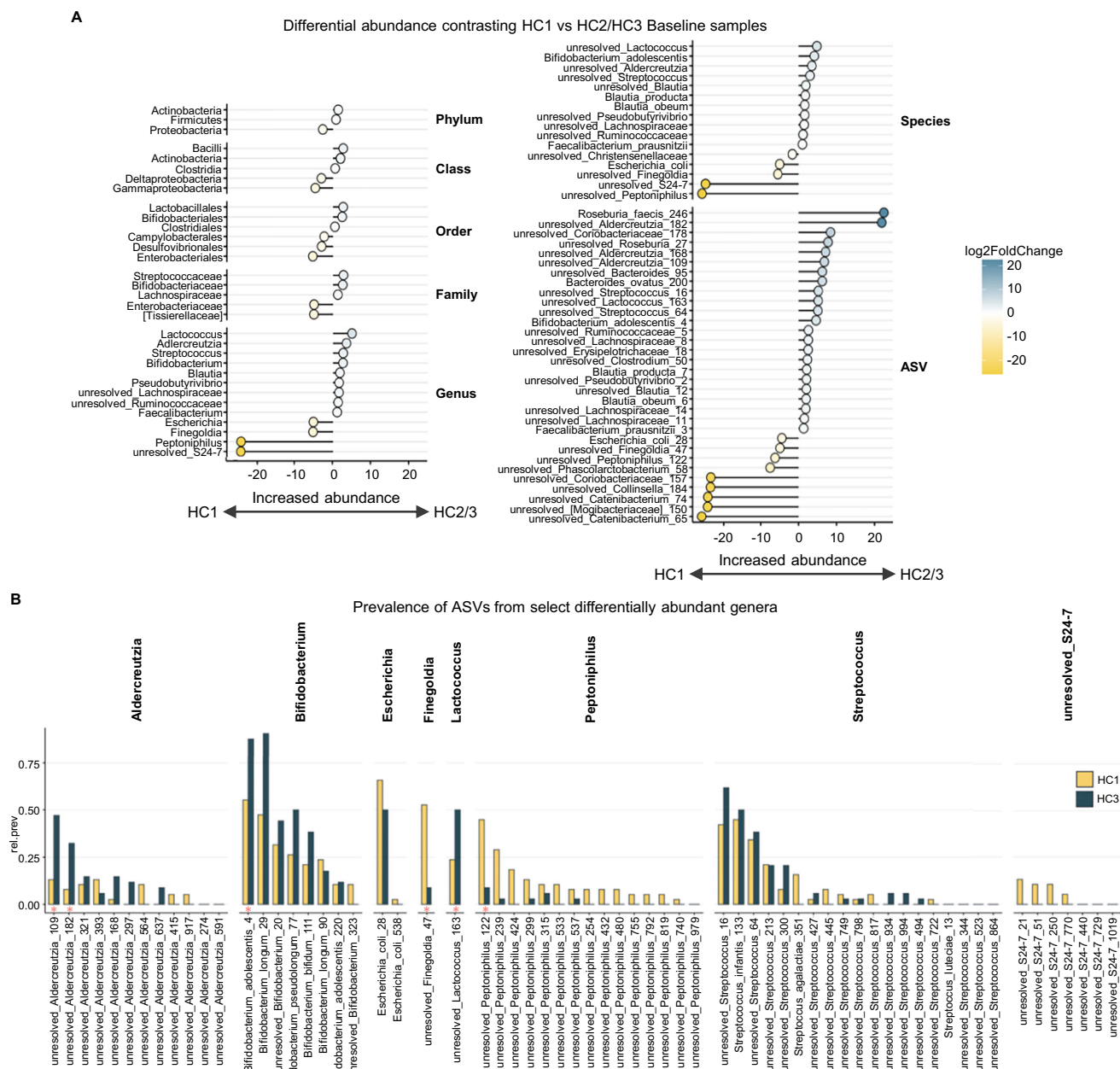
Fig. 8 | Univariate and multivariate assessments identify ASVs and higher taxonomic levels that differ and distinguish microbiota cohorts. A Differential abundance of amplicon sequence variants (ASVs) across microbiota clusters. Baseline samples were compared between HC1 (*n* = 10 participants) and the combined participants from HC2 and HC3 (*n* = 22 participants). Each lolliplot represents the log2fc for ASVs or higher taxonomic levels that differ significantly, as determined by DESeq2 analysis, with points to the left being more prevalent in HC1 and those to the right more prevalent in HC2 and HC3. Taxonomic levels include the phylum, class, order, family, and genus levels. 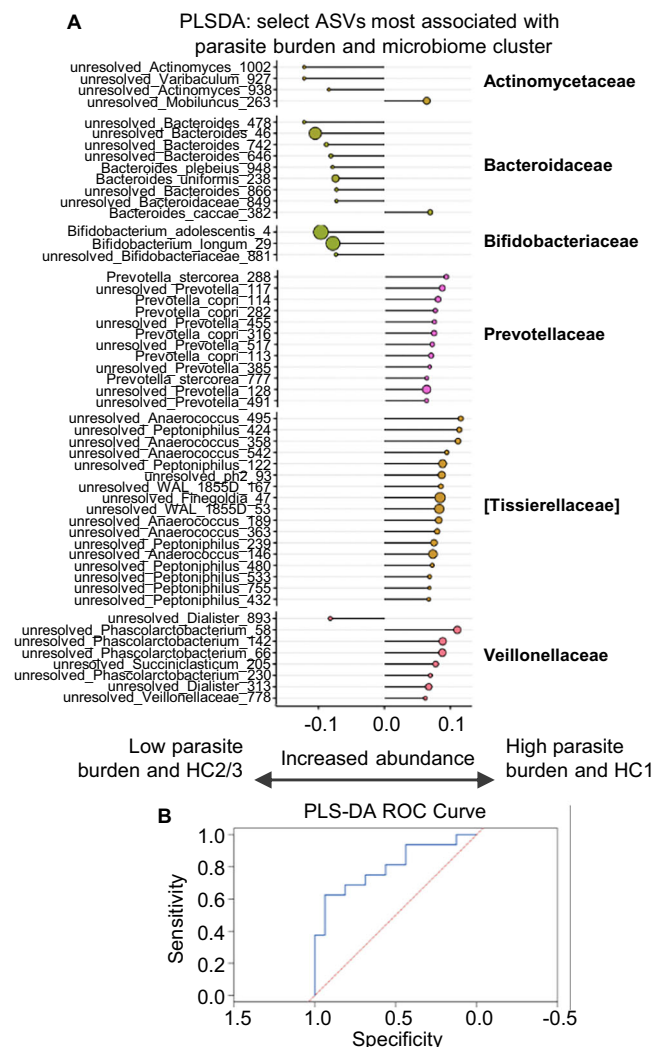The color intensity indicates the degree of log2fc. **B** Bar graph depicting the prevalence levels of select differentially abundant genera in HC1 (*n* = 10 participants) and HC3 (*n* = 9 participants). The height of each bar indicates the proportion of samples for each cohort that were positive for detection of a particular ASV. Significant differences in prevalence as determined using a 2-sided Wilcoxon rank-sum test with Benjamini-Hochberg correction were assessed only for differentially abundant ASVs and are indicated by pink stars (*unresolved_Aldercreutzia_109*, *p* = 0.0057; *unresolved_Aldercreutzia_182*, *p* = 0.0304; *Bifidobacterium_adolescentis_4*, *p* = 0.0099; *unresolved_Finegoldia_47*, *p* = 0.0006; *unresolved_Lactococcus_163*, *p* = 0.0464; *unresolved_Peptoniphilus_122*, *p* = 0.0043).

described above, our panel of 9 ASVs achieved an AUROC of 0.801, suggesting that *Bifidobacteriaceae* levels could have a predictive capability in clinical settings (Fig. 9B).

