

Gut *Bacteroides* act in a microbial consortium to cause susceptibility to severe malaria

Rabindra K. Mandal¹, Anita Mandal¹, Joshua E. Denny², Ruth Namazii³, Chandy C. John¹, and

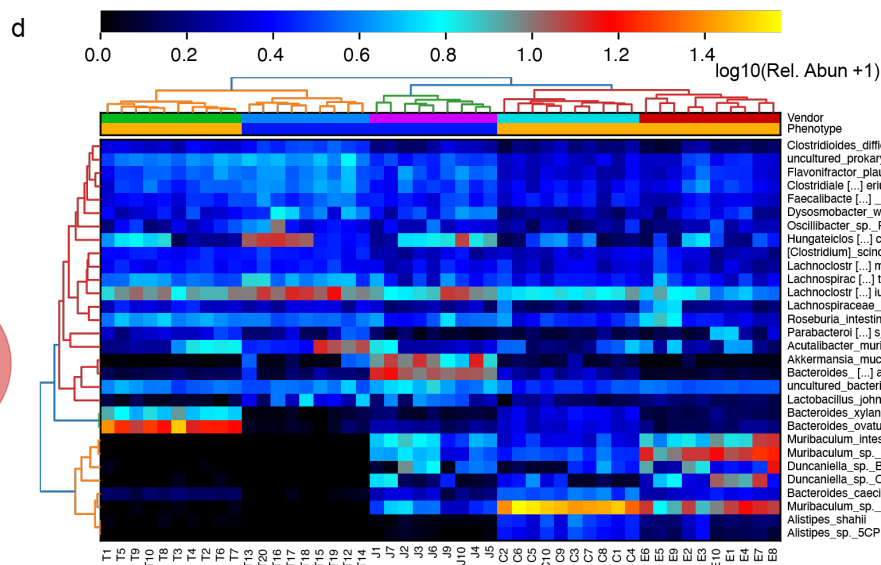
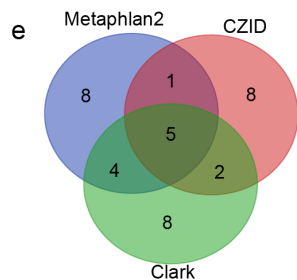
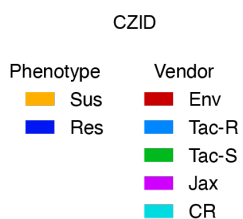
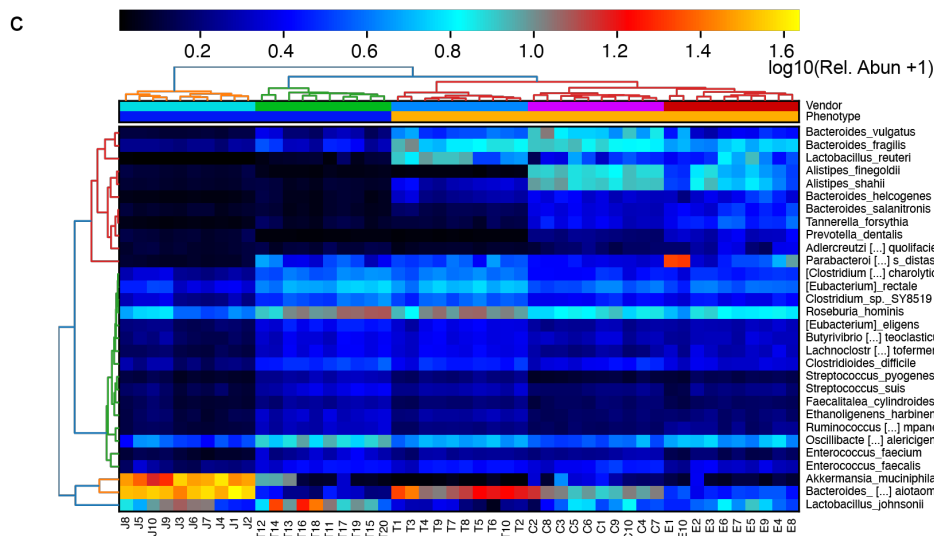
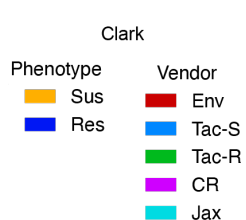
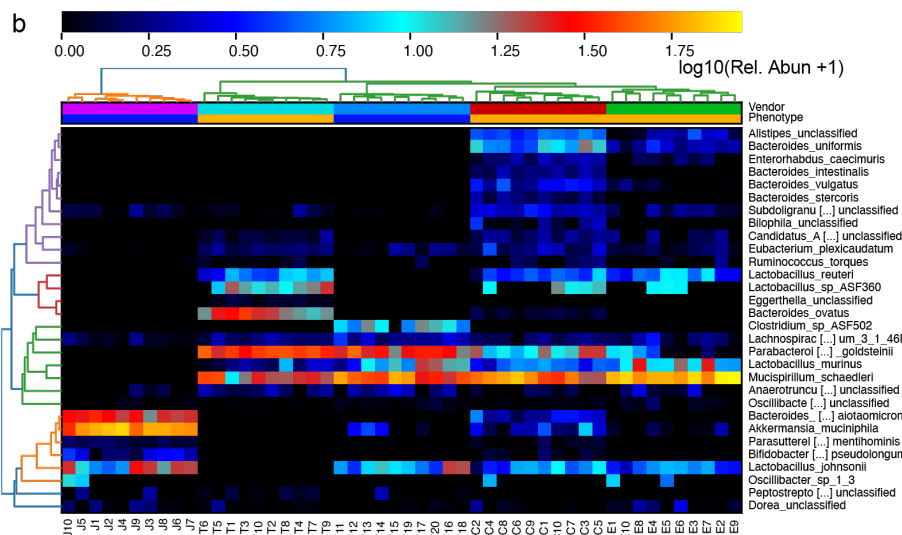
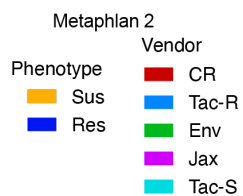
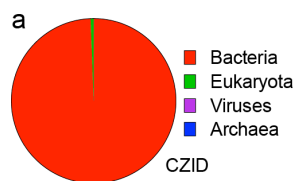
Nathan W. Schmidt^{1,2*}

¹Ryan White Center for Pediatric Infectious Diseases and Global Health, Herman B Wells Center for Pediatric Research, Department of Pediatrics, Indiana University School of Medicine

