# Identification of host genetic factors involved in cerebral malaria resistance: from mouse to human

Joanne Berghout

Department of Biochemistry
McGill University
Montreal, QC, CANADA

Aug 2014

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Doctor of Philosophy

# Table of Contents

# List of Figures

## List of Tables

## Abstract:

Malaria is a parasitic (*Plasmodium spp.*) blood infection that affects over 300 million people annually and has applied strong selective pressure on the human genome in endemic areas. Several protein variants in erythrocyte and immune cells have been associated with protection against uncomplicated and severe disease including malarial anemia and cerebral malaria. We have used an unbiased forward genetic approach to compare inbred mice that display differential susceptibility to *P. berghei* ANKA infection as a means to study genetic determinants of experimental cerebral malaria. Susceptible C57BL/6 mice display early appearance of neurological symptoms and uniform lethality prior to day 12 post-infection, while resistant BALB/c mice succumb to anemia much later in infection without developing significant cerebral pathology. We used quantitative trait locus (QTL) mapping in F2 mice derived from C57BL/6 and BALB/c parental strains, and localized this difference to a 10-20Mb region of chromosome 19 (LOD = 4.69, $p<0.05$), which we termed *Berr5*. Comparing brain RNA expression profiles following infection suggested the interferon-induced proteins with tetratricopeptide repeats gene cluster (*Ifit1-3*) or *Fas* as candidate genes underlying differential survival. *Berr5* also overlapped with three other infection and inflammatory QTLs supporting a model where upregulation of inflammatory molecules may be critical to cerebral malaria pathology. Independently, we also report that BXH2 mice carrying a mutation in Interferon Regulatory Factor 8 ($Irf8^{R294C}$) are completely resistant to cerebral malaria. Using microarray transcript profiling and chromatin immunoprecipitation coupled to sequencing (ChIP-Seq), we found that 92 of the 125 genes upregulated by infection in susceptible controls contained an IRF8 binding site, and further identified a core subset of 53 genes which were also regulated in response to

*Mycobacterium tuberculosis*. Response of these genes in infected BXH2 mice is significantly blunted and therefore, this list may represent key drivers of the pathological response. Third, having previously shown that mutations in pyruvate kinase cause susceptibility to blood stage malaria *in vivo* (*P. chabaudi*) and in culture (*P. falciparum*), we investigated whether loss-of-function variants in human *PKLR* may be retained in human populations living in malaria-endemic areas. As a first step we sequenced all coding regions of *PKLR* in 384 geographically diverse human samples. We observed high diversity in African samples and identified 12 putatively loss-of-function alleles, translating to a mutant allele carrier frequency of 1.5%. Additional analyses revealed a slight, but significant deviation from neutral evolution in both Pakistan and Sub-Saharan Africa, consistent with the possibility that PKLR deficiency may offer some advantage, possibly by modulating malarial infection.

# Resumé

Chaque année, 300 millions de personnes sont victimes du paludisme, une infection parasitaire (*Plasmodium spp.*) du sang, qui s'est aussi révélé la source d'une forte pression sélective sur le génome humain dans les régions endémiques. Plusieurs variants de protéines chez les cellules immunitaires et les érythrocytes ont été associés à un effet protecteur contre la progression et certains symptômes sévères du paludisme, telles anémie et atteintes cérébrales. Par approche génétique directe, nous avons comparé certaines souches de souris démontrant différents degrés de susceptibilité à l'infection par *P. berghei ANKA* afin d'étudier les facteurs génétiques liés au neuropaludisme en laboratoire. La souris C57BL/6 (susceptible) présente très tôt des symptômes neurologiques et succombe généralement avant le 12ième jour suivant l'infection, alors que la souris BALB/c (résistante) cèdera beaucoup plus tard à la maladie des suites d'une anémie. Nous avons adopté une stratégie par cartographie de loci à caractère quantitatif (QTLs) chez les souris F2 obtenues à partir des souches parentales C57BL/6 et BALB/c. Nous y avons distingué un locus de 10-20Mb sur le chromosome 19 (LOD = 4.69, $p<0.05$) responsable de l'écart entre chacun des phénotypes parentaux, lequel fut prénommé *Berr5*. En comparant l'expression des gènes suivant l'infection, le groupe de gènes "interferon-induced proteins with tetratricopeptide repeats" (*Ifit1-3*) et *Fas* furent impliqués comme gènes candidates contrôlant le taux de survie pour cette infection. *Berr5* correspond aussi à trois autres QTLs associés à certaines autres infections et à l'inflammation, ce qui pourrait indiquer qu'une augmentation des facteurs inflammatoires pourrait bien être essentielle à la pathologie du neuropaludisme.

En deuxième lieu, nous faisons aussi état d'une souris (BXH2) porteuse d'une mutation dans le facteur de transcription Facteur régulateur d'interféron 8 (*Irf8*$^{R294C}$) comme étant totalement résistante au neuropaludisme. Par profilage de l'expression des gènes sur micropuces ainsi que par l'immunoprécipitation de chromatine suivie par séquençage (ChIP-Seq), nous avons observé la présence d'une site de liaison de IRF8 dans la séquence de 92 des 125 gènes activés au cours de l'infection par *P. berghei ANKA* et nous avons identifié un noyau de 53 gènes qui sont également activés suivant une infection par *Mycobacterium tuberculosis*. Puisque l'activation de ces gènes est sévèrement réduite chez la souris BXH2, les gènes relevés pourraient représenter les principaux facteurs dans le développement d'une réponse pathologique. Enfin, ayant déjà démontré que les mutations dans le gène codant pour la pyruvate kinase (*Pklr*) sont la cause d'une susceptibilité accrue au stade érythrocytaire du paludisme *in vivo* (*P. chabaudi*) et en culture (*P. falciparum*), nous avons porté notre attention sur les variants déficitaires de PKLR chez l'humain et la sélection positive de ces allèles dans la population des régions du monde où le paludisme est endémique. Nous avons tout d'abord séquencé la séquence codante complète pour le gène PKLR chez 384 individus de différentes origines. Nous avons observé une grande diversité d'allèles dans les échantillons d'origine africaine ainsi qu'un total de 12 variants présumés déficitaires, ce qui se traduit par une fréquence d'individus porteurs d'un allèle muté à 1.5%. En parallèle, nos études ont aussi révélé la déviation légère mais significative en comparaison à une évolution totalement neutre du gène dans deux régions géographiques additionnelles (Pakistan et Afrique sub-saharienne), ce qui consolide la notion qu'une déficience enzymatique pour la protéine PKLR pourrait se traduire en un avantage et influencer l'infection du paludisme de manière protectrice.

## Acknowledgements

First and foremost I would like to thank my advisor Dr. Philippe Gros for giving me the opportunity to work in his lab, for the fascinating projects and for all his guidance, support and mentorship along the way. He provided me with outstanding training in research science, academic integrity, and brought together a group of people who were stimulating, collegial and a lot of fun to work with, as he is himself. I would also like to thank Dr. Elena “Lala” Torban who got my feet wet in molecular biology. As well, Gundula Min-Oo who was not only a close friend, but also an invaluable teacher in the lab for complex trait genetic analysis and malaria biology. All members of the laboratory deserve acknowledgement for their contributions to an excellent workplace, the exchange of ideas across the aisles or in lab meetings and practical suggestions or assistance along the way. For my projects as described in this thesis, I would like to especially thank Susan Gauthier for mouse colony management, Dr. Silayuv Bongfen, Sabrina Torre, Jimmy Kennedy, Patricia D’Arcy, Jeremy Schwartzentruber and the others involved who kept the ENU project running smoothly and productively. Dr. David Langlais and Irena Radovanovic who performed cell assays, FACS, ChIP (David) and were always sharp on statistical and bioinformatic analyses, particularly on the IRF8 project. Dr. Kevin Kain and his laboratory as well as Dr. Philip Awadalla and his laboratory who collaborated on the human side of malarial disease, and Mifong Tam who provided technical prowess with injections, tissue collection and analysis and valuable advice about immune responses in murine malaria. I would like to acknowledge the members of my Research Advisory Committee: Dr. Maya Saleh, Dr. Silvia Vidal, Dr. Anny Fortin and Dr. Mary M. Stevenson for their perspectives and advice on long-range project direction. I am grateful as well for the support of NSERC CGS-M, CIHR PGS-D, the McGill Department of

Biochemistry Graduate Excellence Fellowship, the Faculty of Medicine Principal's Graduate Student Fellowship, the Arthur S. Hawkes Fellowship, and the McKinnell Travel award. And last, but not least, I want to acknowledge the great friendships I formed during graduate school, who kept me sane, motivated, and we even talked a lot of science; in no particular order: Dr. Gundula Min-Oo, Oxana Kapoustina, Dr. Katrina Choe, Dr. Mark Livingstone, Dr. Mike Bidinosti and Dr. Filipp Frank in particular.

## Preface

The work described in Chapters 2, 3 and 4 of this thesis is published as follows:

**Chapter 2:** Berghout J., Min-Oo G., Tam M., Gauthier S., Stevenson M.M., Gros P. 2010. "Identification of a novel cerebral malaria susceptibility locus (Berr5) on mouse chromosome 19." *Genes and Immunity*. Jun;11(4):310-8 Epub 2009 Oct 20.



**Chapter 3:** Berghout J.*, Langlais D.*, Radovanovic, I., Tam M., MacMicking J.D., Stevenson M.M., Gros P. 2013 "Irf8-regulated genomic responses drive pathological neuroinflammation during cerebral malaria" *PLoS Pathogens*; 9(7):e1003491 Epub 2013 July 11

*these authors made equal contributions to the final work



**Chapter 4:** Berghout J., Higgins S., Loucoubar C., Sakuntabhai A., Kain K.C., Gros P. 2012 "Genetic diversity in human erythrocyte pyruvate kinase." *Genes and Immunity.* Jan;13(1):98-102 Epub 2011 Aug 11

# Contribution of Authors:

**Chapter 2:** Dr. Gundula Min-Oo introduced me to working with malaria parasites and trained me in the background required for QTL mapping. She also participated in many valuable discussions about the results and experimental strategies. Mifong Tam (of Dr. Mary M. Stevenson's group) injected all of the mice in this project, and performed some early dissections and ELISA assays. Susan Gauthier managed the mouse breeding for the large intercross as well as normal colony maintenance. Dr. Mary M. Stevenson assisted with immunological expertise and a critical reading of the manuscript. I performed all of the experiments and calculations used in the paper and co-wrote the manuscript with Dr. Philippe Gros who provided guidance and supervised the project.

**Chapter 3:** Dr. David Langlais performed and analyzed all of the ChIP-Seq experiments, queried gene lists for IRF8 binding sites, prepared figure 4 and offered a lot of excellent feedback on the manuscript and other figures. Irena Radovanovic repeated the FACS profiling and analysed all FACS plots and statistics together with David Langlais. Mifong Tam conducted infections with *Il12p40*$^{-/-}$, *Ifng*$^{-/-}$ and appropriate control mice. She also collected serum for cytokine analysis. Dr. John D. MacMicking kindly provided *Irgm1*$^{-/-}$ mice and offered feedback on the manuscript. Dr. Mary M. Stevenson provided *Il12p40*$^{-/-}$ and *Ifng*$^{-/-}$ mice and participated in helpful discussions, especially as the project was beginning. I conducted most of the infections, prepared RNA, analyzed the microarrays and did the gene list analysis. I co-wrote the paper with Dr. Philippe

Gros who also offered many helpful suggestions and advice regarding results and experimental design.

**Chapter 4:** Sarah Higgins and I designed the PKLR sequencing strategy and she performed about half of the PCR reactions. Cheik Loucoubar (of Dr. Anavaj Sakuntabhai's group) performed the coalescent calculations to determine empirical p-values. Dr. Anavaj Sakuntabhai made many helpful suggestions for human sequence data analysis, including suggesting specific statistical tests. I performed PCR reactions, identified SNPs and performed all statistical analysis on the data. I wrote the manuscript. Dr. Philippe Gros conceived and supervised the project, as well as co-wrote the manuscript.

# Original Contributions to Knowledge:

**Chapter 2:** Identification of a novel cerebral malaria susceptibility locus (*Berr5*) on mouse chromosome 19

1. Characterized differential host responses to standardized *Plasmodium berghei* ANKA infection between C57BL/6 and BALB/c inbred strains of mice in terms of survival, cerebral malaria symptoms development, blood brain barrier integrity, parasitemia and serum IFNG production.
2. Identified a locus on mouse Chromosome 19 (*Berr5*) that differentiates this host response in an informative F2 cross. Positional candidate genes include the *Ifit1-3* gene family, *Fas* and 13 others with strain-dependent differential gene expression during infection.

**Chapter 3:** Irf8-regulated genomic responses drive pathological neuroinflammation during cerebral malaria

3. Identification of interferon regulatory factor 8 (*Irf8*) as a critical factor for experimental cerebral malaria pathological processes to occur. Cerebral malaria resistance in hypomorphic $Irf8^{R294C}$ mutants is inherited semi-dominantly and associated with IRF8 activation signaling activity, as heterozygote $Irf8^{R294C/+}$ mice have normal cellular profiles but significantly improved survival versus $Irf8^{+/+}$.
4. *Performed with Irena Radovanovic:* characterized infection-induced myeloid and lymphoid cellular infiltration to brain tissue in IRF8-competent and IRF8-deficient mice.

5. *Performed by Dr. David Langlais, but presented in this work:* generated a genome-wide profile of IRF8 transcription factor binding sites in stimulated and unstimulated mouse macrophages using chromatin immunoprecipitation and sequencing (ChIP-seq).
6. Delineation of a gene expression core shared by mice with experimental cerebral malaria pathology and *Mycobacterium tuberculosis* pathology, where IRF8 transcription factor binding sites are extremely highly represented.

**Chapter 4:** Genetic diversity in human erythrocyte pyruvate kinase

7. Characterization of the genetic diversity in human *PKLR* using a panel of malaria endemic and non-endemic ethnic DNA. This was the first large-scale effort examining pyruvate kinase variation in multi-ethnic populations by sequencing.
8. Identification of 7 putative loss-of-function PKLR alleles in 12 unrelated individuals, including novel variants in DNA samples derived from malaria-endemic populations. These appeared at a total allele frequency of 1.5±0.9%.
9. These population characteristics were suggestive of positive selection for non-synonymous variant retention by neutrality testing.

# Chapter 1 : Introduction and Literature Review

## 1.1 Malaria as a global health threat.

Nearly half of the world's population is at risk for malaria and almost 250 million clinical cases are reported each year, making it one of the "Big Three" infectious disease along with tuberculosis and HIV/AIDS[1]. Malaria is caused by infection with blood-borne *Plasmodium spp*. parasites, notably *P. falciparum* and *P. vivax,* though *P. malariae* and *P. ovale* are also known cause human disease. The greatest disease burden is in sub-Saharan Africa, which accounts for 80% percent of all malarial infections and 90% of deaths, due almost entirely to *P. falciparum* (Figure 1-1)[2]. Through much of Asia and South America, *P. falciparum* co-exists with the less lethal *P. vivax* parasite[3, 4]. Malaria parasites are passed via the bite of parasitized *Anopheles* mosquitoes who inject *Plasmodium* sporozoites along with saliva into the bloodstream of human hosts during a blood meal. In hyperendemic areas, inhabitants may receive between 149-269 infective mosquito bites per year, generally with significant seasonal and geographic variability[5, 6]. One study reported transmission intensity as high as >1,500 infective bites per person per year in a swampy area near the Nile River in Uganda which resulted in more than 80% of children aged 1-9 years old being chronically positive for parasites[7]. Other analyses in multiple African sites have calculated an average entomological inoculation rate (EIR) at 112.2 infective bites per year ranging from 0.6 in one Sudanese site to 418.0 in a Ghanaian location[8]. Either way, an EIR greater than 15 bites/year was associated with >50% malaria prevalence in all sites and EIR greater than 200 was consistently associated with >80% malaria prevalence[9].

Consequently, major efforts at malaria control are often aimed at vector (mosquito) control using insecticide-treated bed nets, indoor residual spraying and environmental management to minimize or eliminate standing water that

serves as mosquito breeding grounds and larval habitats[10]. Intermittent preventive treatment (IPT), which is administering anti-malarial drugs periodically to asymptomatic individuals and pregnant women in endemic areas, additionally aims to reduce parasite reservoirs and curtail development of severe disease[11]. Currently, there is no vaccine though research efforts have been underway for many years.

Despite substantial efforts aimed at both preventing infection and treating patients with anti-malarial drugs, malaria was still responsible for 781 000 deaths in 2009[1]. While this is a decrease from 985 000 deaths in 2000[1], it is still an incredible toll on global health and hampers future development, especially within Africa.

Morbidity as a result of malaria infections is generally due to either severe malarial anemia or cerebral malaria, with the additional risk of placental malaria in pregnant women which risks the health of both mother and developing child[12]. Children are particularly vulnerable to severe manifestations of the disease with the peak of severe cases and morbidity between 1-5 years of age[13, 14]. Severe malarial anemia is responsible for approximately one third of deaths associated with infection and is defined as >10,000 parasites/μl blood with hemoglobin concentration <5 g/dL (or hematocrit <15%)[15]. Disease may be accompanied by defective compensatory erythropoiesis, splenomegaly, thrombocytopenia, respiratory distress and hypoglycemia along with the recurrent fever and malaise common to uncomplicated malaria. Cerebral malaria is heterogeneous, but clinically characterized by unrousable coma, retinopathy, increased intracranial pressure due to oedema, ataxia and neurological lesions, often accompanied by microvasculature occlusion by erythrocytes and leukocytes[16]. Cerebral malaria pathology develops in an estimated 7-9% of clinical *P. falciparum* infections with a cerebral malaria case fatality rate near

16% (range 5-50%)[17-19]. Severe malarial anemia is treated with anti-parasitic drugs (preferentially quinine or artesunate), blood transfusion (if necessary) and iron supplementation, while treatment for cerebral malaria additionally requires management of the unconscious patient for possible hypoglycemia, extremely high fever and other organ involvement, as well as control of convulsions using sedatives and anti-seizure medications[20]. The particular pathology experienced by an individual is influenced by complex interactions between host genetics, environmental variables and parasite specific factors. Many of these factors remain unidentified, while others are partially understood. Combating the disease effectively requires a comprehensive knowledge of malaria parasite biology, host responses to infection and the capacity of specific host:parasite interactions to contribute towards disease resistance or susceptibility.

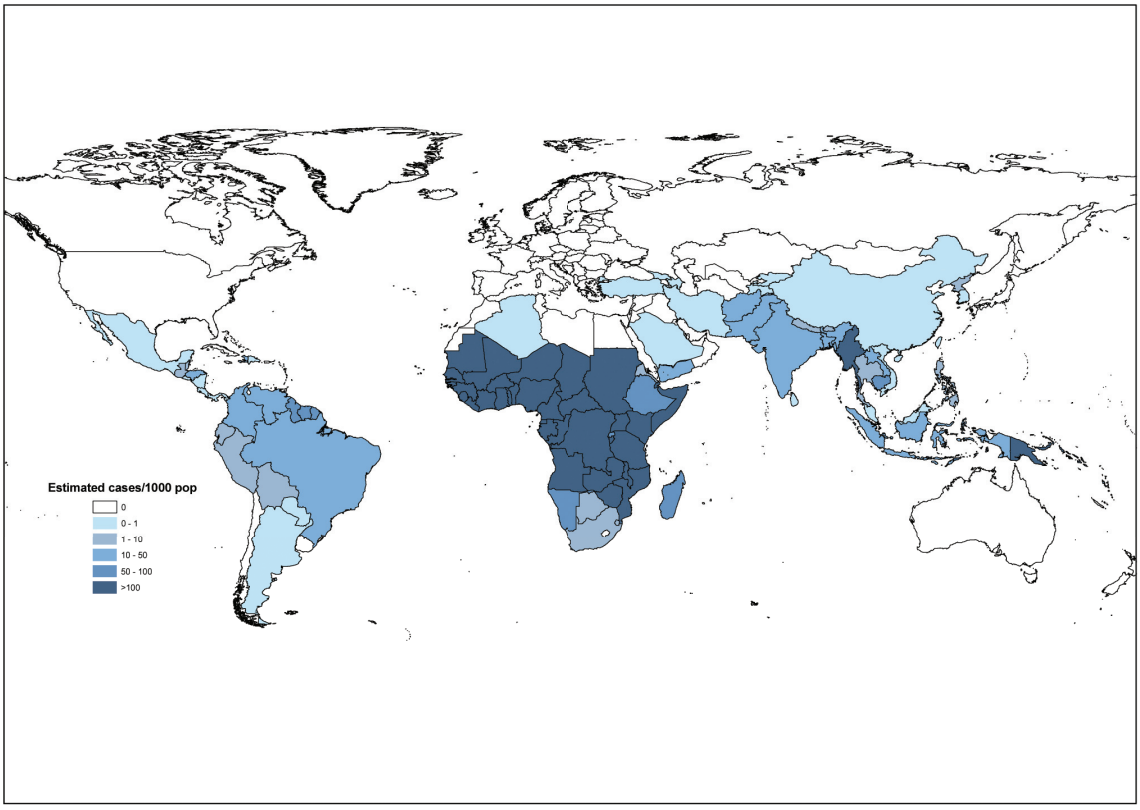



**Figure 1-1.** Estimated number of malaria cases per 1,000 people in the population for 99 endemic countries

## 1.2 Plasmodium parasite life cycle and biology:

### 1.2.1 Asexual (human infective) stage

During a normal infection, *Plasmodium* sporozoites are inoculated into the human bloodstream via the bite of a female *Anopheles* mosquito and rapidly make their way to the liver. Here they enter hepatocytes where they divide and mature into haploid merozoites. In *P. vivax* and *P. ovale* infections, sporozoites may also form hypnozoites – a dormant asymptomatic liver-sequestered stage which may be reactivated into symptomatic blood-stage malaria months or even years later[21]. The biology of dormancy and reactivation is not well understood, however reactivation has been linked to changes in temperature (perhaps indicative of seasonality), host stress, triggers within the saliva of mosquitoes and substrain or individual genetic factors in the parasites themselves[21].

Whether initial infection or reactivation, once merozoites enter the circulation they rapidly invade and multiply within erythrocytes, beginning the symptomatic blood stage portion of the disease (Figure 1-2). Inside an erythrocyte, each merozoite enlarges to form a ring-stage trophozoite, which then asexually replicates its genetic material, becoming a multinucleate schizont. The schizont then divides further to form 8-32 uninucleate merozoites[22, 23]. *P. falciparum* is a synchronous infection with a variable, but substrain-constant 24-48h periodicity, while the replicative cycle for *P. vivax* or *P. ovale* typically lasts 48h and *P. malariae* cycles go as long as 72h each[24]. Specific parasite alleles can alter the timing of the erythrocytic cycle with single-allelic family infections (clonal, or related parasites) behaving more predictably, and multi-allelic family infections (multiple species or strains) appearing more spread out or completely asynchronous. Each cycle represents a synchronized rupture of erythrocytes,

releasing merozoites along with erythrocyte cellular components and debris. This efflux of freely circulating immunogenic compounds provokes a rapid and powerful host immune response in macrophages and other cells which release anti-parasitic and pyrogenic cytokines, inducing the characteristic fever paroxysms of blood stage infection[25]. These reproducible cycles were responsible for the original descriptions of malarial disease as tertian (every third day) or quartan (every fourth day) fever. The synchronized cyclic nature of these parasite bursts is thought to contribute towards the ability of *Plasmodium* parasites to evade constant pressure from the host immune system [26].

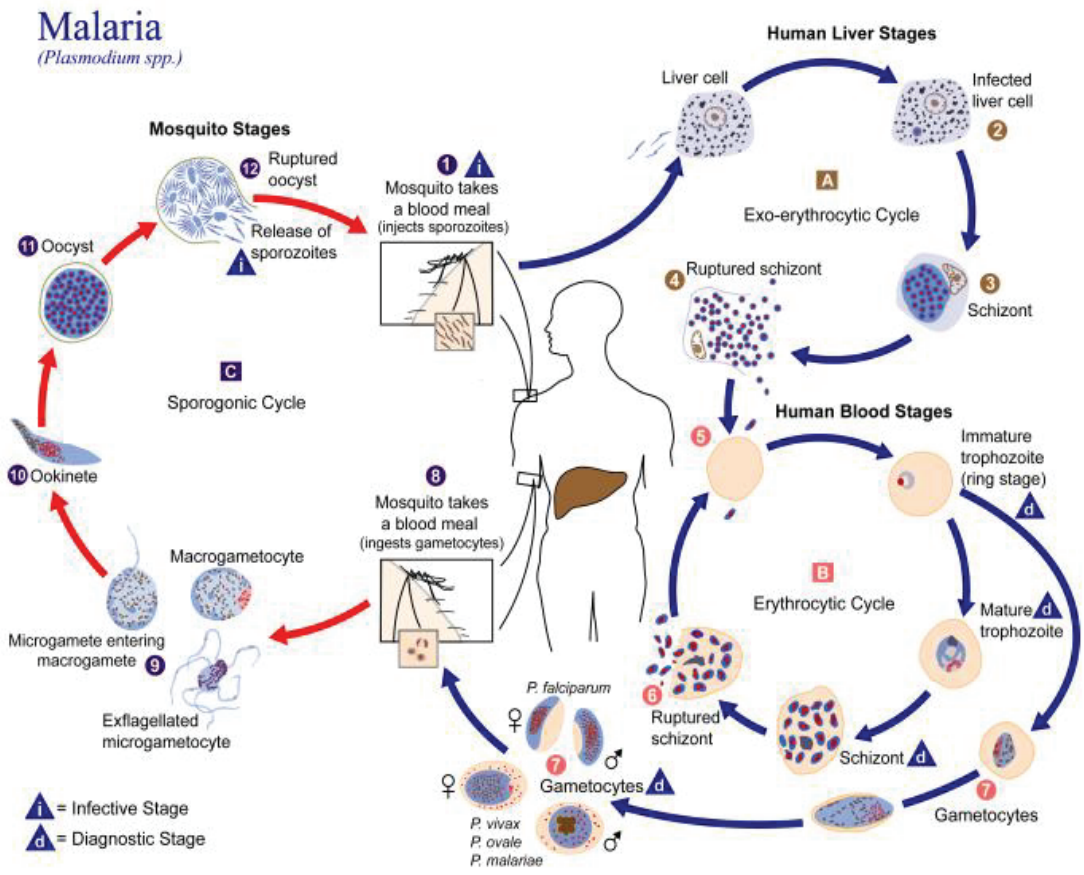



**Figure 1-2** The complex life cycle of malaria parasites

The malaria parasite life cycle involves two hosts. During a blood meal, a malaria-infected female *Anopheles* mosquito inoculates sporozoites into the human host (1) which infect liver cells (2) and mature into schizonts (3). These rupture to release merozoites (4). After this initial replication in the liver (exo-erythrocytic schizogony; [A]), the parasites undergo asexual multiplication in the erythrocytes (erythrocytic schizogony; [B]). Merozoites infect red blood cells (5). The ring stage trophozoites mature into schizonts, which rupture releasing merozoites (6). Some parasites differentiate into gametocytes (7). Gametocytes are ingested by an *Anopheles* mosquito during a blood meal (8) and begin the sexual sporogonic cycle [C]. While in the mosquito's mid-gut, the microgametes penetrate the macrogametes generating zygotes (9) which become motile and elongated (ookinetes) (10), capable of invading the mosquito midgut wall. Here they develop into oocysts (11). The oocysts grow, rupture, and release sporozoites (12), which make their way to the salivary glands in order to be inoculated into a new human host during a subsequent blood meal (1).

From the Centers for Disease Control, used under Wikimedia Commons.

### 1.2.2 Erythrocyte invasion

Extra-cellular merozoites rapidly collide with new, uninfected erythrocytes, adhere using specific ligands, and invade to perpetuate the infection. Merozoite invasion of erythrocytes takes place in four distinct phases: (1) adhesion, (2) parasite reorientation, (3) formation of a tight junction complex and (4) ingress[4]. Parasite adhesion may occur through multiple redundant pathways. *P. vivax* depends almost completely on parasite Duffy binding protein (PvDBP) engaging the Duffy blood group antigen (Fy, coded for by the *DARC* gene) on the surface of human erythrocytes[27-29]. In *P. falciparum*, most attention has focused on the Erythrocyte Binding-Like (EBA) protein family (EBA-175, EBA-140/BAEBL, EBA-181/JESEBL, EBL-1) and the Reticulocyte binding Homolog (Rh) family (PfRh1, PfRh2a, PfRh2b, PfRh3, PfRh4, PfRh5). Ligands belonging to the EBA family govern sialic acid dependent invasion pathways, while the Rh family ligands are sialic acid independent[30]. Erythrocyte receptors for these EBL proteins have been identified as glycophorins A, B and C (for EBA-175, EBL-1 and EBA-140, respectively)[31-34]. The receptors for Rh family ligands are less thoroughly characterized, though complement receptor 1 (*CD35*) has been confirmed as the receptor for PfRh4 and the immunoglobulin basigin (*BSG*) was recently identified as the receptor for PfRh5[35, 36]. Interestingly, while the other interactions appear to be redundant, and disruptions to any or all of the other ligands do not substantially impair parasite invasion, the PfRh5-basigin binding is essential, in *ex vivo* culture at least, making it a promising candidate for intervention in the form of pharmacologically mediated invasion blocking[36].

Once an appropriate parasite ligand-host receptor interaction is established, the invading merozoite reorients itself such that its apical complex (i.e. the pointed end) is in direct contact with the erythrocyte plasma membrane. When this is complete, adhesion molecules facilitate the formation of a tight junction and the contents of the parasite rhoptry – a specialized secretory organelle – are released into the erythrocyte to aid in creating the parasitophorous vacuole[37]. As the merozoite invades, it grasps and drags the erythrocyte plasma membrane over its surface until it is completely enveloped. Migration of the tight junction complex pulling the erythrocyte membrane from the "head" of the parasite to the rear is an ATP-dependent process, relying on an active myosin motor complex[4]. When the ring-like tight junction reaches the far end of the merozoite, it closes up and seals the parasite entirely within the newly formed vacuole. Cloaked within the intact erythrocyte, the merozoite is hidden from the host immune system and able to replicate, consuming up to 80% of the host cell hemoglobin as well as other small molecules for nutrition[38].