Microbial pathways in HC1 communities are enriched for pro-inflammatory products
To conclude our analyses, we utilized our 16S rRNA count data to perform a Picrust2 analysis to identify functions that are potentially enriched in a given sample based on its compositional makeup. Unsupervised analysis of the pathway enrichments via PCA

demonstrated that the majority of HC1 participants who developed elevated parasitemia levels clustered away from participants grouped in HC2 and HC3 (Fig. S11A). Interestingly, when Picrust2 results were assessed through PLS-DA to identify pathways that best distinguished HC1 (Fig. S11B), HC1 participants who developed elevated parasitemia were enriched for key steps in lipopolysaccharide and polyamine synthesis, both of which are known risk factors for epithelial inflammation and exfoliation (Fig. S11C). These analyses also suggested that ASVs found in HC1 were enriched for bacteria with potential for poly-myxin resistance.

A

PLSDA: select ASVs most associated with parasite burden and microbiome cluster

Actinomycetaceae

unresolved_Actinomyces_1002
unresolved_Varibaculum_927
unresolved_Actinomyces_938
unresolved_Mobiluncus_263

Bacteroidaceae

unresolved_Bacteroides_478
unresolved_Bacteroides_48
unresolved_Bacteroides_742
unresolved_Bacteroides_646
Bacteroides_plebeius_948
Bacteroides_uniformis_238
unresolved_Bacteroides_866
unresolved_Bacteroidaceae_849
Bacteroides_caccae_382

Bifidobacteriaceae

Bifidobacterium_adolescentis_4
Bifidobacterium_longum_29
unresolved_Bifidobacteriaceae_881

Prevotellaceae

Prevotella_stercorea_288
unresolved_Prevotella_114
Prevotella_copri_114
Prevotella_copri_282
unresolved_Prevotella_455
Prevotella_copri_316
unresolved_Prevotella_517
Prevotella_copri_113
unresolved_Prevotella_385
Prevotella_stercorea_777
unresolved_Prevotella_128
unresolved_Prevotella_491

[Tissierellaceae]

unresolved_Anaerococcus_495
unresolved_Peptoniphilus_424
unresolved_Anaerococcus_358
unresolved_Anaerococcus_542
unresolved_Peptoniphilus_122
unresolved_ph2_93
unresolved_WAL_1855D_167
unresolved_Finegoldia_47
unresolved_WAL_1855D_53
unresolved_Anaerococcus_189
unresolved_Anaerococcus_363
unresolved_Peptoniphilus_239
unresolved_Anaerococcus_146
unresolved_Peptoniphilus_480
unresolved_Peptoniphilus_533
unresolved_Peptoniphilus_755
unresolved_Peptoniphilus_432

Veillonellaceae

unresolved_Dialister_893
unresolved_Phascolarctobacterium_58
unresolved_Phascolarctobacterium_142
unresolved_Phascolarctobacterium_66
unresolved_Succiniclasticum_205
unresolved_Phascolarctobacterium_230
unresolved_Dialister_313
unresolved_Veillonellaceae_778

Low parasite burden and HC2/3 ← Increased abundance → High parasite burden and HC1

B

PLS-DA ROC Curve

Fig. 9 | *Bifidobacteriaceae* feature set could serve as biomarker panel to predict risk of elevated parasite burden. A Select results from Partial Least Squares Discriminant Analysis (PLS-DA) showing the separation of microbial communities between HC1 ($n = 10$ participants) and HC2/HC3 ($n = 22$ participants), based on mean aggregated CLR-transformed counts of ASVs. The x axis represents the first discriminant component (comp1) which captures the maximum variance between the microbial clusters. Points represent individual ASVs, with the color indicating the family and the size reflecting the relative prevalence of each ASV. The direction of the axis indicates increased abundance, with taxa associating with lower peripheral parasitemia (HC2/HC3; $n = 22$ participants) to the left, and taxa associating with higher peripheral parasitemia (HC1, $n = 10$ participants) to the right. **B** Receiver Operating Characteristic (ROC) curve for Partial Least Squares-Discriminant Analysis (PLS-DA). The ROC curve plots the true positive rate (Sensitivity) against the false positive rate (1-Specificity) to evaluate the diagnostic ability of the PLS-DA model at various threshold settings. The solid line represents the performance of the PLS-DA model, while the dashed red line indicates the line of no-discrimination, which represents a model with no diagnostic ability (random guessing). *Bifidobacteriaceae* levels achieved an AUROC of 0.801, suggesting a utility in clinical settings.

Discussion

Here, we demonstrate that the pre-existing GI microbiota composition in CHMI participants associated with peripheral parasitemia following *P. falciparum* challenge. Our strongest findings in support of this link between microbiome composition and response to *Plasmodium* challenge include significantly lower levels of the core GI commensal *Bifidobacterium adolescentis* in participants that experienced higher peripheral parasitemia. Not only were levels of *B. adolescentis* lower, but their prevalence was significantly reduced, with many participants

who experienced the highest peripheral parasitemia lacking this taxon during the study. These findings extended to trends for several other *Bifidobacterium*, including *B. longum*, *B. pseudolongum*, and *B. bifidum*, and importantly, we show that participants parasitemia levels could be accurately predicted prior to challenge using a biomarker panel comprised of just 9 ASVs within *Bifidobacteriaceae*. Notably, our results are in agreement with observational studies using incident detection in human children from malaria endemic areas[19,20] as well as previous murine studies[22–24], both of which have suggested a link between GI microbiome composition and risk for severe malaria infection. Because our work extends these observations to healthy adult humans in the U.S., it highlights *Bifidobacterium* as a defining microbial feature in limiting risk for elevated *P. falciparum* parasitemia. The agreement across studies, in terms of a role for the microbiome in parasitemia risk, is notable when considering that our CHMI study was performed on a malaria-naïve cohort of adult individuals in Seattle, Washington, whereas the previous human studies were conducted among infants and children in Uganda and Mali, where malaria incidence is seasonal. Taken together, our findings position *Bifidobacterium* as a key taxon for future studies aimed at mitigating malaria infection in endemic regions.

These reductions in a key commensal genus in CHMI participants were not limited to *Bifidobacterium*. Participants who developed the highest peripheral parasitemia were generally lacking in Actinobacteria and Firmicutes, two phyla whose ASVs dominated the core microbiome in CHMI participants. Importantly, our results indicate that decreasing levels of core microbes were significantly associated with increasing peak parasite levels. These absences were potentially compounded by increased levels of Proteobacteria, a phylum frequently associated with inflammation in the human GI tract. Participants with the highest peripheral parasitemia also evidenced a high number of relatively rare ASVs, a finding that agreed with lower GI microbiome stability in this group.

Our data suggest that elevated diversity in the GI microbiota prior to *P. falciparum* exposure could contribute to the development of above median parasitemia in CHMI participants. This finding is notable when considering that the route of CHMI infection used in this study was via infected mosquito bite, meaning that the infection passed through the clinically silent liver stage before progressing to a brief blood stage infection prior to treatment. Taken together, this fact, in combination with our observation of an association between pre-infection GI microbiome structure and very early blood stage parasitemia, further supports a link between pre-infection GI microbiome composition and peripheral parasitemia that is detectable even in the beginning phases of the blood stage of infection. Mechanistically, the loss or reduction of key commensals could result in inappropriate or inadequate priming of host immunity due to disruptions in beneficial, low-level immune stimulation[41,45–47]. Alternatively, their displacement by an expansion of uncommon taxa could disturb GI metabolic networks and subsequently promote pathologic changes to the GI barrier[48]. Intriguingly, this possibility is heightened in light of our Picrust analyses, which showed that participants who experienced higher peripheral parasitemia possessed microbiota enriched in LPS and polyamine synthesis. In either hypothetical, host immune disruptions could manifest locally and systemically such that *Plasmodium* replication is promoted. Recent work has also demonstrated a propensity for certain bacterial species to leverage within host evolution to promote their translocation to the liver, altering immune homeostasis in this tissue[49]. Future investigations focusing on the translocation potential of the differentially abundant GI microbial taxa identified in this study are needed. Alternatively, microbial composition may serve as a proxy for indirectly related facets of overall health that may ultimately play larger roles in malaria progression; our identification of a similar phenomenon in healthy RMs, however, argues against this. Future investigations to determine the precise

mechanisms by which GI microbiota composition influences *P. falciparum* parasitemia in humans are warranted.