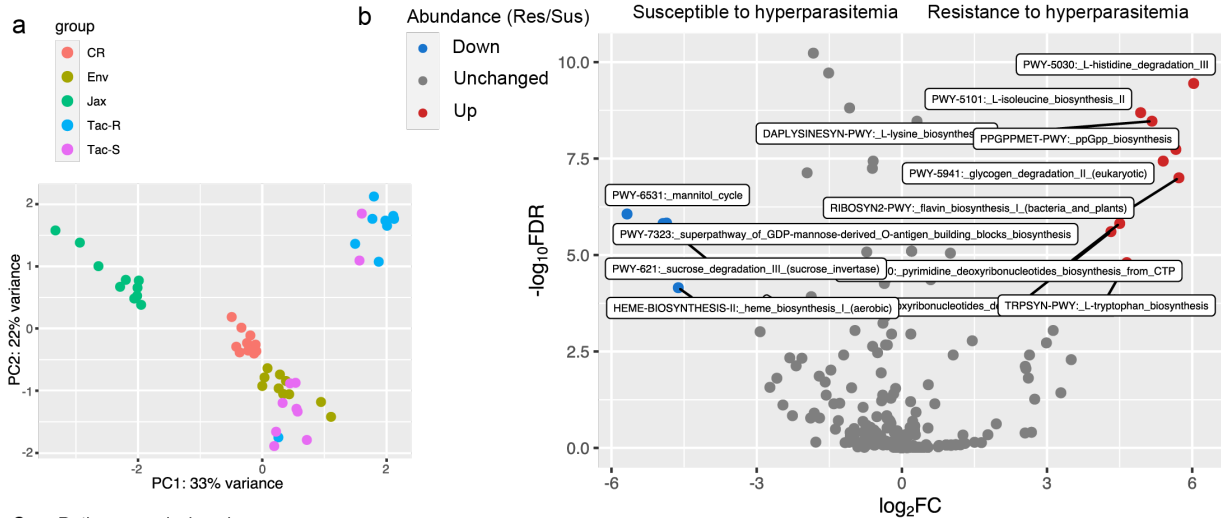
²Department of Microbiology and Immunology, University of Louisville

³Department of Paediatrics and Child Health, Makerere University, Kampala, Uganda

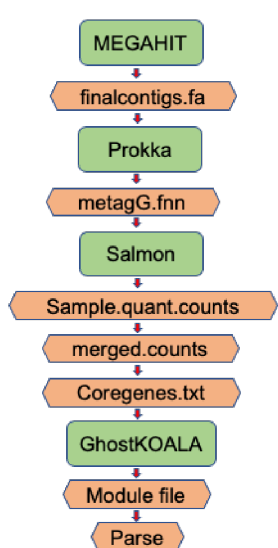
Supplementary Information



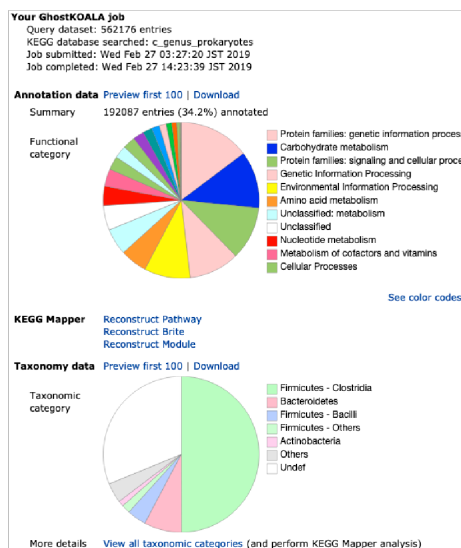
S Figure 1 (Related to Figure 1): Taxonomic profile of naïve mice differentially susceptible to Py hyperparasitemia. Heatmap shows relative abundance of top 30 bacterial species detected using different bioinformatic approaches. **a** Microbiota detected at kingdom level using CZID. **b** Heatmap based on Metaphlan2 output. **c** Heatmap based on Clark output. **d** Heatmap based on CZID. **e** Venn diagram showing overlap at species level between the different bioinformatics approaches. Source data are provided in the Source Data file.



