Transcriptome analysis of blood and spleen in virulent and avirulent mouse malaria infection

Yuancun Zhao1, Caroline Hosking2, Deirdre Cunningham2, Jean Langhorne2✉ & Jing-wen Lin1✉

Background & Summary

Malaria is a devastating infectious disease and the immune response is complex and dynamic during a course of a malarial infection. Rodent malaria models allow detailed time-series studies of the host response in multiple organs. Here, we describe two comprehensive datasets containing host transcriptomic data from both the blood and spleen throughout an acute blood stage infection of virulent or avirulent Plasmodium chabaudi infection in C57BL/6 mice. The mRNA expression profiles were generated using Illumina BeadChip microarray. These datasets provide a groundwork for comprehensive and comparative studies on host gene expression in early, acute and recovering phases of a blood stage infection in both the blood and spleen, to explore the interaction between the two, and importantly to investigate whether these responses differ in virulent and avirulent infections.

Genome-wide expression profiling is being increasingly applied to dissect the complex details of the host response to malaria infection10. As blood is the most accessible tissue in field studies, numerous field studies analyse blood transcriptomes as read-outs for anti-malarial immunity11–13. Therefore, it is very important to understand whether the immune responses detected in the blood serve as a reliable proxy for immune responses occurring in the spleen; if so, to what extent and at which stage of infection are they most closely related. To date, few studies have performed transcriptomic analyses of the blood in the mouse model14 and only one study carried out by us attempted to investigate the similarities between blood and spleen transcriptome14.

Here we describe two comprehensive, time-series analyses of the blood and spleen transcriptomic changes throughout the acute phase of blood stage infection (Fig. 1a) using a well-established rodent malaria model,
Plasmodium chabaudi. This parasite is widely used to study host responses as it mirrors many pathological manifestations associated with *P. falciparum* infection, the most deadly species infecting humans, including parasite sequestration, severe malarial anemia, and chronic infection.

Time-series gene expression analysis is most helpful in identifying genes with transient expression changes and in investigation of gene regulation profiles during an infection. In our previous studies, we showed that pathology signatures can be picked up from blood transcriptome and they are quite distinct in the avirulent *P. chabaudi* AS or virulent *P. chabaudi* CB infection; further analysing the blood and spleen transcriptome from the avirulent *P. chabaudi*
AS infection, we identified only a small set of immune genes shared between them. Here we report a new dataset of spleen transcriptome from the virulent *P. chabaudi* CB infection, which were collected from the same mice as the published blood transcriptome. Our datasets, including the published PcAS/PcCB blood and this new PcCB spleen transcriptome, offer a unique possibility to identify the complete set of activated or suppressed genes during an acute blood stage infection, to infer their rates of change and their causal effects. Further, it would be of high interest to investigate whether the interaction between blood and spleen differ in these two infections or whether more subtle relationship can be unearthed using more elaborate time modeling methods.

**Methods**

**Mice and parasites.** Female C57BL/6 aged 6–8 weeks from the SPF (Specific Pathogen Free) unit at the Francis Crick Institute Mill Hill Laboratory were housed under reverse light conditions (light 19.00–07.00, dark 07.00–19.00 GMT) at 20–22 °C, and were allowed access to diet and water ad libitum. This study was carried out in accordance with the UK Animals (Scientific Procedures) Act 1986 (Home Office license 80/2538 and 70/8326), and was approved by the Francis Crick Institute Ethical Committee.

Cloned lines of *Plasmodium chabaudi chabaudi* AS and CB were originally obtained from David Walliker, University of Edinburgh, United Kingdom. Infections were initiated by intraperitoneal injection of 10⁵ parasitised erythrocytes derived from cryopreserved stocks. The course of infection was monitored on Giemsa-stained thin blood films by enumerating the percentage of erythrocytes infected with asexual parasites (parasitemia). The limit of detection for patent parasitemia was 0.01% infected erythrocytes. During the experiments, mouse condition were closely monitored. Core body temperature was measured with an infrared surface thermometer (Fluke); body weight was calculated relative to a baseline measurement taken before infection; and erythrocyte density was determined on a VetScan HM5 haematology system (Abaxis). The animals were euthanized upon reaching humane end points by showing the following signs: emaciation (more than 25% weight loss), persistent labored breathing, severe hypothermia (body temperature below 28 °C), inability to remain upright when conscious or lack of natural functions, or continuous convulsions lasting more than 5 min.