As in CHMI participants, pre-challenge GI microbiota composition in RMs was significantly associated with risk of or protection from elevated *P. fragile* parasitemia. In particular, pre-challenge differences in the *Lactobacillus:Prevotella* ratio between cohorts remained stable over time, with RMs that developed the highest peripheral parasitemia enriched for *Lactobacillus* taxa and lacking *Prevotella*. Previous studies from our group and others have suggested that an imbalance in ASVs from these genera can disrupt immune homeostasis in the GI tract by altering immune cell abundance, activation potential, and survival[41]. While the alignment between housing/diet cohorts, microbiomes, and *P. fragile* prevents our RM analyses from producing causal associations, the significant association between microbiome composition and peripheral parasitemia in our human study supports our hypothesis and findings that similar interactions occurred in RMs. Given the striking divergence in microbiome composition that developed in our RMs in less than three months after separation and housing in different buildings where they received either high protein or high fiber primate chow, these findings are important for both the malaria research community and the NHP research community more broadly.

Notably, although the GI microbiome was generally resilient to transient changes during *Plasmodium* infection in both CHMI participants and RMs, our data showed that when disruptions occurred, they were observed near peak parasitemia. Similarly, RMs infected with *P. cynomolgi* also experienced differences in rectal community structure that coincided with peak parasitemia, including significantly increased abundances of *Helicobacteraceae* and significantly decreased abundances of *Lactobacillaceae*, *Ruminococcaceae*, and *Prevotellaceae*[50]. These findings in RMs are also consistent with previous work which revealed only minimal differences in stool community structure and composition in Kenyan infants before and after *P. falciparum* infection and treatment[51]. In contrast to data from CHMI participants, RMs did not exhibit differences in pre-infection abundances of *Bifidobacterium* that were subsequently associated with above or below median *P. fragile* parasitemia. Rather, the presence of *Prevotella* in RM was noted as being associated with lower risk for *P. fragile* parasitemia. These findings do not negate our results, as the overall differences in core taxa between species are well documented[52]. Importantly, despite these taxonomic differences in the GI microbiome between species, it is possible that the specific pre-infection microbial species identified in CHMI participants and RMs that associate with above or below median parasitemia risk may have similar functionality that influences risk or protection from *Plasmodium* parasitemia. More work will be needed to test this theory and elucidate the specific mechanism by which the GI microbiome influences parasitemia. Taken together, the overall trends in our RM GI microbiome findings agree with both our CHMI data and prior human observational and murine studies in that pre-infection microbiome composition is associated with *Plasmodium* parasitemia. Moreover, our findings indicate the translational potential of the RM *P. fragile* model. Given the additional similarities between RMs and humans in terms of GI functionality and mucosal immunity, future studies utilizing this model to investigate the impact of pre-infection GI microbial functionality on *Plasmodium*-mediated disruptions in intestinal mucosal barrier integrity and peripheral parasitemia, and the potential utility of dietary composition modulation to mitigate poor disease outcomes are justified.

A major strength of our RM model is that it allowed for a greater degree of control over critical experimental and population variables. Indeed, although our CHMI study employed narrowly defined parameters for recruitment of healthy U.S. adults[42], there remained a potential for unassessed genetic, nutritional, and social factors which may have contributed to microbiota variability between study participants. This created a challenge for determining whether observed differences in baseline GI microbiota composition were a proxy for

some other aspect of *P. falciparum* susceptibility, or alternatively, played a direct role in risk for above median parasitemia. Intriguingly, our RM model allowed us to overcome some of these challenges, including the reduction of social, economic, and medical confounders common among human cohorts. Importantly, here we found that housing of RMs in different facilities with different diets promoted distinct GI microbial profiles. It is well accepted that diet is a major contributor to GI microbiota composition in both humans and RMs[40,53]. In particular, prior work has indicated that dietary changes (e.g., shifts to high protein/fat or high fiber diets) in humans is associated with detectable changes in GI microbial composition[39,40]. For example, a diet high in fiber in humans has been linked with increased abundance of *Prevotella* and *Lactobacillus*[54,55]. Interestingly, RMs naturally have a high abundance of *Prevotella* and *Lactobacillus*, likely due to their primarily pant-based diets[56,57]. This is in line with the data in our study, which indicated that although both cohorts of RMs possessed both *Prevotella* and *Lactobacillus*, they experienced a divergence in the *Lactobacillus* to *Prevotella* ratio upon cohort separation into distant housing facilities where they were fed a standardized diet of commercial primate chow rich in fiber (Building A, crude fiber not more than 10% and crude protein not less than 20%) or high in protein (Building B, crude fiber not more than 6.5% and crude protein not less than 25%), in addition to daily fruits and vegetables. Specifically, the RMs in Building A with the high fiber diet skewed towards a more *Lactobacillus* dominant microbiome with lower *Prevotella* abundance, while Building B RMs with the high protein diet skewed towards a greater abundance of *Prevotella* with lower *Lactobacillus*. Thus, our data agrees with previous literature showing that RMs and a primarily plant-based diet have a GI microbiome dominated by *Prevotella* and *Lactobacillus* and demonstrate that diets that differ in the percentage of fiber and protein included may influence the dominance of one over the other. Additionally, housing may also influence GI microbial community structure. Prior work has demonstrated that even with 80% diet similarity between a free-ranging/semi-wild population of RMs compared to captive facility-based RMs there were significant differences in 22 microbial genera between the two groups, including increased *Prevotella* and *Lactobacillus* in the facility-based RMs as compared to the free-ranging RMs[58]. The degree to which specific diets and housing conditions impact the microbiome, and the speed with which changes in housing/diet can cause compositional shifts requires more study. Notably, while cohort and animal-to-animal alterations in GI microbiota composition prior to *P. fragile* challenge occurred due to housing and dietary composition differences, this variability was unlikely to indicate unappreciated comorbidities or other factors that may have influenced *P. fragile* infection. Therefore, our study provides both a motivation and foundation for future *Plasmodium* experiments in NHP that could directly assess our conclusions through experimental manipulation of microbiomes prior to *P. fragile* challenge. Indeed, previous work by our group and others has shown the ability to promote GI microbiome shifts through antibiotic administration[59,60]. This tool, in combination with dietary composition modifications, could allow for the targeted alteration of microbial community structures and subsequent testing of how these microbial shifts influence *Plasmodium* infection outcome. In sum, our study and results presented here verify the value of this critical model system for future examination of *Plasmodium*-microbiome interactions and identify dietary composition modulation as a potential therapeutic avenue to reduce *Plasmodium* infection risk.