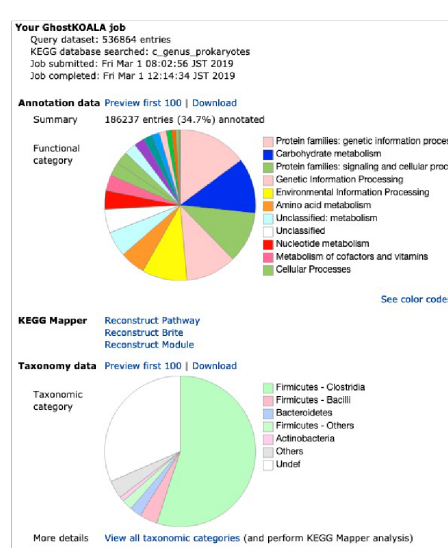
c Pathway analysis using GhostKOALA



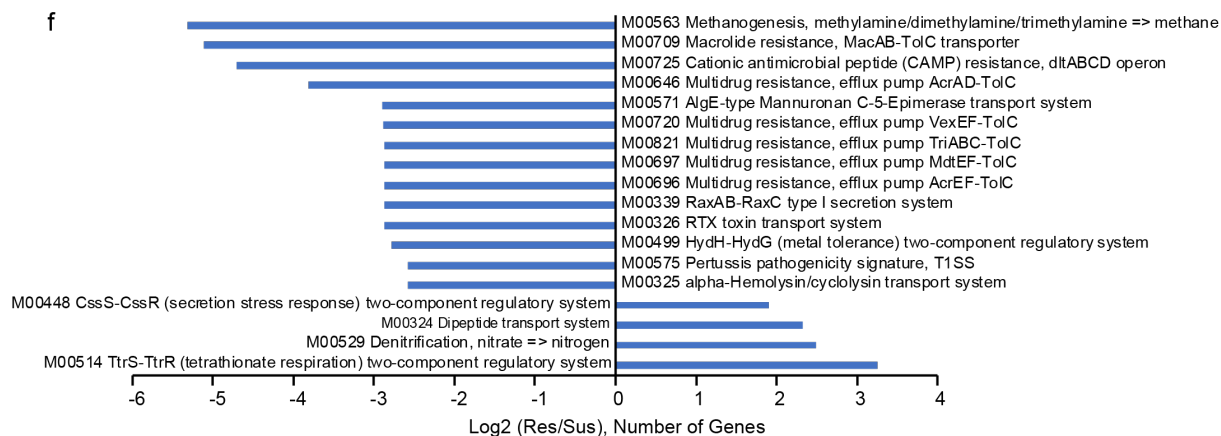
d CR



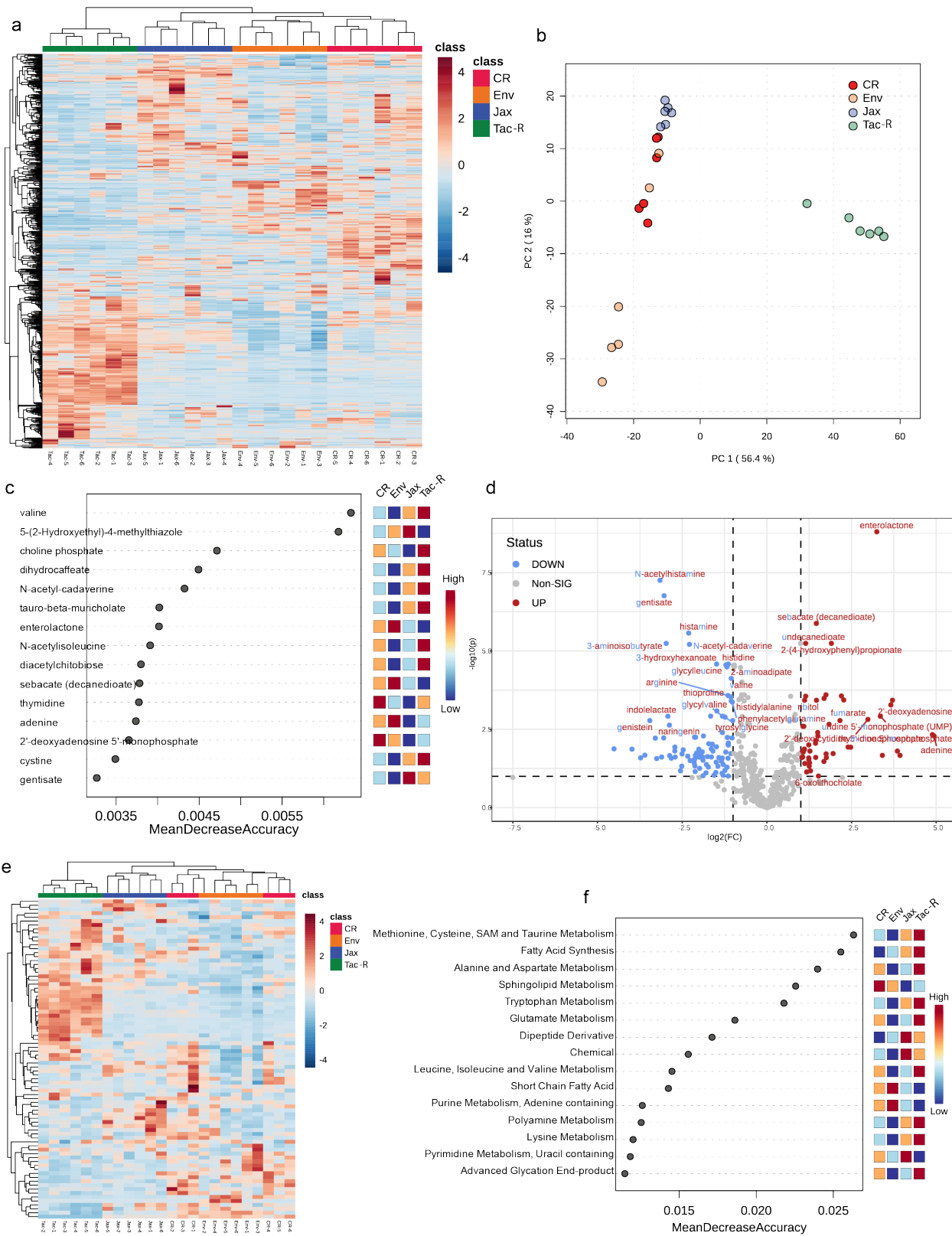
e Tac-Res



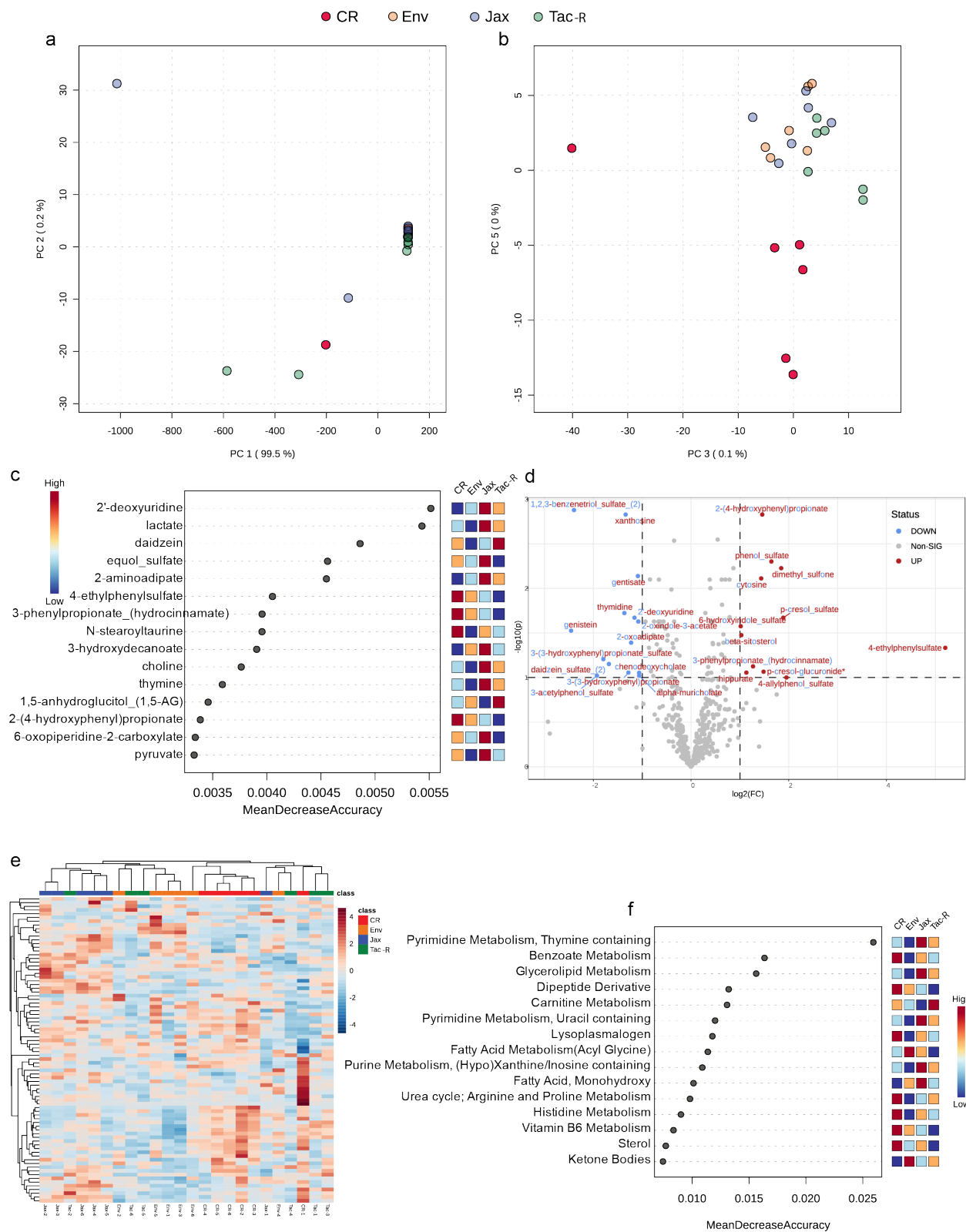
f



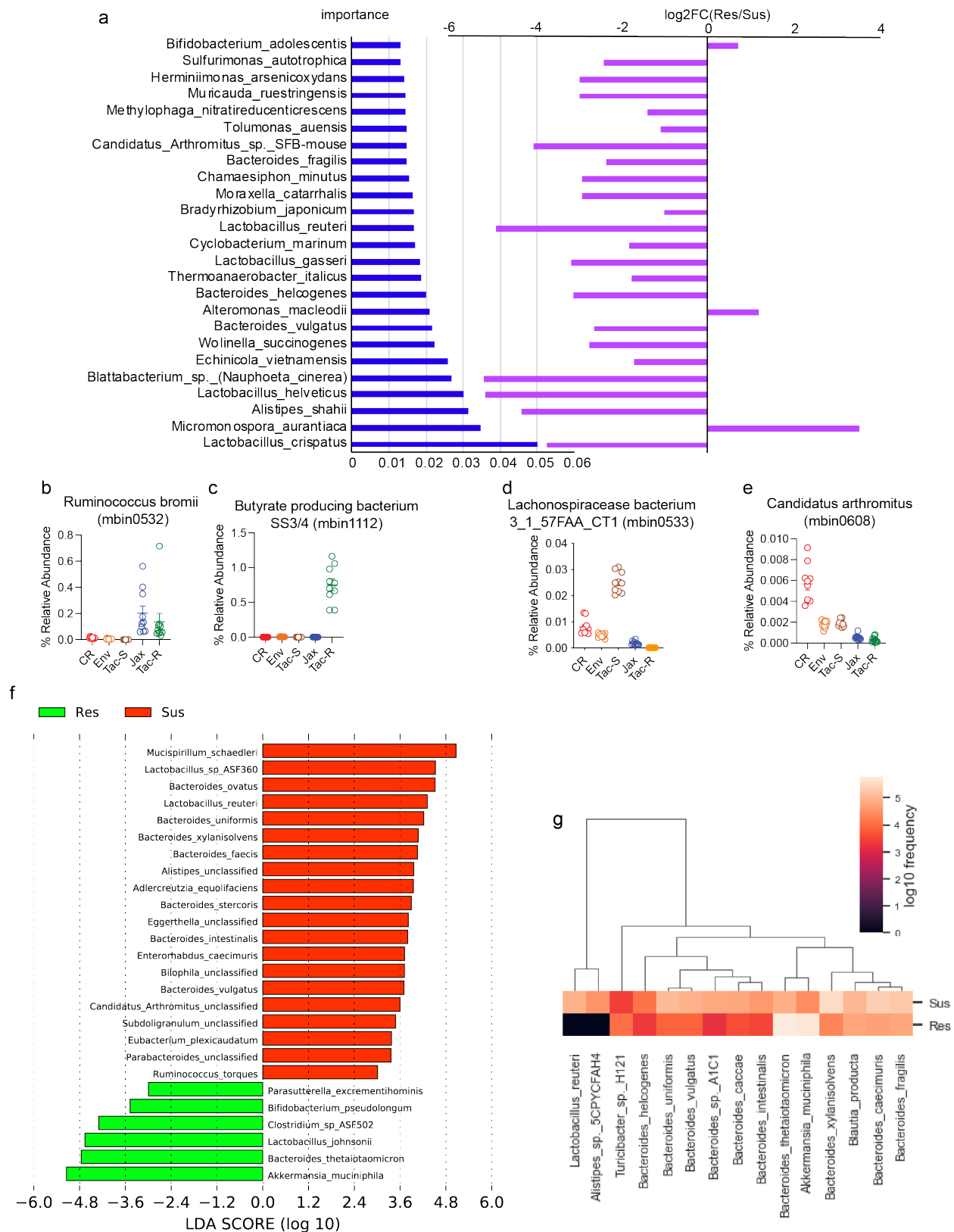
S Figure 2 (Related to Figure 1): Microbial pathway analysis in naïve mice differentially susceptible to Py hyperparasitemia. **a** Principal component analysis plot using HUMAnN 2 at pathway level. Volcano plot showing differentially abundant microbial pathways with $\log_2FC > |3|$ and $FDR < 0.05$. **b** Schematic showing pathway analysis using GhostKOALA. Green box represents bioinformatics package and orange represents the output file. **c** Screenshot of GhostKOALA output of Py hyperparasitemia susceptible CR mice. **d** Screenshot of GhostKOALA output of Py hyperparasitemia resistant Tac mice. **e** Pathways differentially abundant in Py hyperparasitemia susceptible (CR+Env) and resistant (Tac+Jax) mice. Log₂FC is based on the number of microbial genes belonging to a given KEGG module. Source data are provided in the Source Data file.