### 1.2.3 Gametocytogenesis: the sexual stage of the *Plasmodium* life cycle

During each erythrocytic cycle, a portion of merozoites develop into "male" (micro) and "female" (macro) gametocytes which allow the parasite to be transmitted back to the mosquito vector and complete the sexual portion of its life cycle. New gametocytes will be generated every round of erythrocyte replication, but only make up a small portion of the total parasite burden with asexual forms outnumbering sexual in a ratio ranging from 10:1 to 156:1[39]. The "decision" to enter gametocytogenesis is not well understood, though it has been linked to external environmental variation, host cell pressure and parasite specific

programmes. Changes in temperature, pH, exposure to anti-malarial drugs and increasing anemia (erythropoietin, presence of reticulocytes) have all been shown to increase gametocyte number, consistent with the hypothesis that gametocytogenesis can be amplified by a worsening blood environment for the parasite[40-42]. Many of these strong cues are also correlated with malarial disease symptoms such as fever.

*P. falciparum* gametocytes typically appear in circulation 7 to 15 days after the first round of asexual division, and mature gametocytes can remain in the bloodstream for an average of 6.4 days[39]. Commitment to gametocyte maturation is determined soon after erythrocyte derived merozoites invade a new cell, with each trophozoite developing into either asexual merozoites, male gametocytes or female gametocytes, but never a combination of the three. Female gametocytes (or “macrogametocytes”) are more prevalent in circulation by approximately four-fold compared to male gametocytes (“microgametocytes”).

When a mosquito takes a blood meal and ingests gametocytes, they leave the erythrocyte membrane shell and finalize their differentiation into gametes. This phase is triggered by xanthurenic acid in mosquito saliva, a drop in temperature as the parasite leaves mammalian circulation, and a rise in pH as the ingested parasites travel into the mosquito stomach[43]. Female gametocytes become larger and more spherical, while male gametocytes divide to form eight flagellated microgametes containing little more than the nucleus and flagellum. Macrogametes and microgametes fuse within the mosquito midgut to form a zygote, then a motile ookinete[44]. The ookinetes invade mosquito midgut epithelial cells to form oocysts in the intercellular space between the epithelial layer and basal lamina. Over a period of one to three weeks, oocysts undergo multiple rounds of nuclear division, maturing into clusters that release thousands of haploid sporozoites into the circulating mosquito haemolymph. In

the haemolymph, sporozoites make their way to the mosquito salivary glands where they remain until injected into their next human host during a subsequent blood meal.

## 1.3 Human Malarial Disease

Symptomatic malarial disease in humans occurs during the blood stage of the parasite life cycle when merozoites are found within erythrocytes; approximately 8 – 25 days post initial infection. The first signs and symptoms resemble those of flu or multiple other viral and bacterial infections, with fever, chills, headache, body aches and loss of appetite. This may also be accompanied by abdominal pain, nausea, jaundice and hemoglobinurea[20]. The appearance of these symptoms along with confirmation of malaria parasites in circulation by either blood smear microscopy or antigen-based rapid diagnostic testing is sufficient to diagnose uncomplicated malaria, which accounts for the majority of infections – especially in semi-immune adults.

For reasons that are still poorly understood, a fraction of patients go on to develop severe manifestations of disease, which can be broadly subdivided into severe malarial anemia or cerebral malaria. Prognosis for these patients is less certain and these pathologies account for the majority of fatal outcomes. Patient age, generalized health status including nutrition, prior exposure to malaria parasites and co-infection, as well as a variety of genetic factors have been established as important variables or drivers in this phenomenon though many unknowns still remain.

### 1.3.1 Pathogenesis of severe malarial anemia

Severe malarial anemia is the primary clinical manifestation of severe disease following infection with *P. falciparum* or *P. vivax*[45]. Anemia is due to a huge increase in erythrocyte clearance along with an inadequate erythropoetic response by the bone marrow. While as a syndrome, severe malarial anemia is defined specifically by hemoglobin levels below 5g/dL (normal hemoglobin in developing countries is >11g/dL), patients may also suffer respiratory distress, hypoglycemia, kidney failure, metabolic acidosis and extremely high fever. The precise sequelae leading to death are not always consistent. Pro- and anti-inflammatory innate immune processes are dysregulated, which additionally contributes to disease pathophysiology and medical treatment is required. Studies in endemic areas of Africa have demonstrated that the annual rate of hospital presentation for severe malaria was 7.6 admissions per 1000 children (ages 0 – 4), with a case fatality rate of 9.7%[46].

Parasite driven hemolysis during merozite egress from erythrocytes and phagocytosis of infected cells are obvious contributors to low hemoglobin levels, but these are only a part of the story. In fact, risk of severe malarial anemia is not necessarily directly correlated with high parasite burden[47]. It has been estimated that for every infected erythrocyte that is eliminated, eight to ten uninfected cells are also removed from circulation during the acute phase of a severe malaria infection which contributes significantly towards the anemia[48]. One of the pleiotropic effects of malarial infection is increased rigidity within the entire red blood cell population, possibly as a response to free serum hemin (a byproduct of hemoglobin digestion) and hydrogen peroxide (produced by activated phagocytes)[49]. These less flexible cells are more likely to obstruct blood flow through microcirculation and will be cleared when passing through the

spleen. Shear stresses can damage membranes, additionally shortening cellular lifespan[50]. Furthermore, receptor-mediated clearance is enhanced by deposition of immunoglobulin and complement on the cell surfaces of both infected and uninfected cells[48]. Red cell surface IgG is elevated in children suffering severe versus uncomplicated malaria, which leads these cells to be more susceptible to phagocytosis as can be demonstrated *in vitro*. The work of Layez and colleagues has pointed towards parasite rhoptry ring surface protein (RSP2) as a contributor to this effect, as it was shown to 'tag' the surface of uninfected erythrocytes and erythrocyte precursors during failed invasion attempts, increasing their antigenicity and likelihood of phagocytosis[47, 51]. Declining levels of complement receptor 1 (CR1, aka CD35) have been associated with severe malaria, consistent with decreased ability to inactivate the C3b membrane attack complex[48].

Ineffective compensatory erythropoiesis is the other half of severe malarial anemia pathology. Examination of bone marrow in children suffering from severe anemia reveals damaged bone parenchyma, hypercellularity and abnormal features in erythroid progenitors. Production and maturation of new reticulocytes is impaired compared to the uninfected state[46, 52]. Erythropoetic suppression can be linked to an imbalance of inflammatory mediators, particularly tumor necrosis factor alpha (TNF-$\alpha$), interferon gamma (IFN-$\gamma$), nitric oxide (NO) and interleukin 23 (IL-23) which are upregulated in response to parasitemia and also directly cause dyserythropoesis[46]. Other cytokines and chemokines such as chemokine (C-C motif) ligand 5 (CCL5) and transforming growth factor beta (TGF-$\beta$) are able to promote erythrocyte proliferation, but it has been shown that their expression can be suppressed by leukocyte phagocytosis of parasite hemozoin[46].

### 1.3.2 Pathogenesis of cerebral malaria

Cerebral malaria is the most grievous potential complication of *Plasmodium* infection. It is clinically defined as any malaria parasite positive infection that includes coma, but may also be accompanied by retinopathy, seizures, altered sensorium, and other organ involvement[20, 53]. While parasites do not directly invade brain parenchyma during cerebral malaria, they are capable of setting in motion a cascade of parasite and host driven processes including sequestration and adhesion of parasites and host cells within brain and other organ microvasculature, dysregulated cytokines and mediators, and altered permeability of the blood brain barrier (Figure 3)[54].

When it comes to cerebral malaria, the host immune response to infection is of critical importance. Cerebral malaria is a complex pathology that involves mechanical blockage of microvasculature by adherent cells, pro-inflammatory dysregulation of the innate immune system, and endothelial activation – though the details of pathogenesis are incompletely understood and remain debated[55]. While there have been scattered reports of cerebral malaria caused by *P. vivax*, this pathology is considered overwhelmingly specific to *P. falciparum* infection, and is most common in children living in low transmission areas. Non-immune adults may also be affected during their first exposure[45]. One of the major features of cerebral disease is increased cytoadherence of infected and uninfected cells, primarily erthyroid, to endothelial microvasculature, typically in small venules and capillaries such as those in the brain where the blood flow rate is less forceful[56]. For infected cells, this allows them to avoid passage through the spleen where they risk immune clearance.

In order for parasitized erythrocytes to adhere to one another and endothelial vasculature, *P. falciparum* expresses variant erythrocyte membrane

protein 1 (PfEMP1) at the erythrocyte cell surface membrane in knob structures. PfEMP1 antigens are encoded by the *var* multigene family and are extremely polymorphic. In addition to their innate genetic diversity, parasites are able to switch PfEMP1 classes during an infection to circumvent antibody-mediated immune responses[56, 57]. PfEMP1 molecules are capable of binding to multiple receptors on human endothelial cells, notably CD36 and intracellular adhesion molecule 1 (ICAM1)[58, 59]. ICAM1 receptors are found on leukocytes and endothelial tissue, while CD36 is highly expressed everywhere but the brain. Cerebral malaria has been specifically associated with PfEMP1–ICAM1 binding and researchers have hypothesized that CD36 may act as a decoy receptor to sequester any parasitized erythrocytes in less delicate tissues[59]. Adherent parasitized erythrocytes within brain microvasculature will begin to clog capillaries and may encourage additional rossetting or clumping as passing uninfected cells can bind to the adherent cells via PfEMP interactions with CR1, blood cell antigens and glycosaminoglycans. Platelets, monocytes, and other leukocytes can bind to activated epithelium or adherent infected erythrocytes as well, exacerbating vessel occlusion. Blocked blood flow leads to ischemia, tissue hypoxia and cellular stress which can amplify inflammation and, in the brain, cause coma[60].

Once bound to brain microvasculature, cytoadherent cells initiate a series of endothelial activation signals that can lead to a breakdown of the blood-brain barrier tight junctions. Signs of this can be observed by clinicians examining patients for retinopathy, and leaks have been observed deep within patient brains using fluorescein dye angiography[60]. $CD8^+$ T-cell recruitment leading to high local levels of inflammatory cytokines and effector molecules have been implicated in this process[55]. Attempts to correlate serum cytokines with disease severity have yielded conflicting results and little prognostic value, but autopsy

samples, genetic studies, animal models and functional biochemistry have repeatedly pointed towards leukocyte recruitment and dysregulation of inflammation for roles in triggering or exacerbating cerebral disease. Markers of endothelial activation (soluble ICAM-1, van Willebrand factor, angiopoietin 1) are elevated in cerebral versus uncomplicated malaria, lending support to the argument that abnormal vascular endothelium is a hallmark of pathology[61].

A single, universal tipping point event leading to cerebral pathology seems unlikely. The current simplified model (Figure 3) includes contributions from both host and parasite factors, often with multiple possible avenues towards the same pathogenic outcome. However, many gaps still remain to be closed and pathways delineated in order to better understand the pathology and target interventions.

**Figure 1-3.** *(see next page)* A simplified model of cerebral malaria pathogenesis

(a) Sporozoites migrate from a site of mosquito inoculation to the liver, where they multiply for 10–14 days before (b) thousands of merozoites erupt and invade red blood cells (iRBC). (c) Intra-erythrocytic merozoites and developing ring stage parasites circulate in blood and pass through the spleen. Circulating iRBCs and debris from hemolytic burst cycles will activate innate immune responses, including dendritic cells (DCs) and splenic macrophages, which will clear parasites (as well as uninfected RBCs) via phagocytosis. These cells will also release inflammatory cytokines and mediators involved in enhancing the immune response and direct killing of parasites (TNF-$\alpha$, IFNs NO). (d) iRBCs that evade direct killing or phagocytosis are capable of activating vascular endothelium, upregulating adhesion molecules such as ICAM-1 and CD36. iRBC will bind to these receptors and sequester in microvascular endothelial beds. (e) Activated DCs will present parasite antigens to T-cells, activating them in turn and amplifying both cytokine and leukocyte production. (f) Adherent iRBCs bind additional infected and uninfected RBCs as well as platelets which obstructs blood flow through small vessels. (g) Presence of parasite antigens, endothelial activation and cellular hypoxic stress may recruit leukocytes to sites of iRBC sequestration. (h) iRBCs may also adhere to ICAM-1 expressed on brain microvasculature, activating endothelium, which compromises the integrity of the blood brain barrier. Recruitment and sequestration of cytokine-producing leukocytes can lead to immunopathological tissue damage as well as additional vessel blockages. Activated cerebral endothelium will be inundated with dysregulated cytokines, upregulated coagulation, additional activating molecules leading to focal disruption via apoptosis or direct damage.

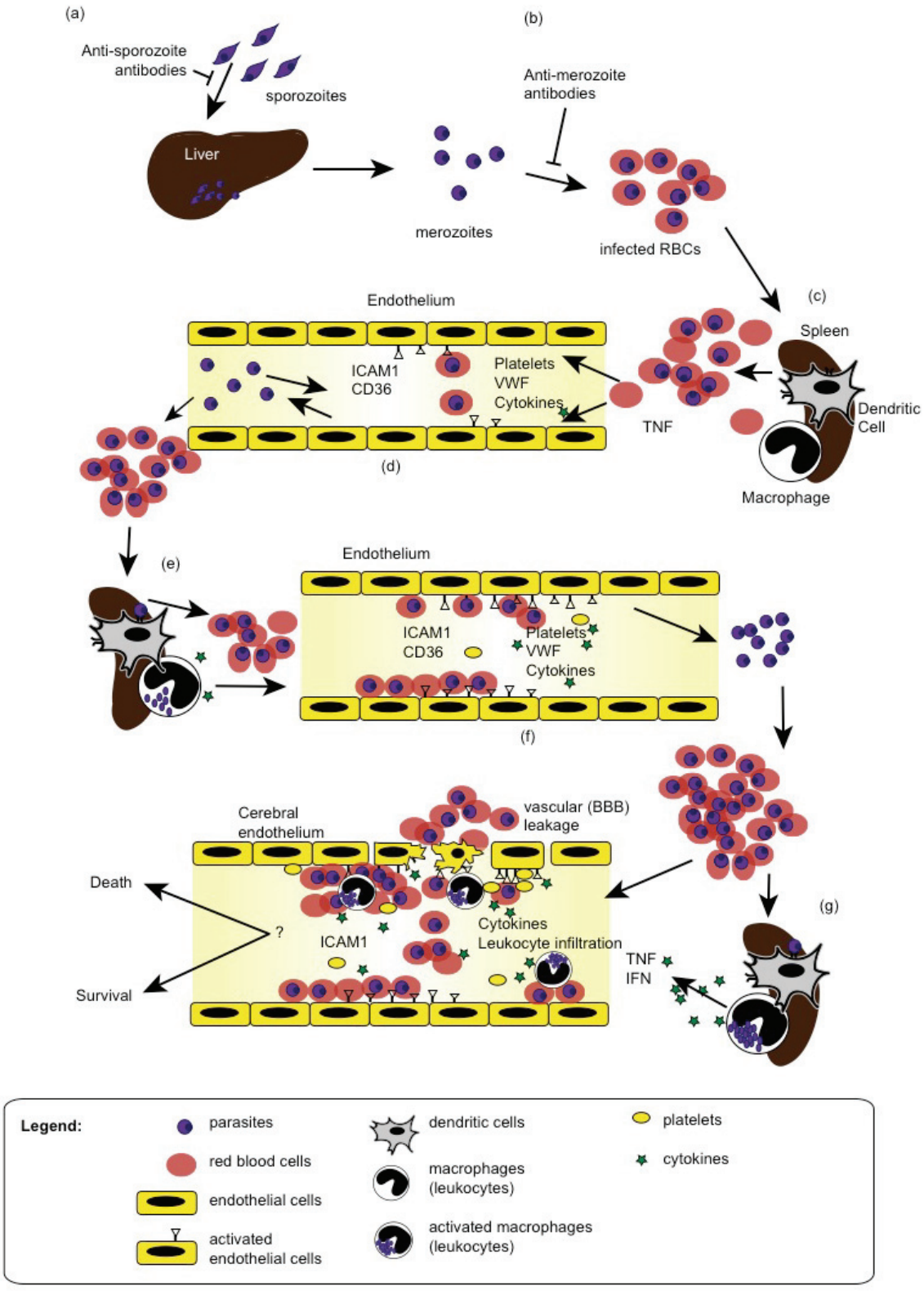



Figure 1-3. A simplified model of cerebral malaria pathogenesis

## 1.4 Human Genetic Factors Influence Malarial Disease

As was alluded to above, the host response to malaria plays a major role in determining disease susceptibility, progression and outcome. Malaria is one of the best known examples of an infectious agent placing significant selective pressure on the human genome, shaping multiple aspects of cell biology and immune response as both human and parasite co-evolve to genetically outwit one another. J.S. Haldane first proposed the hypothesis that immunity to infectious disease has been a driving force in evolution, refining his hypothesis in 1949 with the statement “the struggle against disease, and particularly infectious diseases, has been a very important evolutionary agent”[63].

*Plasmodium* parasites have coexisted with humans for millions of years, becoming particularly important following the widespread development of agriculture approximately 10,000 years ago. Stable community living, irrigation and proximity to standing water used by breeding mosquitoes led to increased exposure, and increased child mortality in particular strengthened the evolutionary impact of infection[64, 65]. *P. falciparum* infection has been detected in the mummified remains of King Tutankhamun (circa 1550 – 1479 BC), possibly contributing to his death; and periodic fevers with concomitant splenomegaly have been described since time immemorial, appearing in ancient Chinese, Sumerian and Indian writings as well as making appearances in myth and legend[66, 67]. This long-standing relationship has left indelible signatures of selection across the entire human genome, particularly in genes involved in the immune response or relating to erythrocyte structure and function.

### 1.4.1 The case for sickle cell hemoglobin (HbS)

The most striking and well-supported example of malaria shaping the human genome is that of sickle cell hemoglobin (HbS). The sickle cell variant allele of beta-hemoglobin (*HBB*) involves a single point mutation at the sixth amino acid, substituting a valine for glutamic acid (E6V, rs334). HbS is soluble in the oxygenated state, but in the deoxygenated state is prone to polymerize into long fibrous aggregates as a result of the hydrophobic interaction between valines[68]. Erythrocytes containing these polymerized hemoglobin chains become deformed and rigid. Loss of elasticity causes problems when these cells attempt to travel through narrow capillaries and they may become stuck, causing painful vaso-occlusive crises, reduced blood flow to organs and increased risk of stroke and heart disease[69]. Periodic deformation and cell membrane damage lead to a shortened erythrocyte lifespan, and HbS carriers may be chronically anemic with varying degrees of severity. People homozygous for HbS suffer from sickle cell anemia with the majority of patients dying under the age of five, or requiring splenectomy and regular medical treatments including blood transfusion[70]. Considering the pathological nature of this condition, the high frequency of HbS – exceeding 15% allele frequency in some populations – would then be very surprising in the absence of some form of balancing positive selective pressure[70].

The first observation of a possible malaria protective role of HbS came from Anthony Allison in 1949 who observed a much higher frequency of sickle cell heterozygote carriers (near 20%) in coastal areas of Kenya where malaria is endemic, compared to low frequencies in the highlands where malaria is not transmitted[71]. Following up on that hypothesis, he found that children carrying

an HbS allele had lower parasite densities and reduced morbidity compared to children with only the wild type HbA (hemoglobin: adult) allele. Further investigation revealed that the geographic distribution of HbS mirrors that of *P. falciparum* endemicity extremely well with the highest allele frequency in sub-Saharan African populations and significant polymorphic frequency throughout Africa, the Middle East and into Southeast Asia[68, 70]. The E6V mutation appears on four independent haplotypes (Senegal, Benin, Central African and Indian/Arabic types) which further supports functionality and positive selection based on this specific allele[71]. Retention of the HbS allele within these populations is an example of a balanced polymorphism with heterozygote advantage. The HbAS carrier state is protective against all major forms of severe *P. falciparum* malaria[72, 73]. HbAS is significantly associated with fewer clinical malarial episodes in general (Odds ratio range = 0.08 to 0.27 vs. HbAA, $p<0.001$), reduction in severe malarial anemia episodes (OR range = 0.09 to 0.49, $p<0.001$), reduced parasite density during infection (OR = 0.49 to 0.73, $p<0.0001$), and fewer instances of cerebral malaria (OR = 0.07 to 0.12, $p<0.001$)[71, 73-78]. Protection from these severe symptoms correlates with increased survival, especially in pediatric cohorts. While estimates for the protective effect range based on ethnic background, age at infection, HbS allele frequency within the study population (statistical power), clinical definitions and study design, the general trend of HbS protection against clinical malaria is extremely reproducible. Meta-analysis of forty-four independent case-control studies calculated a 10-fold decreased risk of severe malaria in HbAS carriers following *P. falciparum* infection at (OR = 0.09, 95% Confidence Interval, range 0.06 – 0.12) and an approximately 2-fold decreased risk of uncomplicated clinical malaria in prospective cohort studies (Incidence Rate Ratio = 0.69, 95% CI, range 0.61-79)[78].

The mechanism for the protective effect of HbS against malaria has been extensively studied though the exact nature of protection remains unclear and is likely multifactorial. Functional studies have demonstrated decreased intra-erythrocytic parasite growth in HbAS erythrocytes, decreased invasion, as well as increased sickling and clearance of infected erythrocytes[79]. Other studies have focused on the specific protective effect of HbAS against the development of severe symptoms including cerebral malaria, and have identified impaired cytoadherence of HbAS blood cells leading to reduced endothelial activation and microvasculature blockage as a result of abnormal PfEMP1 (*Plasmodium falciparum* erythrocyte membrane protein)-mediated knob display and reduced rosetting of infected erythroctyes[80]. Cohort studies in Uganda have also shown that children carrying HbAS have improved acquired immunity to second or third malarial infections compared to HbAA genotypes. This results in a lower force of infection and fewer malarial episodes as the child matures[79]. Increased circulation, rather than sequestration, of PfEMP1 expressing infected cells may also lead to increased leukocyte exposure and production of anti-PfEMP1 antibodies. This boosts acquired immune responses, allowing patients to clear infected cells more quickly and improving natural resistance to infection[81].

### 1.4.2 Malaria protective effect of other hemoglobinopathies

Sickle cell hemoglobin is not the only hemoglobin abnormality that has been linked to protection against malaria. Two other point mutations in *HBB* termed HbC (E6K) and HbE (E26K) have also been identified in African and Asian populations with a distribution and frequency suggestive of positive selection due to malarial pressure[82]. HbE, which is extremely frequent in Southeast Asia, has been associated with reduced prevalence of malaria and

lower odds of severe symptoms[83]. *In vitro*, heterozygote HbAE erythrocytes restrict *P. falciparum* invasion to approximately 25% that of HbAA cells, which may prevent high parasite burdens[82, 84]. Evidence for HbC has been contradictory, though studies have demonstrated a malaria protective effect in certain populations[85]. For HbC, the effect is generally associated with allele homozygosity (HbCC), and specific protection against severe manifestations of disease, with little to no effect in the heterozygote state[78].

Normal adult hemoglobin exists as a tetramer of two $\alpha$-globin and two $\beta$-globin chains ($\alpha_2\beta_2$) complexed to four $Fe^{+2}$ containing heme groups. Alpha globin chains are encoded by two adjacent genes on telomeric chromosome 16 (*HBA1* and *HBA2*), while $\beta$ chains are encoded by *HBB* on chromosome 11. Proteins produced by *HBA1* and *HBA2* are identical with a few functionally irrelevant polymorphisms in the 3' and 5'-untranslated regions of the mRNA[86]. Impaired protein production by disruption of *HBA1*/*HBA2* or *HBB* leads to $\alpha$-thalassaemia or $\beta$-thalassaemia, respectively. In the case of the HBA cluster, loss of one $\alpha$ allele with three normal alleles remaining is asymptomatic, loss of two alleles causes microcytic anemia (abnormally small erythrocytes), and loss of three alleles results in severe underproduction of the $\alpha$ chain, severe microcytic anemia, and the aggregation of tetrameric $\beta$ chains ($\beta_4$ also termed HbH or Heinz bodies). Loss of all four alleles is perinatally lethal[87, 88]. $\alpha$-Thalassemia is genetically heterogeneous and may be the result of deletions in either *HBA1* or *HBA2* ($\alpha^0$) and/or polymorphisms reducing, but not eliminating protein expression ($\alpha^+$)[88]. The highest frequency of $\alpha$-thalassemia is in Southeast Asia where carrier frequency for $\alpha^0$ can reach 10% and $\alpha^+$ carrier frequency ranges from 2-70%, though it is also common throughout Africa, the Middle East, and Mediterranean regions[88].

β-thalassemia can be found in about 1.5% of the world's population with the highest frequencies in Mediterranean populations (10.8%), and throughout the Middle East, central Asia, and China[89, 90]. Over 200 different $\beta^0$ or $\beta^+$ mutations have been identified (including population specific alleles), most of which are point mutations or deletions within the promoter region or splice sites. Certain structural mutations (including HbE, $\beta^E$) also affect the amount of β-chain protein that is produced and can produce a thalassemic effect[87, 90].

A modest protective effect against clinical malaria incidence and severe malarial anemia has been described for patients with α-thalassemia. Several reports have shown that thalassemic patients are significantly less likely to require hospitalization following malarial infection[91, 92]. Other case-control and prospective studies have yielded inconsistent results however, and clinical heterogeneity, small effect sizes and population-specific variables continue to be a challenge for association studies[78, 92-94]. Aside from $\beta^E$, β-thalassemia has not been subject to malaria association studies in the field. In culture, parasites replicated more slowly in thalassemic erythrocytes, and infected erythrocytes were more readily phagocytosed than normal cells[87].

Although not all of these variants can be conclusively linked to malaria resistance at this point, there is no doubt that selective pressure exerted by *Plasmodium* parasites has had a substantial impact on hemoglobin allelic distribution or that variation within hemoglobin genes impacts on disease progression and outcome.

### 1.4.3 Erythrocyte membrane and structural variants

Erythrocyte disorders are not limited to hemoglobinopathies, and multiple other genes and proteins expressed within or on the surface of red cells have been linked to malaria resistance.

An expression silencing polymorphism in the promoter of glycophorin D, also called Duffy antigen (DARC/Fy), leading to a Duffy-negative blood type ($FY*B^{ES}$) has been implicated in *P. vivax*'s exclusion from large regions of West and Central Africa where the mutation is near fixation[71, 95]. Duffy-negative homozygotes are completely refractory to *P. vivax* parasite infection, as merozoites require cell surface expression of this antigen in order to attach and invade. Widespread distribution of Duffy-negative genotypes and the specificity of this interaction suggests that malaria pressure has been the agent of selection for this polymorphism[96, 97]. Mutations in other glycophorins have also been reported at high frequencies in malaria endemic populations such as the Gerbich glycophorin C variant (GYPC, ex3del) which has reached 47% allele frequencies in Papua New Guinea[87].

The most prevalent protein on erythrocyte membranes is Band 3 (solute carrier family 4, anion exchanger, member 1; SLC4A1), an integral glycoprotein involved in inorganic anion exchange and cell structure. Heterozygous deletion in the N-terminus of the protein (*SLC4A1Δ27*) causes a syndrome called southeast Asian ovalocytosis (SAO) which is strongly protective against both *P. vivax* and *P. falciparum*, especially with respect to cerebral disease[93, 98]. SLC4A1Δ27 proteins bind more tightly to ankyrin, another scaffolding protein within erythrocytes. The cells are unusually shaped (elliptical), more rigid, ATP-depleted, and express fewer antigens at the outer membrane. The majority of cultured parasite lines are not able to invade SAO cells efficiently *in vitro*, with a

handful of exceptions[98]. Specific protection of *SLC4A1Δ27* against cerebral disease (OR = 0 – 0.85) may also be due to increased adhesion of infected SAO cells to CD36 and decreased adherence to ICAM-1. Since CD36 is ubiquitously expressed on vascular endothelial cells except for those in the brain (which express primarily ICAM-1), this decoy binding could exclude cerebral pathogenesis[99, 100]. ATP-depletion is also thought to contribute to a hostile replicative environment for intracellular parasites[101].

One of the more recent erythrocyte disorders that has been linked to malaria is pyruvate kinase (liver and red cell isoform, *PKLR*) deficiency. Pyruvate kinase (PK) catalyses the rate limiting step in glycolysis (the transfer of a phosphate group to ADP, generating ATP), which is the sole means for mature erythrocytes to generate cellular energy as they lack mitochondria. Unlike many of the other observations which grew out of a tantalizing geographical overlap between mutation frequency and malaria endemicity, evidence for *PKLR* mutations in malaria resistance came from laboratory experiments using animal models (see section 1.5)[102]. Briefly, different inbred mouse strains exhibit differential susceptibility to experimental infection with Plasmodial parasites. During a complex backcross breeding scheme to generate a panel of recombinant inbred mice from the A/J and C57BL/6 parental strains, a spontaneous mutation arose in *Pklr* which was then fixed and passed down to a subset of daughter lines. Profiling the entire panel of recombinant inbred lines for peak parasitemia as a measure of susceptibility to blood stage malaria following infection with *P. chabaudi* AS revealed two strains (AcB55 and AcB61) with outlier resistance phenotypes, despite a majority susceptible genetic background. The genetic effect responsible for this resistance was mapped to an I90N substitution in *Pklr*. A second, more severe $Pklr^{G338D}$ mutation in an independent mouse strain further confirmed the malaria-protective effect of PK-deficiency in

mice and suggested a correlation between mutation severity and resistance, possibly due to more rapid erythrocyte turnover[103].

Translating these findings to human studies demonstrated that human PKLR mutant erythrocytes were resistant to *P. falciparum* infection in culture via the dual mechanisms of reduced invasion and increased clearance by macrophages[104]. Reducing cellular ATP in wild type and *PKLR*$^{-/-}$ erythrocytes was able to reproduce a similar phenotype indicating that ATP and cellular energetic mechanisms have a powerful impact on intracellular parasite success and therefore, disease progression[101]. Pyruvate kinase deficiency is the second most common erythrocyte enzyme disorder globally, and it is possible that malaria has allowed the retention of certain PKLR mutations within malaria-endemic populations. Genetic and disease surveillance leaves much to be desired in many of these areas of the world, however, and the geographic-specific prevalence of pyruvate kinase deficiency is not known. Chapter 4 seeks to investigate this possibility in more detail[105].

### 1.4.4 Non-erythrocyte genes associated with malarial resistance

While the erythrocyte is home to parasites during blood stage infection and therefore erythrocyte polymorphisms will logically have a substantial impact on disease; adhesion and immune polymorphisms have also been shown to have major consequences for disease progression. Many of these molecules have been linked to the difference between uncomplicated malaria versus severe pathology, rather than infection per se. The genes of the major histocompatibility complex (MHC), including human leukocyte antigens (HLA) are key determinants of the host response to multiple infections, other stimuli and autoimmunity[106]. The hallmark of this gene family is its extraordinary diversity and variation which

is suggestive of balancing selection favoring heterozygosity, allows a wide range of possible responses to a broad spectrum of threats, and is instrumental for cells to differentiate “self” from “non-self”[106]. HLA class I antigens are found on all nucleated cells and present peptides to cytotoxic T-cells, while HLA class II molecules are only found on lymphocytes and specific antigen presenting cells which are involved in antibody-mediated immunity. A large-scale case-control study in West Africa found that the resistance allele HLA-Bw53 (class I) and the DRB1*1302-DQB1*0501 haplotype (class II) accounted for a significant proportion of the variance in clinical cases, severe malarial anemia and cerebral malaria (relative risk for severe malaria = 0.59 for Bw53 and 0.49-0.52 for DRB1*1302-DQB1*501)[72].