A critical aspect of our study was the use of orthogonal approaches to segregate microbiome cohorts at baseline and assess for key features that distinguished these cohorts. Specifically, the use of divisive hierarchal clustering, as opposed to agglomerative, enhanced our ability to segregate microbiota communities into meaningful clusters. The ability to then identify stepwise increases in parasitemia across these microbiota clusters allowed us to verify previous studies

suggesting an association between the microbiome and *Plasmodium* parasitemia and extend these findings to a human adult cohort in the U.S., as well as a physiologically relevant RM model system. Importantly, our use of multivariate assessments via PLS-DA, which identify and rank the ASVs that best distinguished cohorts, fully supported our findings using canonical univariate techniques, such as differential abundance. In sum, the data presented here and our specific bioinformatic approach suggests the generalizability of our observations beyond specific CHMI participant populations, as well as to other host and *Plasmodium* species.

A caveat of our study is that our observed association between pre-infection microbial community structure and risk for *P. falciparum* parasitemia was obtained using samples from a CHMI study conducted among healthy adults in the US. Given that the GI microbiome is different between individuals with disparate diets, as emphasized by our RM cohort findings, and that reside in varied geographical locations[61], our identification of specific microbes associated with risk or protection from *P. falciparum* parasitemia in healthy US adults may not fully mirror the dominant microbial communities that could be linked with *P. falciparum* parasitemia risk in individuals in areas of moderate-to-high malaria transmission. In addition, because exclusion criteria for CHMI participants included a history of malaria and residence in a malaria-endemic area for more than 6 months, it is possible that a lack of regular exposure to malaria antigens, as would be expected in malaria endemic areas, could lead to differences between the outcomes we observed among CHMI participants and what would be seen in individuals that reside in malaria endemic areas. These potential differences are notable in light of prior work indicating that compositional differences in the microbiome are linked to vaccine efficacy and immunogenicity[62]. A further limitation of our CHMI study is participants were treated immediately after detection of blood-stage infection, as required by the parent protocol[42], which likely contributed to the low levels of parasitemia observed. Although this means that the parasitemia levels in these participants are lower than what is likely to occur in real-world settings, this also speaks to the magnitude of the associations we observed here between pre-infection microbial community structure and post-infection parasitemia levels and provides rationale for future studies that assess prospective clinical cohorts where individuals in malaria endemic areas are observed at point-of-care. A similar limitation in our RM study exists in that it is possible that any differences that may have occurred in the timing of sample collection for parasitemia checks between the Building A/high fiber diet and Building B/high protein diet cohorts could have influenced parasitemia readouts. Because data on the timing of blood collection for parasitemia checks in RMs are unavailable, our results should be interpreted with this caveat in mind and future studies should account for the impact of timing of sample collection on parasitemia. Finally, CHMI participant stool samples in our study were frozen within 24 h of collection. Similarly, in our RM study, animals were moved into protected contact the night before sampling and fecal material was collected in the morning. Because stool may have been produced any time between separation and sampling, it is possible that several hours may have passed between collection and freezing. Although the length of time between collection and freezing could have impacted microbial communities, prior work has shown that room temperature storage of stool samples for 72 h has been reported to cause less shifts in detected communities than chemical storage methods[63]. Moreover, an additional study showed that variation within donors was less than variation between samples frozen immediately or kept at room temperature for 1 week[64]. Taken together, it will be important for future studies to consider how microbial differences between adult CHMI participants in the US and individuals in malaria endemic areas could intersect to influence both malaria vaccine effectiveness as well as *Plasmodium* burden and disease progression. Moreover, steps should be taken to ensure that rigorous

specimen collection timing and storage methods are used to ensure robust results in both CHMI and RM studies.

An additional caveat of our study is that CHMI participants and RMs were infected with *Plasmodium* via different routes. Indeed, the CHMI participants were inoculated with *P. falciparum* via infected mosquito bite while RMs were inoculated by i.v. injection of *P. fragile*-infected erythrocytes. This means that the *Plasmodium* infection in CHMI participants passed through the clinically silent liver stage before progressing to a brief blood stage infection prior to treatment, while the RM infection bypassed the liver stage. Several studies have demonstrated that the overarching links between GI microbiome structure and malaria severity that have been observed in children in Mali and Uganda infected with *P. falciparum* via mosquito bite also exist in mice inoculated with *Plasmodium* via the i.v. route[19,20,22,24,65,66]. Therefore, although it is possible that differences in the route of infection may contribute to variations in the observed outcomes between CHMI participants and RMs, further work will be necessary to identify the precise ways in which the GI microbiome may impact, and in turn be impacted by, *Plasmodium* infection when exposed via different routes. Notably, the strain of *P. fragile* used in the studies described here has been shown to retain the ability to produce infective gametocytes and *P. fragile* sporozoite transmission via mosquito bite has been successfully conducted[27], thus future pre-clinical assessments of microbiome-targeting interventions to reduce parasitemia in RMs in the context of vector transmission of *P. fragile* parasites are feasible and warranted.

In summary, the data presented here demonstrates that pre-existing GI microbiome composition is indicative of *Plasmodium* peripheral parasitemia in CHMI participants in the U.S. and RMs. Our findings support the existence of a broad link between GI microbiome communities and *Plasmodium* parasitemia and represent a critical extension of previous murine and human studies[19,20,22–24] into an adult human population in the U.S. The recapitulation of this phenomenon in a relevant RM model expands the ability to perform mechanistic investigations of microbiota-influenced anti-malarial immunity. While our current study is limited in its ability to make mechanistic determinations on how the GI microbiome contributes to protection against above median *Plasmodium* parasitemia, the identification of high-level taxonomic differences in both CHMI participants and RMs suggests that microbiome targeting therapies, such as dietary composition changes (e.g., high fiber or high protein), microbiome modulation (e.g., probiotics), or small molecule supplementation (e.g., microbial-derived bioactive molecules), could be investigated to mitigate *Plasmodium* infection. Taken together, our data provide a foundation for future RM and CHMI studies to identify the functional role of distinct microbial communities in controlling malaria parasitemia, and pre-clinically test preventative and interventional strategies leveraging the microbiome to reduce malaria burden.

Methods

Nonhuman primate study animals, approval, and sample collection

Male Indian-origin rhesus macaques (*Macaca mulatta*; RMs; $n = 16$) between the ages of 3.09–3.41 years (median = 3.2 years; Table S1) were housed and cared for at the Washington National Primate Research Center (WaNPRC) under a protocol reviewed and approved by the University of Washington Office of Animal Welfare (OWA) Institutional Animal Care and Use Committee (IACUC; Protocol 4266-13; Animal Welfare Assurance Number D16-00292; see Source Data File 5 for RM metadata). The housing, care, and procedures performed in this study were done in AAALAC-accredited facilities, in accordance with the regulations put forth by the US Department of Agriculture, including the Animal Welfare Act (9 CRF) and the Animal Care Policy Manual, as well as with the guidelines outlined by the National Research Council in the Guide for the Care and Use of Laboratory Animals and the

Weatherall Report. RMs in this study were raised, purchased, and transported together to the WaNPRC and housed in WaNRC Building A. RMs were then divided into two cohorts of $n = 8$ RMs each. One cohort of $n = 8$ RMs remained in Building A while the second cohort of $n = 8$ RMs was relocated to WaNPRC Building B and housed in this location for 6 weeks prior to the first stool collection, which occurred two weeks prior to *P. fragile* inoculation. All RMs were provided with a commercial primate chow (Building A: Laboratory Fiber-Plus Monkey Diet 5049, Catalog #0005326; Building B: High Protein Monkey Diet 5045, Catalog #0001335, both from Lab Diet, PMI Nutrition International, St. Louis, MO) twice per day, supplemented with daily fruits and vegetables and water ad libitum. Composition of each diet is given in Table S2. Novel food items, foraging opportunities, and destructible and indestructible manipulanda were provided as environmental enrichment. All RMs were housed in stainless steel cages with a 12/12 light cycle. All cage pans and animal rooms were cleaned daily and sanitized at least every two weeks. Stool was collected from each RM two weeks prior to *P. fragile* inoculation (Week −2), the day of *P. fragile* inoculation (Week 0), and during *P. fragile* infection and treatment (Weeks 1, 2, 3, 4, and 6). The evening prior to stool collection, RMs in full social contact were shifted into protected (grooming) contact only. The following morning, stool was collected from each animal's cage pan and immediately stored at -80 °C. Following cage pan stool collection, RMs were returned to full social contact. Upon thawing, DNA was extracted from NHP stool samples and used for 16S rRNA sequencing analysis.