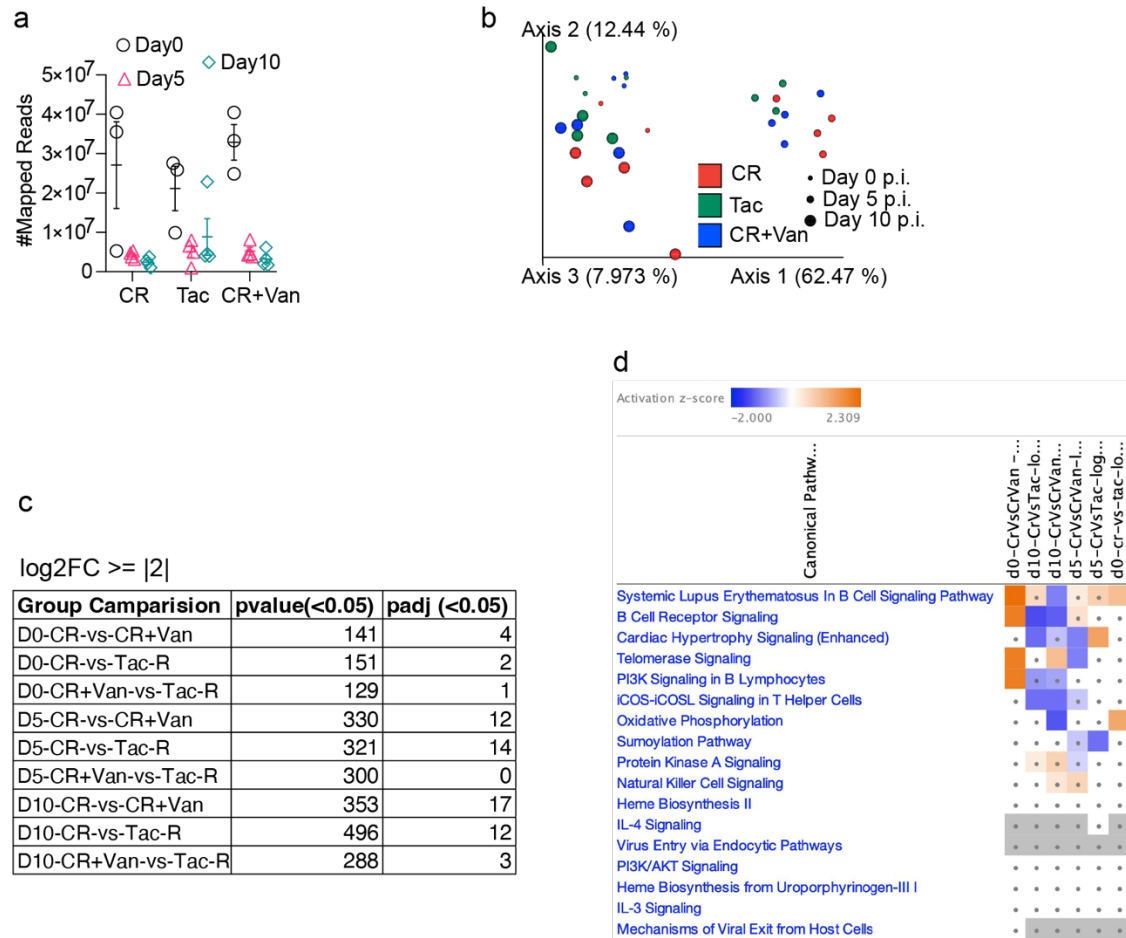
S Figure 3 (Related to Figure 1): Ceca metabolomics. Ceca contents were collected from hyperparasitemia susceptible (CR and Env) and resistant naïve mice (Tac-R and Jax) n=6/group. Global metabolomics were performed using LC-MS/MS. **a** Heatmap showing clustering of samples and metabolites. **b** Principal component analysis (PCA) plot showing clustering of samples by groups. **c** Features (metabolites) ranked by their contribution to classification accuracy (Mean Decrease Accuracy) among four groups using random forest classifier. **d** Volcano plot showing differentially abundant metabolites (Sus/Res). Metabolites with higher abundance in susceptible group are shown in red and in resistant group in blue with $\log_2FC > |2|$ and false discovery rate < 0.05 . **e** Heatmap showing clustering of samples and sub-pathway. **f** Features (sub-pathway) ranked by their contribution to classification accuracy (Mean Decrease Accuracy) among four groups using random forest classifier. Source data are provided in the Source Data file.