Association and case control studies have also indicated a differential response to infection for polymorphisms in *TNF*, *IFNG*, *ICAM1*, *NOS2A*, *IL1*, *IL12*, *LTA*, *FCGR2A* and *IRF1*, though for the most part, these have been difficult to replicate across populations[82, 87, 96]. TNF has been extensively investigated following the results of a case control study which reported a seven fold increase in relative risk for developing cerebral malaria in children homozygous for a polymorphism at -308G/A in the promoter, which increases TNF-$\alpha$ transcription following *in vitro* stimulation[107, 108]. As this SNP is associated with more severe outcomes, this finding offers valuable information for understanding malaria pathogenesis, but it is unlikely that malaria pressure was the selective force driving its retention within African populations.

Endothelial adhesion and activation molecules have been extensively studied for involvement in cerebral malaria. Circulating or cell-expressed ICAM1, vascular cell adhesion endothelial molecule 1 (VCAM1), platelet endothelial cell adhesion molecule 1 (PECAM1), von Willebrand factor (VWF), angiopoietin 2 (ANG2), p-selectin (SELP) and e-selectin (SELE) are all upregulated during

malaria infections and are presumed to facilitate the pathological processes of pRBC/RBC/leukocyte/platelet sequestration and increased vascular (including blood brain barrier) permeability[109-112]. Parasitized erythrocytes have been demonstrated to activate endothelium via direct physical contact of PfEMP1 to host cell receptors and co-localization of sequestered parasites with endothelium expressing high levels of ICAM1 in particular has been demonstrated in patients exhibiting cerebral malaria[110, 113, 114]. The most convincing evidence for the importance of ICAM1 in cerebral pathology comes from autopsy studies and *in vitro* stimulation assays as circulating soluble ICAM1, though typically elevated, has not been consistent enough to function as a biomarker for cerebral disease specifically – and may not be a good surrogate for local conditions within the brain microvasculature. Researchers examining the genetics of *ICAM1* in human populations found the counterintuitive association of a derived African allele (*ICAM1*$^{Kilifi}$, L29M, local population frequency >30%) with increased susceptibility to cerebral malaria versus uncomplicated or severe, non-cerebral malaria in a cohort of Kenyan children (OR=2.23)[115]. As noted by the authors, while this finding strengthens the evidence linkage between ICAM1 and cerebral malaria, some non-malarial counterbalancing selection force must be responsible for its retention within the population, similar to the Gambian *TNF* promoter SNP described by McGuire and colleagues[107]. Follow up studies have suggested a reduction in tissue damage and death due to nonmalarial fever and sepsis in people with ICAM1$^{Kilifi}$ [116]. This failure to find a reproducible cerebral-malaria specific DNA signature in immune genes beyond HLA may be due to the diversity and burden of infection in endemic areas, as many candidates are demonstrably responsive to infection.

### 1.4.5 Impact of host genetic factors on gametocytogenesis

Human host genetic background has been shown to play a role in provoking sexual stage development of *P. falciparum* parasites during infection, which can have significant consequences in terms of disease transmission rates. Percentage of circulating gametocyte as a within a host's bloodstream are partially heritable, however, the specific genetic factors responsible for this trait have not been identified[40]. In a study demonstrating this, blood smears from Senegalese families living in areas of high malaria transmission intensity were microscopically inspected for the presence of gametocytes throughout several seasons regardless of clinical or symptomatic infection status. Heritability of gametocyte-positive smears during asymptomatic infections was estimated between 16% (villages of Dielmo and Ndiop) and 50% (village of Gouye Kouly) while heritability of gametocyte carriage during symptomatic infections could not be demonstrated. This is not necessarily to say that genetics do not influence gametocytogenesis during symptomatic cases however, as administration of antimalarial drugs will also alter parasite life cycles and are a known confound. Other studies have shown differences in gametocyte carriage between ethnic groups (Mossi versus Fulani) within the same environment in Burkina Faso, further supporting the idea that human genetics can influence parasite life cycle[117]. Erythrocyte or immune polymorphisms (including acquired immunity) are the strongest logical candidates for this effect, though sickle cell hemoglobin (HbS) has only a modest impact and the remainder of the variance is yet to be determined[40].

### 1.3.5 Genome Wide Association Studies

Most studies investigating the genetic factors influencing malaria susceptibility in humans have focused on candidate genes, however advances in genetic tools have allowed researchers the ability to identify novel associations in a systematic fashion. Genome wide association studies (GWAS) profile hundreds of thousands to millions of SNPs throughout the entire human genome and use unbiased statistics to correlate chromosomal regions with disease phenotypes, similar to the idea of performing a case-control association study with every gene. To date, one major GWAS has been done in Africa useing 958 severe malaria cases and 1382 controls in Gambia, genotyping them across ~500,000 SNPs. An independent replication study within the same report added 1087 more cases and 2376 controls[74]. Unsurprisingly, this analysis identified HbS as the most significant genetic determinant in malaria disease resistance ($p = 4.5 \times 10^{-14}$ using an imputed association signal, $p = 1.3 \times 10^{-28}$ using genotypic validation at rs334 itself)[74]. What was surprising to many is that this was the only gene/locus that met or exceeded the significance threshold across the entire genome. Nineteen additional chromosomal regions showed suggestive linkage ($p < 10^{-5}$), with the strongest non-*HBB* effects observed near SCO cytochrome oxidase deficient homolog 1 (*SCO1*) and in dopa decarboxylase (*DDC*) genes. Both of these are novel and require further validation before strong conclusions can be drawn. It is also possible that these indicate regulatory regions or have long-range effects on more distant genes. An expression SNP (eSNP) in *SCO1* was also detected at genome wide significance ($p = 8.91 \times 10^{-14}$; rs201621) in separate study that measured transcript expression levels in a genotype-by-infection interaction though this did not address any mechanistic questions

either[118]. Biochemical, mechanistic or case-control confirmation of an effect still remains to be seen.

Further within this first GWAS, additional attention was paid to a subset of genes (*CD36*, *CD40LG*, *CR1*, *ICAM1*, *IL22*, *NOS2*, *TNF*) that had been previously associated with malaria resistance. None of these candidates reached genome-wide significance within the populations tested here at the power levels of this study, though a suggestive association was detected for TNF ($p = 0.02$). No other SNPs (novel or candidate) reached significance in this study, possibly a result of inadequate genome coverage and haplotype imputation, population structure, and low linkage disequilibrium all of which are known issues when doing genetic analyses within African ethnic populations. Additionally, categorizing individuals within cases may be clinically variable and dependent on non-genetic factors such as exposure and patient history. Nonetheless, this report attempts a new approach and presents two new candidate loci. It also highlights the complex and heterogeneous nature of malaria resistance and lends support to the hypothesis that multiple genetic loci with small or moderate effects explain the majority of phenotypic variance.

## 1.5 Mouse models of infection

Genetic association studies involving disease responsiveness in humans, while extremely valuable, have a number of limitations including ethical considerations, treatment requirements, variable environmental factors, population structure and differing genetic architecture, parasite heterogeneity, and the inability to control for numerous personal variables within subjects (age, diet, co-infection, prior exposure, other medical history, etc). In order to counter some of these problems, researchers can take advantage of animal models,

which have proven extremely valuable at elucidating the complex etiology of malarial pathogenesis.

The laboratory mouse has been essential to genetic research since “fancy” mouse breeding by zoologists began in the 19th century. Since then, a number of fully inbred strains have been developed which exhibit differential biological characteristics, including response to infection with a wide variety of pathogens[119]. Comparisons between resistant and susceptible inbred strains allow the dissection of complex genetic traits, while enabling researchers to control confounding variables within experimental conditions. In addition, strategic breeding and genetic engineering techniques allow more flexibility than ever to explore the genetics underlying simple or multifaceted traits[96, 120].

### 1.4.1 Mouse Pathology

Several species of *Plasmodium* parasites have the ability to cause malaria in rodents, with *P. chabaudi* commonly used to model blood stage infection and *P. berghei* ANKA to induce cerebral disease. Draft genome sequences show a high degree of conservation between *Plasmodium spp.* with relatively few differences unique to primate parasites. The most diverse genes are those responsible for antigenic variation and immune evasion; nevertheless, patterns of molecular evolution between murine and human parasites are similar[121-123].

*P. chabaudi* infection shares a number of features with *P. falciparum* infection, including severe anemia, splenomegaly, intense erythrophagocytosis, haemolysis, and immune activation. Under laboratory conditions where the parasite and dose can be held constant, inbred strains of mice respond differently to infection, indicating a major role for genetic background in determining the course and outcome of infection. Susceptible mouse strains

such as A/J will develop extremely high levels of blood parasitemia (perhaps >50%) and may succumb to severe anemia at the peak of infection, typically around day 7. Resistant mouse strains such as C57BL/6 still develop parasitemia, peaking around the same time but to a much lower degree and with near 100% survival. They are able to clear the infection entirely in approximately two weeks. *P. chabaudi* infection models have been used to break down many of the pathological features of malarial anemia, to confirm and investigate the mechanism of candidate genes and test anti-malarial drug regimens.

Replication of *P. chabaudi* parasites in mouse erythrocytes leads to hemolysis, erythrophagocytosis (with compensatory erythropoiesis), and splenomegaly[96]. Parasites are capable of sequestering within spleen and liver, making these, along with bone marrow, the primary target tissues of infection. Unlike what is seen in cerebral malaria (see also *P. berghei* induced pathology below), *P. chabaudi* parasites are not capable of sequestering in the brain and there are no signs of neurological involvement, even in the face of very high peripheral parasitemia. A strong early innate immune response (*Ifng*, *Tnf*, *Irf8*, *Irak4*)[120, 124-126] is necessary to control initial replication, and adaptive immunity (*Cd28*, *Il2*, *Il12*, *Il15*) is required to successfully clear the infection post-peak (Table 1)[127-129].

A different parasite species, *P. berghei* ANKA, is used to recapitulate cerebral malaria in an experimental setting. Parasites are able to sequester within microvasculature, cause permeability to the blood brain barrier, and reproduce neurological symptoms such as ataxia, paralysis, seizures and coma[120]. Susceptible mice develop these neurological symptoms between 5 and 9 days post infection and then rapidly deteriorate, with morbidity expected less than 24h after onset of symptoms despite relatively low levels of peripheral parasitemia (generally <25%). Resistant mice do not exhibit neurological

symptoms, but will go on to develop high parasitemia (>60%) and eventually succumb to anemia between days 13 and 21. The two phenotypes are easily differentiated, and it is possible to classify response to experimental cerebral malaria as a binary trait. Within a genetically identical inbred strain, the observed phenotype is very stable. Without treatment, infection is eventually lethal in all inbred strains of mice. In contrast to *P. chabaudi* infections, *P. berghei* pathology has been linked to excessive involvement of innate immunity and Th1 polarized responses, with CM-susceptible mice succumbing to host-mediated immunopathology, rather than excessive parasite replication or toxicity.

The model has been criticized for having different cellular sequestration behavior than *P. falciparum* parasites and the relevance of pathological neuroinflammation to human samples has also been disputed[130]. No animal model will ever be able to perfectly reproduce a complex human disease, however ethical constraints restrict human sampling to blood or autopsy samples which may not accurately reflect local dynamic processes throughout infection either. Rigorous studies from both human and animal researchers have identified numerous parallels along with the differences (Table 2) which supports a valuable role, even if imperfect, for both to contribute towards greater understanding, particularly for mechanistic information of disease processes[131].

## Table 1-1. Selected single gene knockouts associated with host response to murine malaria.

| Gene Name | KO effect on blood stage disease [b] | | KO effect on cerebral disease | Refs |
|---|---|---|---|---|
| | Parasitemia | Survival | | |
| Ank1, Ankyrin 1 | low | ↑ survival | ↑ survival | [132, 133] |
| Asc, PYD/CARD domain containing | | | none | [134] |
| Casp1, Caspase 1 | | | none | [134-136] |
| Ccr5, Chemokine (C-C) receptor 5 | | | ↑ survival | [137] |
| Cd1, CD1 antigen | | | CM develops in BALB/c | [138] |
| Cd8, CD8 antigen | | | ↑ survival | [139] |
| Cd40, CD40 antigen | | | ↑ survival | [140] |
| Cxcr3, Chemokine (C-X-C) receptor 3 | | | ↑ survival | [141] |
| Cxcl9, Chemokine (C-X-C) ligand 9 | | | ↑ survival | [141] |
| Cxcl10, Chemokine (C-X-C) ligand 10 | | | ↑ survival | [141] |
| Darc, Duffy blood group, chemokine receptor | none | none | none | [142] |
| Fas, $Fas^{lpr}$ hypomorph | none | none | ↑ survival | [143, 144] |
| Fcgr2, Fc receptor, IgG low affinity IIb | high | | ↑ survival | [145] |
| Hmox1, Heme oxygenase 1 | none | ↓ survival | CM develops in BALB/c | [146, 147] |
| Hp, Haptoglobin | high | none | none | [148] |
| Icam1, Intracellular adhesion molecule | | | ↑ survival | [149] |
| Ifnar1, Interferon (alpha and beta) receptor 1 | high | | | [150] |
| Ifng, Interferon gamma | high | ↓ survival | ↑ survival | [129, 151, 152] |
| Il1b, Interleukin 1 beta | low | ↑ survival | none | [136, 153] |
| Il10, Interleukin 10 | none | ↓ survival | | [129] |
| Il12rB2, Il12 receptor beta 2 | | | ↑ survival | [154] |
| Irak4, Interleukin 1 receptor associated kinase 4 | high | ↓ survival | ↑ survival | [126] |
| Irf1, Interferon regulatory factor 1 | | | ↑ survival | [155] |
| Lta, Lymphotoxin A | impaired clearance | none | ↑ survival | [156-158] |
| Ltbr, Lymphotoxin β receptor | none | ↓ survival[a] | ↑ survival | [157, 159] |
| Myd88, Myeloid differentiation primary response | none | none | ↑ survival (controversial) | [136, 160-162] |
| Nlrp3, Nod like receptor 3 | low | ↑ survival | delayed onset | [134, 153] |
| Nos2, Nitric oxide synthase 2 (iNos) | low | none | none | [163, 164] |
| Nos3, Nitric oxide synthase 3 (eNos) | none | none | none | [164, 165] |
| Pf4 (Cxcl14), Platelet factor 4 | | | ↑ survival | [166] |
| Pklr, Pyruvate kinase | low | ↑ survival | none | unpub, [102] |
| Rag2, recombination activating gene 2 | | | ↑ survival | [167] |

| | | | | |
|---|---|---|---|---|
| Selp, P-selectin | | | ↑ survival | [168] |
| Socs1, Suppressor of cytokine signaling 1 | | | ↑ survival | [169] |
| Spna1, Spectrin | low/absent | | low/absent | [133] |
| Stat6, Signal transducer and activator 6 | high | | | [170] |
| Tnf, Tumor necrosis factor | high | ↑ mortality | none | [152, 156, 158] |
| Tnfr1a TNF receptor superfamily, 1a | high | none | none | [157, 171, 172] |
| Tnfr1b TNF receptor superfamily, 1b | | | ↑ survival | [157, 171, 172] |
| Vnn3, Vanin3 | low | ↑ survival | | [173] |

[a] Indicates a noted sex effect; generally, males are more susceptible than females.
[b] Most knockouts are created in, or crossed onto, a C57BL/6 background which would be expected to survive a *P. chabaudi* infection and succumb to *P. berghei*.

*Portions of this table have been published previously. Adapted from [120] with permission.*

To look more closely at one of the criticisms of the murine model, both leukocytes and parasitized erythrocytes have been reported within the brains of humans and mice suffering from cerebral disease, though experimental cerebral malaria features predominantly leukocytes and human cerebral malaria is largely erythrocyte sequestration[167, 174-178]. Franke-Fayard and colleagues showed that fluorescent *P. berghei* schizonts do sequester within deep tissue, including the brain, though lungs, adipose tissue and spleen are the most heavily sequestered organs. Variations in observed brain sequestration may reflect differences in total parasite burden, parasite life stage or mouse strain[174, 175]. Building on this, it was shown that brain sequestration of parasitized erythrocytes in experimental cerebral malaria is controlled primarily by the presence of $CD8^+$ T-cells, as $CD8^{-/-}$ mice had significantly reduced accumulation of parasites within brain tissue compared to other strains, despite roughly equal overall parasite burdens[167]. Coincidental segregation of both parasitized erythrocytes and

$CD8^+$ T-cells in the brain is required for pathology suggesting a model where continuous stimulation of T-cells by parasites causes fatal immunopathology[176]. Migration, activation and sequestration (possibly via ICAM-1 upregulation on endothelial cells) of both cell types as well as platelets are mediated by IFN-$\gamma$ signaling, initiated by Natural Killer (NK) cells and $CD4^+$ T-cells[167, 176, 179].

Table 1-2. Sequestration properties of blood stages of *P. falciparum* in humans and *P. berghei* in rodents.[174]

| | *P. berghei ANKA* | *P. falciparum* |
|---|---|---|
| **Sequestration of blood stages** | | |
| Stage of sequestration | Schizonts (from the onset of nuclear division). Maturing trophozoites do NOT sequester and all gametocyte stage appear to remain in circulation | Maturing trophozoites, schizonts and immature gametocytes |
| Host cell receptors for adherence | CD36, Identified by in vivo analysis of sequestration in CD36-/- mice. | Multiple receptors including CD36, ICAM1, PCAM1/CD31, CR1, CSA identified using in vitro binding assays |
| Sites of sequestration | Abundant CD36-mediated schizont sequestration in lungs, adipose tissue and spleen. | Many tissues including brain, lung, spleen, intestine, skin, fat tissue, bone marrow, skeletal and cardiac muscle. Organ sequestration assessed by examining post-mortem tissue obtained from individuals that died of severe malaria. |
| **Sequestration of blood stages associated with severe disease** | | |
| *Stage of sequestration* | Schizonts. In addition, sequestration/ accumulation of other asexual blood stages | *See above.* In addition, evidence has been presented for sequestration of young trophozoites during severe disease. |
| *Host cell receptors* | CD36 for schizonts (*see above*). Receptors for other blood stages have not been identified | *See above.* During severe disease the distribution of schizont sequestration may change as a result of changes in expression of host cell receptors |
| *Parasite ligands* | *See above.* No putative ligand identified for CD36-mediated sequestration nor any other ligand involved in sequestration/ accumulation of other blood stages during severe disease | *See above.* During severe disease and pregnancy-associated malaria, selection favors parasites expressing specific variants of PfEMP1 |
| *Sites of sequestration* | CD36-mediated schizont sequestration in lungs and adipose tissue. Accumulation of parasitized erythrocytes in different tissues including the brain and placenta of pregnant mice. | See above. Abundant sequestration in brain tissue in many patients that died from cerebral complications and in the placenta of pregnant women. |

*Franke-Fayard et al (2010)* PloS Pathogens*. Used with permission[174]*

### 1.4.2 Forward genetic approaches in mouse models

One of the advantages of experimental models is the ability to monitor an infection through pre-symptomatic early stages and follow the natural progression of pathology within an individual in the absence of intervention. Phenotype driven forward genetic experiments have allowed the identification of pathological drivers of severe malarial disease by comparing strains of mice that respond differently to identical experimental infections. Similar to what is seen in humans, the genetic control of murine susceptibility to malaria is complex, with common inbred mouse strains showing a wide range of reproducible differences under controlled conditions[180]. These genetically based differences can affect infection dynamics including peak parasitemia, susceptibility to anemia, cerebral disease, other facets of the host response, and ultimately, survival. Investigating these different phenotypes across strains enables the discovery of key features regulating susceptibility in an unbiased, non-candidate driven fashion.

Quantitative trait locus (QTL) mapping seeks to correlate regions of parental chromosomes that specifically co-segregate with resistance or susceptibility in the progeny of an informative cross between resistant and susceptible strains. The most commonly used breeding schemes to generate an informative cross are F2 ($2^{nd}$ filial) intercross and backcross designs (Figure 4). F2 intercrosses begin by pairing males (AA) with females (BB) from strains that have a different phenotype to generate a fully heterozygous F1 hybrid generation (AB). F1 mice are then intercrossed via brother-sister mating to produce the F2 generation which will carry 50% of its genetic material from each original parent in a randomly assorted 25% AA, 50% AB, 25% BB Mendelian ratio. In a backcross, the F1 generation is bred back with one of the parents to generate a population of mice that are 50% homozygous for that parental genotype and

50% heterozygous. Intercrosses are required to study recessive traits and preferred for most multigenic traits where interactions may be important.

QTL mapping has been applied to experimental malaria using *P. chabaudi*, *P. berghei* and *P. yoelii* infections. Using *P. chabaudi* as a blood-stage model of infection, SJL, C3H/H3, NC/Jic, DBA/2, CBA/J, SM/J, MOLF/Ei and A/J strains are all susceptible with high peak parasitemia and death while C57BL/6, C57BL/10, BALB/c, DBA/1, SPRET/Ei and 129SvJ are resistant[120]. Degree of parasitemia and/or survival has been used to map ten *chabaudi* resistance loci (*Char1-10*) with genome wide significance (Figure 5)[102, 173, 181-186]. Most of these QTLs overlay a large chromosomal segment with hundreds of potential candidate genes. In the case of *Char3,* which is involved in parasite clearance, a precise candidate polymorphism may be impossible to tease out as it maps within the mouse major histocompatibility cluster (MHC), a site rich in regulators and mediators of the immune response analogous to the human HLA complex[87, 183]. Congenic and subcongenic lines, gene expression analysis, and sequencing studies have been used to narrow down the intervals for other QTLs and two genes have been positively identified: *Pklr* for *Char4,* and the Vanin cluster (*Vnn1* and *Vnn3*) for *Char9*[102, 173].

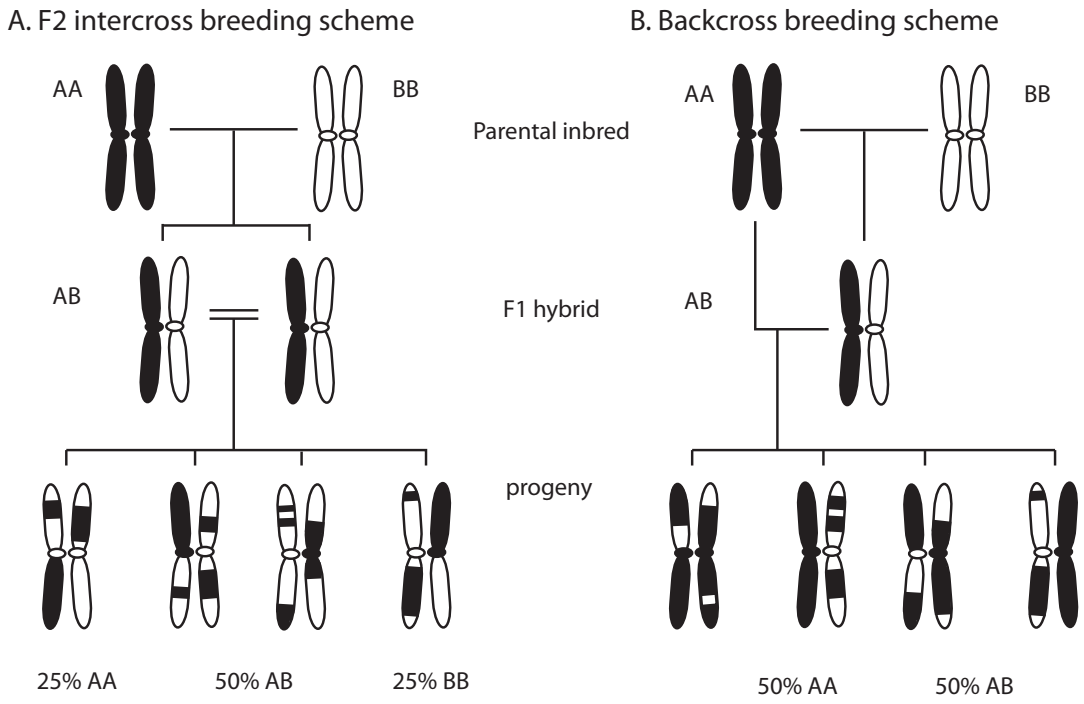



**Figure 1-4.** Breeding scheme for informative F2 intercross and backcross designs illustrating parental genomic contribution to the progeny

*P. berghei* has been used to map eleven QTL loci (*Berr1-9*, *cmsc*, and an unnamed locus on Chr18) regulating susceptibility to cerebral malaria, typically with survival post-infection as the binary or quantitative trait (Figure 5)[187-193]. Most of inbred strains of mice that have been tested are susceptible to the development of experimental cerebral malaria including 129, A/J, C57BL/6, C57BL/10, CAST/Ei, CBA, DBA/1, SJL[87, 194, 195]. Onset of neurological symptoms may vary by +/- 2 days (between d5 and d9 by strain), however, all mice from these strains develop characteristic seizures, coma and neuroinflammation while parasitemia levels remain under 25%. Resistant strains include AKR/J, BALB/c, DBA/2 (males) and wild derived (WLA) mice, where more than 80% of mice survive through the cerebral window and go on to develop hyper-parasitemia and anemia late during infection[191, 194, 195].

Peroxisome proliferator activated receptor gamma (*Pparg*) has been proposed as the candidate for *Berr6* based on prior association, and though positional candidates have been suggested for other QTL loci, each of these requires additional validation and follow up. While each of these QTLs was calculated to have a significant impact on infection outcome in the mice, *Berr1* and *Berr2* as well as *Berr7* and *Berr8* were additionally shown to interact in a two-locus model of resistance. The *Berr1/Berr2* interaction follows a model where susceptibility is controlled by independent action of recessive alleles derived from the C57BL/6 strain, with a homozygous genotype at either single locus able to confer an effect[187, 196]. In contrast, *Berr7* and *Berr8* appear to interact epistatically with a major effect from *Berr8*. *Berr7* then provides a gene-dosage dependent protective boost in the context of *Berr8* heterozygosity or 129S1 homozygosity. Unlike *Berr8*, *Berr7* does not appear to provide a significant effect on survival when examined alone[192].

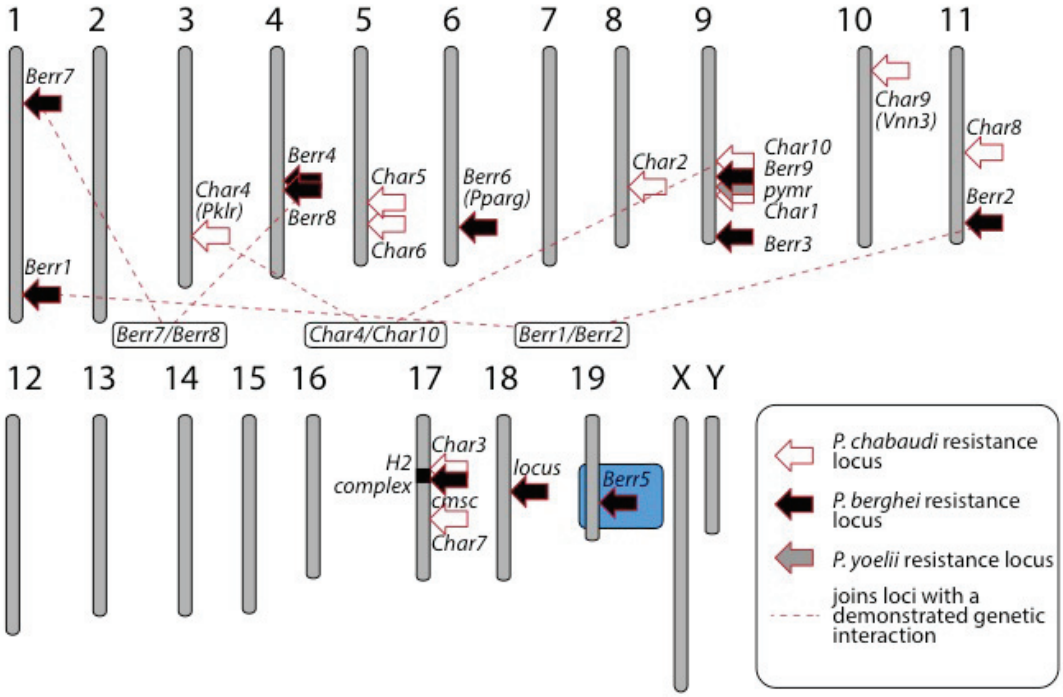



**Figure 1-5. Char (Plasmodium chabaudi) and Berr (Plasmodium berghei) resistance loci described in the scientific literature**

Chromosomal position of quantitative trait loci controlling murine response to blood stage (*P. chabaudi*) or cerebral (*P. berghei*) malaria infection. The *Berr5* locus described in Chapter 2 is highlighted in blue.

In addition to exploiting the genetic diversity present in inbred strains of laboratory mice, it is possible to introduce random germ line mutations and screen for pedigrees that have a reproducible, heritable response to infection that is different from the parental background indicating a “hit” to an important pathway. Random mutagenesis can introduce genetic variability into genes that are not functionally divergent within the limited set of common inbred strains, and using an experimentally blind approach can highlight a role for unexpected genes in pathology. If a phenodeviant pedigree is identified, the *de novo* mutation responsible for the trait can be identified by outcrossing and QTL mapping or by deep sequencing of resistant mouse mutants.

*N*-ethyl-N-nitrosurea (ENU) is a potent alkylating agent capable of inducing mutations at a frequency of one per 2.7Mb of chromosomal sequence[197]. When injected into male mice, this translates to approximately 1000 de novo mutations per gamete, with ~20 estimated to be in functional coding or regulatory regions. The overwhelming majority of changes will be single base pair substitutions, and ENU mutagenesis has a bias towards T-A and A-T transversions (44%)[197]. Outcrossing mutagenized males and then backcrossing daughters permits homozygous mutant offspring to be bred in a 25% ratio (assuming viability) and allows researchers to screen for recessive traits. The advent of whole genome and whole exome sequencing have made ENU mutagenesis exponentially more amenable to gene discovery and the technique is gaining popularity in spite of requiring a significant investment in mouse colony maintenance[198].