P. fragile inoculation, post-infection monitoring, and treatment

In this study, we used the Sri Lanka strain of *P. fragile*[26,27] to experimentally inoculate RMs. To prepare *P. fragile*-infected RM erythrocytes for inocculation[67], frozen erythrocytes from a *P. fragile*-infected RM were thawed rapidly at 37 °C and transferred into a sterile tube. Erythrocytes were resuspended in 12% NaCl while gently shaking, followed by a 5 min incubation at room temperature without shaking. Next, 1.6% NaCl was added dropwise while gently shaking, followed by centrifugation at 228 x g, room temperature, 10 min. After centrifugation, the supernatant was aspirated and discarded, and the pellet resuspended in a solution of 0.9% NaCl and 2% dextrose, followed by centrifugation at 228 x g, room temperature, 10 min. Finally, the supernatant was aspirated and discarded, and the pellet resuspended in 0.9% NaCl solution. RMs were intravenously inoculated with 20×10^6 *P. fragile*-infected erythrocytes in 1 ml 0.9% NaCl solution.

Giemsa staining of thin blood smears was used to assess *P. fragile* parasitemia in each RM. Blood for the smears was collected from non-sedated RMs using positive reinforcement conditioning. Parasitemia checks occurred three times per week following *P. fragile* inoculation, with an additional check occurring at week 6 post-*P. fragile* infection. Parasitemia was quantified using a light microscope and 1:10 ratio Miller Disk Reticle (Microscope World, Carlsbad, CA)[68]. Briefly, the total number of all red blood cells (RBCs), including both infected RBCs (iRBCs) and non-infected RBCs, inside the minor square are enumerated, followed by counting of all iRBCs in the major square. A total of fifteen fields were counted in three separate areas of the slide, for a total of 45 fields. The % parasitemia was calculated in each of the three areas as follows: (# of iRBC in major square / # of all RBC in minor square) x 10. The % parasitemia in each of the three areas was then averaged together to determine the final % parasitemia for each RM at each timepoint.

The % parasitemia at each time point was used to determine the need for anti-malarial treatment. RMs were treated using the following guidelines: if parasitemia was between 0.5–1%, RMs were treated orally with a suboptimal dose of quinine sulfate (150 mg; NDC: 53489-0700-07; Sun Pharmaceuticals, Princeton, NJ)[69,70] with the goal of establishing a chronic malaria infection that would model the prolonged exposure to *Plasmodium* infection that occurs in

malaria-endemic areas[71,72]. If parasitemia was above 1%, RMs were treated orally with 150 mg quinine sulfate twice per day. If parasitemia was above 0.5% for two consecutive weeks despite quinine sulfate treatment, RMs were switched to oral chloroquine treatment (10–20 mg/kg; NDC: 64980-0178-02; Rising Pharmaceuticals, East Brunswick, NJ) once per day. Of note, two RMs (A18028 and A18021) had three consecutive weeks of parasitemia above 0.5% before switching to chloroquine treatment. Quinine sulfate and chloroquine treatments were halted as soon as RMs were observed to be below the 0.5% parasitemia threshold. Treatment information for all RMs included in this study, including timings and dose, are shown in Table S1.

Controlled human malaria infection (CHMI) study and approval

Healthy male and female adult participants ($n = 35$) in the US were enrolled in a previously conducted, 2-part, randomized, double-blinded, placebo-controlled, single-center study completed at the Seattle Malaria Clinical Trials Center to test the prophylactic efficacy, safety, and pharmacokinetics of KAF156, a novel imidazolopiperazine class of antimalarial drug[42] (see Source Data File 6 for CHMI participant metadata). The study was approved by the Western Institutional Review Board with reliance agreements for the University of Washington and Fred Hutchinson Cancer Research Center (clinical trials registration number: NCT04072302)[42]. All participants provided written informed consent for the present study. As reported previously, KAF156 was found to be safe, well-tolerated, and resulted in a high degree of protective efficacy from *P. falciparum* infection at higher doses, which was tested through a controlled human malaria infection (CHMI) model[42]. A subset of participants that received either placebo ($n = 8$), 20 mg ($n = 13$), or 50 mg ($n = 14$) of KAF156 post-CHMI exposure to *P. falciparum*-infected mosquitoes also provided stool samples at four distinct time points: (1) within 1−2 days of CHMI; (2) during a three-day course of treatment with the antimalarial drug Malarone (atovaquone/proguanil); (3) at the end of Malarone treatment; and (4) during CHMI and treatment follow-up (within D28-D42). All $n = 35$ participants received the Malarone treatment prior to collection of the final follow-up sample. *P. falciparum* infection in study participants was monitored in whole peripheral blood using a *Plasmodium* 18S ribosomal RNA (rRNA) qRT-PCR that quantified Pf A-type 18S rRNA[42,73,74]. All participants in the placebo and 20 mg KAF156 treatment groups, and all but three of the 50 mg KAF156 treatment groups developed detectable *P. falciparum* infection. Of note, the three participants that did not develop detectible *P. falciparum* parasitemia by qPCR were not included in the sequencing analysis performed here. Stool samples were frozen at -80 °C within 24 h of collection. Upon thawing, DNA was extracted from human stool samples and used for 16S rRNA sequencing analysis, as detailed below.

16S rRNA sequencing

DNA was extracted from both NHP and CHMI participant stool samples using the PowerFecal DNA Isolation Kit (Qiagen, Valencia, CA). Sequencing of the 16S small subunit ribosomal ribonucleic acid (SSU rRNA) gene was performed at the University of Minnesota Genomics Center. Briefly, a quality control qPCR was performed in triplicate to verify that sufficient quantities of the 16S SSU rRNA gene were present in the sample for sequencing. After confirmation that adequate levels of the 16S SSU rRNA gene were extracted, the following primers were used to sequence the V3/V4 region of the 16S SSU rRNA gene: 357 F (5'-CCTACGGGNGGCAGCAG-3') – 806 R (5'-GGACTACNVGGGTWTCT AAT- 3')[75–78]. Next, a dual-indexing qPCR was used to barcode and add indexes for gene amplicon sequencing[79]. The resulting amplicons were normalized, pooled, cleaned, and quantified (KAPA quantification, KAPA Biosystems, Wilmington, MA). An Illumina MiSeq (Illumina, San Diego, CA) was used to sequence the library with a 2 × 300 bp cycle with 15% PhiX (Illumina).

Bioinformatics, statistics, and reproducability

16S rRNA amplicon sequencing reads were demultiplexed in Illumina Basespace. Fastq files were assessed for read quality, trimmed (trimLeft = 12, truncLen = 150), dereplicated, and filtered using the dada2 pipeline in R studio (Version 4.2). Following removal of chimeric sequences, reads were assigned a taxonomy using greengenes (gg_13_8_train_set_97.fa.gz) and a phylogenetic tree was built using MSA and Phangorn. The reads, metadata, and tree were then combined into a phyloseq object for additional quality controls and filtering, which included removing amplicon sequence variants (ASVs) mapping to mitochondria, chloroplast and cyanobacteria reads, as well as any ASV detected in only one sample. The resulting phyloseq object was used for downstream analyses.