| mouse No. | Blood (GSE93631) | Spleen (GSE123391) |
|-----------|------------------|---------------------|
| naive_D0_m1 | GSM2459164 | 8762536135_E | GSM3502544 | 9440690022_B |
| naive_D0_m2 | GSM2459165 | 8762536084_A | GSM3502545 | 9440690030_C |
| naive_D0_m3 | GSM2459166 | 8762536052_F | GSM3502546 | 9440690035_B |
| naive_D12_m1 | GSM2459167 | 8762536072_E | GSM3502568 | 9440690030_F |
| naive_D12_m2 | GSM2459168 | 8784170061_E | GSM3502569 | 9440690037_C |
| naive_D12_m3 | GSM2459169 | 8784170059_F | GSM3502570 | 9440690042_A |
| AS_D2_m1 | GSM2459172 | 8762536135_C | GSM3502547 | 9440690022_C |
| AS_D2_m2 | GSM2459171 | 8762536084_B | GSM3502548 | 9440690035_A |
| AS_D2_m3 | GSM2459173 | 8784170059_G | GSM3502549 | 9440690037_D |
| AS_D2_m4 | GSM2459170 | 8762536072_F | GSM3502550 | 9440690042_C |
| AS_D4_m1 | N/A | N/A | GSM3502551 | 9440690022_D |
| AS_D4_m2 | GSM2459175 | 8762536084_F | GSM3502552 | 9440690030_A |
| AS_D4_m3 | GSM2459174 | 8762536052_B | GSM3502553 | 9440690037_E |
| AS_D4_m4 | GSM2459176 | 8784170061_A | GSM3502554 | 9440690042_B |
| AS_D6_m1 | N/A | N/A | GSM3502555 | 9440690022_E |
| AS_D6_m2 | GSM2459177 | 8762536052_E | GSM3502556 | 9440690035_C |
| AS_D6_m3 | GSM2459179 | 8784170059_B | GSM3502557 | 9440690042_D |
| AS_D6_m4 | GSM2459178 | 8762536135_A | GSM3502558 | 9440690037_A |
| AS_D8_m1 | GSM2459182 | 8784170059_E | GSM3502559 | 9440690022_F |
| AS_D8_m2 | GSM2459181 | 8762536135_F | GSM3502560 | 9440690030_B |
| AS_D8_m3 | GSM2459180 | 8762536084_D | GSM3502561 | 9440690035_D |
| AS_D10_m1 | GSM2459185 | 8784170061_B | GSM3502562 | 9440690035_E |
| AS_D10_m2 | GSM2459184 | 8762536135_D | GSM3502563 | 9440690030_D |
| AS_D10_m3 | GSM2459183 | 8762536052_C | GSM3502564 | 9440690037_F |
| AS_D12_m1 | GSM2459186 | 8762536072_C | GSM3502565 | 9440690030_E |
| AS_D12_m2 | GSM2459188 | 8784170061_C | GSM3502566 | 9440690035_F |
| AS_D12_m3 | GSM2459187 | 8762536084_E | GSM3502567 | 9440690037_B |

Table 1. GEO accession numbers of blood or spleen samples that were derived the same PcAS infected mouse (Data Source: GSE93631 and GSE123391).
RNA isolation and preparation for microarray analysis. The sample collection and processing workflow is summarised in Fig. 1. These methods are expanded versions of descriptions in our related studies. Female C57BL/6 mice aged between 6–8 weeks were injected intraperitoneally with 10^5 infected red blood cells of *P. chabaudi* AS or CB strain. At 2, 4, 6, 8, 10 and 12 days post infection (dpi), 0.5 mL of blood was collected via cardiac puncture into 1 mL Tempus RNA stabilising solution (Applied Biosystems). Spleens were aseptically removed and were homogenised immediately in TRI reagent (Ambion) by pulsing with a Polytron homogenising unit (Kinematic). An extra day 9 group was collected from *PcCB* infected mice that had reached humane end points. Naïve control samples were also collected on day 0 (the day of infection) and day 12 (the end of the experiment). Samples were snap frozen in dry ice and stored at −80 °C until RNA isolation.