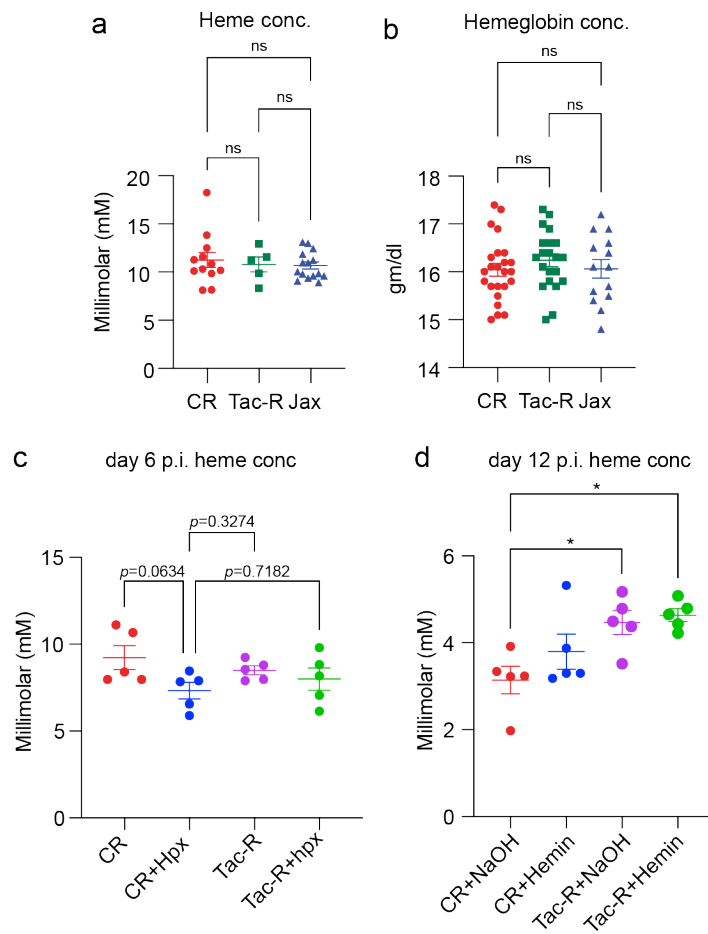
S Figure 4 (Related to Figure 1): Serum metabolomics. Sera were extracted from hyperparasitemia susceptible (CR and Env) and resistant naïve mice (Tac-R and Jax) n=6/group. Global metabolomics were performed using LC-MS/MS. **a** PCA plot showing clustering of samples by groups using PC1 and PC2. To note, 99.7% of the variation is explained by only 5 samples out of 24 samples. **b** PCA plot showing clustering of samples by groups using PC3 and PC5 where only 0.1% of variation is explained. **c** Features (metabolites) ranked by their contribution to classification accuracy (Mean Decrease Accuracy) among four groups using random forest classifier. **d** Volcano plot showing differentially abundant metabolites (Sus/Res) in sera. Metabolites with higher abundance in susceptible group are shown in red and in resistant group in blue with $\log_2FC > |2|$ and false discovery rate < 0.05 . **e** Heatmap showing clustering of samples and sub-pathway. **f** Features (sub-pathway) ranked by their contribution to classification accuracy (Mean Decrease Accuracy) among four groups using random forest classifier. Source data are provided in the Source Data file.



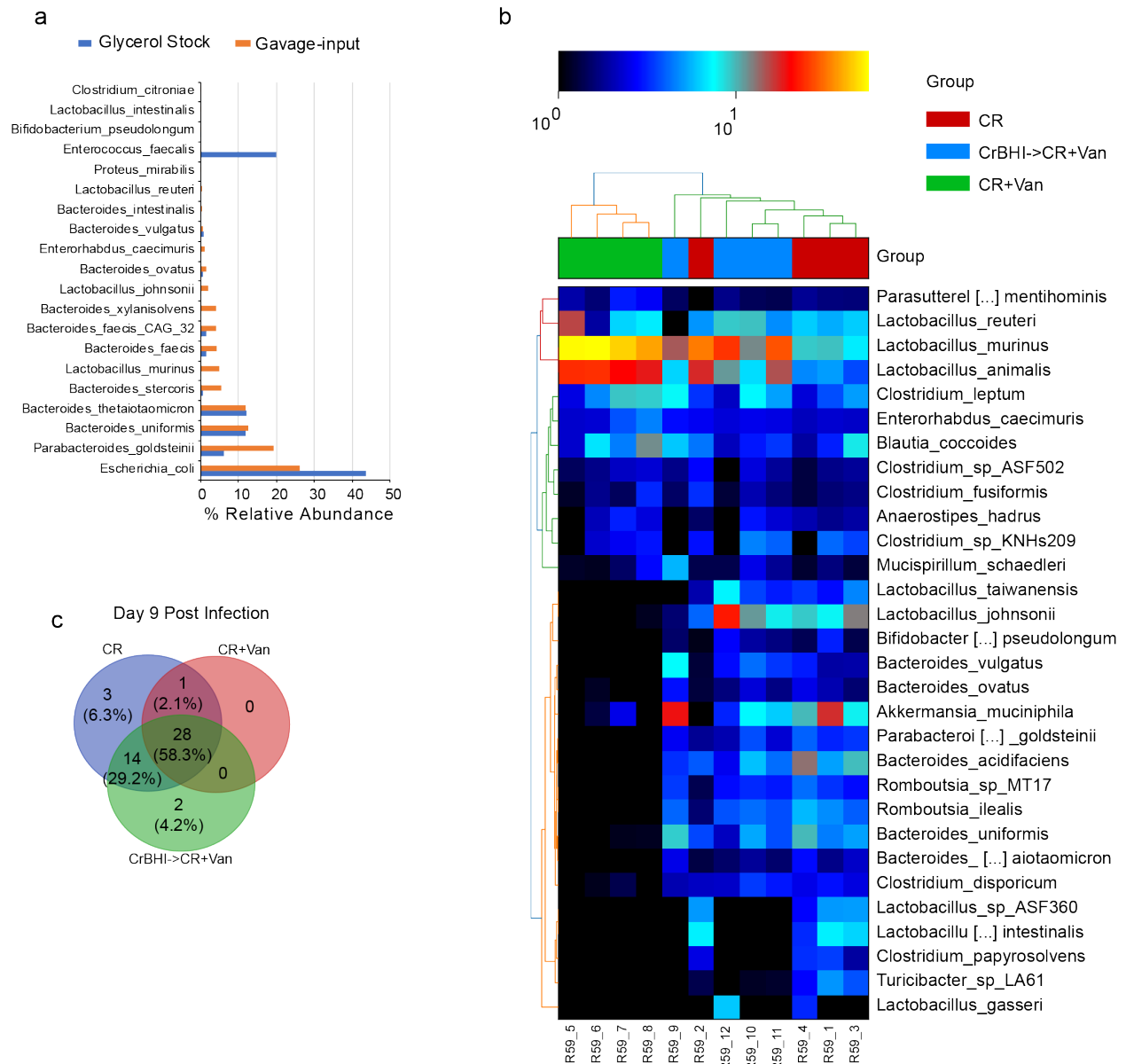
S Figure 5 (Related to Figure 2): Differential abundance analysis of gut bacterial species in naïve mice differentially susceptible to Py hyperparasitemia. **a** Importance of bacterial features able to classify samples based on malaria severity phenotype. Clark output at species level was fed to random forest classifier. Log2FC indicates directionality of bacterial abundance in Py hyperparasitemia resistant or susceptible mice. **b** Percent relative abundance of mbin0532 classified as *Ruminococcus bromii*. **c** Percent relative abundance of mbin1112 classified as Butyrate producing bacterium SS3/4. **d** Percent relative abundance of mbin0533 classified as *Lachnospiraceae bacterium 3_1_F7FAA_CT1*. **e** Percent relative abundance of mbin0608 classified as *Candidatus arthromitus*. **f** LEfSe analysis showing differentially abundant species identified by MetaPhlan2. LDA score cutoff were $>|3.5|$. **g** Heatmap showing the top bacterial features able to classify samples to Py hyperparasitemia resistance or susceptible mice using CZID. Features were identified using random forest. All data are mean \pm SE (standard error). Source data are provided in the Source Data file.