Unsupervised hierarchical clustering was conducted on center log ratio (CLR)-transformed read counts, using a divisive approach (Cluster::diana). CLR transformation is a widely recognized method for analyzing compositional data[80–82]. The CLR transformation mitigates compositional constraint by converting proportions to log-ratios, thus enabling the comparison of relative abundances across samples in a mathematically coherent and interpretable manner. The same CLR-transformed read counts were used to calculate Aitchison distance and visualize these distances in principal coordinate analysis (PCoA) space. Alpha diversity metrics, including Shannon diversity, Pielou's evenness, and core abundance, were calculated using microbiome::alpha to assess raw reads. Prevalence figures and corresponding Fisher's exact tests were assessed using custom scripts. Differential abundance was performed using DESeq2 to assess raw reads on ASVs, and higher level taxonomic classifications were assessed using objects generated via phyloseq::tax_glom. Partial Least Squares Discriminant Analysis (PLSDA) was performed using the sparse option (mixomics::plsda). Area under the receiver operating curve (AUROC) was utilized to assess performance of feature sets for stratifying parasitemia risk using a leave-one-out cross validation (LOOCV) approach and a model trained through the Caret package using the ROC metric.

All statistical comparisons across groups (peripheral parasitemia, Aitchison distance travelled, PCoA axis coordinates, Shannon diversity, core abundance, observed ASVs (richness), sequencing depth, Pielou's evenness, ASV prevalence, BMI, and age) were compared using a Wilcoxon rank-sum test (Mann-Whitney U test) unless otherwise noted. Regression analysis of peripheral parasitemia and core microbiota abundance was performed using stats::lm. Odds ratios were calculated using the finalfit package. Significant differences in prevalence were assessed by Fisher's exact test. Significant differential abundance was performed using DESeq2 with betaPriors=FALSE to assess raw reads.

For both the RM and CHMI studies presented here, no statistical method was used to predetermine sample size. As previously reported[42], male and female participants in the CHMI study (n = 35) were randomly assigned to experimental groups receiving various doses of KAF156 following *P. falciparum* infection (50 or 20 mg KAF156) or control groups. In the current study, all participants that progressed to detectable *P. falciparum* infection (n = 32) were first considered together as a single group, then assigned to human microbiome clusters (HC1, HC2, and HC3) based on the results from unsupervised hierarchical clustering of 16S rRNA gene sequencing data. For the RM study, all RMs (n = 16) were inoculated with *P. fragile* and were first considered together as a single group, then were subsequently analyzed as two building-dependent groups based on the results from unsupervised hierarchical clustering of 16S rRNA gene sequencing data and supported by housing records. Although the parent CHMI study included both male and female participants, due to an insufficient number of female participants, post hoc sex based analyses were not conducted in the present analyses. All RMs included in this study were male, thus post hoc sex based analyses were not possible. In the CHMI study, sequencing data from three participants were excluded because they did not progress to detectable parasitemia. No data were excluded from the RM study. Finally, for both the RM and CHMI study, the data analyst was blinded to experimental infection groups and outcomes during the analysis.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

16S rRNA gene sequence data generated in this study have been deposited in the NCBI Sequence Read Archive (SRA) under accession code PRJNA918700. All other data that support the findings of this study are provided in the Supplementary Information and the Source Data file. Source data are provided with this paper.

Code availability

All code and custom scripts are available at https://github.com/taproot-informatics/malaria_microbiome; https://doi.org/10.5281/zenodo.16768601[83].

References

1. WHO. World Malaria Report *2023*. Vol. License: CC BY-NC-SA 3.0 IGO (World Health Organization, Geneva, Switzerland, 2023).
2. Sowunmi, A., Ogundahunsi, O. A., Falade, C. O., Gbotosho, G. O. & Oduola, A. M. Gastrointestinal manifestations of acute falciparum malaria in children. *Acta Tropica* **74**, 73–76 (2000).
3. Seydel, K. B., Milner, D. A. Jr., Kamiza, S. B., Molyneux, M. E. & Taylor, T. E. The distribution and intensity of parasite sequestration in comatose Malawian children. *J. Infect. Dis.* **194**, 208–205 (2006).
4. Pongponratn, E., Riganti, M., Punpoowong, B. & Aikawa, M. Microvascular sequestration of parasitized erythrocytes in human falciparum malaria: a pathological study. *Am. J. Tropical Med. Hyg.* **44**, 168–175 (1991).
5. Mooney, J. P. et al. Inflammation-associated alterations to the intestinal microbiota reduce colonization resistance against nontyphoidal Salmonella during concurrent malaria parasite infection. *Sci. Rep.* **5**, 14603 (2015).
6. Milner, D. A. Jr. et al. Quantitative Assessment of Multiorgan Sequestration of Parasites in Fatal Pediatric Cerebral Malaria. *J. Infect. Dis.* **212**, 1317–1321 (2015).
7. Olsson, R. A. & Johnston, E. H. Histopathologic changes and small-bowel absorption in falciparum malaria. *Am. J. Tropical Med. Hyg.* **18**, 355–359 (1969).
8. Wilairatana, P., Meddings, J. B., Ho, M., Vannaphan, S. & Looareesuwan, S. Increased gastrointestinal permeability in patients with Plasmodium falciparum malaria. *Clin. Infect. Dis.: Off. Publ. Infect. Dis. Soc. Am.* **24**, 430–435 (1997).
9. Karney, W. W. & Tong, M. J. Malabsorption in Plasmodium falciparum malaria. *Am. J. Tropical Med. Hyg.* **21**, 1–5 (1972).
10. Molyneux, M. E. et al. Reduced hepatic blood flow and intestinal malabsorption in severe falciparum malaria. *Am. J. Tropical Med. Hyg.* **40**, 470–476 (1989).
11. Church, J. A., Nyamako, L., Olupot-Olupot, P., Maitland, K. & Urban, B. C. Increased adhesion of Plasmodium falciparum infected erythrocytes to ICAM-1 in children with acute intestinal injury. *Malar. J.* **15**, 54 (2016).
12. Chau, J. Y. et al. Malaria-associated L-arginine deficiency induces mast cell-associated disruption to intestinal barrier defenses against nontyphoidal Salmonella bacteremia. *Infect. Immun.* **81**, 3515–3526 (2013).
13. Olupot-Olupot, P. et al. Endotoxaemia is common in children with Plasmodium falciparum malaria. *BMC Infect. Dis.* **13**, 117 (2013).
14. Berkley, J. A. et al. HIV infection, malnutrition, and invasive bacterial infection among children with severe malaria. *Clin. Infect. Dis.: Off. Publ. Infect. Dis. Soc. Am.* **49**, 336–343 (2009).

15. Bronzan, R. N. et al. Bacteremia in Malawian children with severe malaria: prevalence, etiology, HIV coinfection, and outcome. *J. Infect. Dis.* **195**, 895–904 (2007).

16. Alamer, E. et al. Dissemination of non-typhoidal Salmonella during Plasmodium chabaudi infection affects anti-malarial immunity. *Parasitol. Res.* **118**, 2277–2285 (2019).

17. Potts, R. A. et al. Mast cells and histamine alter intestinal permeability during malaria parasite infection. *Immunobiology* **221**, 468–474 (2016).

18. Roth, A. N., Grau, K. R. & Karst, S. M. Diverse Mechanisms Underlie Enhancement of Enteric Viruses by the Mammalian Intestinal Microbiota. *Viruses* **11**, 760 (2019).

19. Mandal, R. K. et al. Dynamic modulation of spleen germinal center reactions by gut bacteria during Plasmodium infection. *Cell Rep.* **35**, 109094 (2021).

20. Yooseph, S. et al. Stool microbiota composition is associated with the prospective risk of Plasmodium falciparum infection. *BMC Genomics* **16**, 631 (2015).