Total blood RNA was extracted using PerfectPure RNA Blood Kit (5 PRIME), and Globin mRNA was removed from 2 µg of total isolated RNA using GLOBINclear 96-well Mouse/Rat Whole Blood Globin Reduction Kit (Ambion) according to the manufacturer's instructions. Total splenic RNA was extracted using RiboPure RNA Purification Kit (Ambion) following the manufacturer's protocol. All RNA samples derived from the same experiment were isolated altogether at the end of the experiment, in 2–3 batches within a day. Globin mRNA reduction was performed in 2 batches, one for all *PcAS* blood samples and the other for all *PcCB* blood. Batch information for RNA isolation and subsequent processing was provided in Online-only Tables 1–4.

To test whether parasite RNA gives signals in microarray analysis of mouse gene expression, an independent experiment was performed using naïve and infected mouse blood RNA, and purified parasite RNA. RNA from naïve or infected blood was processed as described above. Parasite purification and RNA extraction methods were performed as described previously. Briefly, infected blood collected at day 8 post infection were depleted of leukocytes by filtration through Plasmodipur filters (EuroProxima) followed by erythrocyte lysis using 0.15%
Saponin (Sigma) in ice-cold PBS and extensive washes with PBS. Purified parasite pellets were then resuspended in 1 ml TRI reagent (Ambion), snap-frozen on dry ice and kept at −80 °C. Parasite RNA was extracted using RiboPure RNA Purification Kit (Ambion) according to the manufacturer’s protocols. Parasite RNA was also subjected to Globin mRNA removal. Batch information for RNA isolation and processing was provided in Online-only Table 5.

Biotinylated, amplified antisense complement RNA (cRNA) samples were prepared from 300 ng of either globin reduced blood/parasite RNA, or splenic total RNA using Illumina TotalPrep RNA Amplification Kit (Ambion). cRNA was prepared in 4 batches: PcAS blood, PcCB blood, all spleen samples and blood/parasite RNA. At each step, the quantity of the RNA samples was measured using NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific) and the quality of RNA was verified using Agilent 2100 bioanalyzer (Agilent Technologies) or Caliper LabChip GX (Caliper Life Sciences), provided as RNA Integrity Number (RIN) or RNA Quality Score (RQS), respectively. Only RNA samples with RIN/RQS above 7 were used for subsequent treatment and analysis; however, for blood samples with high parasite load, parasite rRNA affected RIN/RQS determination, and the quality of these samples were determined by examining electropherograms. RNA concentration and RIN/RQS of each sample were also provided in Online-only Tables 1–5.

**Microarray hybridisation and raw data export.** The following procedures for microarray hybridisation and data acquisition was done for each sample. Briefly, 1.5 µg of labelled cRNA was hybridised to Illumina Mouse WG-6 v2.0 Expression BeadChip (consisting of 45,281 probe sets representing 30,854 genes) according to the manufacturer’s protocols. The arrays were then washed, blocked, stained and scanned on an Illumina iScan, following the manufacturer’s instructions. Illumina BeadStudio/GenomeStudio 1.8.0 software was used to generate

---

**Fig. 2** Quality check of BeadChip gene expression data of PcAS blood samples. (a) Box plot showing distribution of 3,000 randomly sampled probe signals for normalised PcAS infected blood expression data. The median, two hinges, two whiskers and outlying points were shown. (b) Principal component analysis of normalised expression data of naïve and infected blood samples. (c) Hierarchical clustering plot of normalised intensity data among the samples was generated using agglomerative clustering with average linkage. (d) Heatmap of Euclidean distance. A full feature set was used for (b–d). This dataset was submitted to GEO (GSE93631).
signal intensity values, quality control values, and to subtract background. Hybridisation was performed in 4 batches: PcAS blood, PcCB blood, all spleen samples and blood/parasite RNA.