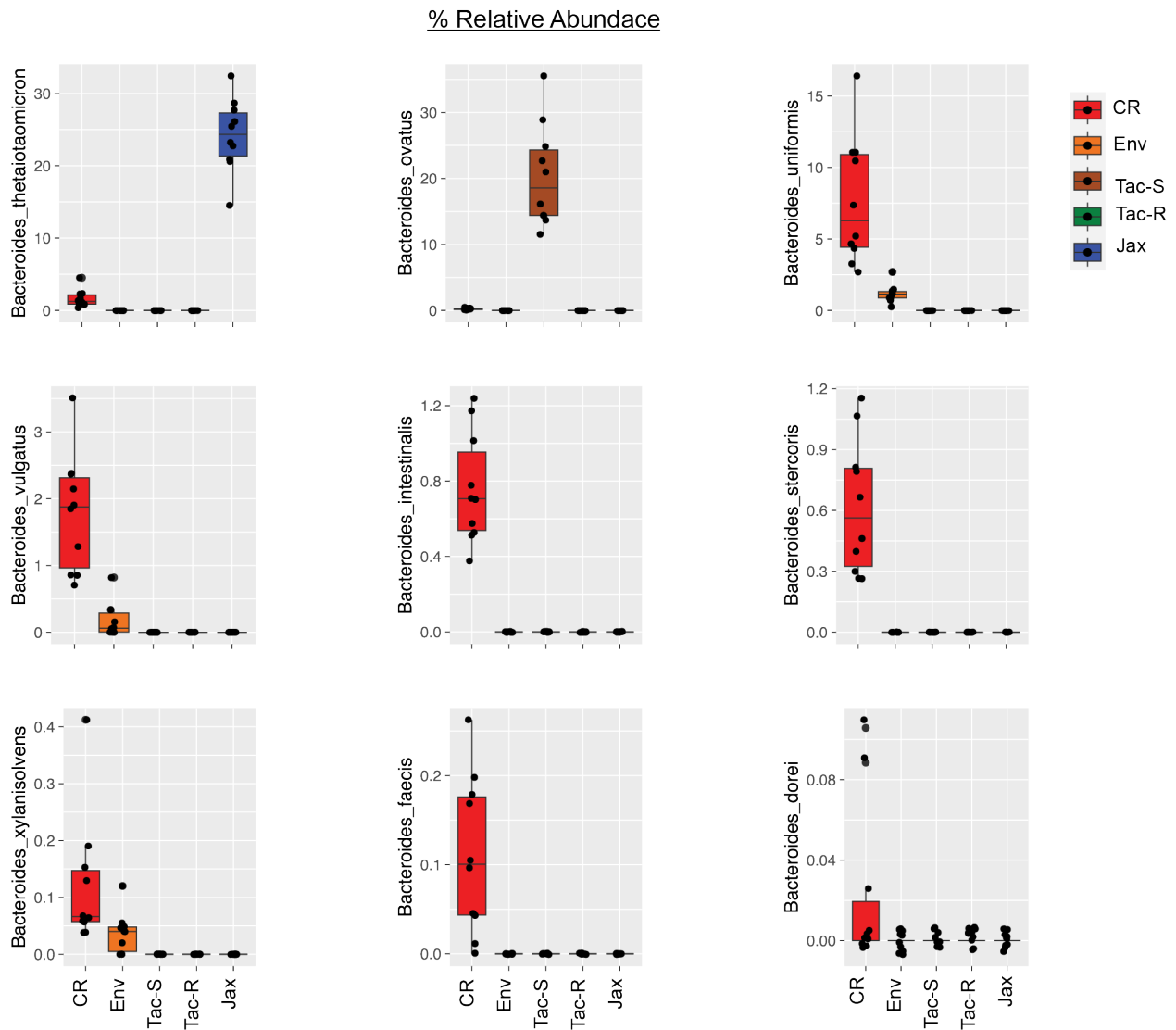
S Figure 6 (Related to Figure 3): Whole blood metatranscriptomics. a Number of quality reads mapped to mouse transcriptome. **b** PCoA plot shows global transcriptome response of mice. **c** Differentially abundant transcripts with log2FC >= 2. D0: day 0 p.i., D5: day 5 p.i.; D10: day 10 p.i., CR+Van: CR treated with vancomycin. **d** Canonical pathway analysis using IPA. gray dots: insignificant pathway ($p \geq 0.1$). Blue and orange are enriched in hyperparasitemia resistant and susceptible group, respectively. Column names at the top indicate group comparisons based on day p.i. between susceptible (CR) and resistant group (Tac and CR+Van). Source data are provided in the Source Data file.