One study has been published using ENU mutagenesis as a platform for the identification of cerebral malaria associated genes[199]. Susceptible C57BL/6 were used as the background and a complex backcross scheme was used to generate G3 progeny for phenotypic testing by *P. berghei* infection. A

pedigree was identified with 31% resistant mice, and QTL mapping together with full genome sequencing identified and confirmed a W81R mutation in *Jak3* as responsible for the observed resistance. The mechanism of for $Jak3^{W81R}$ mice was attributed to depletion and dysfunction of $CD8^+$ T-cells, preventing their recruitment to the brain and the onset of cerebral pathology. This study also functions as a proof of principle, and future investigations into ENU mutagenized pedigrees promise to be interesting.

### 1.4.3 Reverse genetic approaches to study malaria

As the field advances, numerous pathways involved in malarial disease are uncovered or hinted at by forward genetics and human studies. Targeted gene knockout mice can be used to validate these findings and build on them to develop more comprehensive models and delineate pathological mechanisms. To this end, multiple transgenic mice have been infected with either *P. chabaudi* or *P. berghei* (Table 1).

Current evidence suggests that experimental cerebral malaria is an encephalitis with pronounced neuroinflammation as a result of sequestered parasitized erythrocytes and leukocytes. Mice with targeted knockout mutations in cytokines, chemokines and receptors such as *Ccr5*, *Cxcl9*, *Cxcl10*, *Cxcr3*, *Icam1*, *Ifng*, *Lta*, and *Ltbr* are all highly resistant to the development of neurological symptoms (Table 1)[141, 151, 156, 200]. Similarly, deletion of cytokine responsive genes including *Irf1, Irak4* and *Socs1* also causes resistance[126, 155, 169]. Surprisingly $Tnf^{-/-}$ mice are as susceptible to cerebral malaria as wildtype parental mice indicating that experimental cerebral malaria is independent of TNF-$\alpha$ signaling (or at least, that blocking *Tnf* signaling is itself not sufficient to prevent pathology)[108, 156]. The Nlrp3 inflammasome ($Nlrp3^{-/-}$, $P2x7r^{-/-}$, $Asc^{-/-}$, $Casp1^{-/-}$, $Il1r^{-/-}$) is also extraneous to pathology despite including

multiple genes which are upregulated during a wild type infection[134]. These studies suggest that while inflammation is clearly an important element of experimental cerebral malaria pathology, there are specific inflammatory pathways that are critical, while others may be redundant or insufficient to alter the outcome of this host:parasite interaction.

Gene knockouts that are capable of eliminating entire cell populations have also been extremely valuable to understanding pathogenesis. *Cd1*$^{-/-}$ mice which have normal Th2 responses but lack Natural Killer-like T-cells (NKT) fail to develop experimental cerebral malaria, defining a role for NKT cells in disease promotion, possibly via antigen presentation and/or altering the Th1/Th2 balance required for non-pathogenic immunity[138]. $CD4^+$ and $CD8^+$ T-cells have also been shown to be absolutely essential for experimental cerebral malaria (*Cd8*$^{-/-}$, *Rag1*$^{-/-}$, *Jak3*$^{-/-}$), but not B-cells (B6.μMT)[139, 179, 199].

Human studies in malaria have yielded many genetic loci associated with protection from malarial disease, but cannot be used to establish a mechanism of resistance. By the same token, complex interactions between multiple cell types including erythrocytes, immune and epithelial cells are impossible to accurately reproduce in a culture dish. While experimental infections in mice may not display all the same associations due to a number of host or parasite specific factors, gene knockout mice that reproduce the same phenotype can be used to confirm a role for certain genes and pathways within the complex etiology of disease. These gene-knockout mice can be used to study the mechanism of genetic resistance to malaria in a dynamic way, allowing pre-terminal tissue sampling across multiple organ systems, and/or using interventions that would be intrusive or unethical in human subjects.

Similar to humans, mice genetically engineered to express sickle cell hemoglobin (*Hb*$^{SAD}$) are resistant to *Plasmodium* infection. They do not develop

cerebral malaria in response to *P. berghei* infection, and survive *P. chabaudi* infection with moderate anemia compared to controls (personal observations and [201]). In these mice, resistance to cerebral malaria can be tied to increased levels of heme oxygenase-1 (*Hmox1*) expression and reduced activation and population expansion of pathogenic CD8+ T-cells. The reduction in T-cell number is independent of parasite burden and may involve a reduction in free heme (heme is scavenged by HO-1) or antigen presenting[201]. Knocking out *Hmox1* in resistant BALB/c mice can induce cerebral malaria, and the phenotype can be reversed by the administration of inhaled carbon monoxide, zinc protoporphyrin-IX or cobalt protoporphyrin-IX, bolstering confidence for this mechanism and opening up new doors for therapeutic intervention[202].

ICAM1 is another gene regularly associated with resistance to cerebral disease in humans with the presumptive mechanism of binding infected erythrocytes within microvasculature, which has support from *in vitro* studies[58, 203, 204]. Infection of *Icam1*$^{-/-}$ mice confirms this model with increased survival and fewer adherent cells within brain venules – particularly fewer infected erythrocytes and platelets[149, 205].

Overall, communication between malarial researchers working with human patients and mouse models has proven beneficial to both parties, with each contributing towards a more comprehensive understanding of malaria pathogenesis. The discovery of genes and mechanisms involved in host resistance responses allow the development of specifically targeted drugs aimed at blocking interactions required by the parasites for invasion, immune avoidance or homeostasis, as well as therapeutics designed to combat host mediators of acute pathogenesis. The second benefit is that unlocking novel ways to combat malaria or malarial disease sidesteps the increasing problem of parasite drug resistance, which has been observed in the field against all current

treatments[206]. These potential therapeutics themselves must also go through rigorous safety and efficacy characterization which requires the use of well-developed model infections as a primary or intermediate step before beginning clinical field studies.

# Chapter 2 Identification of a novel cerebral malaria susceptibility locus (*Berr5*) on mouse chromosome 19

## Preface to Chapter 2:

Well established observations of differential susceptibility to cerebral malaria infection between inbred strains of mice infected with *P. berghei* ANKA indicated a genetic basis to this trait, however the identity of those factors remains unknown.

In the following chapter we sought to use forward genetics and QTL mapping in order to identify the genetic basis of resistance between the common lab strains C57BL/6 (susceptible) and BALB/c (resistant). We were able to identify a locus on chromosome 19 where parental genotype has a significant effect on survival, explaining between 6-9% of the variance. The position of this *Berr5* QTL overlapped with three other inflammatory QTLs, suggesting that the locus is involved in inflammatory response, which is known to be a major contributor to murine cerebral malaria. Using expression analysis, we identified a subset of genes that are the most likely positional candidates to underlie the trait.

## 2.1 Abstract

Cerebral malaria (CM) is an acute, generally lethal condition characterized by high fever, seizures and coma. The genetic component to CM can be investigated in mouse models that vary in degree of susceptibility to infection with *Plasmodium berghei* ANKA. Using survival time to measure susceptibility in an informative F2 cross (*n*=257), we identified linkage to chromosome 19 (*Berr5* (*Berghei* resistance locus 5), LOD=4.69) controlling, in part, the differential response between resistant BALB/c and susceptible C57BL/6 progenitors. BALB/c alleles convey increased survival through the cerebral phase of infection but have no quantitative effect on parasitemia during the later, anemic phase. The *Berr5* locus colocalizes with three other immune loci, including *Trl-4* (tuberculosis resistance), *Tsiq2* (T-cell secretion of IL-4) and *Eae19* (experimental allergic encephalitis 19), suggesting the possibility of a common genetic effect underlying these phenotypes. Potential positional candidates include the family of *Ifit1–3* (interferon-inducible protein with tetratricopeptide repeats 1–3) and *Fas*.

## 2.2 Introduction

Malaria is a major global health problem with approximately 40% of the world's population living in at risk areas. In Africa, malaria is responsible for 20% of childhood deaths with a cumulative mortality of 1–3 million people each year[207]. The most severe complication of *Plasmodium falciparum* infection is cerebral malaria (CM). CM is caused by infected erythrocytes becoming trapped in brain microvasculature, leading to an excessive, localized host-driven inflammatory response and resulting in hypoxia, loss of blood–brain barrier (BBB) integrity and cell death[200, 208, 209]. Patients develop seizures, paralysis, ataxia and unrousable coma that often ends in death. The cellular mechanisms of the pathological response are poorly understood and must be better characterized to develop strategies for prevention and treatment of this condition.

The host genetic background has a major effect on response to malaria in humans (reviewed by Kwiatkowski[65] and Mazier et al.[210]). The genetic component of susceptibility to CM is multigenic and further modified by parasite virulence and other environmental determinants. Inbred strains of mice exhibit different responses to *P. berghei* ANKA (PbA), with susceptible strains such as C57BL/6 (B6) or CBA developing an acute cerebral syndrome within 6–7 days characterized by ataxia, paraplegia, seizures and coma, leading to uniform lethality by day 8–10 after infection. On the other hand, resistant mice including the BALB/c and WLA strains do not develop neurological symptoms, but die 2–4 weeks after infection due to severe anemia caused by high levels of blood parasitemia. Current evidence suggests that CM in mice infected with PbA is caused in part by a robust but detrimental localized inflammation in response to parasitized erythrocytes trapped in the brain microvasculature[208]. Transcript profiling experiments have documented a substantial pro-inflammatory response

that is more pronounced in susceptible mice[211-213]. In addition, loss-of-function mutations in genes coding for pro-inflammatory molecules such as interferon-γ (*Ifng*), lymphotoxin (*Lta*), hemolytic complement (*C5a*) and interferon regulatory factor 1 (*Irf1*) have been shown to protect against CM[155, 156, 194, 214]. These observations support the hypothesis of a central function for immune-mediated pathology in CM, with T cells, platelets and resident macrophages implicated in inflammatory response and tissue injury[215-217].

Quantitative trait locus (QTL) mapping studies by whole-genome scanning in mice have detected five chromosomal loci affecting the host response to PbA infection. Using survival as a phenotypic trait – an unnamed locus on chromosome 18 in F2 mice derived from susceptible B6 and resistant DBA/2 was first identified[190]. Subsequently, QTLs designated *Berr1* (chromosome 1), *Berr2* (chromosome 11), *Berr3* (chromosome 9) and *Berr4* (suggestive linkage to chromosome 4) affecting the presence and severity of neurological symptoms following infection were mapped in F2 mice derived from C57BL/6 (susceptible) and the resistant wild-derived WLA strain[187, 188]. Finally, the *cmsc* locus mapping to the H-2 region of chromosome 17 was found to affect susceptibility in progeny of CBA mice crossed with the resistant DBA/2 strain[189]. In all cases, the detected QTLs span large chromosomal segments, and the causative genes underlying their effect on survival time in *P. berghei*-infected animals have not been identified.

In this study, we have investigated the genetic basis of differential susceptibility of the commonly compared B6 strain and the BALB/c strain to infection with *P. berghei* ANKA.

## 2.3 Materials and Methods

*Animals*

Pathogen-free, 8- to 10-week-old inbred C57BL/6 (B6) and BALB/c mice were purchased from Charles River Laboratories (Wilmington, MA, USA). 257 (C57BL/6 × BALB/c)F2 progeny (CB6F2) were bred by systematic brother–sister mating of (C57BL/6 × BALB/c)F1 progenitors (Charles River Laboratories). All animal experiments were carried out according to Canadian Council on Animal Care guidelines and approved by the Animal Care Committee of McGill University.

*Parasite and infection*

*P. berghei* ANKA was obtained from the Malaria Reference and Research Reagent Resource Centre (MR4) and maintained as frozen stocks at −80 °C that were passaged in B6 mice before use in infection. All experimental mice were injected intraperitoneally with $5 \times 10^4$ parasitized red blood cells. Daily blood parasitemia was determined on Diff-Quick (Dade Behring, Newark, DE, USA) stained thin blood smears. Neurological symptoms and survival were monitored and CM was diagnosed by the presence of ataxia, weak gripping and righting responses, paralysis, seizures and coma between days 6 and 10. Serum was collected at day 5 after infection from all CB6F2s and IFNγ levels were measured using a commercial enzyme-linked immunosorbent assay kit (eBiosciences, San Diego, CA, USA).

*Evan's Blue dye extravasation assay*

To examine BBB permeability during infection *in vivo*, we injected control and PbA-infected C57BL/6 and BALB/c mice (at day 8 after infection) intraperitoneally with 0.2 ml of 1% Evan's Blue dye (E2129; Sigma-Aldrich, Oakville, ON, Canada) in phosphate-buffered saline (PBS). Ninety minutes later, mice were killed by $CO_2$ inhalation, perfused with PBS and brains were harvested and photographed.

*Genotyping*

All 257 CB6F2 mice were genotyped at the McGill University and Genome Quebec Innovation Centre (Montreal, QC, Canada) using the low-density Illumina GoldenGate mouse panel that contained 246 informative markers distributed across the genome. Fine mapping and filling gaps larger than 40 Mb was carried out using microsatellite markers obtained from the Mouse Genome Informatics Database (www.informatics.jax.org)[218].

*Data analysis*

Survival of CB6F2 mice was analyzed by the Kaplan–Meier log-rank test and survival fractions were compared using the $\chi^2$ statistic (GraphPad Prism 5, San Diego, CA, USA). Blood parasitemia and serum IFNγ levels were analyzed using ANOVA, followed by a Bonferroni *post hoc* test. LOD scores were calculated with R/qtl software version 2.7.0 using survival time as a phenotype. Because the data were not normally distributed, we used the two-part model[219, 220] that was developed for phenotypes with a spike in the distribution. This method has been used to analyze similar data sets[188]. Data were also analyzed using the binary model, segregating mice into two categories based on survival past day 13. The cutoff used for genome-wide significance ($P<0.05$) was

4.19 for the two-part model and 3.50 for the binary model based on 1000 permutations.

*Microarray analysis*

Gene expression data sets published by Lovegrove et al[213] were downloaded (GEO database; record GSE7814) and analyzed with the GeneSifter software package. These transcript profiles correspond to total brain RNA from B6 and BALB/c mice at day 0, 1, 3 and 6 after PbA infection, which had been hybridized to Affymetrix GeneChip Mouse Genome 430A 2.0 Array chips (Affymetrix, Santa Clara, CA, USA). Expression was compared using two-way ANOVA with a 1.4-fold cutoff, Bonferonni and Hoch correction. In the *Berr5* interval, 15 genes showed a difference in expression level due to either strain-specific variation or over the course of infection. Ten genes (*Ifit1*, *Ifit2*, *Ifit3*, *Cd274*, *Ch25h*, *Blnk*, *Fas*, *Il33*, *Cutc* and *Vldlr*) were regulated by the interaction of genetic background and infection. Genes that were differentially regulated between strains either during infection (day 6) or at basal levels (day 0) were sorted according to chromosomal location to identity positional candidates.

*Semiquantitative RT-PCR*

Uninfected control and day 6 PbA-infected C57BL/6 and BALB/c mice were killed and brains were harvested. RNA from four replicate brains was pooled and 2 µg of the pooled sample was reverse transcribed using oligo d(T) and MMLV RT (Invitrogen). PCR amplification of target genes was carried out using Taq DNA polymerase (Invitrogen, Burlington, ON, Canada) with 5 min denaturing (94 °C) and 24 cycles amplification (30 s denaturing at 94 °C, 30 s annealing at 55 °C and 30 s extension at 72 °C) with a final 7 min extension. The RT-PCR primers used were as follows: *Ifit1*, 5t1 Taq DNA polymerase (Invitrogen,

Burlington, ON, C*Ifit2*, 5t2 Taq DNA polymerase (Invitrogen, Burlington, ON, Canad*Ifit3*, 5′-GCCGTTACAGGGAAATACTGG and 3′-CCTCAACATCGGGGC TCT. Relative gene expression was quantified using Quantity One software following overnight hybridization of a [α-32P]dATP-labeled probe specific for each amplicon as described by Marquis et al[221].

## 2.4 Results

We compared the response of the inbred C57BL/6 and BALB/c strains of mice to infection with PbA. B6 were found to be susceptible with the majority of animals dying during the early cerebral phase of infection (Figures 2-1a and b). Neurological symptoms and death in these mice occurred while levels of blood parasitemia were <20% (data not shown). Conversely, BALB/c mice were resistant to PbA-induced CM, with >75–80% of the animals surviving the cerebral phase. Surviving mice did not develop any subsequent cerebral symptoms. Instead, they developed high levels of blood parasitemia (>60%; data not shown) and died of severe malarial anemia.

CM in susceptible PbA-infected mice has been previously associated with disruption of the integrity of the BBB[222], which we monitored using a dye extravasation assay (Figure 2-1c). Examination of brains from Evan's Blue-injected mice revealed that in B6 mice the entire tissue was dyed blue with several darker foci whereas BALB/c brains showed very little uptake of the blue dye, with levels similar to those seen in uninfected controls (data not shown). These results confirm increased PbA-induced BBB permeability as a distinguishing feature between B6 and BALB/c animals.

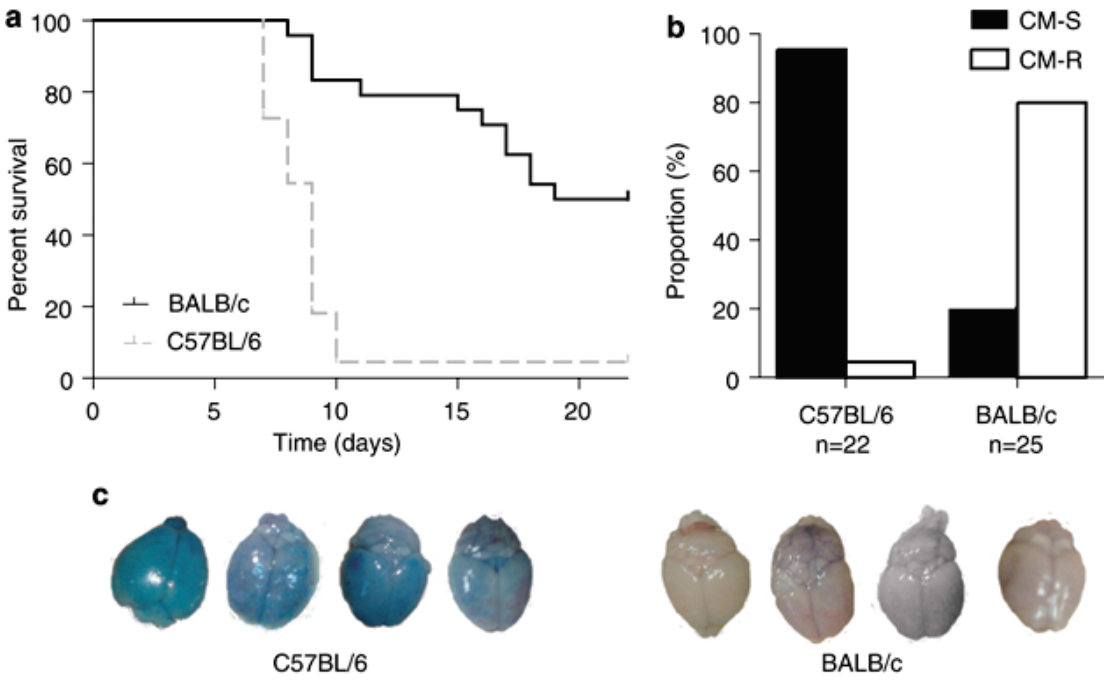



**Figure 2-1. Differential susceptibility of inbred mouse strains to Plasmodium berghei-induced cerebral malaria**

(a) Mice were infected by the intraperitoneal route with $5 \times 10^4$ *P. berghei* ANKA parasitized erythrocytes and survival was monitored. Neurological symptoms began to appear at day 6 in susceptible (CM-S) C57BL/6 mice. The majority of BALB/c mice were categorized as resistant (CM-R), surviving well past day 12. (b) Proportion of CM-S (filled bars; dead before day 10) and CM-R (empty bars; survive past day 12) mice detected in each strain. (c) Integrity of the blood–brain barrier in PbA-infected B6 and BALB/c mice was evaluated 8 days after infection using an Evan's Blue dye permeability assay *in vivo* (see Materials and methods section).

To study the genetic complexity of the B6 vs. BALB/c differential susceptibility to CM, we infected 257 (C57BL/6 × BALB/c)F2 mice (CB6F2) with PbA. CB6F2 mice segregated into two groups, with a majority of animals (~75%) dying in the early cerebral phase (before day 13), and the remainder surviving past day 17. The experiment was terminated at day 22 (Figures 2-2a and b). All animals dying from early CM (susceptible, CM-S) had low parasitemia (<20%), whereas those dying late (resistant, CM-R; including those killed at day 22) developed high parasitemia (40–90%). In addition, the predictive value of blood parasitemia at day 6 (Figure 2c) and serum IFNγ levels at d5 (Figure 2-2d) on overall susceptibility was determined. We observed that CM susceptibility as measured by survival was significantly correlated ($p = 0.0001$) with both higher early parasitemia values and higher IFNγ levels.

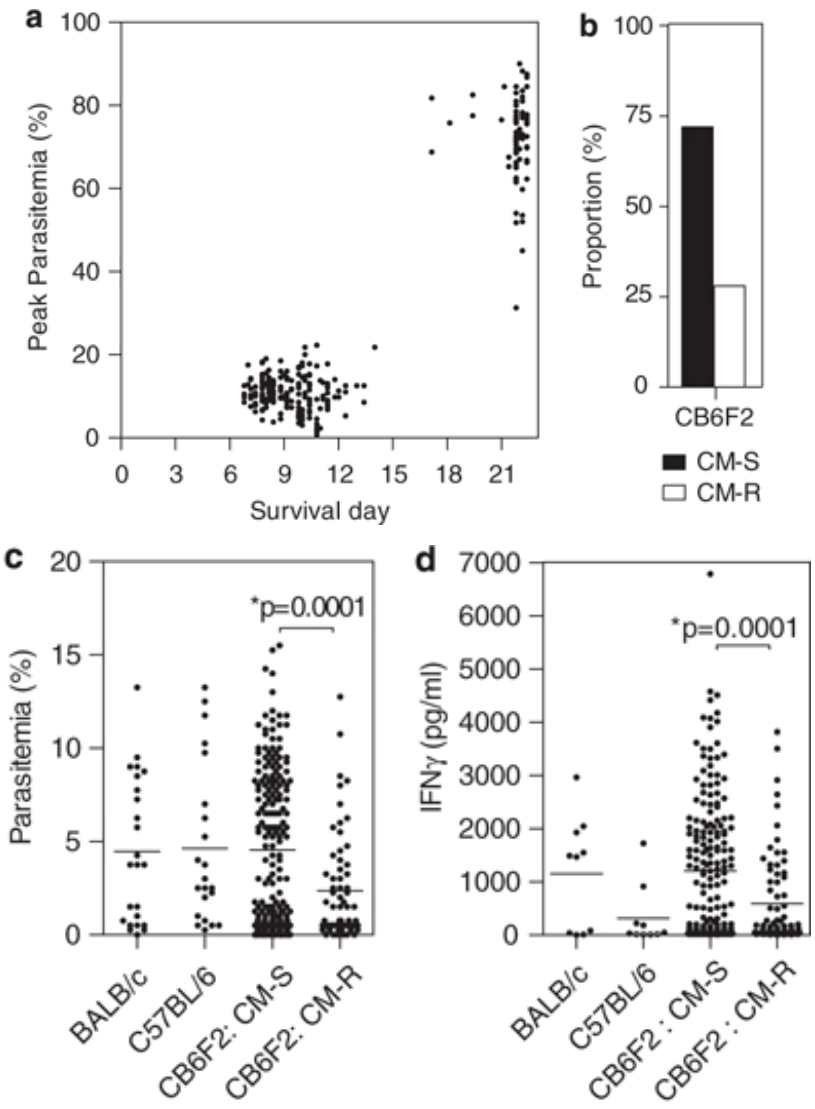



**Figure 2-2.** Segregation analysis of susceptibility to P. berghei induced cerebral malaria in (C57BL/6 x BALB/c)F2 mice.

Daily survival and level of blood parasitemia were recorded for (C57BL/6 × BALB/c)F2 mice (n=257) infected with PbA. (a) The level of parasitemia at the peak of infection, or immediately before moribund animals were killed, along with the time of survival are shown. (b) The fraction of CB6F2 mice showing either a CM-S phenotype (dying before day 13 with <25%parasitemia) or a CM-R phenotype (surviving past day 13 with high 40–90% blood parasitemia) is shown. (c) Day 6 blood parasitemia values in parental and F2 mice, classified as either CM-S or CM-R as described in b. CM-S F2s showed significantly higher blood parasitemia at day 6 than CM-R F2s. (d) Day 5 serum interferon-γ (IFNγ) levels in parental and F2 mice, classified as either CM-S or CM-R.

A genome-wide scan was carried out on all 257 CB6F2 mice, using 252 informative markers with survival as a marker of CM susceptibility. Both a binary (CM-S vs CM-R) and a parametric model of survival time (two-part model) were used in this analysis that yielded similar results. The two-part model takes into account both the binary trait (probability of survival to the experimental end point) and the approximately normally distributed survival within the cerebral phase (survival time contingent on not reaching the end point)[219]. Both analyses identified a single genetic effect on chromosome 19, with maximum linkage to the CEL-19_32349880 marker at chromosomal position 32.3 Mb (Figures 2-3a and c). Using the two-part model, we found that the peak LOD score was 4.69 ($p < 0.05$) whereas the binary analysis yielded a peak LOD score of 3.51 at the same marker ($p < 0.05$). This locus controlling PbA-induced CM was given the temporary identification *Berr5* (*Berghei* resistance locus 5). The 95% confidence interval for the *Berr5* QTL was estimated to be approximately 10 – 20 Mb (25.6 – 45.1 Mb based on a 1.5 LOD decrease; 29.3–41.3 Mb using the 95% Bayesian credible interval) (Figure 2-5a). *Berr5* was estimated to explain between 5.9 and 8.8% of the phenotypic variance. We also used day 6 parasitemia as a quantitative measure in whole-genome scanning. This analysis revealed a single suggestive linkage peak on chromosome 19 in the *Berr5* interval (LOD = 3.35; Figures 2-3b and c). These results show that *Berr5* alleles can affect both early parasitemia and CM susceptibility in CB6F2 mice. This colocalization agrees with the results in Figure 2-2c, showing an association between day 6 parasitemia levels and CM-associated death.

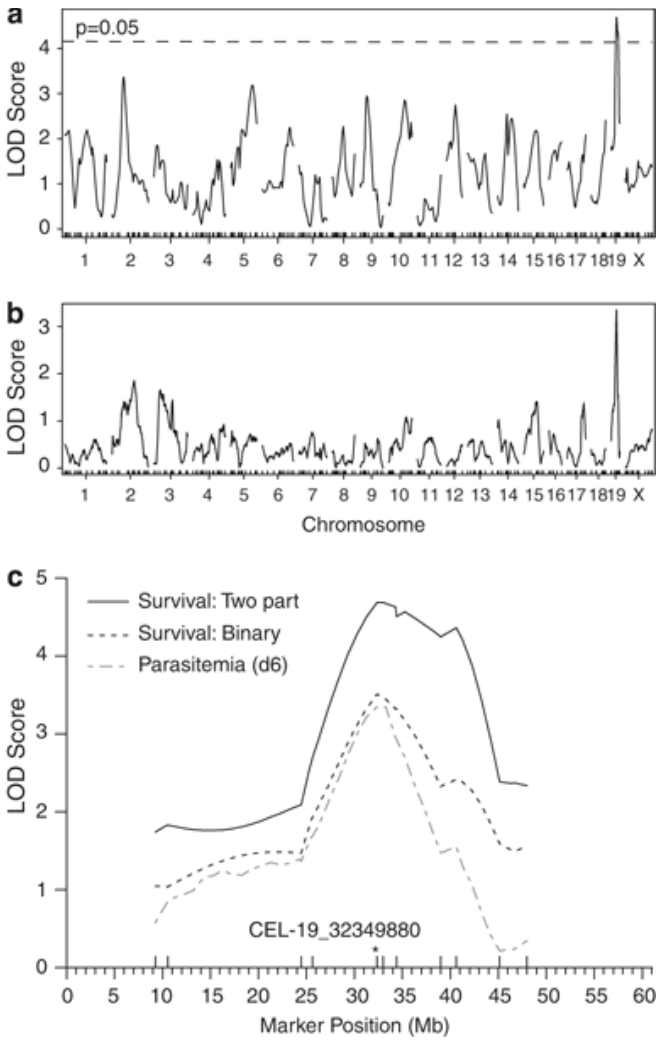



**Figure 2-3.** Mapping of genetic effects (*Berr5*) influencing susceptibility to *P. berghei*-induced cerebral malaria

Whole-genome scanning was carried out in 257 (C57BL/6 × BALB/c)F2 mice, using time of survival after infection with *P. berghei* ANKA or day 6 parasitemia. Interval mapping was carried out using the R/qtl software package, and results for all chromosomes are plotted. (a) Survival time was analyzed according to the two-part model of Boyartchuk et al.[219] A significant linkage peak (LOD=4.69) was detected on chromosome 19. Permutation testing (1000 tests) revealed the cutoff for genome-wide significance (P<0.05) in this model to be 4.19 (dashed line). (b) A suggestive linkage to the same portion of chromosome 19 was also detected using day 6 parasitemia as a quantitative trait. (c) LOD score plots were generated for chromosome 19, using either time of survival in the two-part model (black line) or binary analysis of CM-S and CM-R phenotypes (dashed line), and day 6 parasitemia (gray line), which all point to the same linkage peak.

The effect of *Berr5* alleles on the phenotypic markers of CM susceptibility was examined in the CB6F2 mice (Figure 2-4). Homozygosity for B6 (susceptible) parental alleles at the peak marker was associated with CM-S, with only 9% of these animals surviving the cerebral phase whereas homozygosity for BALB/c (resistant) alleles was associated with increased CM-R and 39% survival through the cerebral phase ($p < 0.005$, Kruskal–Wallis). Heterozygosity was also associated with increased resistance (31% survival), suggesting that resistance alleles at *Berr5* are inherited dominantly (Figures 2-4a and c). Similarly, B6-derived susceptibility alleles at *Berr5* were associated with higher day 6 parasitemia in CB6F2s when compared with either homozygotes or heterozygotes for BALB/c-derived resistance alleles at *Berr5* ($p < 0.01$, t-test; Figure 2-4b).