21. Van Den Ham, K. M. et al. The gut microbiome is associated with susceptibility to febrile malaria in Malian children. *Nat. Commun.* **15**, 9525 (2024).

22. Villarino, N. F. et al. Composition of the gut microbiota modulates the severity of malaria. *Proc. Natl. Acad. Sci. USA* **113**, 2235–2240 (2016).

23. Taniguchi, T. et al. Plasmodium berghei ANKA causes intestinal malaria associated with dysbiosis. *Sci. Rep.* **5**, 15699 (2015).

24. Waide, M. L. et al. Gut Microbiota Composition Modulates the Magnitude and Quality of Germinal Centers during Plasmodium Infections. *Cell Rep.* **33**, 108503 (2020).

25. Coatney, G. R., Collins, W. E., Warren, M. & Contacos, P. G. The primate malarias. (1971).

26. Dissanaike, A., Nelson, P. & Garnham, P. Two New Malaria Parasites, *Plasmodium cynomolgi ceylonesis* subsp. nov and *Plasmodium fragile* sp. nov., from Monkeys in Ceylon. *Ceylon J. Med. Res.* 717–721 (1965).

27. Collins, W. E. et al. Studies on sporozoite-induced and chronic infections with Plasmodium fragile in Macaca mulatta and New World monkeys. *J. Parasitol.* **92**, 1019–1026 (2006).

28. Collins, W. E., Chin, W. & Skinner, J. C. Plasmodium fragile and Macaca mulatta monkeys as a model system for the study of malaria vaccines. *Am. J.Tropical Med. Hyg.* **28**, 948–954 (1979).

29. Fremount, H. N. & Miller, L. H. Deep vascular schizogony in Plasmodium fragile: organ distribution and ultrastructure of erythrocytes adherent to vascular endothelium. *Am. J. Tropical Med. Hyg.* **24**, 1–8 (1975).

30. David, P. H., Handunnetti, S. M., Leech, J. H., Gamage, P. & Mendis, K. N. Rosetting: a new cytoadherence property of malaria-infected erythrocytes. *Am. J. Tropical Med. Hyg.* **38**, 289–297 (1988).

31. Handunnetti, S. M., Mendis, K. N. & David, P. H. Antigenic variation of cloned Plasmodium fragile in its natural host Macaca sinica. Sequential appearance of successive variant antigenic types. *J. Exp. Med.* **165**, 1269–1283 (1987).

32. Fujioka, H. et al. A nonhuman primate model for human cerebral malaria: rhesus monkeys experimentally infected with Plasmodium fragile. *Exp. Parasitol.* **78**, 371–376 (1994).

33. Tripathi, R., Vishwakarma, R. A. & Dutta, G. P. Plasmodium fragile: efficacy of arteether (alpha/beta) against cerebral malaria model. *Exp. Parasitol.* **87**, 290–292 (1997).

34. Puri, S. K. & Dutta, G. P. Blood schizontocidal activity of WR 238605 (Tafenoquine) against Plasmodium cynomolgi and Plasmodium fragile infections in rhesus monkeys. *Acta Tropica* **86**, 35–40 (2003).

35. Awoke, N. & Arota, A. Profiles of hematological parameters in Plasmodium falciparum and Plasmodium vivax malaria patients

36. attending Tercha General Hospital, Dawuro Zone, South Ethiopia. *Infect. Drug Resist.* **12**, 521–527 (2019).

36. Elkhalifa, A. M. E. et al. Hematological indices and abnormalities among patients with uncomplicated falciparum malaria in Kosti city of the White Nile state, Sudan: a comparative study. *BMC Infect. Dis.* **21**, 507 (2021).

37. Kotepui, M., Phunphuech, B., Phiwklam, N., Chupeerach, C. & Duangmano, S. Effect of malarial infection on haematological parameters in population near Thailand-Myanmar border. *Malar. J.* **13**, 218 (2014).

38. Maina, R. N. et al. Impact of Plasmodium falciparum infection on haematological parameters in children living in Western Kenya. *Malar. J.* **9**, S4 (2010).

39. Udomkarnjananun, S. et al. Dietary composition modulate gut microbiota and related biomarkers in patients with chronic kidney disease. *Sci. Rep.* **15**, 36274 (2025).

40. Wu, G. D. et al. Linking long-term dietary patterns with gut microbial enterotypes. *Science* **334**, 105–108 (2011).

41. Hensley-McBain, T. et al. Increased mucosal neutrophil survival is associated with altered microbiota in HIV infection. *PLoS Pathog.* **15**, e1007672 (2019).

42. Kublin, J. G. et al. Safety, Pharmacokinetics, and Causal Prophylactic Efficacy of KAF156 in a Plasmodium falciparum Human Infection Study. *Clin. Infect. Dis.: Off. Publ. Infect. Dis. Soc. Am.* **73**, e2407–e2414 (2021).

43. Topçuoğlu, B. D., Lesniak, N. A., Ruffin, M. T., Wiens, J. & Schloss, P. D. A Framework for Effective Application of Machine Learning to Microbiome-Based Classification Problems. *mBio* **11**, https://doi.org/10.1128/mbio.00434-00420 (2020).

44. Armour, C. R., Topçuoğlu, B. D., Garretto, A. & Schloss, P. D. A Goldilocks Principle for the Gut Microbiome: Taxonomic Resolution Matters for Microbiome-Based Classification of Colorectal Cancer. *mBio* **13**, e03161–03121 (2022).

45. Abt, M. C. et al. Commensal bacteria calibrate the activation threshold of innate antiviral immunity. *Immunity* **37**, 158–170 (2012).

46. Brown, R. L., Sequeira, R. P. & Clarke, T. B. The microbiota protects against respiratory infection via GM-CSF signaling. *Nat. Commun.* **8**, 1512 (2017).

47. Ganal, S. C. et al. Priming of natural killer cells by nonmucosal mononuclear phagocytes requires instructive signals from commensal microbiota. *Immunity* **37**, 171–186 (2012).

48. Rooks, M. G. & Garrett, W. S. Gut microbiota, metabolites and host immunity. *Nat. Rev. Immunol.* **16**, 341–352 (2016).

49. Yang, Y. et al. Within-host evolution of a gut pathobiont facilitates liver translocation. *Nature* **607**, 563–570 (2022).

50. Farinella, D. N. et al. Malaria disrupts the rhesus macaque gut microbiome. *Front. Cellular Infection Microbiol.* **12**, https://doi.org/10.3389/fcimb.2022.1058926 (2023).

51. Mandal, R. K. et al. Longitudinal Analysis of Infant Stool Bacteria Communities Before and After Acute Febrile Malaria and Artemether-Lumefantrine Treatment. *J. Infect. Dis.* **220**, 687–698 (2019).

52. Chen, Z. et al. Diversity of macaque microbiota compared to the human counterparts. *Sci. Rep.* **8**, 15573 (2018).

53. Zmora, N., Suez, J. & Elinav, E. You are what you eat: diet, health and the gut microbiota. *Nat. Rev. Gastroenterol. Hepatol.* **16**, 35–56 (2019).

54. Makki, K., Deehan, E. C., Walter, J. & Bäckhed, F. The Impact of Dietary Fiber on Gut Microbiota in Host Health and Disease. *Cell host microbe* **23**, 705–715 (2018).

55. So, D. et al. Dietary fiber intervention on gut microbiota composition in healthy adults: a systematic review and meta-analysis. *Am. J. Clin. Nutr.* **107**, 965–983 (2018).

56. Cui, Z. et al. Seasonal diet and microbiome shifts in wild rhesus macaques are better correlated at the level of nutrient components than food items. *Integr. Zool.* **17**, 1147–1161 (2022).