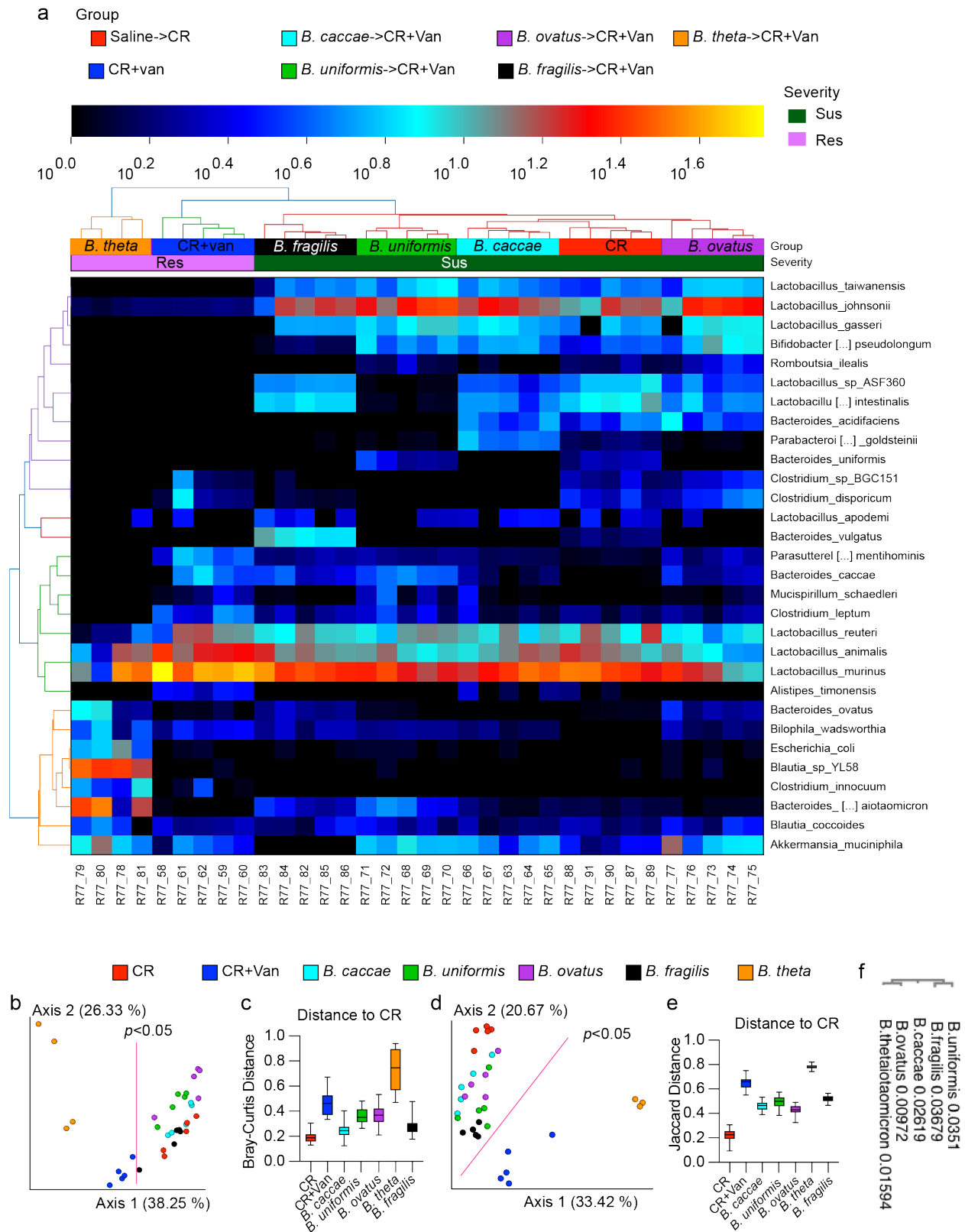
S Figure 7 (Related to Figure 3): Measurement of heme and hemoglobin. **a** Heme concentration in blood of naïve mice. **b** Hemoglobin concentration in blood of naïve mice. **c** Heme concentration in blood of mice treated with hemopexin on day 6 p.i. **d** Heme concentration in blood of mice treated with Hemin on day 12 p.i. All data are mean \pm SE (standard error). Statistical analyses were performed using one-way ANOVA. * $p<0.05$. ns: non-significant. Data in **a** and **b** are from multiple shipments of mice. Heme concentration in **c** and **d** were measured from one experiment.



S Figure 8 (Related to Figure 4): Taxonomic profile of CR+Van mice gavaged with CrBHI. a Relative abundance of culturable bacteria of CR ceca grown on BHI agar (CrBHI). Gavage input is the bacterial mass obtained after 48 hrs of incubation in BHI medium using CrBHI glycerol stock as inoculum. Bacterial profiling was performed using shotgun metagenomics. **b** Heatmap shows the abundance of bacterial species on day 9 post Py infection. **c** Venn diagram shows overlap of fecal microbiota on day 9 post Py infection. Source data are provided in the Source Data file.

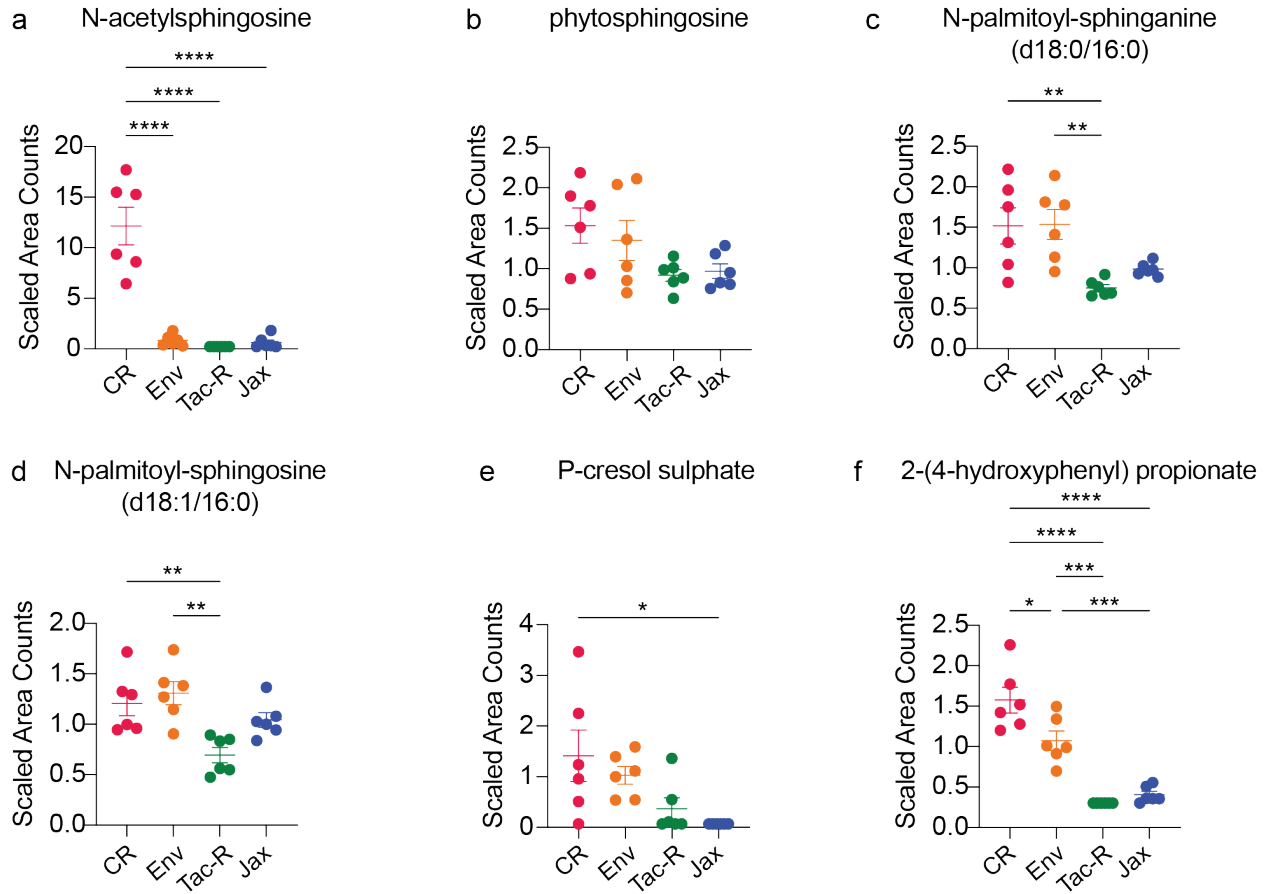


S Figure 9 (Related to Figure 5): Relative abundance (%) *Bacteroides* species detected in ceca of mice using shotgun metagenomics. Relative abundance of bacteria using CLARK pipeline is shown. Lower and upper end of boxplot corresponds to first and third quartile. Horizontal line through the box is median. Data beyond the end of whiskers are outliers. Source data are provided in the Source Data file.