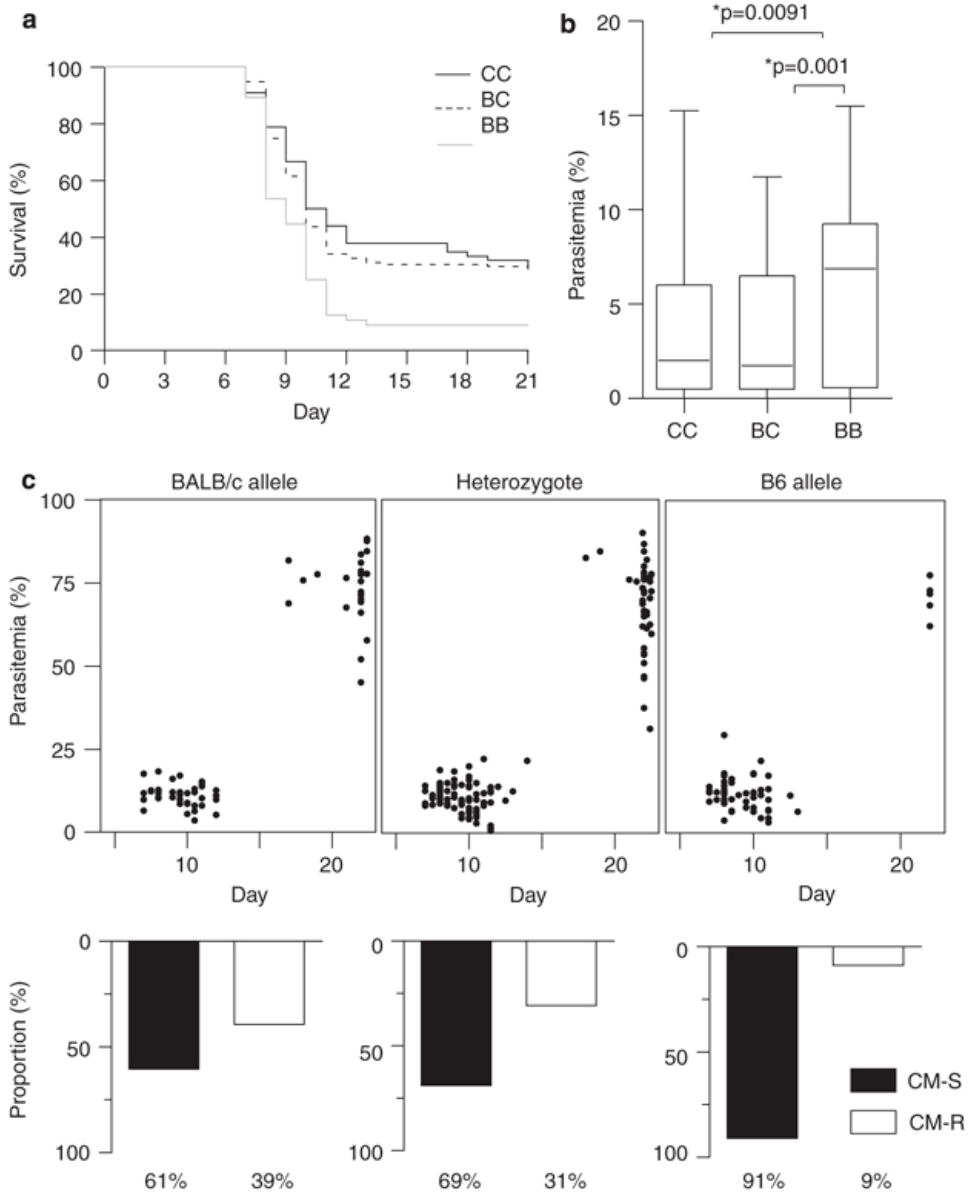



**Figure 2-4.** Effect of haplotype combination at Berr5 on susceptibility to P. berghei-induced cerebral malaria.

Effect of haplotype combination at *Berr5* on survival (a) and day 6 parasitemia (b) in *P. berghei*-infected CB6F2 mice. In panels a and b, 'C' refers to the BALB/c allele at *Berr5* and 'B' refers to the B6 allele. (c) Distribution of CM-S and CM-R phenotypes in CB6F2 mice according to their *Berr5* haplotype. Mice homozygous for the C57BL/6 allele (91%) were CM-S, whereas only 61% of mice carrying the BALB/c allele were CM-S ($P<0.0001$, $\chi^2$-test).

The *Berr5* interval spans 25 Mb and contains 167 annotated genes (Supplementary Table A) that could be reduced to 14 Mb and 84 genes by superimposing the linkage peak obtained using day 6 parasitemia (Figure 2-5a). We analyzed possible modulation of mRNA expression in response to PbA infection for the genes in the minimal *Berr5* interval using transcript profiling information from B6 and BALB/c brains published by Lovegrove et al.[213] Examination of their data showed that the expression of 15 genes within the 95% confidence interval for *Berr5* interval are differentially modulated in brain tissue in response to PbA infection and/or in a strain-specific manner (two-way analysis of variance (ANOVA) with a 1.4-fold cutoff, Bonferroni correction). Figure 2-5b compares the level of expression of these genes in B6 and BALB/c in response to infection (presented as a ratio compared to levels in uninfected BALB/c controls). Of particular interest are the three members of the interferon-induced protein with tetratricopeptide repeats (TPRs; *Ifit*) gene family. By microarray, these three genes show a dramatic upregulation (between 8- and 30-fold) in infected B6 brains compared to BALB/c. To confirm this observation, we performed semiquantitative reverse transcription (RT)-PCR for *Ifit1–3* (Figure 2-6), which verified significant differential induction of the genes controlled by strain-specific genetic factors over the course of infection. In addition, the tumor necrosis factor (TNF) receptor family member-encoding gene, *Fas*, maps within the *Berr5* interval. These two groups of genes constitute excellent candidates for the *Berr5* effect, and thus to explain the differential host response of the B6 and BALB/c strains.

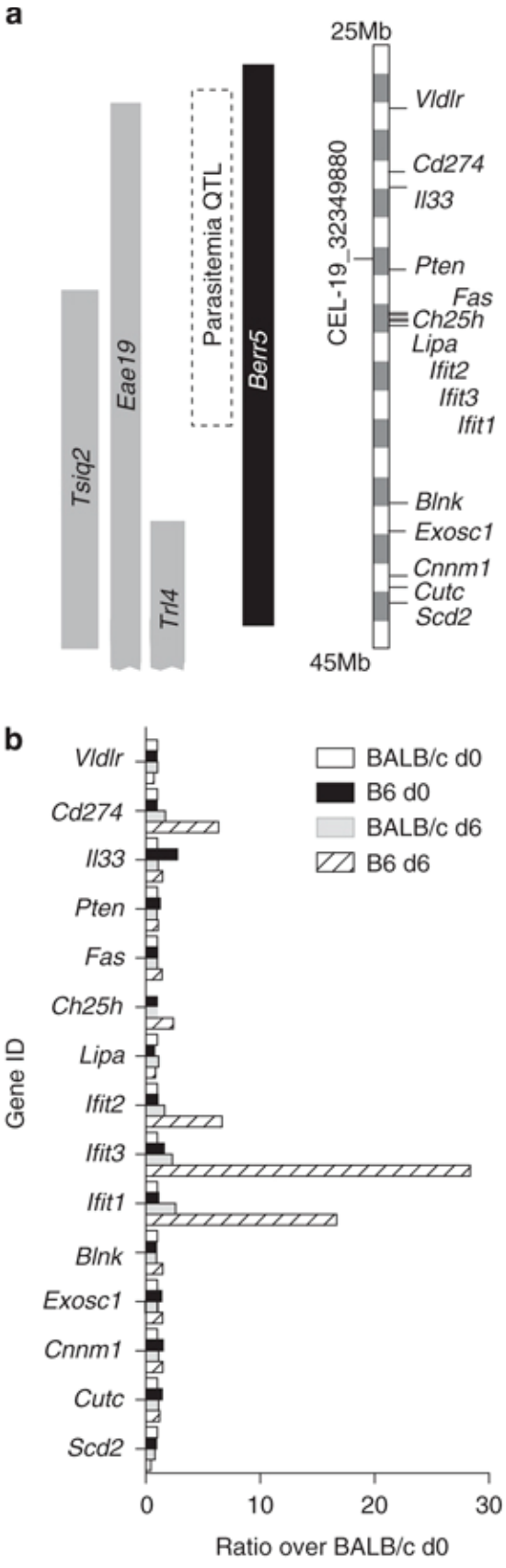

a
25Mb
Vldlr
Cd274
Il33
Pten
Fas
Ch25h
Lipa
Ifit2
Ifit3
Ifit1
Blnk
Exosc1
Cnnm1
Cutc
Scd2
45Mb
CEL-19_32349880
Parasitemia QTL
Berr5
Eae19
Tsiq2
Trl4
b
BALB/c d0
B6 d0
BALB/c d6
B6 d6
Gene ID
Vldlr
Cd274
Il33
Pten
Fas
Ch25h
Lipa
Ifit2
Ifit3
Ifit1
Blnk
Exosc1
Cnnm1
Cutc
Scd2
0
10
20
30
Ratio over BALB/c d0

**Figure 2-5.** Analysis of positional candidate gene expression in the minimum genetic interval for *Berr5* and associated quantitative trait loci.

(a) Approximate position of three previously mapped loci associated with host response to inflammatory or infectious stimuli, including *Tsiq2*[223], *Eae19*[224], *Trl-4*[225] and *Berr5* (this study). The interval mapped using survival is indicated with a black box whereas the suggestive interval determined using day 6 parasitemia is marked in the stippled box. The marker that gave the strongest linkage for all analysis (CEL-19_32349880) is also plotted. Differentially expressed positional candidate genes are also identified along with their chromosomal positions. (b) Among the 167 genes in the *Berr5* interval, the expression of 15 of them in brain was regulated in a C57BL/6J vs BALB/c manner after infection with *P. berghei* (two-way analysis of variance (ANOVA), Bonferroni and Hoch correction with a 1.4-fold cutoff). The relative expression by microarray of each of these 15 genes is shown calculated as a ratio (fold expression) over expression levels measured in uninfected brains from BALB/c mice.

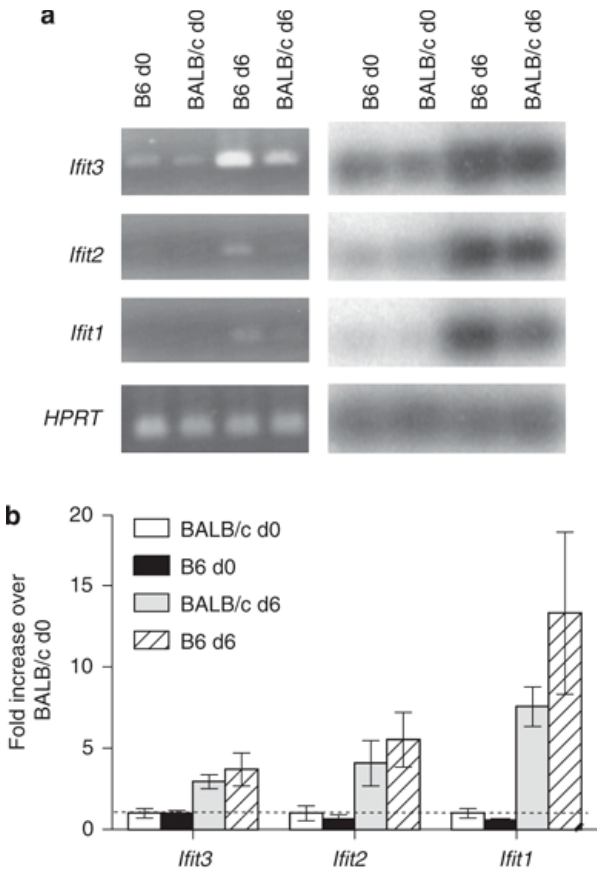



**Figure 2-6. Confirmation of different *Ifit1-3* upregulation in the brains of PbA-infected mice**

Semiquantitative reverse transcription (RT)-PCR was carried out for each transcript in control and PbA-infected brains as described in the Materials and methods section. (a) Amplicons for each target were run on agarose gels and then hybridized with specific [$^{32}$P]dATP-labeled probes. The intensities of the hybridization signals were quantified using a phosphorimager. (b) Each replicate was normalized to *Hprt* and then expressed as a relative ratio over the signal for BALB/c brains at day 0.

## 2.5 Discussion

CM is a complex pathology that involves interactions between host genetics, the Plasmodium parasite and the type and magnitude of the host response to infection (reviewed by Haldar et al.[226]). Although CM symptoms may appear while levels of blood parasitemia are still low, host-driven inflammatory response at the site of infection in the form of increased pro-inflammatory mediators (TNFα, IFNγ), has been proposed to contribute to CM pathology (reviewed by Hunt and Grau[227]). Genetic studies in humans have shown that functional promoter polymorphisms affecting the level of expression of TNFα mRNA can affect the incidence of CM in field studies[228]. Likewise, studies in the PbA mouse model of CM have shown that loss-of-function mutations in pro-inflammatory molecules or their receptors have protective effects against CM[141, 156].

Here, we have studied the differential susceptibility of inbred mouse strains B6 (CM-S) and BALB/c (CM-R) to PbA-induced CM. We have identified a locus on chromosome 19 that we have designated *Berr5*, which influences both survival time (two-part model; LOD=4.69) and the early (day 6) blood parasitemia levels (LOD=3.35). Although maximum LOD scores are obtained for both phenotypes with the same marker (CEL-19_32349880), the strength of the linkage is greater when survival is used as the marker of CM susceptibility. In both cases, B6-derived *Berr5* alleles are associated with increased susceptibility among the CB6F2s. The weaker linkage observed when using parasitemia as a quantitative trait may be the result of the fairly narrow numerical range for this trait, and the asynchronous aspect of the rise of blood parasitemia in infected mice. Therefore, it is possible that survival and parasitemia may be influenced by the same gene but with different penetrance/expressivity or differences in

ascertainment for the two traits. It is also possible that both traits are influenced by distinct but closely linked genes. Additional experimentations will be required to distinguish the two possibilities. Nevertheless, the fact that *Berr5* alleles appear to influence both early parasitemia and early death merits further discussion. One explanation would be that *Berr5* influences the extent PbA replication early during infection, with a threshold effect triggering cerebral symptoms and early death. Although possible, this explanation is incompatible with the observation that all mice surviving the cerebral phase go on to develop high parasitemia without CM. A second explanation would be that *Berr5* has little effect on the early rate of replication, but rather influences the host response to rising parasitemia. In this case, host inflammatory response would be proportional to the extent of early parasitemia (for example, measured at day 6), and would be associated with CM in mice homozygous for B6 alleles at *Berr5* at a much higher rate than mice carrying the BALB/c allele. Finally, it is possible that *Berr5* haplotype somehow influences the balance between the beneficial and immunopathological effects of the early inflammatory response.

There are a total of 167 positional candidates in the estimated genetic interval for *Berr5*. The information in Supplementary Table A shows that 46 of these genes are expressed in the brain and 27 in the hemolymphoid system (22 expressed in both; Mouse Genome Informatics Gene Expression Database; http://www.informatics.jax.org/expression.shtml), two systems known to be involved in CM. Examination of transcript profiling data[213] revealed strain-specific factors and/or PbA infection regulated the expression of 15 of these 167 genes (Figure 2-5; 1.4-fold threshold). When analyzing both the absolute level of infection-induced modulation, and interstrain differences in this modulation between B6 (CM-S) and BALB/c (CM-R), most differences are modest, with the notable exception of the *Isg56*/*Ifit* gene family (*Isg56*/*Ifit1*, *Isg54*/*Ifit2* and

*Isg49*/*Ifit3*) that were much more highly upregulated in B6 mice. The Ifit gene family has three members in mice and four members in humans, each clustered on small chromosomal segments, likely having arisen by two successive gene duplication events. In the mouse, the three *Ifit* genes are each organized into two exons and code for highly similar proteins[229]. IFIT proteins are highly hydrophilic cytoplasmic proteins[230] containing multiple TPR motifs that represent degenerate protein interaction modules. All three mouse *Ifit* genes show two functional interferon-stimulated response element motifs in their proximal promoters (−60 to −100 position), which have a critical function in the regulation of mRNA expression[229]. Expression of *Ifit* genes is part of an early response to infectious agents or products derived from them, and can be induced by many stimuli including viral infection (Sendai virus, herpes simplex virus, human cytomegalovirus, West Nile virus), double-stranded RNA, type I interferon (α, β), lipopolysaccharide (LPS), Toll-like receptor ligands and several others[231]. Their molecular function is not fully understood, but they have been reported to inhibit translation by binding to subunits of the eIF3 complex[231]. In addition, it has been shown that overexpression of *Ifit2* in mouse macrophages (RAW264.7) modulates LPS-induced expression of TNFα, interleukin-6 (IL-6) and MIP-2[232]. These data suggest that Ifit proteins may have a function in early host response to different types of infection, which supports the possibility that they may be responsible for the *Berr5* effect on CM.

Also mapping within the *Berr5* interval is the *Fas* gene. Although the differences in *Fas* mRNA levels noted were not as impressive as those noted for *Ifit* genes, the known functions of the FAS protein (member of the TNF-receptor superfamily) in inflammation and the pathogenesis of various malignancies of the immune system make it an attractive positional candidate for the *Berr5* effect[233]. Hypomorphic $Fas^{lpr}$ mouse mutants show increased resistance to

PbA-induced CM indicating a pathological function for this molecule[144]. We have noted lower Fas RNA expression in PbA-infected brains from BALB/c (CM-R) compared with B6 (CM-S) (Figure 2-5). Interestingly, a functional promoter polymorphism (rs50344448; pst −375) has been detected in the mouse Fas gene: B6 carries an A at that position whereas BALB/c carries a G. The latter allele alters a putative AP-1 binding site[234] and reduces Fas mRNA expression.

Finally, *Berr5* overlaps three previously mapped loci affecting host response to infectious or inflammatory stimuli (Figure 2-5a). The *Trl-4* (tuberculosis resistance; 41–52 cM) locus affects susceptibility to pulmonary tuberculosis and modulates replication of *Mycobacterium tuberculosis* in the lungs of (C57BL/6 × DBA/2)F2 mice[225]. B6 alleles at *Trl-4* are associated with increased resistance to infection in the form of a robust early immune response to a mounting bacillary load. Another QTL overlapping *Berr5* is the *Tsiq2* (T-cell secretion of IL-4) locus (24–43 Mb) that regulates the capacity of spleen CD4+ T cell to secrete IL-4 in response to stimulation[223]. BALB/c alleles at *Tsiq2* are associated with high IL-4 production (Th2 polarization of early T-cell response) whereas B6 alleles are associated with lower IL-4 production (pro-inflammatory, Th1 polarization). Th1 response is characterized by expression of pro-inflammatory mediators including IFNγ, TNFα, lymphotoxin and IL-12 all previously implicated in CM immunopathology[156, 235, 236]. In contrast, IL-4-driven Th2 response is associated with low-level expression of pro-inflammatory cytokines but elevated levels of IL-4, IL-6, IL-10 and TGF-β, which are thought to have an immune-regulatory, protective function against CM[223, 227, 235]. It is interesting to speculate that, by analogy with *Tsiq2*, *Berr5* may modulate CM susceptibility by altering the Th1/Th2 balance after infection with PbA. The third overlapping QTL (*Eae19* (experimental allergic encephalitis 19); 26–53 cM) affects the severity of EAE that is used as a mouse model of multiple sclerosis, a

known inflammatory disease of the central nervous system (CNS). *Eae19* (26–53 cM) is specifically associated with brain demyelination in response to injection of heterologous CNS homogenates, a process thought to involve inflammatory cytokines, demyelinating antibodies and apoptosis of affected cells[224]. It is interesting to speculate that *Trl-4*, *Tsiq2*, *Eae19* and *Berr5* may be pleiotropic manifestations of the same gene effect. The formal identification of the gene responsible for *Berr5* and associated QTLs on this portion of chromosome 19 will involve the creation and characterization of congenic and subcongenic lines to narrow down the physical interval of *Berr5*, as well as the formal evaluation of positional candidates, including the creation and testing of loss-of-function mutations in vivo.

## 2.6 Acknowledgements

Work in PG's lab was supported by a Team Grant (CTP 79842) from the Canadian Institutes of Health Research (CIHR) to K Kain (U Toronto) and PG (McGill University), and by a CIHR operating Grant (FRN79343) to MM Stevenson, and to PG. PG is a James McGill Professor of Biochemistry, and JB is supported by a Doctoral Award from CIHR.

## 2.7 Supplementary Information

Supplementary Table A which lists the 167 genes under the *Berr5* interval.

## Table 2-1. Supplementary Table A. All annotated genes in the 95% confidence interval for *Berr5*

Information on expression in brain and haemolymphoid systems was obtained from the Gene Expression Database (www.jax.informatics.org/expression). Expression differences by microarray was based on analysis of the data sets published by Lovegrove and colleagues [[213], see Materials and Methods].

| Chr | Genome Coordinates (Mb) (strand) | Gene ID | Brain | Haemo-lymphoid | Detection Technique | Microarray Differences |
|---|---|---|---|---|---|---|
| 19 | 25.68 – 25.69 (+) | Dmrt3 | present, regional | | RNA in situ | |
| 19 | 25.74 – 25.75 (+) | Dmrt2 | Present | | RNA in situ | |
| 19 | 26.67 – 26.85 (+) | Smarca2 | Present | present | Western blot | |
| 19 | 27.28 – 27.32 (+) | Vldlr | Present | absent | Northern blot | yes |
| 19 | 27.39 – 27.4 (+) | Kcnv2 | | | | |
| 19 | 27.83 – 28.08 (-) | Rfx3 | present, regional | | RNA in situ | |
| 19 | 28.33 – 28.61 (-) | Glis3 | Moderate | not specified | Northern blot | |
| 19 | 28.9 – 28.98 (+) | Slc1a1 | Strong | trace | RT-PCR | |
| 19 | 28.97 – 29.03 (-) | 4430402I18 Rik | | | | |
| 19 | 29.03 – 29.03 (+) | Ppapdc2 | | | | |
| 19 | 29.06 – 29.08 (+) | Cdc37l1 | | | | |
| 19 | 29.09 – 29.11 (-) | Ak3 | | | | |
| 19 | 29.17 – 29.21 (+) | Rcl1 | | | | |
| 19 | 29.2 – 29.2 (+) | Mirn101b | | | | |
| 19 | 29.32 – 29.38 (+) | Jak2 | Strong | moderate | RNA in situ | |
| 19 | 29.39 – 29.39 (-) | Insl6 | | | | |
| 19 | 29.4 – 29.4 (-) | Rln1 | | | | |
| 19 | 29.43 – 29.45 (+) | Cd274 | | | | yes |
| 19 | 29.48 – 29.54 (+) | Pdcd1lg2 | | | | |
| 19 | 29.68 – 29.71 (-) | Ermp1 | | | | |
| 19 | 29.76 – 29.77 (+) | Mlana | | | | |
| 19 | 29.87 – 29.88 (-) | Ranbp6 | | | | |
| 19 | 29.99 – 30.03 (+) | Il33 | | | | yes |
| 19 | 30.07 – 30.07 (+) | Trpd52l3 | | | | |
| 19 | 30.1 – 30.16 (+) | Uhrf2 | present, regional | | RNA in situ | |
| 19 | 30.16 – 30.24 (-) | Gldc | Present | | RNA in situ | |
| 19 | 30.23 – 30.23 (+) | EG625917 | | | | |
| 19 | 30.3 – 30.31 (+) | Mbl2 | | | | |
| 19 | 30.61 – 30.62 (-) | Dkk1 | Present | weak | In situ reporter | |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | (knock in) | |
| 19 | 30.63 – 31.83 (-) | Prkg1 | | | | |
| 19 | 31.15 – 31.15 (+) | Cstf2t | | | | |
| 19 | 31.93 – 32.02 (+) | A1cf | Present | | RNA in situ | |
| 19 | 32.05 – 32.17 (-) | Asa2 | | | | |
| 19 | 32.19 – 32.45 (-) | Sgms1 | | | | |
| 19 | 32.55 – 32.58 (+) | Minpp1 | | | | |
| 19 | 32.69 – 32.73 (+) | Papss2 | Present | | RNA in situ | |
| 19 | 32.74 – 32.78 (-) | Atad1 | | | | |
| 19 | 32.82 – 32.89 (+) | Pten | | present, regional | RNA in situ | yes |
| 19 | 34.03 – 34.04 (+) | Lipf | | | | |
| 19 | 34.07 – 34.11 (+) | Lipk | | | | |
| 19 | 34.13 – 34.15 (+) | Lipn | | | | |
| 19 | 34.17 – 34.19 (+) | Lipm | | | | |
| 19 | 34.19 – 34.23 (-) | Ankrd22 | | | | |
| 19 | 34.26 – 34.31 (+) | Stambpl1 | | | | |
| 19 | 34.31 – 34.32 (-) | Acta2 | not specified | blood vessels | RT-PCR | |
| 19 | 34.36 – 34.39 (+) | Fas | Moderate | moderate | RT-PCR | yes |
| 19 | 34.54 – 34.54 (-) | Ch25h | | | | yes |
| 19 | 34.56 – 34.59 (-) | Lipa | Strong | strong | Northern blot | yes |
| 19 | 34.62 – 34.64 (+) | Ifit2 | Present | absent | RNA in situ | yes |
| 19 | 34.65 – 34.65 (+) | Ifit3 | Present | absent | RNA in situ | yes |
| 19 | 34.71 – 34.72 (+) | Ifit1 | Present | absent | RNA in situ | yes |
| 19 | 34.73 – 34.81 (-) | Slc16a12 | | | | |
| 19 | 34.88 – 34.95 (-) | Pank1 | | | | |
| 19 | 34.89 – 34.89 (-) | Mirn107 | | | | |
| 19 | 34.99 – 35.04 (+) | Kif20b | | | | |
| 19 | 36.03 – 36.12 (-) | Htr7 | Present | present | In situ reporter (knock in) | |
| 19 | 36.15 – 36.17 (+) | Rpp30 | | | | |
| 19 | 36.18 – 36.18 (-) | Ankrd1 | | | | |
| 19 | 36.41 – 36.52 (+) | Pcgf5 | | | | |
| 19 | 36.62 – 36.69 (+) | Hectd2 | | | | |
| 19 | 36.8 – 36.8 (-) | Pp1r3c | | | | |
| 19 | 36.83 – 36.96 (+) | Tnks2 | | present | RNA in situ | |
| 19 | 36.98 – 36.98 (-) | Fgfbp3 | | | | |
| 19 | 36.99 – 37.08 (+) | Btaf1 | | | | |
| 19 | 37.09 – 37.27 (-) | Cpeb3 | | | | |
| 19 | 37.27 – 37.29 (+) | March5 | | | | |
| 19 | 37.34 – 37.4 (-) | Ide | Present | present, regional | RNA in situ | |
| 19 | 37.44 – 37.49 (+) | Kif11 | Present | present | RNA in situ | |
| 19 | 37.5 – 37.51 (+) | Hhex | Present | weak | Northern blot | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 19 | 37.62 – 37.75 (+) | Exoc6 | | | | |
| 19 | 37.75 – 37.76 (+) | Cyp26c1 | present, regional | absent | RNA in situ | |
| 19 | 37.76 – 37.77 (+) | Cyp26a1 | present, regional | present | RNA in situ | |
| 19 | 37.85 – 37.85 (+) | EG668251 | | | | |
| 19 | 37.94 – 37.94 (+) | EG209281 | | | | |
| 19 | 37.96 – 38.11 (-) | Fer1l3 | | | | |
| 19 | 38.12 – 38.14 (+) | Cep55 | | | | |
| 19 | 38.16 – 38.18 (+) | Gpr120 | | | | |
| 19 | 38.18 – 38.19 (-) | Rbp4 | | | | |
| 19 | 38.2 – 38.25 (+) | Pde6c | | | | |
| 19 | 38.33 – 38.37 (+) | Lgi1 | Present | | RT-PCR | |
| 19 | 38.46 – 38.47 (+) | Tmem20 | | | | |
| 19 | 38.59 – 38.85 (+) | Plce1 | Present | weak, regional | RT-PCR | |
| 19 | 38.85 – 38.88 (-) | Noc3l | Present | present, thymus | RT-PCR | |
| 19 | 38.9 – 38.98 (+) | Tbc1d12 | | | | |
| 19 | 39 – 39.04 (+) | Hells | Present | present, thymus | Northern blot | |
| 19 | 39.07 – 39.11 (+) | Cyp2c55 | | | | |
| 19 | 39.13 – 39.16 (+) | Cyp2c65 | | | | |
| 19 | 39.18 – 39.24 (+) | Cyp2c66 | | | | |
| 19 | 39.29 – 39.3 (+) | Cyp2c53-ps | | | | |
| 19 | 39.34 – 39.38 (+) | Cyp2c29 | Absent | present | RNA in situ | |
| 19 | 39.44 – 39.52 (-) | Cyp2c38 | | | | |
| 19 | 39.56 – 39.62 (+) | Cyp2c39 | | | | |
| 19 | 39.66 – 39.7 (-) | Cyp2c67 | | | | |
| 19 | 39.74 – 39.79 (-) | Cyp2c68 | | | | |
| 19 | 39.82 – 39.87 (-) | Cyp2c40 | | | | |
| 19 | 40.05 – 40.07 (+) | Cyp2c37 | | | | |
| 19 | 40.09 – 40.13 (-) | Cyp2c54 | | | | |
| 19 | 40.14 – 40.17 (+) | Cyp2c50 | | | | |
| 19 | 40.21 – 40.24 (-) | Cyp2c70 | | | | |
| 19 | 40.27 – 40.32 (-) | Pdlim1 | Present | moderate | RT-PCR | |
| 19 | 40.35 – 40.57 (-) | Sorbs1 | | | | |
| 19 | 40.6 – 40.64 (-) | Aldh18a1 | | | | |
| 19 | 40.65 – 40.67 (-) | Tctn3 | | | | |
| 19 | 40.76 – 40.79 (+) | Entpd1 | | | | |
| 19 | 40.8 – 40.88 (+) | Cc2d2b | | | | |
| 19 | 40.88 – 40.9 (+) | Ccnj | | | | |
| 19 | 40.95 – 40.97 (+) | Zfp518 | | | | |
| 19 | 40.98 – 41.05 (-) | Blnk | Weak | strong | Northern blot | yes |
| 19 | 41.08 – 41.11 (+) | Dntt | | weak, regional | Northern blot | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 19 | 41.12 – 41.13 (-) | Tmem10 | | | | |
| 19 | 41.14 – 41.26 (-) | Tll2 | | | | |
| 19 | 41.26 – 41.32 (-) | Tm9sf3 | | | | |
| 19 | 41.33 – 41.44 (-) | Pik3ap1 | | | | |
| 19 | 41.58 – 41.62 (+) | Lcor | present, regional | | RNA in situ | |
| 19 | 41.63 – 41.64 (+) | Gm340 | | | | |
| 19 | 41.65 – 41.65 (+) | Al606181 | | | | |
| 19 | 41.65 – 41.8 (-) | Slit1 | Present | | RNA in situ | |
| 19 | 41.82 – 41.86 (-) | Arhgap19 | | | | |
| 19 | 41.88 – 41.89 (+) | Frat1 | Weak | weak | In situ reporter (knock in) | |
| 19 | 41.9 – 41.9 (-) | Frat2 | present, regional | present, regional | RNA in situ | |
| 19 | 41.92 – 41.95 (-) | Rrp12 | | | | |
| 19 | 41.97 – 41.97 (+) | Pgam1 | Present | present | RNA in situ | |
| 19 | 41.98 – 41.99 (-) | Exosc1 | | | | yes |
| 19 | 41.99 – 42 (+) | Zdhhc16 | | | | |
| 19 | 42 – 42.03 (-) | Mms19 | present, regional | | RNA in situ | |
| 19 | 42.04 – 42.09 (+) | Ubtd1 | | | | |
| 19 | 42.09 – 42.1 (+) | Ankrd2 | | | | |
| 19 | 42.14 – 42.18 (+) | Pi4k2a | | | | |
| 19 | 42.18 – 42.18 (-) | Avpi1 | | | | |
| 19 | 42.2 – 42.21 (+) | Marveld1 | | | | |
| 19 | 42.22 – 42.25 (+) | Zfyve27 | | | | |
| 19 | 42.25 – 42.26 (-) | Sfrp5 | present, regional | | RNA in situ | |
| 19 | 42.34 – 42.48 (-) | Crtac1 | | | | |
| 19 | 42.65 – 42.67 (-) | Loxl4 | | | | |
| 19 | 42.81 – 42.83 (-) | Hps1 | | | | |
| 19 | 42.96 – 43.07 (-) | Hpse2 | | | | |
| 19 | 43.49 – 43.55 (+) | Cnnm1 | | | | yes |
| 19 | 43.55 – 43.58 (-) | Got1 | Present | present, regional | Northern blot | |
| 19 | 43.67 – 43.67 (+) | Nkx2-3 | | | | |
| 19 | 43.72 – 43.73 (-) | Slc25a28 | Weak | weak | Northern blot | |
| 19 | 43.74 – 43.79 (+) | Entpd7 | | | | |
| 19 | 43.79 – 43.81 (-) | Cox15 | | | | |
| 19 | 43.81 – 43.82 (+) | Cutc | | | | yes |
| 19 | 43.84 – 43.89 (+) | Abcc2 | | | | |
| 19 | 43.9 – 43.97 (-) | Dnmbp | | | | |
| 19 | 44.01 – 44.04 (-) | Cpn1 | | | | |
| 19 | 44.06 – 44.08 (-) | Cyp2c44 | | | | |
| 19 | 44.09 – 44.12 (-) | Erlin1 | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 19 | 44.13 – 44.16 (-) | Chuk | Strong | strong | RT-PCR | |
| 19 | 44.16 – 44.19 (-) | Cwf19l1 | | | | |
| 19 | 44.19 – 44.2 (-) | Bloc1s2 | | | | |
| 19 | 44.2 – 44.25 (-) | Pkd2l1 | | | | |
| 19 | 44.26 – 44.3 (+) | Scd3 | | | | |
| 19 | 44.35 – 44.36 (+) | Scd2 | Present | | RNA in situ | yes |
| 19 | 44.39 – 44.4 (+) | Scd4 | | | | |
| 19 | 44.45 – 44.46 (-) | Scd1 | | | | |
| 19 | 44.55 – 44.57 (+) | Wnt8b | present, regional | absent | RNA in situ | |
| 19 | 44.57 – 44.6 (-) | Sec31b | | | | |
| 19 | 44.6 – 44.61 (-) | Ndufb8 | | | | |
| 19 | 44.62 – 44.63 (+) | Hif1an | | | | |
| 19 | 44.81 – 44.89 (+) | Pax2 | Present | | In situ reporter (knock in) | |
| 19 | 45.04 – 45.06 (+) | Sema4g | Present | | Northern blot | |
| 19 | 45.06 – 45.06 (-) | Mrpl43 | | | | |
| 19 | 45.06 – 45.07 (+) | Peo1 | | | | |
| 19 | 45.07 – 45.08 (+) | Lzts2 | | | | |
| 19 | 45.09 – 45.1 (-) | Pdzd7 | | | | |
| 19 | 45.1 – 45.11 (+) | Sfxn3 | | | | |