57. Lv, L. et al. Seasonal Variations in the Structure and Function of the Gut Flora in Adult Male Rhesus Macaques Reared in Outdoor Colonies. *Microorganisms* **13** https://doi.org/10.3390/microorganisms13010117 (2025).

58. Kuthyar, S. et al. Limited microbiome differences in captive and semi-wild primate populations consuming similar diets. *FEMS Microbiol Ecol* **98**, https://doi.org/10.1093/femsec/fiac098 (2022).

59. Manuzak, J. A. et al. Antibiotic-induced microbiome perturbations are associated with significant alterations to colonic mucosal immunity in rhesus macaques. *Mucosal Immunol.* **13**, 471–480 (2020).

60. Ortiz, A. M. et al. Experimental bacterial dysbiosis with consequent immune alterations increase intrarectal SIV acquisition susceptibility. *Cell Rep.* **42**, 112020 (2023).

61. Gupta, V. K., Paul, S. & Dutta, C. Geography, Ethnicity or Subsistence-Specific Variations in Human Microbiome Composition and Diversity. *Front. Microbiol.* **8**, 1162 (2017).

62. Lynn, D. J., Benson, S. C., Lynn, M. A. & Pulendran, B. Modulation of immune responses to vaccination by the microbiota: implications and potential mechanisms. *Nat. Rev. Immunol.* **22**, 33–46 (2022).

63. Choo, J. M., Leong, L. E. & Rogers, G. B. Sample storage conditions significantly influence faecal microbiome profiles. *Sci. Rep.* **5**, 16350 (2015).

64. Song, S. J. et al. Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies. *mSystems* **1**, https://doi.org/10.1128/mSystems.00021-16 (2016).

65. Denny, J. E. et al. Differential Sensitivity to Plasmodium yoelii Infection in C57BL/6 Mice Impacts Gut-Liver Axis Homeostasis. *Sci. Rep.* **9**, 3472 (2019).

66. Stough, J. M. et al. Functional Characteristics of the Gut Microbiome in C57BL/6 Mice Differentially Susceptible to Plasmodium yoelii. *Front. Microbiol.* **7**, 1520 (2016).

67. *Methods in malaria research.* 6th Edition edn, (EVIMalaR and MR4/ATCC, 2013).

68. Lim, C. et al. Improved light microscopy counting method for accurately counting Plasmodium parasitemia and reticulocytemia. *Am. J. Hematol.* **91**, 852–855 (2016).

69. Trott, K. A. et al. Evidence for an increased risk of transmission of simian immunodeficiency virus and malaria in a rhesus macaque coinfection model. *J. Virol.* **85**, 11655–11663 (2011).

70. Trott, K. A., Richardson, A., Hudgens, M. A. & Abel, K. Immune activation and regulation in simian immunodeficiency virus-Plasmodium fragile-coinfected rhesus macaques. *J. Virol.* **87**, 9523–9537 (2013).

71. Felger, I. et al. The Dynamics of Natural Plasmodium falciparum Infections. *PloS one* **7**, e45542 (2012).

72. Lindblade, K. A., Steinhardt, L., Samuels, A., Kachur, S. P. & Slutsker, L. The silent threat: asymptomatic parasitemia and malaria transmission. *Expert Rev. Anti-Infective Ther.* **11**, 623–639 (2013).

73. Murphy, S. C., Daza, G., Chang, M. & Coombs, R. Laser cutting eliminates nucleic acid cross-contamination in dried-blood-spot processing. *J. Clin. Microbiol.* **50**, 4128–4130 (2012).

74. Seilie, A. M. et al. Beyond Blood Smears: Qualification of Plasmodium 18S rRNA as a Biomarker for Controlled Human Malaria Infections. *Am. J. Tropical Med. Hyg.* **100**, 1466–1476 (2019).

75. Sim, K. et al. Improved detection of bifidobacteria with optimised 16S rRNA-gene based pyrosequencing. *PloS one* **7**, e32543 (2012).

76. Thompson, L. R. et al. A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**, 457–463 (2017).

77. Apprill, A., McNally, S., Parsons, R. & Weber, L. Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* **75**, 129–137 (2015).

78. Caporaso, J. G. et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624 (2012).

79. Gohl, D. M. et al. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat. Biotechnol.* **34**, 942–949 (2016).

80. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* **8**, https://doi.org/10.3389/fmicb.2017.02224 (2017).

81. K. Gerald van den Boogaart, R. T.-D. *Analyzing Compositional Data with R.* 1 edn, (Springer, 2013).

82. Quinn, T. P. et al. A field guide for the compositional analysis of any-omics data. *GigaScience* **8**, https://doi.org/10.1093/gigascience/giz107 (2019).

83. Gustin, A. T. et al. *Distinct gastrointestinal microbial signatures predict parasite levels in controlled Plasmodium infections in both rhesus macaques and humans, Github,* https://doi.org/10.5281/zenodo.16768602 (7 Aug 2025).

Author contributions

A.T.G. assisted with sample collection from the RM study, analyzed and interpreted the 16S rRNA sequencing data, and wrote the paper. C.A.B. assisted with sample collection from the RM study and performed the DNA extractions and 16S rRNA sequencing for both the RM and CHMI studies. K.H. coordinated the collection of fecal samples from the CHMI study. E.C. provided support for sample processing from the RM study. S.W., N.I., C.Y.A., and W.D.G. provided research support for the R.M. study. K.A.G. provided veterinary support for the RM study. K.D.P. provided the initial *P. fragile* stock. B.L.R., M.G. and N.R.K. assisted with data interpretation and paper preparation. J.G.K. led and provided samples from the KAF156 CHMI study, assisted with data interpretation and paper preparation. J.A.M. conceived of the study, oversaw the planning of the project, performed experiments, interpreted the data, and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-67241-2.

Correspondence and requests for materials should be addressed to Jennifer A. Manuzak.

Peer review information *Nature Communications* thanks Andrea Berry, and the other, anonymous, reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Andrew T. Gustin [1,2,3], Courtney A. Broedlow[2,4,5], Kevin Hager [6], Ernesto Coronado[2,4], Solomon Wangari[2], Naoto Iwayama[2], Chul Y. Ahrens[2], William D. Garrison[2], Kathryn A. Guerriero[2,7], Kristina De Paris[8], Berlin Londono-Renteria [9], Michael Gale Jr. [1,2,10], Nichole R. Klatt [2,4,5], James G. Kublin [6] & Jennifer A. Manuzak [2,4,11,12] ✉

[1]Department of Immunology, University of Washington, Seattle, WA, USA. [2]Washington National Primate Research Center, University of Washington, Seattle, WA, USA. [3]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [4]Department of Pharmaceutics, University of Washington, Seattle, WA, USA. [5]Division of Surgical Outcomes and Precision Medicine Research, Department of Surgery, University of Minnesota, Minneapolis, MN, USA. [6]Seattle Malaria Clinical Trials Center, Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [7]Veterinary Medicine and Resources Branch, National Institute of Mental Health, Bethesda, MD, USA. [8]Department of Microbiology and Immunology, University of North Carolina School of Medicine, Chapel Hill, NC, USA. [9]Department of Tropical Medicine and Infectious Disease, Tulane University Celia Scott Weatherhead School of Public Health and Tropical Medicine, New Orleans, LA, USA. [10]Department of Microbiology and Immunology, and Institute on Infectious Diseases, University of Minnesota School of Medicine, Minneapolis, MN, USA. [11]Division of Immunology, Tulane National Biomedical Research Center, Tulane University, Covington, LA, USA. [12]Department of Microbiology and Immunology, Tulane University School of Medicine, New Orleans, LA, USA. ✉e-mail: jmanuzak@tulane.edu