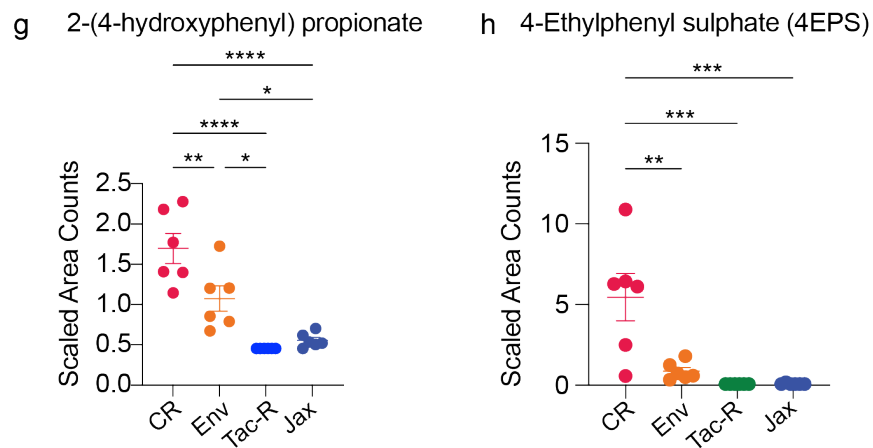


S Figure 10 (Related to Figure 5 and 6): Taxonomic profile of CR+Van mice gavaged with individual *Bacteroides* species. Experiments from Figures 5 and 6 were performed with overlapping control (CR) and CR+Van mice. Gut microbiota profiling was done using MVRSION on day 1 p.i. which is 48 hours after the last bacterial gavage. One of the experiments was chosen to profile fecal pellet microbiota. **a** Heatmap shows the abundance of top 30 bacterial species on day 1 Py infection in CR mice treated with vancomycin for two weeks followed by gavage with single *Bacteroides* species. **b** Beta diversity analysis shown by PCoA plot using Bray-Curtis distance of fecal pellet bacteria at species level. **c** Bray-Curtis distance of CR mice to rest of the groups. **d** Beta diversity analysis shown by PCoA plot using Jaccard distance of fecal pellet bacteria at species level. **e** Jaccard distance to CR mice to rest of the groups. **f** Phylogenetic tree based on 16S rRNA gene sequence linked to ATCC strains. This is a neighbor-joining tree without distance corrections using real branch length constructed by Clustal Omega. Line between the graph separates hyperparasitemia susceptible and resistant groups of mice (**b** and **d**). Pairwise PERMANOVA showed all groups were significantly different from each other (**b** and **d**). Box ends indicate first and third quartile. Box plot whiskers represent minimum and maximum value and horizontal line inside the box is mean (**c** and **e**). Source data are provided in the Source Data file.

Ceca metabolites



Sera metabolites



S Figure 11 (Related to Figure 7, S Figure 3 and S Figure 4): Differentially abundant metabolites related to *Bacteroides* genus in ceca and serum of naïve mice. a-f Ceca metabolites. g and h

Serum metabolites. N=6/group. Data are mean \pm SE (standard error). Statistical analyses were performed using ordinary one-way ANOVA. * p <0.05; ** p <0.01; *** p <0.001; **** p <0.0001. Source data are provided in the Source Data file.

Supplementary Table 1: Culturable Bacteria identified using MALDI-TOF and sanger sequencing of 16S rRNA.

Colony#	NCBI (refseq_RNA) (16S rRNA)	MALDI-TOF	Media
CrBH1	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH2	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH3	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH4	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH5	Shigella sonnei	Escherichia coli	BHI
CrBH6	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH7	Shigella sonnei	Escherichia coli	BHI
CrBH8	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH9	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH10		Enterococcus faecalis	BHI
CrBH11	Escherichia fergusonii	Escherichia coli	BHI
CrBH12	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH13	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH14	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH15	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH16	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH17	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH18	Escherichia fergusonii	Escherichia coli	BHI
CrBH19	Shigella sonnei	Escherichia coli	BHI
CrBH20	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH21	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH22	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH23	Shigella sonnei	Escherichia coli	BHI
CrBH24	Shigella sonnei	Escherichia coli	BHI
CrBH25	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH26	Shigella sonnei	Escherichia coli	BHI
CrBH27	Bacteroides thetaiotaomicron	Escherichia coli	BHI
CrBH28	Enterococcus faecalis	Enterococcus faecalis	BHI
CrBH29	Shigella dysenteriae	No organism identification	BHI
CrBH30		Enterococcus faecalis	BHI
CrBH31		Enterococcus faecalis	BHI
CrBH32	Escherichia fergusonii	Escherichia coli	BHI
CrBH33	Pseudocitrobacter anthropi	Escherichia coli	BHI
CrBH34		NA	BHI
CrBH35		Enterococcus faecalis	BHI
CrBH36	Shigella dysenteriae	Escherichia coli	BHI
CrBH37	Shigella sonnei	Escherichia coli	BHI
CrBH38	Shigella sonnei	Escherichia coli	BHI
CrBH39	Shigella sonnei	NA	BHI
CrBH40		NA	BHI
CrBH41	Shigella sonnei	Escherichia coli	BHI
CrBH42	Escherichia coli	Escherichia coli	BHI
CrBH43		Enterococcus faecalis	BHI
CrBH44		Enterococcus faecalis	BHI
CrBH45	Shigella sonnei	Escherichia coli	BHI
CrBH46		Enterococcus faecalis	BHI
CrBH47		Enterococcus faecalis	BHI
CrBH48	Enterobacter hormaechei	Escherichia coli	BHI
CrBH49	Shigella sonnei	Escherichia coli	BHI
CrBH50	Shigella sonnei	Escherichia coli	BHI
CrBH51		NA	BHI
CrBH52	Shigella sonnei	Escherichia coli	BHI
CrBH53		Enterococcus faecalis	BHI
CrBH54		Escherichia coli	BHI
CrBH55		Enterococcus faecalis	BHI
CrBH56		Enterococcus faecalis	BHI
CrBH57		Lactobacillus murinus	MRS, Ph 5.5
CrBH58		no peaks found	MRS, Ph 5.5
CrBH59		no peaks found	MRS, Ph 5.5
CrBH60		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH61		no peaks found	MRS, Ph 5.5
CrBH62		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH63		Lactobacillus reuteri	MRS, Ph 5.5
CrBH64		Lactobacillus reuteri	MRS, Ph 5.5
CrBH65		no peaks found	MRS, Ph 5.5
CrBH66		No organism identification	MRS, Ph 5.5
CrBH67		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH68		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH69		no peaks found	MRS, Ph 5.5
CrBH70		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH71		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH72		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH73		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH74		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH75		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH76		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH77		no peaks found	MRS, Ph 5.5
CrBH78		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH79		Lactobacillus intestinalis	MRS, Ph 5.5
CrBH80		Lactobacillus reuteri	MRS, Ph 5.5