# Chapter 3 *Irf8*-Regulated Genomic Responses Drive Pathological inflammation during Experimental Cerebral Malaria

## Preface to Chapter 3

Previous work in the lab had identified BXH2 mice as extremely susceptible to *Mycobacterium bovis* BCG and *M. tuberculosis* which was attributed to a *de novo* mutation in the transcription regulator Interferon Regulatory Factor 8 ($Irf8^{R294C}$). Inflammation, particularly *Ifng* and *Il12p40*, is important early during a tubercular infection for bacterial control.

Evidence from the work of others and our own QTL mapping in Chapter 2 paint experimental cerebral malaria as a disease of pathological neuroinflammation. To determine if IRF8 mediated mechanisms contribute towards this pathology, we infected BXH2 mice with *P. berghei* and found that this did indeed seem to be the case. In order to further explore the mechanism of resistance in BXH2 mice, we used expression profiling. This indicated a number of IRF8-dependent indicators of pathology placing *Irf8* as a central regulator of pathological inflammation.

## 3.1 Abstract

Interferon Regulatory Factor 8 (IRF8) is required for development, maturation and expression of anti-microbial defenses of myeloid cells. BXH2 mice harbor a loss-of-function allele at *Irf8* (*Irf8*$^{R294C}$) that causes susceptibility to infection with intracellular pathogens, including *Mycobacterium tuberculosis*. We report that BXH2 are completely resistant to the development of cerebral malaria (ECM) following *Plasmodium berghei* ANKA infection. Comparative transcriptional profiling of brain RNA as well as chromatin immunoprecipitation and high-throughput sequencing (ChIP-seq) was used to identify IRF8-regulated genes whose expression is associated with pathological acute neuroinflammation. Genes increased by infection were strongly enriched for IRF8 binding sites, suggesting that IRF8 acts as a transcriptional activator in inflammatory programs. These lists were enriched for myeloid-specific pathways, including interferon responses, antigen presentation and Th1 polarizing cytokines. We show that inactivation of several of these downstream target genes (including the *Irf8* transcription partner *Irf1*) confers protection against ECM. ECM-resistance in *Irf8* and *Irf1* mutants is associated with impaired myeloid and lymphoid cells function, including production of IL12p40 and IFNg. We note strong overlap between genes bound and regulated by IRF8 during ECM and genes regulated in the lungs of *M. tuberculosis* infected mice. This IRF8-dependent network contains several genes recently identified as risk factors in acute and chronic human inflammatory conditions. We report a common core of IRF8-bound genes forming a critical inflammatory host-response network.

## 3.2 Author Summary

Cerebral malaria is a severe and often lethal complication from infection with *Plasmodium falciparum* which is driven in part by pathological host inflammatory response to parasitized red cells' adherence in brain microvasculature. However, the pathways that initiate and amplify this pathological neuroinflammation are not well understood. As susceptibility to cerebral malaria is variable and has been shown to be partially heritable, we have studied this from a genetic perspective using a mouse model of infection with *P. berghei* which induces experimental cerebral malaria (ECM). Here we show that mice bearing mutations in the myeloid transcription factor IRF8 and its heterodimerization partner IRF1 are completely resistant to ECM. We have identified the genes and associated networks that are activated by IRF8 during cerebral malaria. Loss-of-function mutations of several IRF8 targets are also shown to be protective. Parallel analysis of lungs infected with *Mycobacterium tuberculosis* show that IRF8-associated core pathways are also engaged during tuberculosis where they play a protective role. This contrast illustrates the balancing act required by the immune system to respond to pathogens and highlights a lynchpin role for IRF8 in both. Finally, several of the genes in these networks have been individually associated with chronic or acute inflammatory conditions in humans.

## 3.3 Introduction

IRF8 is a member of the Interferon Regulatory Factor (IRF) family of transcription factors that plays a central role in interferon signaling, response to infection and maturation of dendritic cells (DCs) and other myeloid lineages[237, 238]. IRF8 regulates elements of constitutive gene expression in both myeloid and lymphoid cells and can also activate or suppress pathogen responsive transcription programs following exposure to type I or type II interferons, lipopolysaccharides, and a range of additional microbial products [237, 238]. Heterodimerization of IRF8 with members of the IRF (IRF1, IRF4) or ETS (PU.1) families leads to DNA binding and transcriptional regulation of target genes containing ISRE (GAAAnnGAAA) and EICE-type (GGAAnnGAAA) canonical motifs in their promoters [239-241].

During hematopoiesis, IRF8 promotes differentiation of myeloid progenitors towards the mononuclear phagocyte lineages (monocytes, macrophages, DCs) by acting as an antagonist of the polymorphonuclear granulocyte pathway [242-245]. This is accomplished through positive regulation of pro-apoptotic signals (*Cdkn2b, Nf1, Bax*), and negative regulation of pro-survival signals (*Bcl2*, *Bcl-XL*) in $CD11b^{+}$ myeloid precursors [246-248]. Mice bearing either a targeted null allele [243] or a severely hypomorphic mutation ($Irf8^{R294C}$) [245] show either complete absence of all $CD11c^{+}CD8\alpha^{+}$ dendritic cell subsets ($Irf8^{-/-}$) while BXH2 mice ($Irf8^{R294C}$) retain some plasmacytoid DCs ($B220^{+}/CD11c^{low}$) [249-251]. Furthermore, both mutants display a chronic myeloid leukemia-like phenotype dominated by expansion of $Gr1^{+}/CD11b^{+}$ immature myeloid cells [242-245]. Additionally, IRF8 is required for B lymphocyte lineage specification, commitment, and differentiation, including expression of

certain biochemical pathways that play a key role in the specialized functions of these antigen-presenting cells (APCs) [238, 241, 252].

During infection, IRF8 activates antimicrobial defenses in myeloid cells, propagates pro-inflammatory signals and is required to amplify early immune responses by these cells. IRF8 activity is essential to express *Il12p40*, *Il12p35* and *Il18* in response to IFNγ [242, 253-257] and is therefore required for APC-mediated Th1 polarization of early immune responses [237, 238]. *Irf8*-deficient mice, then, display defective Th1 responses (absence of antigen specific and IFNγ producing $CD4^+$ T cells) [258], enhanced Th17 responses [259], and are susceptible to *in vivo* infection with many intracellular pathogens [254, 260-262] including tuberculosis [258] and blood-stage malaria [124]. *Irf8*-deficient macrophages are also extremely susceptible to *ex vivo* infection by intracellular pathogens [254, 263-265]. Studies using genome-wide transcript profiling, chromatin immunoprecipitation [239-241] and individual gene targets [237] show that IRF8 regulates multiple aspects of antimicrobial defenses in mononuclear phagocytes. These include antigen recognition and processing, phagosome maturation, production of lysosomal enzymes and other cytoplasmic microbicidal pathways.

*IRF8* mutations in humans cause pathologies remarkably similar to those observed in *Irf8* mutant mice, affecting the myeloid compartment in general and DCs in particular [266]. We have shown that homozygosity for a DNA-binding incompetent and transcriptionally inactive human *IRF8* mutant variant ($IRF8^{K108E}$) is associated with severe recurrent perinatal bacterial and fungal infections, with absence of blood monocytes and DCs, and a lack of IL-12 and IFNγ production following *in vitro* stimulation of blood cells [266]. We also reported a milder autosomal dominant form of IRF8-deficiency ($IRF8^{T80A}$) in two patients suffering from *Mendelian susceptibility to mycobacterial disease* (MSMD) with recurrent

episodes of mycobacterial infections following perinatal vaccination with *M. bovis* BCG. These patients showed selective depletion of the $CD11c^+$ $CD1c^+$ DC subset and impaired production of IL-12 by circulating peripheral blood cells. The $IRF8^{T80A}$ variant displays negative dominance and can suppress the trans-activation potential of wild type IRF8 for known transcriptional targets such as *NOS2* and *IL-12* [266]. Interestingly, T cells raised in the absence of an intact DC compartment in patient $IRF8^{K108E}$ show impaired function (production of IFNg and other cytokines in response to non-specific stimuli) [266]. Finally, recent results from genome-wide association studies (GWAS) have pointed to a role for *IRF8* in the complex genetic etiology of several human inflammatory diseases. Strong and independently replicated associations have been detected between polymorphic variants within or near the *IRF8* gene for systemic lupus erythematosus [267], ulcerative colitis [268], Crohn's disease [269, 270], and multiple sclerosis (MS) [271-273]. Details of one study in MS patients showed that the *IRF8* susceptibility allele (rs17445836) was associated with higher expression of both *IRF8* mRNA and downstream IFNβ-responsive targets [271].

We have used an experimental model of murine cerebral malaria (ECM) induced by infection with *Plasmodium berghei* ANKA (PbA) to investigate the role of *Irf8* in pathological inflammation. In this model, adherence of PbA-infected erythrocytes to brain microvasculature leads to acute and rapidly fatal neuroinflammation. Symptoms such as tremors, ataxia and seizures appear between d5 and d8 in susceptible mice, progressing to morbidity and death within hours. *Irf8*-deficient BXH2 mice ($Irf8^{R294C}$) did not develop any neurological symptoms and were found to be completely resistant to PbA-induced ECM. Comparative transcript profiling of PbA-infected wild-type C57BL/6 (B6) and BXH2 mice, together with IRF8 chromatin immunoprecipitation coupled to high-throughput DNA sequencing (ChIP-seq) have identified a list of key *Irf8* targets

whose expression is associated with acute ECM-associated neuroinflammation. This list has substantial overlap with genes activated in mouse lungs following infection with *M. tuberculosis* (Mtb), suggesting a shared core inflammatory response to infection that is protective against Mtb infection but deleterious in ECM. These studies identify IRF8 as a key regulator of acute neuroinflammation during ECM and a major inflammatory mediator.

## 3.4 Materials and Methods

*Ethics statement:* All mice were kept under specific pathogen free conditions and handled according to the guidelines and regulations of the Canadian Council on Animal Care. Mice experimentation protocol was approved by the McGill Facility Animal Care Committee (protocol number: 5287).

*Mice:* C57BL/6J (B6), BXH2, Il12p40$^{-/-}$, Irf1$^{-/-}$, and Isg15$^{-/-}$ mutant mice were obtained from The Jackson Laboratory (Bar Harbor, ME). Stat1$^{-/-}$ mutant mice were purchased from Taconic Farms (Germantown, NY). Ifng$^{-/-}$ deficient mice were obtained from Dr. M. M. Stevenson (Montreal General Hospital Research Institute), Ifit1$^{-/-}$ mutant mice were obtained from Dr. M. Diamond (Washington University School of Medicine, St-Louis), Irgm1$^{-/-}$ from Dr. J. D. MacMicking (Yale, New Haven, CT) and Nlrc4$^{-/-}$ from Millenium Pharmaceuticals, Inc. and Dr. R. A. Flavell (Yale, New Haven, CT). [BXH2 X B6]F2 mice were generated by inter-crossing [BXH2 X B6]F1 mice.

*Parasites and infection: P. berghei ANKA* (PbA) was obtained from the Malaria Reference and Research Reagent Resource Centre (MR4), and was stored frozen at −80°C. Prior to experimental infections, PbA was passaged in B6 mice until peripheral blood parasitemia levels reached 3-5%, at which point animals were euthanized by $CO_2$ inhalation, exsanguinated and an infectious stock was prepared. All experimental infections were done via intraperitoneal (i.p.) injection of $10^6$ parasitized red blood cells (pRBC). Blood parasitemia was monitored during infection by microscopic examination of thin-blood smears stained with Diff-Quick (Dade Behring, Newark, DE, USA). The appearance of neurological symptoms (shivering, tremors, ruffled fur, seizures, paralysis) associated with

cerebral malaria (ECM) was monitored closely, and affected animals were immediately sacrificed as previously described [191]. Survival curves were compared using Kaplan-Meier statistics.

*Evans Blue dye extravasation assay:* To monitor the integrity of the blood brain barrier during experimental ECM, groups of control and PbA-infected B6 and BXH2 mice were injected intraperitoneally with 0.2 ml of 1% Evan's Blue dye (E2129; Sigma-Aldrich, Oakville, ON, Canada) in sterile phosphate-buffered saline (PBS) on d7 and d16 (BXH2 only) post-infection (n=3 mice/condition). The dye was allowed to circulate for 1h, then the mice were sacrificed by $CO_2$ inhalation, perfused with PBS and the brains were dissected and photographed.

*Cytokine determination:* B6, BXH2, [BXH2 X B6]F1 and *Irf1*$^{-/-}$ mice were infected with PbA and sacrificed at day 6 post-infection. Whole blood was collected by cardiac puncture and serum separated by centrifugation. Levels of circulating IL12p40 and IFNγ were measured using a commercially available ELISA kit, according to manufacturer's instructions (Biolegend and eBioscience, respectively). Spleens from PbA-infected mice were collected, single-cell suspensions prepared and re-suspended in complete RPMI. $4x10^6$ cells were plated in 48-well tissue culture plates and were stimulated with PMA/Ionomycin (eBioscience) or IL12p70 (Biolegend) for 48 hours. Culture supernatants were collected and assayed for IL12p40 and IFNγ production.

*Flow Cytometry Analysis:* Naïve or PbA infected B6, BXH2, and [BXH2 X B6]F1 mice were sacrificed, exsanguinated and perfused with PBS containing 2mM EDTA. Spleens were collected, single-cell suspensions prepared and

resuspended in FACS buffer (PBS, 2% FBS, 0.02% $NaN_3$). Infiltrating brain leukocytes were enriched by Percoll gradient centrifugation, as previously described [274]. Briefly, brains were gently disrupted using a dounce homogenizer, and then separated over discontinuous 70/30% Percoll gradients. Cells from the interface were collected, washed and resuspended in FACS buffer. Cells were surface stained for 30 minutes in the dark at 4˚C with the following cocktails: APC-eFluor780 anti-CD45 (30-F11), APC anti-TCRβ (H57-597), PE anti-CD4 (GK1.5), FITC anti-CD8a (53-6.7) to stain lymphoid cells, or APC-eFluor780 anti-CD45 (30-F11), PE-Cy7 anti-CD11b (M1/70), PE anti-Ly6C (HK1.4), BV421 anti-Ly6G (1A8, Biolegend), eFluor660 anti-F4/80 (BM8), FITC anti-MHCII (M5/114.15.2), and PerCP-Cy5.5 anti-CD11c (N418) to stain myeloid cells. Acquisition was performed using an eight-color FACS Canto II flow cytometer (BD Biosciences) and data analyzed using FlowJo software (Tree Star). For the spleen, between $5x10^4$ and $10^5$ live cells were acquired; for the brain, a maximum number of cells was acquired, but did not exceed $5x10^3$. Cell aggregates were gated out based on the forward scatter (FSC)-height versus FSC-area plot and the live cell gate established based on the side scatter (SSC)-area versus FSC-area. Leukocytes were gated as $CD45^+$ cells in the spleen and as $CD45^{hi}$ cells in the brain; microglia were excluded based on $CD45^{low}$ staining. Data were expressed as the percentage of total $CD45^+$ cells or absolute numbers according to total splenocyte numbers (Figure S2B). All antibodies are from eBioscience, unless otherwise stated.

*Transcript Profiling:* Whole brains were dissected from B6 and BXH2 mice either prior to (d0) or d7 post infection (n = 3/condition). Total brain RNA was isolated using TRIzol reagent (Invitrogen, Burlington, Canada) according to the manufacturer's instructions, followed by further purification with Rneasy columns

(Qiagen, Toronto, Canada) and hybridized to Illumina MouseWG-6 v2.0 microarrays at Genome Quebec Innovation Centre, Montreal, Canada. Unsupervised principal components analysis was done in R, using the lumi [275] package to transform with vst (variance stabilizing transformation) and to perform quartile normalization. For other analyses, microarray expression data was log2-transformed, median normalized and analyzed using GeneSifter (Geospiza) software. Groups were compared using either pairwise t-tests (≥ 2-fold cutoff, Benjamini-Hochberg corrected $p_{adj}$-values < 0.05) or two-factor ANOVA (≥ 2-fold cutoff, Benjamini-Hochberg corrected $p_{adj}$-values < 0.05) to identify genes whose expression is modulated in a strain-dependent, infection-dependent and/or interactive fashion. Lists of genes that were differentially expressed were clustered according to fold change using Multi Experiment Viewer [276]. Raw data can be accessed through the Gene Expression Omnibus.

*Chromatin immunoprecipitation (ChIP):* The J774 mouse macrophage cell line was grown to 80% confluence in complete Dulbecco's modified Eagle's medium (DMEM). The cells plated in 150mm tissue culture-grade Petri dishes (Corning Inc., Corning, NY) were treated with 400U/ml IFNγ (Cell science, Canton, MA) and CpG DNA oligonucleotides (5'-TCCATGACGTTCCTGACGTT-3') for 3h. Chromatin immunoprecipitations were performed as previously described with few modifications [277]. Briefly, treated cells were crosslinked for 10min at 20$^{o}$C with 1% formaldehyde in culture medium. Crosslink was stopped with ice-cold PBS containing 0.125M glycine for 5 min. Nuclei were prepared and chromatin was sonicated with a Branson Digital Sonifier (Branson Ultrasonics, Danbury, CT) to an average size of 250bp. Sonicated chromatin was incubated overnight on a rotating platform at 4ºC with a mixture of 20µl Protein A and 20µl Protein G Dynabeads (Invitrogen, Carlsbad, CA) pre-bound with 6µg of normal goat IgG

(sc-2028) or IRF8 (sc-6058x) antibodies (Santa Cruz Biotechnologies, Santa Cruz, CA). Immune complexes were washed sequentially for 2 min at room temperature with 1ml of the following buffers: Wash B (1% Triton X-100, 0.1% SDS, 150 mM NaCl, 2 mM EDTA, 20 mM Tris-HCl pH 8), Wash C (1% Triton X-100, 0.1% SDS, 500 mM NaCl, 2 mM EDTA, 20 mM Tris-HCl pH 8), Wash D (1% NP-40, 250 mM LiCl, 1 mM EDTA, 10 mM Tris-HCl pH 8), and TEN buffer (50 mM NaCl, 10 mM Tris-HCl pH 8, 1 mM EDTA). After decrosslinking, the DNA was purified with QIAquick PCR purification columns following manufacturer's procedure (Qiagen, Mississauga, Ca). IRF8 ChIP efficiency relative to the IgG control was assessed by qPCR using the Perfecta SYBR green PCR kit (Quanta Bioscience, Gaithersburg, MD) for known IRF8 binding sites [239] (oligonucleotide sequences are available upon requests).

*ChIP-seq preparation and analysis:* A total of 8 independent ChIPs were pooled for each condition (IRF8 and IgG). Libraries and flow cells were prepared following Illumina's recommendation (Illumina, San Diego, CA), with a size selection step targeting fragments between 250 and 500bp. The ChIP libraries were sequenced on Illumina HiSeq 2000 sequencer. The sequencing yielded 86 and 79 million 50bp sequence reads for IgG control and IRF8 samples, respectively. The reads were mapped to the mouse mm9 genome assembly using Bowtie with the following parameters: -t –solexa1.3-qual –sam –best mm9 [278]. The mapping efficiency was 91.7% for IgG and 91.9% for IRF8 samples. To identify IRF8 binding peaks, we used the MACS 1.4.1 peak finder with the following parameters: --bw 250 –mfold 7,30 –pvalue 1e-5 –g mm [279]. This analysis yielded 11216 genomic regions bound by IRF8 with p-values under the threshold of $10^{-5}$. The genes identified as affected by PbA infection in the expression profiling experiment were queried for the presence IRF8 binding

peaks in a 20kb interval around the gene transcription start site (TSS). This analysis was also performed for all the genes represented on the Illumina mouse WG-6 v2.0 array used in the microarray experiments, to assess the background association of IRF8 peaks with surrounding genes (Figure 3-3B).

*Gene network analysis:* The list of all genes differentially regulated in B6 mice during infection (B6 d7/d0 pairwise) was uploaded into Ingenuity Pathway Analysis and networks were generated based on known direct or indirect interactions from published reports and the IPA databases. Seventeen networks were constructed and the three most significant were re-drawn in Adobe Illustrator CS4 14.0.0 (Adobe Systems Inc.). IRF8 binding sites from ChIP-seq data were cross-referenced and genes were colored according to their pairwise fold change during infection for each strain (Figure S1).

A second set of networks was generated according the same procedure using the list of 53 genes detailed in Table S1 possessing at least one IRF8 binding site and up-regulated by both PbA and Mtb infection (Figure 3-4).

## 3.5 Results

### 3.5.1 Inactivation of *Irf8* causes resistance to cerebral malaria

BXH2 is a recombinant inbred mouse strain derived from B6 and C3H/HeJ (C3H) parents which displays a myeloid defect in the form of immature myeloid hyperplasia and susceptibility to multiple infections. We have previously used high resolution linkage analysis, positional cloning and candidate gene sequencing to demonstrate that myeloid hyperplasia and susceptibility to infections are caused by a severely deleterious hypomorphic allele at *Irf8* (*Irf8*$^{R294C}$) that spontaneously arose during the breeding of this strain[244, 248].

To assess the contribution of *Irf8* to pathological inflammation, we infected BXH2 mice with PbA parasites, the murine agent of experimental cerebral malaria (ECM). Parasite replication in the blood, appearance of neurological symptoms and overall survival were recorded over 18 days (Figure 3-1). While all B6 mice developed ECM and succumbed by d9, BXH2 mice were completely resistant to the ECM phase, succumbing later to hyperanemia caused by uncontrolled blood-stage replication of the parasite (Figures 3-1A and 3-1B). [BXH2 X B6]F1 mice showed significant resistance to PbA-induced ECM when compared to susceptible B6 parental control ($p<0.0001$), with approximately 50% of the animals surviving past d9 (Figure 3-1A). Additional phenotyping of a small group of segregating [BXH2 X B6]F2 mice (n = 24) identified ECM-resistance only in mice either homozygote or heterozygote for the $Irf8^{R294C}$ allele, confirming that the protective effect we observed is due to the $Irf8^{R294C}$ mutation with minimal or no contribution of the mixed B6/C3H genetic background of BXH2 (Figure 3-1C). These data show that a) partial or complete loss of IRF8 function protects mice against lethality in an ECM-associated neuroinflammation model, and b) that the ECM-protective effect of the $Irf8^{R294C}$ mutation is inherited in a co-dominant fashion. Resistance to ECM in BXH2 mice was not associated with decreased parasite burden, as B6, BXH2 and [BXH2 X B6]F1 mice showed similar levels of circulating blood parasitemia at d5, d7 and d9 post-infection ($p > 0.1$)(Figure 3-1B). However, as the infection progressed, some BXH2 mice developed extremely high levels of blood parasitemia between d12-d21 in sharp contrast to any surviving controls or [BXH2 X B6]F1s. This high parasitemia, rather than cerebral inflammation, was responsible for the observed mortality.

Lethal ECM in PbA-infected mice is associated with endothelial dysfunction, including loss of integrity of the blood brain barrier (BBB) [178, 191, 280]. Using an Evans Blue dye extravasation assay (Figure 3-1D), PbA-infected

B6 mice display obvious BBB permeability by d7, indicated by the blue color, while infected BXH2 mice retained BBB integrity both early (d7) and late (d16) during infection. The BBB integrity was also assessed by flow cytometry cellular profiling of perfused brains from day 6 infected mice. In B6 brains, ECM was associated with presence of both $CD4^+$ and $CD8^+$ lymphocytes, as well as $CD11b^+/Ly6C^+$ granulocytes and monocytes (Figure 3-1E). Interestingly and despite the immature myeloid hyperplasia in peripheral myeloid and lymphoid organs, infiltration of T cells and myeloid cells in the brain was not seen in PbA-infected BXH2, nor was it seen in [BXH2 X B6]F1 (Figure 3-1E). Together, these results indicate that ECM susceptibility in B6 is associated with infiltration of myeloid and lymphoid cells at the site of pathology. Such infiltration is not seen in ECM-resistant mice that are either homozygous (BXH2) or heterozygous ([BXH2 X B6]F1) for *Irf8*$^{R294C}$.

These results demonstrate that although IRF8 dysfunction is protective against ECM, functional IRF8 is required to control blood stage replication of PbA late in infection. Partial IRF8 activity in [BXH2 X B6]F1 is sufficient to protect against high blood-stage replication.