**Microarray data preparation and analysis.** Data input, quality control, variance stabilisation, log transformation and quantile normalisation were performed using the lumi package\(^\text{18}\). The full feature set (a total of 45,281 probes) of each sample was used for the following analyses including hierarchical clustering, principle component analysis (PCA) and Euclidean distance, all conducted using R 3.6.0 (www.r-project.org). For hierarchical clustering, agglomerative clustering with average linkage was used.

**Data Records**

Gene expression data were deposited at the Gene Expression Omnibus database (GEO) under the following accession numbers: GSE93631\(^\text{19}\) (AS and CB blood) and GSE123391\(^\text{20}\) (AS spleen) which were published previously\(^\text{8,14}\), GSE145781\(^\text{21}\) (CB spleen) and GSE145634\(^\text{22}\) (the raw data of parasite RNA experiment) which were new datasets.

GEO accession numbers of blood or spleen samples that were derived from the same mouse were provided in Tables 1 and 2. Batch information, RNA quality and concentration and related GEO accession numbers were provided in Online-only Tables 1–5.

**Technical Validation**

**Sample preparations and quality control.** Several aspects of the experiment were designed to ensure the quality of the data. For example, the control naïve mice were randomly selected from the same batch of age-matched mice, 3 of which were sacrificed at the same day of infection, and 3 of which were housed under the same conditions as the infected mice and were sacrificed along with mice after 12 days of infection. All mice in

---

*Fig. 3* Quality check of BeadChip gene expression data of PcAS spleen samples. (a) Box plot showing distribution of 3,000 randomly sampled probe signals for normalised PcAS spleen expression data. The median, two hinges, two whiskers and outlying points were shown. (b) Principal component analysis of normalised expression data of naïve and infected spleen samples. (c) Hierarchical clustering plot of normalised intensity data among the samples was generated using agglomerative clustering with average linkage. (d) Heatmap of Euclidean distance. A full feature set was used for (b-d). This dataset was submitted to GEO (GSE123391).
the infected group were infected at the same time and were randomly selected for sample collection at each time point. Overall, both blood and spleen samples collected from either PcAS or PcCB infections showed uniformed normalised intensities (Figs. 2a and 3a). Importantly, high similarities were observed between biological replicates (Figs. 2 and 3).

Quality check of time dependent responses. In the naïve control group, mice collected at day 0 or day 12 clustered together in all 4 datasets. Interestingly, samples collected at 2 dpi at which time point the infection rate was below microscopic detection level, also cluster with naïve groups; and this was observed in both the blood and spleen in either infection (Figs. 2–5). In the avirulent PcAS infection, from day 4 onwards, the expression profiles changed significantly, showing clear time-dependent responses in both the blood and the spleen. At 4 dpi, spleen showing longer distance from the naïve groups than the blood, 46.4 vs 21.6 distance on PC2 (Figs. 2b and 3b), which indicates higher host responses in the spleen than in the blood. Similar responses took place in the virulent PcCB infection, showing 4 dpi-naïve distance on PC2 of 18.5 in the blood vs 32.0 in the spleen (Figs. 4b and 5b). This is in line with the current view that parasite-host interaction mainly take place in the spleen.

An interesting difference between the blood and spleen is the divergence between day 6 and 8 post infection. In the avirulent PcAS infection the distances between 6 and 8 dpi on PC1 were 36.4 in the blood and 91.8 in the spleen (Figs. 2b and 3b). This is slightly less striking in the virulent PcCB infection, with 21.1 in the blood and 66.8 in the spleen. These differences are also apparent in hierarchical clustering and heatmaps of euclidean distance (Figs. 2 and 3).

The striking differences between PcAS and PcCB infections were the responses took place between day 10 and 12 post infection. In the avirulent PcAS infection the distances between 6 and 8 dpi on PC1 were 36.4 in the blood and 91.8 in the spleen (Figs. 2b and 3b). This is slightly less striking in the virulent PcCB infection, with 21.1 in the blood and 66.8 in the spleen. These differences are also apparent in hierarchical clustering and heatmaps of euclidean distance (Figs. 2 and 3).
in the virulent PcCB infection the two days differed in both PC1 and PC2 (Figs. 4b and 5b), and the heatmaps of Euclidean distance showed that 12 dpi clearly separate from other samples (Figs. 4d and 5d).