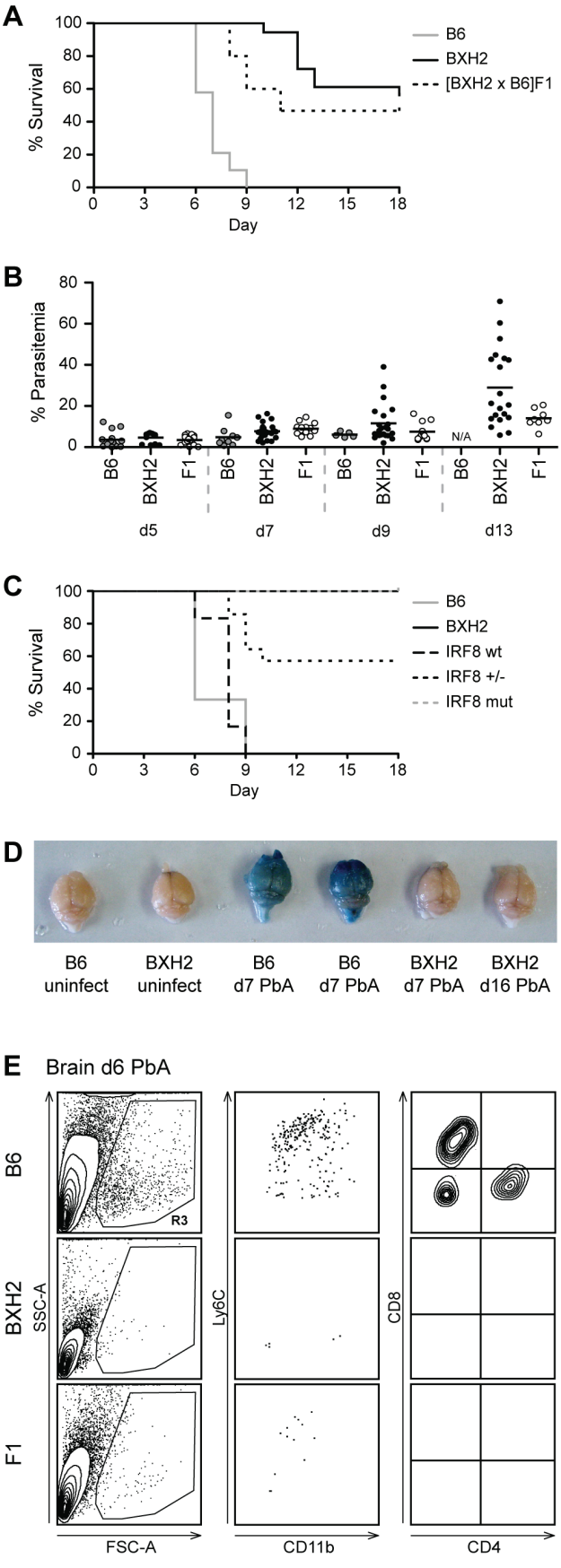

A
% Survival
Day
B6
BXH2
[BXH2 x B6]F1
B
% Parasitemia
N/A
d5
d7
d9
d13
C
% Survival
Day
B6
BXH2
IRF8 wt
IRF8 +/-
IRF8 mut
D
B6 uninfect
BXH2 uninfect
B6 d7 PbA
B6 d7 PbA
BXH2 d7 PbA
BXH2 d16 PbA
E
Brain d6 PbA
B6
BXH2
F1
R3
SSC-A
FSC-A
Ly6C
CD11b
CD8
CD4

**Figure 3-1. BXH2 mice do not develop cerebral malaria following *Plasmodium berghei* infection.**

(A) Survival plots of *Plasmodium berghei* (PbA) infected BXH2 mice (n = 18), heterozygous [BXH2×B6]F1 offspring (n = 15), and ECM-susceptible B6 (n = 19) parental controls. Shown are the combined results of four experiments. (B) Blood parasitemia levels during infection for mice in Panel A following infection with PbA. (C) *Irf8* genotype-specific survival curves for 24 [BXH2=B6]F2 mice along with parental controls. (D) Qualitative comparison of representative Evans blue dyed brains from uninfected and infected B6 and BXH2 mice indicating breakdown of the blood-brain barrier in infected B6 (d7 PbA), but not BXH2 (d7 PbA or d16 PbA) mice. (E) Perfused brains from ECM-susceptible B6 and ECM-resistant BXH2 and [BXH2×B6]F1 mice (n = 2 to 4) were collected six days following infection with PbA. Infiltrating leukocytes were enriched by Percoll gradient and stained with CD45, Ly6C and CD11b, or CD45, TCR, TCRin barrier in infected B6 (d7 PbA), but nooid and lymphoid infiltrates is observed in the brain of B6 mice compared to BXH2 or F1. Gate R3 denotes infiltrating cells gated by side scatter (SSC-A) and forward scatter (FSC-A).

### 3.5.2 Differential transcript profiles associated with *Irf8*-dependent resistance to cerebral malaria

To gain further insight into the genes, proteins and pathways that play a role during pathological neuroinflammation, and identify those whose expression is regulated by *Irf8*, we used transcript profiling of BXH2 and B6 brains either prior to or during PbA-infection. Principal components analysis (PCA) clustered the samples along two axes: component 1, which explained 39.4% of the variance and was associated with infection status (infection component), and component 2, associated with mouse strain (genetic component), which explained 24.4% of the variance (Figure 3-2A). PCA also indicates that PbA

infection had a much stronger impact on transcriptional profiles in B6 mice than in BXH2, with the B6 d7 infected samples forming a remote out-group. In contrast, the BXH2 d7 cluster was only moderately shifted by infection and remained much closer to the BXH2 d0 group (compared to B6 d0 vs. B6 d7) indicating more modest response to infection by BXH2.

We used paired t-tests to assess the infection-induced transcriptional responses in both B6 and BXH2 mice as a way to extract gene lists relevant to ECM susceptibility. As suggested by the PCA, B6 response to infection was robust, with 292 genes showing statistically significant differences in expression (Figure 3-2B, d0 vs. d7; fold change ≥ 2, $p_{adj} < 0.05$). On the other hand, response to infection in BXH2 was more modest with 81 genes reaching statistical significance. More than half of the genes (n = 48) regulated by infection in BXH2 were common to the B6 set and may indicate *Irf8*-independent regulatory mechanisms. This analysis also identified a set of 117 genes that show significantly different levels of expression in B6 and BXH2 mice prior to infection. Only ~10% of these genes (n = 16) were further significantly modulated by PbA infection (Figure 3-2B). Importantly, this analysis further identified a key subset of 231 genes that were specifically regulated in B6, but not BXH2, mice during infection, therefore associated with pathology.

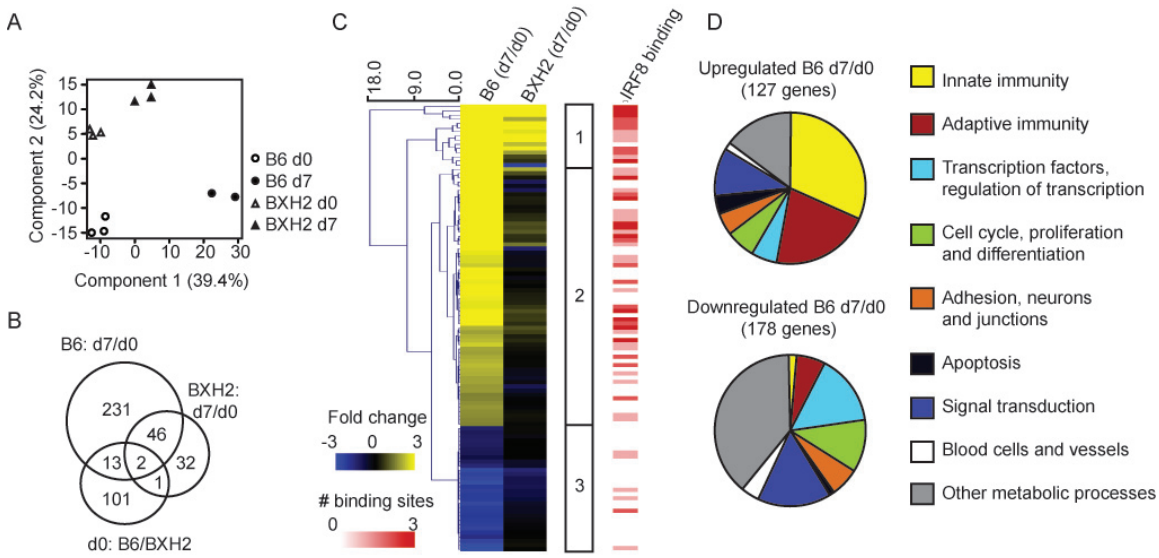



**Figure 3-2. Transcript profiling of PbA infected B6 and BXH2 brains reveals strain and infection specific differences**

(A) Unsupervised principal components analysis clusters samples according to mouse strain and infection status. (B) Intersection of gene lists generated by pairwise comparisons between infected and uninfected B6 and BXH2 transcript profiles. (C) Euclidean clustered heat map of transcripts regulated in both a strain and infection specific manner (two-factor ANOVA, $p_{adj}$-interaction<0.05) illustrated as infection-induced fold change in each strain (d7/d0). Each row represents a unique gene, and in cases where two or more transcript probes for a gene were significant, the average fold change was used. Differential expression patterns clustered into three groups with Group 1 genes being up-regulated by infection in both strains, Group 2 genes up-regulated by infection in B6 mice and unresponsive in BXH2 and Group 3 genes down-regulated by infection, typically more so in B6 than BXH2. See Table S1 for details. Red shaded heat map indicates the presence of one or more IRF8 binding sites within 20 kb of the gene transcription start site as determined by ChIP-seq. (D) Gene ontology for transcripts differentially regulated by infection during ECM pathology in B6 mice (as determined by B6 d7/d0 pairwise analysis, $p_{adj}$<0.05). Up-regulated genes are predominantly involved in innate and adaptive immunity processes, while down-regulated genes are not, and rather include a variety of homeostatic biological and metabolic processes.

We also performed a two-factor ANOVA test accounting for both differences in basal level of gene expression in the brain (B6 vs. BXH2 at day 0), and infection-induced transcriptional response to PbA (Figure 3-2C). This identified a total of 107 genes (123 probes; fold change ≥ 2, $p_{adj}$< 0.05) that were strongly regulated by infection in an *Irf8*-dependent fashion ($p_{adj}$-interaction < 0.05) (Figure 3-2C, Table S1). Euclidean hierarchical clustering of this gene list identified three major categories of transcripts. Group 1 genes (n = 15) were up-regulated by infection in both strains (more pronounced in B6 than BXH2), group 2 genes were up-regulated by infection in B6 mice but not significantly induced in BXH2 (n = 62) and group 3 genes (n = 30) were down-regulated by infection (stronger repression in B6 compared to BXH2). Using the online Database for Annotation and Integrated Discovery (DAVID) tool to examine the list of genes regulated by infection in B6 mouse brains indicated substantial enrichment for immune response (4.4-fold enrichment above Illumina WG-6 v2.0 chip background, $p_{adj} = 3.3x10^{-12}$), antigen processing and presentation (11.0-fold enrichment, $p_{adj} = 1.4x10^{-9}$), defense response (3.8-fold enrichment, $p_{adj} = 1.7x10^{-8}$), chemotaxis (5.2-fold enrichment, $p_{adj} = 3.4x10^{-3}$) and inflammatory response (3.6-fold enrichment, $p_{adj} = 3.7x10^{-3}$). Increased genes on these lists include potent pro-inflammatory chemokines that recruit myeloid and lymphoid cells to the site of infection and/or tissue injury such as *Cxcl9, Cxcl10, Ccl4*, and *Ccl12*; myeloid cell receptors associated with phagocytosis of microbes (*Fcgr3*, *Fcgr4*) and maturation of phagosomes (small GTPases *Igtp, Irgm1, Gbp2, Gbp3*); Irf8's hetero-dimerization partner (*Irf1*), and early type I interferon response (*Oasl2, Ifit3*). Genes under these categories were expressed more highly in B6 than in BXH2, consistent with the notion that resistance to ECM-associated neuroinflammation in BXH2 is linked to reduced *Irf8*-dependent inflammatory and innate immune responses, with a strong involvement of the myeloid compartment.

Organizing the list of genes associated with ECM pathology (B6 d7/d0 pairwise) using Ingenuity Pathway Analysis revealed enrichment for several significant biological networks. The top-scoring network included down-regulation of the hematological response and up-regulation of intercellular signaling (Figure S1A). Several immune cell chemoattractants (*Ccl4*, *Ccl7*, *Cxcl9*) are featured in this network, suggesting recruitment of monocytes, NK cells and T-cells to the site of brain inflammation. Very little regulation of either of these pathways is seen in BXH2 during infection, with most components not making the 2 fold change threshold (Figure S1B). The second network highlights significant up-regulation of interferon-responsive genes and inflammatory mediators, with type 1 interferon, *Irf7* and *Stat1* regulation occupying highly connected nodes (Figure S1C). Most of these inflammatory processes are increased to a lesser degree in BXH2 mice (Figure S1D and details in Table S1). The third network focuses on antigen processing and presentation with immunoglobulins and Fc receptors increased at the center of the B6 infection-responsive image (Figure S1E). Similar to what was seen in network 1, significant up-regulation of these genes was largely missing from the BXH2 response (Figure S1F).

### 3.5.3 Chromatin immunoprecipitation and sequencing (ChIP-seq) highlights IRF8 bound genes activated during acute neuroinflammation

Since total brain RNA was used in our studies, genes differentially regulated in response to infection in an *Irf8*-dependent fashion may represent direct transcriptional targets of IRF8 or may be secondary targets corresponding

to markers of cell populations differentially recruited to the site of infection in B6 and BXH2 mice. To distinguish between these possibilities, we mapped genome-wide IRF8 binding sites using ChIP-seq. The resulting sequence reads were mapped to mouse mm9 genome assembly and IRF8 binding peaks were identified using MACS peaks finding algorithm [279]. In order to validate the ChIP-seq results, we confirmed IRF8 recruitment on known target sites [239] by independent ChIP-qPCR experiments (Figure 3-3A). IRF8-bound genes were identified as those containing a binding peak within a 20kb window from their transcriptional start site (TSS). The list of IRF8-bound genes by ChIP-seq was intersected with the list of genes regulated by PbA in a strain, infection and *Irf8*-associated fashion from the two-factor ANOVA analysis (Figure 3-2C, Table S1). This intersection revealed a strong enrichment of IRF8 binding sites in genes up-regulated during infection, with IRF8 binding sites detected in 85% of Group 1 genes (13/15) and 50% of Group 2 genes (31/62) (Figure 3-2C, Table S1). Differentially down-regulated genes did not show any enrichment with only 13% (4/30) of Group 3 genes containing IRF8 peaks, lower than background peak association (21% of all genes represented on the Illumina array; Figure 3-3B). These results strongly suggest that during neuroinflammation, IRF8 functions as a direct transcriptional activator of multiple genes coding for key pro-inflammatory pathways.

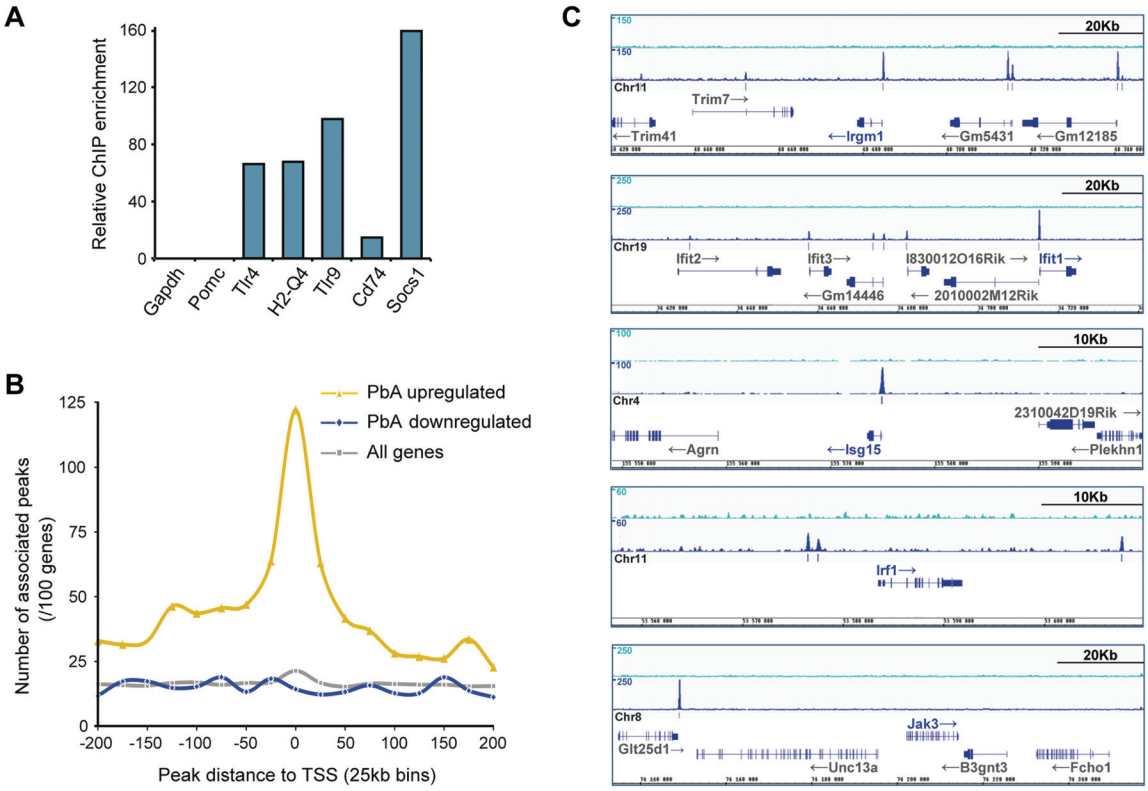



**Figure 3-3. Genes up-regulated during ECM pathology in B6 mice are significantly enriched for IRF8 binding sites.**

(A) Quantitative PCR was used to validate ChIP results using known binding targets of IRF8. Targets were highly enriched in IRF8-immunoprecipitated DNA when compared to control IgG preparations. Representative data from one of five independent experiments is shown. (B) The list of genes regulated by infection in B6 mice (d7/d0 pairwise) was interrogated for IRF8 binding sites within 225 kb of their transcription start site. The graph represents the abundance of IRF8 binding sites in each 25 kb segment. (C) IRF8 and control IgG ChIP-seq sequence reads were mapped to the mouse genome and significant IRF8 binding sites were identified. Light blue (top) track indicates non-specific (IgG) sequencing profile and dark blue track (below) displays IRF8 binding sites. Genes were considered to have an IRF8 binding site if a peak was found within 20 kb of the transcription start site.

To identify IRF8-bound genes associated with ECM neuropathology, we queried the list of all genes whose expression is regulated by infection in *Irf8*-competent B6 mice for the presence of IRF8 binding sites (pair-wise comparison of B6 d7/d0). This analysis (Figure 3-3B) showed very strong enrichment for IRF8 binding sites in the vicinity of genes increased by infection, with 74% of up-regulated genes (92/125; $p<0.0001$, Fisher's Exact test) bearing one or more IRF8 binding sites within 20kb of the TSS (Figure 3-3C for ChIP-seq profile examples and complete list in Table S2). Genes decreased in response to infection did not show enrichment above background (Figure 3-3B and Table S2). Using the AmiGO gene ontology annotation tool [281], this list (Figure 3-2D and see details in Table S3) was found to contain numerous genes involved in inflammatory and innate immune responses, a finding most pronounced in the subset (74%) of genes with IRF8 binding sites. This included inflammatory cytokines and chemokines involved in chemotaxis of myeloid and lymphoid cell types to the sites of infection (*Ccl4, Ccl5, Ccl7, Ccl12, Cxcl9, Cxcl10*), early innate immune recognition and responses (*Nlrc5, Ifi205*), response to viral infections (*Oasl2, Mx2, Oas1g*), type I interferon responsive genes and pathways (*Ifit2, Ifit3, Isg15, Rsad2*), antigen capture (*C1q, C4b, Fcerg1*), phagosome maturation (*Irgm1, Irgm2, Igtp, Gbp2, Gbp3*), antigen processing (*Tap1, Tap2*) and Class I and Class II MHC-dependent antigen presentation in myeloid cells (*B2m, H2-Ab1, H2-D, H2-K, H2-L, H2-Q, H2-T22* ). Other IRF family members implicated in early response to antigenic stimuli or danger signals (*Irf1, Irf7, Irf9*) were also induced (Table S3). These IRF8-regulated pro-inflammatory pathways appear linked primarily to the myeloid cellular compartment.

### 3.5.4 Activation of a common IRF8 transcriptome in cerebral malaria and during pulmonary tuberculosis

In both humans [266] and mice [258], mutations in *IRF8* cause susceptibility to mycobacterial infections. Thus, we compared the list of genes regulated by infection in B6 brains during PbA cerebral malaria (d7/d0) with the list of genes contributing to the protective response in the lungs of *M. tuberculosis* infected B6 mice (d30/d0) [258] (Table S2). B6 and BXH2 mice have opposite phenotypes in the two disease models, with B6 being susceptible to ECM and resistant to *M. tuberculosis*, while BXH2 succumb rapidly to pulmonary tuberculosis [258]. Strikingly, out of the 123 genes up-regulated more than 2-fold during ECM, 66 followed the same trend during *M. tuberculosis* infection ($p < 0.0001$, Fisher's Exact Test). There was minimal overlap in the down-regulated genes (21 *M. tuberculosis*-regulated genes overlapping with the 170 PbA-regulated, including 6 in the opposing direction). The overwhelming majority (80%) of genes increased during both infections contained at least one IRF8 binding site, again highlighting IRF8 as having a central role during inflammation and host response to infections (Table S2).

To visualize the core networks engaged in this common IRF8-regulated host response, we analyzed the 53 genes containing IRF8 binding sites, and which were increased during both ECM and in pulmonary tuberculosis (Figure 3-4 and see Table S2 for details). Several networks with a clear focus on type 1 and type 2 interferon pathways characteristic of a pro-inflammatory innate immune response were detected. All these networks are clearly less activated in PbA infected BXH2 mice. One obvious network (Figure 3-4A) includes three Irfs (Irf1, Irf7, Irf9) and other hallmarks of type 1 interferon and MHC class 1 response. A second network (Figure 3-4B) centers on IFNγ/STAT1 signaling,

with interferon-inducible GTPases, pro-inflammatory cytokines and chemokines forming the downstream response. A third, classically myeloid network (Figure 3-4C) features Fc-γ receptors, immunoglobulins and MHC class 2 molecules, as well as T-cell activating IL-12, CCL4 and CCL5. Across all networks, the most highly connected genes were *Irf1*, *Irf7*, *Ifng*, and *Stat1*, consistent with a key role for IRF8 in regulating pro-inflammatory innate immune responses in myeloid cells.

**Figure 3-4**: (*see over*) Networks were generated for the 53 genes up-regulated in B6 mice by both infections which also possess an IRF8 binding site within 20 kb of the TSS (see Table S2 for details). (A) Top scoring network highlighting IRF signaling and MHC class I antigen presentation with gene circles colored according to fold change during PbA infection in B6 d7/d0 (left), BXH2 d7/d0 (center) and *M. tuberculosis* infection in B6 d30/d0 (right). (B) Second top scoring network highlights *Ifng* and *Stat1* signaling. (C) Third top scoring network highlights MHC Class II and Fc receptors. Genes included in the list of 53 are represented by black circles, while networked genes added by the software are outlined in gray. IRF8-bound genes are outlined in light blue. Direct (black) and indirect (gray) connections between genes are shown by arrows.

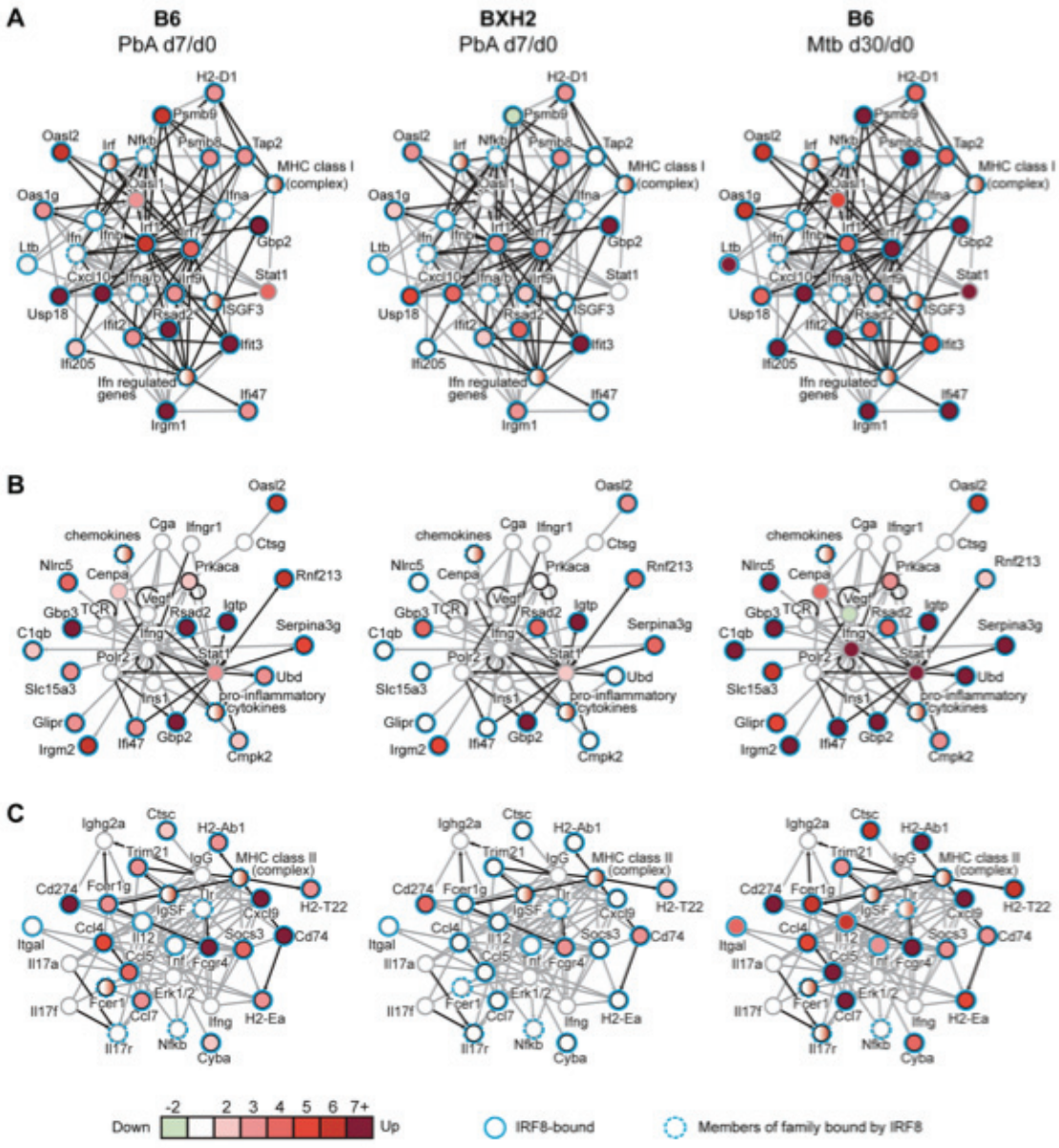



Figure 3-4. IRF8-regulated pro-inflammatory networks commonly activated during cerebral malaria and pulmonary tuberculosis

### 3.5.4 Validation of IRF8 targets as key mediators of neuroinflammation during cerebral malaria

To validate the role of several identified IRF8 targets and associated pathways in pathological neuroinflammation, we phenotyped targeted knock-out mice for several of these loci. These included infection-regulated genes bearing IRF8 binding sites (*Irf1, Ifit1, Isg15, Nlrc4, Il12p40* and *Irgm1*; examples of IRF8 ChIP-seq profiles are shown in Figure 3-3C), and genes known to play key roles in early innate immune response (*Ifng, Jak3, Stat1, Il12p40*). Results from these experiments (Figure 3-5) show that *Ifng*$^{-/-}$, *Jak3*$^{-/-}$ and *Stat1*$^{-/-}$ mutant mice were completely resistant to ECM following PbA infection, highlighting key roles for these molecules in the progression or amplification of the pathological inflammatory response and supporting their central positions in the network models (Figures S1C, 3-4A and 3-4B) [199, 282]. Loss of *Irf1* [283], *IL12p40* [284] and *Irgm1* delayed appearance of neurological symptoms and prolonged survival of PbA-infected mice but did not ultimately confer complete protection (Figure 5). These results validate that certain of IRF8's transcriptional targets during PbA infection in lymphoid and myeloid cells play critical downstream roles in pathology. On the other hand, *Ifit1*$^{-/-}$, *Isg15*$^{-/-}$ and *Nlrc4*$^{-/-}$ mutant mice remained susceptible to PbA-induced ECM, suggesting that although these proteins may play important roles in neuroinflammation, their inactivation is not sufficient to induce protection.

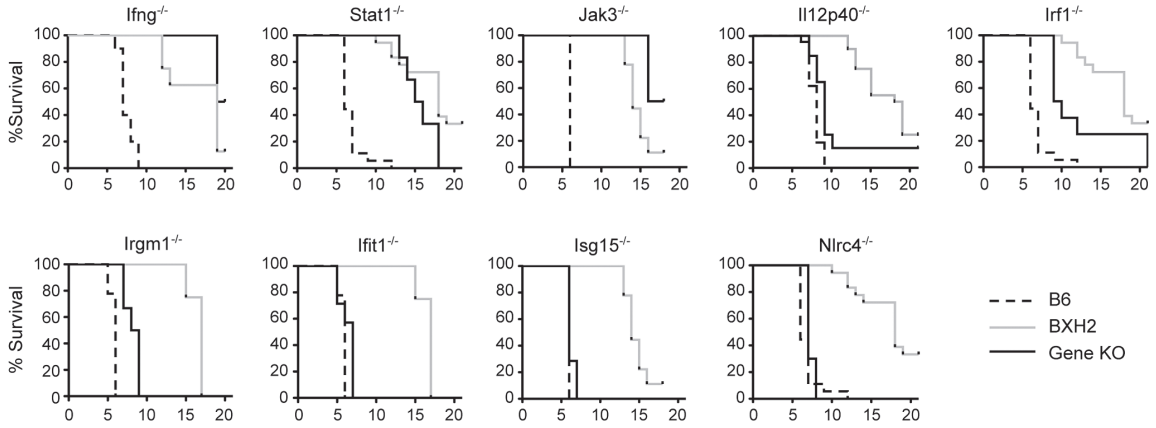



**Figure 3-5. Effect of genetic deletion of Ifng, Stat1, Jak3, Irf1, Irgm1, Il12p40, Ifit1, Isg15 and Nlrc4 on suceptiblity to PbA induced cerebral malaria**

Control and mutant mice were infected with $10^6$ *P. berghei* parasites and survival was monitored. Cerebral malaria susceptible mice succumbed between d5 and d10 post-infection with neurological symptoms, while no mice that survived longer than 13 days developed signs of ECM and these were categorized as resistant. Infection specific B6 and BXH2 controls (n>5 per infection) are plotted alongside each mutant strain.

### 3.5.5 ECM resistance in IRF1 and IRF8 deficient mice is associated with altered IL12p40 and IFNγ production

To clarify the immunological basis of ECM resistance in BXH2, and in [BXH2 X B6]F1, cellular immunophenotyping of spleen cells was carried out, both at steady state (d0) and 6 days following infection with PbA, immediately preceding the appearance of neurological symptoms (Figure 3-6 and 3-7). In uninfected spleens, BXH2 were characterized by a dramatic expansion of $CD11b^+/Ly6C^+$ immature myeloid cells comprised of monocyte-like $F4/80^+$ (Figure 3-6A Day 0 – Gate R1 and Figure S2A Gate R1) and granulocyte-like

Ly6G$^+$ cells (Figure 3-6A Day 0 – Gate R2 and Figure S2A Gate R2), which resulted in reduced proportions of lymphocytes (Figure 3-7A Day 0). In contrast, lymphoid and myeloid cell populations are highly similar in spleens of ECM-susceptible B6 and ECM-resistant [BXH2 X B6]F1 mice prior to PbA infection (Figure 6A and 7A). During *P. berghei* infection, we noted that ECM-resistance in BXH2 and [BXH2 X B6]F1 compared to B6 controls was associated with reduced numbers of splenic myeloid DCs (Fig. 3-6D; Ly6C$^-$CD11b$^+$CD11c$^+$MHCII$^+$), which was concomitant to reduced but not absent serum levels of IL12p40 (Figure 6E) and reduced production of IL12p40 by spleen cells *ex vivo* 6 days post-infection (Figure 6F). We also noted that during infection, the ECM-protective effect of the *Irf8* mutation is phenotypically expressed in the lymphoid compartment of BXH2 and [BXH2 X B6]F1. Indeed, at day 6 post-infection, ECM-resistant BXH2 and F1 mice did not show the significant increase in CD8$^+$ T cells that was detected in ECM-sensitive B6 mice (2X decrease; Figure 3-7C and S3), while CD4$^+$ T cell numbers were increased in all groups by the same factor. Together, these changes were accompanied by a significantly reduced production of IFNγ (6 days post-infection) by splenocytes *ex vivo* measured under non-stimulated conditions (4-5X decrease), in response to PMA/ionomycin, or to IL12p70 treatment (Figure 3-7E). We also noted a 16 and 3 fold reduction in the level of serum IFNg in infected ECM-resistant BXH2 and F1 mice compared to B6 controls, respectively (Figure 3-7D). These studies suggest that heterozygosity and homozygosity for a loss of function allele at *Irf8* (*Irf8$^{R294C}$*) impacts the activity of both myeloid cells and T cells during *P. berghei* infection (resulting in anergy of CD8$^+$ T cells).