Parasite RNA does not affect BeadChip gene expression results. Because the malaria parasite infects erythrocytes, RNA isolated from the infected blood contains both mouse and Plasmodium RNA. We therefore performed an independent experiment to rule out the interference of parasite RNA in downstream analysis using Mouse WG-6 v2.0 Expression BeadChip. We prepared purified *P. chabaudi* AS parasite RNA by passing infected blood through a leukocyte filter, usually removing more than 99% leukocytes, followed by erythrocyte lysis and extensive washes. Globin mRNA removal was also performed as for infected blood samples. As shown in Fig. 6, the numbers of detectable probes in parasite samples were significantly lower (Fig. 6a), and this hindered the normalisation step. Moreover, the non-normalised expression data of parasite samples showed very different density or cumulative density profiles (Fig. 6b,c). After removing parasite data from the dataset, the subsequent analyses can be easily performed and it was clear that the infected blood collected at 8 dpi significantly differed from naïve blood (Fig. 6d), validating our previous finding.

Usage Notes
One major advantage of this study is that we collected both the blood and spleen simultaneously from the same mouse (GEO accession numbers of blood or spleen samples that were derived from the same mouse were provided in Tables 1 and 2) throughout the acute phase of blood stage infection, from as early as day 2 post infection when the infection rate was below microscopic detection, till day 12 post infection when the parasite load was controlled. Moreover, the samples were collected at 2-day intervals to allow a more detailed analysis of the time-dependent transcriptional changes. It is hoped that this will facilitate the users to investigate in detail the interaction between the blood and the spleen. It would also provide some answers to the question of whether some of the responses in the blood happen before or after the spleen responses, for example using time series
modelling. And importantly, we collected samples from both the virulent PcCB and the avirulent PcAS infections. It would be of high interest to investigate whether the interaction between blood and spleen differ in these two infections.

**Code availability**

R scripts for raw data reading, normalisation, QC, and plotting were available at https://github.com/LuChenLab/Rscript_for_BeadChip.git.

Received: 18 March 2020; Accepted: 3 July 2020;

Published online: 04 August 2020

**References**

1. Buffet, P. A. *et al.* The pathogenesis of *Plasmodium falciparum* malaria in humans: insights from splenic physiology. *Blood* **117**, 381–392 (2011).
2. Del Portillo, H. A. *et al.* The role of the spleen in malaria. *Cellular microbiology* **14**, 343–355 (2012).
3. Cunnington, A. J., Riley, E. M. & Walther, M. Stuck in a rut? Reconsidering the role of parasite sequestration in severe malaria syndromes. *Trends Parasitol* **29**, 585–592 (2013).
4. Van den Steen, P. E. *et al.* Pathogenesis of malaria-associated acute respiratory distress syndrome. *Trends Parasitol* **29**, 346–358 (2013).
5. Deroost, K. *et al.* Hemozoin induces lung inflammation and correlates with malaria-associated acute respiratory distress syndrome. *Am J Respir Cell Mol Biol* **48**, 589–600 (2013).

---

**Fig. 6** Validation of parasite RNA does not affect BeadChip gene expression results. (a) Bar chart showing the number of probes detected in each sample. (b) Density plot of non-normalised expression data showing the signal density distribution. (c) Cumulative distribution function plot of non-normalised expression data of each sample. Arrowheads indicate parasite samples. (d) PCA plot of normalised expression data from infected and naïve blood samples after excluding parasite samples. This dataset was submitted to GEO (GSE145634).
Acknowledgements
The authors would like to thank Dr. Lu Chen for critical assistance in data analysis. This work was supported by the Wellcome Trust (WT102907) and Francis Crick Institute (FC001101), and a Wellcome Trust Senior Investigator award to Jean Langhorne (104777/Z/14/Z).

Author contributions
Y.-c. Z. analysed the data, created the figures and wrote the manuscript; J.-w. Lin designed the project, performed the experiment and wrote the manuscript; C.H. and D.C. assisted in the animal experiments and sample maintenance; J.L. directed the project and edited the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to J.L. or J.-w.L.

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

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

The Creative Commons Public Domain Dedication waiver http://creativecommons.org/publicdomain/zero/1.0/ applies to the metadata files associated with this article.

© The Author(s) 2020