IRF1 cooperates with IRF8 to regulate gene expression [237], and *Irf1$^{-/-}$* mice are resistant to ECM (Figure 3-5). Therefore, we also conducted immunophenotyping studies of *Irf1$^{-/-}$* mice before (d0) and during *P. berghei*

infection. As previously described [285], *Irf1*$^{-/-}$ mice display defective lymphocyte maturation, with normal CD4$^{+}$ and greatly reduced CD8$^{+}$ T cells numbers (Figure 3-7A to 3-7C). Following *P. berghei* infection, *Irf1*$^{-/-}$ mice exhibit slight reduction in IL12p40 serum level ($p$=0.0585) and a significant decrease in IL12p40 production by splenocytes (Figure 3-6E to 3-6F). Altered IL12p40 production and reduced CD8$^{+}$ T cell numbers (Figure 3-7A, C) is associated with severe reduction in IFNγ production, both in serum and splenocyte culture supernatants (Figure 3-7D-E). *Irf1*$^{-/-}$ splenocytes produce much less IFNγ than ECM-susceptible B6 mice in response to external secondary triggers such as PMA/ionomycin (Figure 3-7E). Taken together, the combined analysis of *Irf1* and *Irf8* mutant mice associates reduced pro-inflammatory cytokines production (IL12p40 and IFNγ) to resistance to ECM.

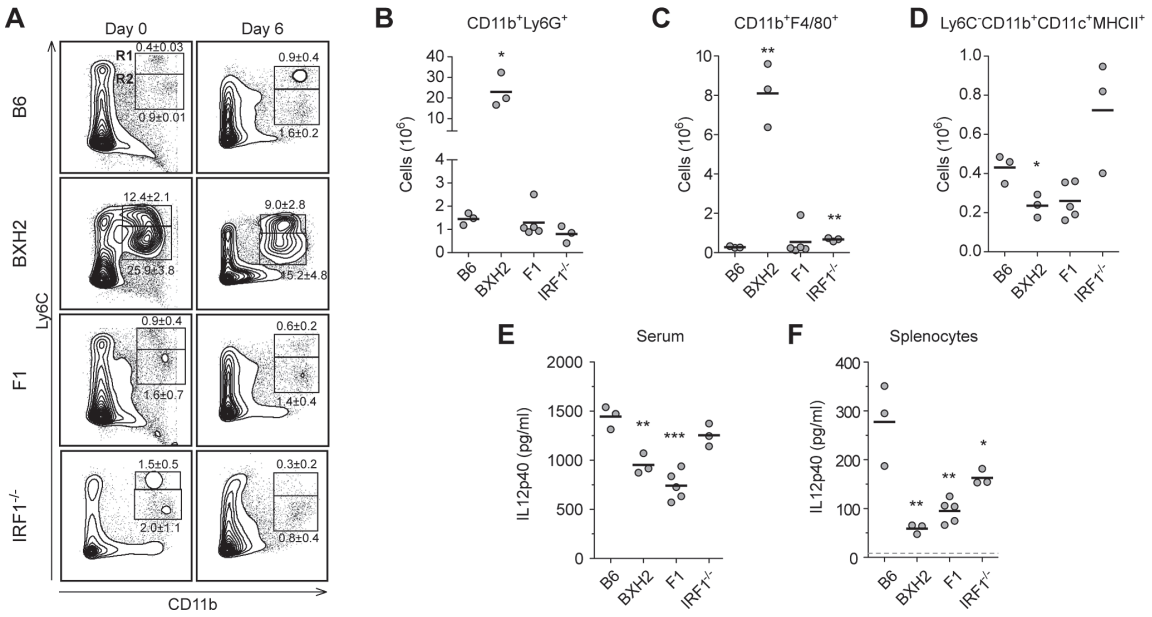



**Figure 3-6. Characterization of the myeloid compartment in PbA-infected B6, BXH2, [BXH2 x B6]F1 hybrids and *Irf1*$^{-/-}$ mice**

Spleens from B6, BXH2, [BXH2×B6]F1 and IRF1$^{-/-}$ mice were harvested prior to, or six days following, PbA infection and processed by flow cytometry. $2\times10^6$ splenocytes were stained with CD45, CD11b, Ly6C, Ly6G, F4/80, CD11c and MHCII, and representative cellular profiles are shown for each strain (A). The numbers within contour plots refer to gates R1 (monocytes) or R2 (granulocytes) and are reported as mean ± SD (gated as percentages of CD45$^+$ cells). Absolute numbers of day 6 PbA-infected mice are shown for (B) CD11b$^+$Ly6G$^+$ and (C) CD11b$^+$F4/80$^+$ populations, indicating an expansion of the myeloid lineage in the BXH2 strain, compared to B6, [BXH2×B6]F1 and IRF1$^{-/-}$ animals. Reduced numbers of myeloid dendritic cells (D) along with lower serum IL12p40 levels (E) are noted in ECM-resistant BXH2 and [BXH2×B6]F1 compared to ECM-susceptible B6 mice. (F) Levels of secreted IL12p40 were determined in culture supernatants from splenocytes of infected mice. Dashed gray line represents IL12p40 detection limit. Differences were considered significant when $p<0.05$ and calculated compared to the B6 strain (Student's t-test: *$p<0.05$, **$p<0.01$, ***$p<0.001$).

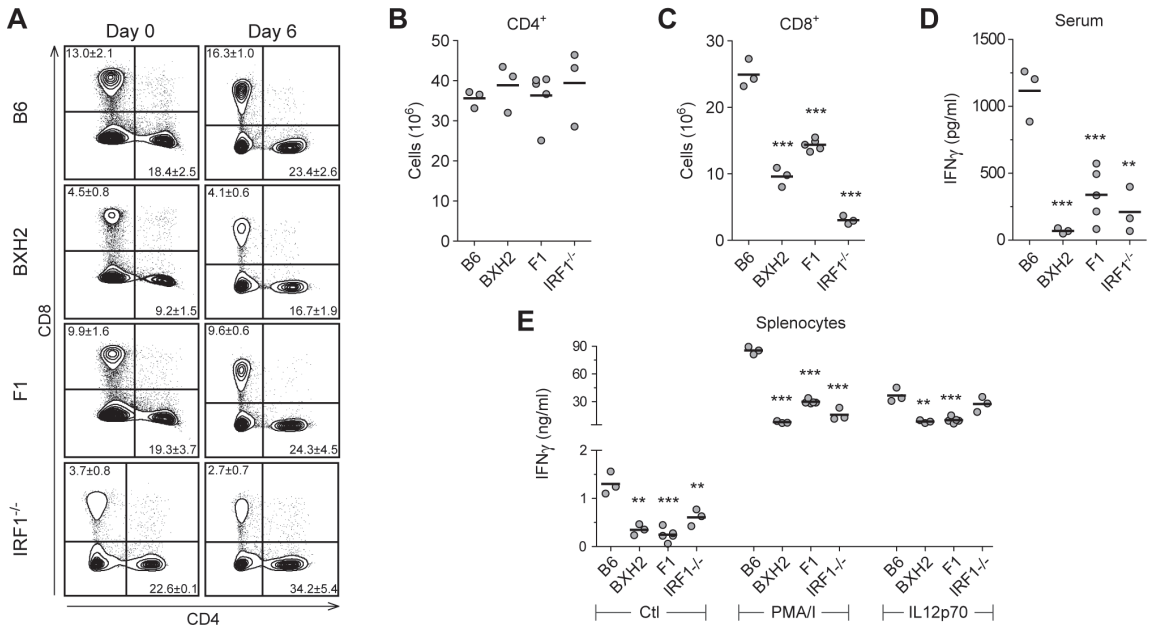



**Figure 3-7. Characterization of the lymphoid compartment in PbA-infected B6, BXH2, [BXH2 x B6]F1 and *Irf1*$^{-/-}$ mice.**

Characterization of the lymphoid compartment was carried out as described in the legend of Figure 3-6. Splenocytes were stained with CD45, TCRβ, CD4 and CD8 antibodies and representative cellular profiles are shown for each strain (A), where numbers indicate mean ± SD (gated as percentages of CD45$^+$ cells). Absolute numbers are indicated in dot plots for total spleen CD4$^+$ cells (B) and CD8$^+$ cells (C). (D) Serum IFNγ were significantly lower in ECM-resistnat BXH2, [BXH2 X B6]F1 and *Irf1*$^{-/-}$ mice. (E) IFNγ production was assayed *in vitro* in culture supernatants from infected mice with or without stimulation with PMA/Ionomycin or with IL12p70. The p-values were calculated relative to B6 controls with Student's t-test (**p<0.01, ***p<0.001).

## 3.6 Discussion

The demonstrated role of IRF8 in the ontogeny of the myeloid lineage, its known role in defense against infectious pathogens and the growing body of evidence from GWAS in humans linking IRF8 variants to chronic inflammatory conditions such as multiple sclerosis [271], systemic lupus erythematosus [267] and inflammatory bowel disease [269, 270, 286] prompted us to investigate a possible role for IRF8 in acute pathological inflammatory reactions. To do so, we used a mouse model of acute cerebral malaria encephalitis caused by infection with PbA, which involves lethal neuroinflammation associated with recruitment of inflammatory mononuclear and polymorphonuclear leukocytes, and loss of integrity of the blood brain barrier (BBB). We found that the loss of *Irf8* function in BXH2 mice completely protects against this pathology, preventing the development of neurological symptoms and prolonging survival post infection. Interestingly, the protective effect was inherited in a co-dominant fashion as 50% of *Irf8*$^{R294C/+}$ F1 heterozygotes survived through the cerebral phase when infected with PbA (Figure 3-1A and B). These finding establish that *Irf8* is critical to the development of acute lethal neuroinflammation and further implicate *Irf8* as a major regulator of this pathological response. Moreover, results from *Irf8*$^{R294C/+}$ F1 heterozygotes indicate that *Irf8* regulates key pro-inflammatory cells and pathways in a gene dosage dependent fashion.

In addition to its established role in ontogeny and function of myeloid cells, *Irf8* is required for certain aspects of B lymphocyte development and T lymphocyte function including Th1 and Th17 responses [241, 259, 287]. Cellular immunophenotyping was conducted in naïve and PbA-infected tissues to identify the cell population(s) associated with and most likely responsible for pathological inflammation, whose absence is associated with ECM-resistance in BXH2 and in

[BXH2 X B6]F1 mice. These studies showed that a) ECM-resistance is independent of the immature myeloid hyperplasia ($CD11b^+/Ly6C^+$) characteristic of BXH2, as this trait is absent from resistant [BXH2 X B6]F1 (Figure 3-6A), b) is linked to brain infiltration of both myeloid ($CD11b^+/Ly6C^+$) and lymphoid cells ($TCR\beta^+/CD4^+$ and $TCR\beta^+/CD8^+$ T cells) in B6 mice (Figure 3-1E), and c) is concomitant with reduced production of IL12p40 by myeloid cells and impaired production of IFNg by T cells following infection with *P. berghei*.

To identify the gene-dosage dependent pathways that are activated by *Irf8* during neuroinflammation, we compared brain transcript profiles from PbA-infected B6 and BXH2 mice and extracted a list of genes that are induced by infection in an *Irf8*-dependent and independent fashion (two-factor ANOVA and pairwise analyses) (Figure 3-2, Table S1). In parallel, we carried out ChIP-seq experiments to map genome-wide IRF8 binding sites. We compared the position of these binding sites to the gene lists generated by transcript profiling and identified both IRF8-bound genes (Table S2), and IRF8-bound genes regulated in an *Irf8*-allele specific fashion (Tables S1). There was substantial overlap between these gene sets, which were similarly dominated by markers of antigen-presenting cells (APC) including antigen processing and presentation, production of type I interferon, pro-inflammatory cytokines/chemokines and others. These analyses confirm that *Irf8* plays a prominent role in the unique functions of APCs including antigen capture and microbial phagocytosis (*C1q, C4b, Fcgr4, Fcgr1*), cytoplasmic inflammasome platforms (*Nlrc5*, *Ifi205*), phagosome maturation and recruitment of small GTPases (*Irgm1, Irgm2, Igtp, Gbp2, Gbp3*), endoplasmic reticulum membrane associated antigen transport (*Tap1, Tap2*), and both Class I and Class II MHC-dependent antigen presentation in APCs (*B2m, H2-A, D, K, L, Q, T* molecules). These gene lists also featured a number of inflammatory cytokines and chemokines involved in chemotaxis of myeloid and lymphoid cell

types to the sites of infection (*Ccl4, Ccl5, Ccl7, Ccl12, Cxcl9, Cxcl10*). Network analysis of the IRF8 targets bound and activated in response to PbA infection confirms that *Irf8* directly regulates several myeloid-specific, pro-inflammatory pathways that are ultimately responsible for pathological inflammation (Figure S1).

These findings are compatible with a simple functional model where myeloid cells (including APCs) and lymphoid cells are rapidly recruited in large numbers to the site of PbA infection and associated tissue injury – namely capillaries of the blood brain barrier. This initiates a robust IRF8-dependent pro-inflammatory cascade. Local amplification of this response by recruited cells leads to excessive production of immunopathological soluble mediators such as IFNγ and TNFα by T lymphocytes and induces other transcription factors including *Stat1* and other IRF family members (*Irf1, Irf7, Irf9*). Damage to the BBB causes further infiltration of pro-inflammatory cytokine-secreting cells (Figure 3-1E), and ultimately appearance of lethal neurological symptoms in susceptible mice. Absence of *Irf8* blunts this pathological response and allows mutant BXH2 mice to avoid developing neuroinflammation during ECM, thus surviving the critical acute phase. Results from immunophenotyping studies (Figures 3-6 and 3-7) indicate that the protective effect of *Irf8* (and *Irf1*) mutations is associated with simultaneously diminished pro-inflammatory responses by both the myeloid and the lymphoid compartments; this is most evident for the blunted production of IFNγ in ECM-resistant BXH2, *Irf1*$^{-/-}$ or [BXH2xB6]F1 mice, when compared to ECM-susceptible B6 (serum level, and by CD4$^{+}$ and CD8$^{+}$ T cells with or without stimulation with PMA/ionomycin).

At steady-state, BXH2 mice show a severe depletion of myeloid DC subsets but display expansion of immature myeloid progenitors, including granulocyte-like cells (Cd11b$^{+}$/Ly6C$^{+}$) in peripheral tissues [248-251]. Literature

in support of [288, 289] or arguing against [290] a role for granulocytes/neutrophils in the pathogenesis of ECM has been published. However, myelocytic hyperplasia in BXH2 is not responsible for ECM resistance, since [BXH2 X B6]F1 mice are also resistant and do not display this myeloid expansion (Figures 3-1A and 3-6). Although the depletion of DCs, along with a concomitant reduction in IL-12 production and antigen-specific T-cell priming in BXH2 is likely to account for a significant component of ECM-resistance, we propose that even in the context of normal myeloid cell numbers, reduced IRF8-dependent transcriptional activation of APC-specific and T-cell specific pathways is sufficient to significantly blunt inflammatory response and protect against acute pathological inflammation. This is based on several observations. First, IRF8 behaves primarily as a transcriptional activator, not a repressor, in myeloid cells as can be seen by the enrichment of IRF8 binding sites in increased genes only (Figure 3-3B). Table S1 also highlights that for each gene regulated in a strain and infection specific way, *Irf8*-competent mice invariably show a higher magnitude fold change than BXH2, and the majority of these genes are increased (Group 1 and 2) rather than decreased (Group 3). Second, *Irf8*$^{R294C/+}$ F1 heterozygotes show normal numbers of myeloid cells (DCs, macrophages) and lymphoid cells, but still exhibit significant resistance to PbA-induced ECM (Figure 3-1). Thirdly, the inactivation of several direct transcriptional targets of IRF8 (identified in our study as bound and regulated by IRF8 during PbA infection) including the phagosome-associated small GTPase *Irgm1* (Figure 3-5), the pro-inflammatory cytokines *Il12p40* (Figure 3-5 and [284]), *Cxcl9*, and *Cxcl10* [141], the Cd11b receptor *Icam1* [149, 291], and the transcriptional activator and IRF8 dimerization partner *Irf1* [155] have been shown to increase resistance to ECM in mouse knockout mutants (Figure 3-5 and Refs [292, 293]). Finally, inactivation of additional IRF8 targets, detected herein by ChIP-seq, have

previously been shown to protect against PbA induced ECM, including *Ifng* [129], *Jak3* [199], *Cd8, Cd14, Cd40, Hc, Fcgr2, Lta* and *Ltbr* [120]. Together, these results highlight the role of IRF8 in regulating pro-inflammatory pathways in myeloid and lymphoid cells during ECM-associated neuroinflammation.

Although IRF8-dependent activation of pro-inflammatory pathways in myeloid cells has detrimental and pathological consequences during PbA infection, it clearly plays a protective role in other infections including pulmonary tuberculosis. Indeed, using the same analysis and stringent statistical parameters, we noted a strong overlap between the list of genes increased in brains of B6 mice in response to PbA infection and up-regulated in lungs 30 days following aerosol infection with *M. tuberculosis* (Table S2). Of the 123 genes increased in PbA-infected brains, more than half (n = 66) were also up-regulated in *M. tuberculosis*-infected lungs, and nearly three quarters (90/123) of the PbA regulated genes harbored an IRF8 binding site. Amongst the 66 genes up-regulated during both *M. tuberculosis* infection and during ECM, a striking 80% (n = 53) display one or more IRF8 binding sites. Furthermore, inactivating mutations in several of these common genes including *Irf8* [258], *Irf1* [294], and *Irgm1* [295] cause susceptibility to pulmonary tuberculosis, while conveying some degree of protection against ECM. Mutations in additional ECM increased genes such as *Tap1* and *B2m* also cause susceptibility to tuberculosis [296], while their direct effects on ECM susceptibility have yet to be tested. We propose that this set of 53 genes represents the core *Irf8*-dependent pro-inflammatory response pathways that play key roles in protection against TB, and pathological inflammation associated with ECM. Network analysis of these direct Irf8 targets highlights central nodes specific to myeloid cells in general, and dendritic cells in particular, including pro-inflammatory cytokines, type 1 interferon response and their transcriptional regulators.

Inactivation of *Irf8* or certain of its transcriptional targets leads to complete protection against ECM. On the other hand, reduced *Irf8* function causes partial protection (in $Irf8^{R294C/+}$ F1 heterozygotes). Such gene-dosage dependent effects raise the possibility that even small changes in expression or activity of IRF8 may have phenotypic consequences, with increased *Irf8* expression possibly associated with enhanced and/or chronic inflammation. Results from GWAS have pointed to *IRF8* as one of the genetic factors implicated in the complex genetic architecture of several human inflammatory conditions. For example, SNPs ~60kb downstream of *IRF8* are associated with increased risk for both ulcerative colitis [268] and Crohn's disease (lead SNP: rs16940202) [270, 286]. In agreement with our hypothesis, the same set of SNPs exhibit a remarkably similar association pattern with *IRF8* expression levels in colon and rectum (Gros and Georges, unpublished). The findings are consistent with the notion that one or more regulatory variants increase IBD risk by enhancing intestinal *IRF8* expression. In addition, *IRF8* is a key risk factor in multiple sclerosis (MS), and its association with this disease has been validated in multiple GWAS and meta-analyses [271, 273]. In MS, disease risk is associated with an expression SNP (rs17445836) mapping 61kb downstream of *IRF8* which is associated with both disease susceptibility and higher peripheral blood mononuclear cell *IRF8* mRNA levels [271]. Finally, a SNP near *IRF8* was found associated with systemic lupus erythematosus [297], a disease where production of type I interferon is central to pathogenesis. These results not only support a role for *IRF8* in human chronic inflammatory conditions but further suggest that, in agreement with our results in mice, even modest changes in expression or activity of *Irf8* in the context of persistent microbial or autoimmune stimulus, may lead to chronic or pathological inflammation. Furthering this proposal, we note that several IRF8 targets regulated during neuroinflammation in PbA-infected mice have also been

detected as genetic risk factors in GWAS of human chronic inflammatory conditions, including the MHC (type 1 diabetes, rheumatoid arthritis, lupus, MS, psoriasis), *CCL7* (IBD), *IRF1* (IBD), *IRF7* (Lupus) and *ICAM1* (IBD) (Table S3). This highlights the role of IRF8 and its regulated pathways in pathological inflammation in humans.

The mouse model of acute neuroinflammation induced by PbA infection has proven valuable to identify novel genes, proteins and pathways involved in pathological inflammatory conditions. This model may help prioritize genes identified in human GWAS for therapeutic development, including assessing activity of novel anti-inflammatory drug candidates for use in common human inflammatory conditions.

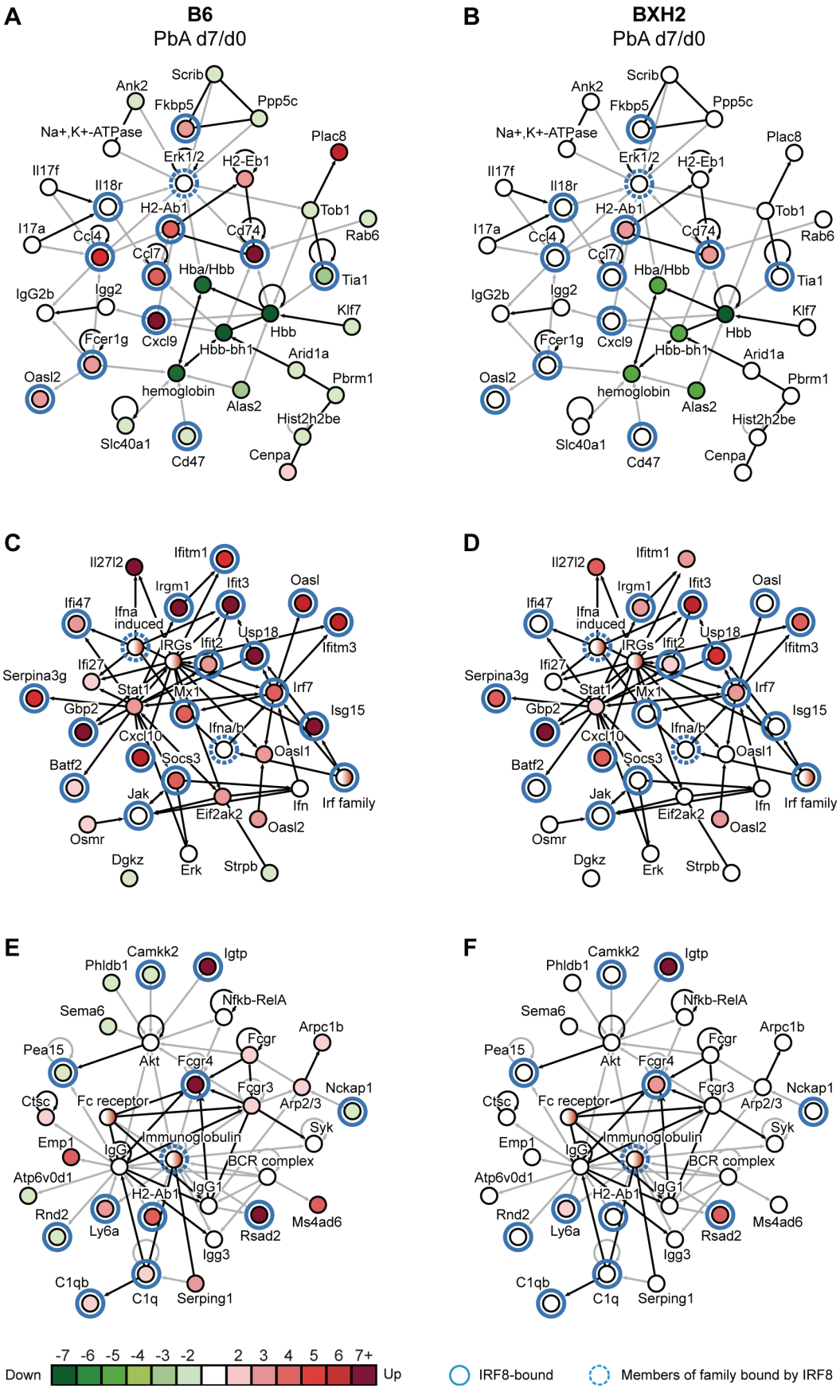
## 3.7 Acknowledgements:

We thank Dr. Michel Georges (University of Liège, Belgium) for insightful comments on the association of *IRF8* alleles in inflammatory bowel diseases, and for critical reading of this manuscript. JB was supported by a CIHR doctoral award and DL was supported by a postdoctoral fellowship from the CIHR/FRSQ training grant in cancer research FRN53888 of the McGill Integrated Cancer Research. This work was supported by a research grant to PG from the Canadian Institutes of Health Research, and the National Institutes of Health (USA). PG is the recipient of a James McGill Professorship.

## 3.8 Supplementary material

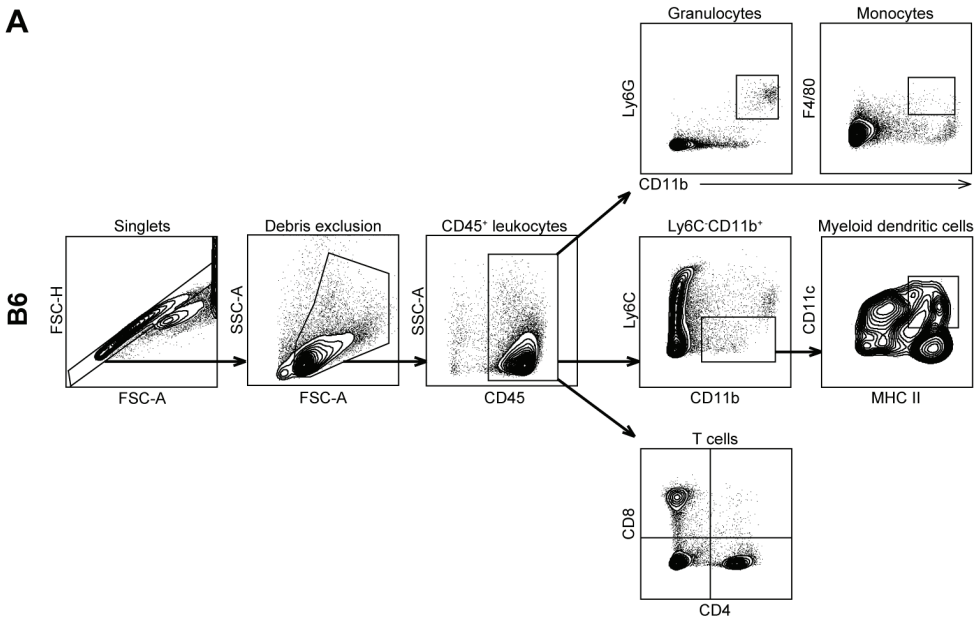
### Figure 3-8. (Supplementary S1) Network analysis of genes regulated by PbA in B6 mice

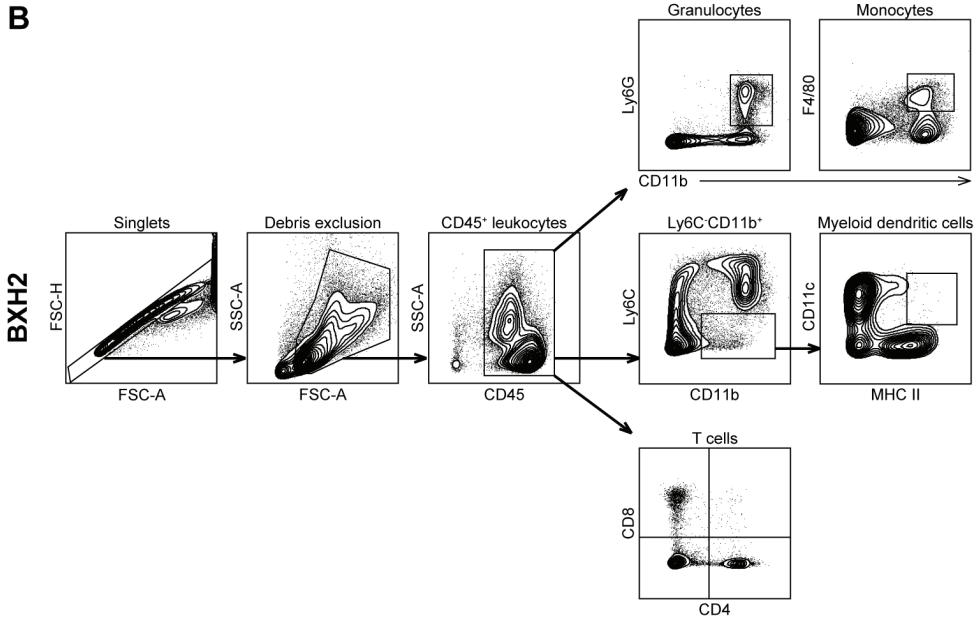
Gene interaction networks were adapted from those generated using Ingenuity Pathway Analysis and the three top scoring networks are depicted. Genes are indicated by circles with a blue ring to indicate those with IRF8 binding sites while arrows represent direct (black) or indirect (gray) biological connections within the networks. Left panels are colored according to fold change during infection in B6 mice while the right panels are colored by fold change in BXH2 mice. Multi-member gene families are indicated by a gradient rather than a single color. (A and B) Cellular signaling and hematology network, (C and D) interferon signaling and inflammatory network and (E & F) antigen presentation response. For clarity, indirect connections are not shown in panels C and D which feature highly interconnected gene sets.

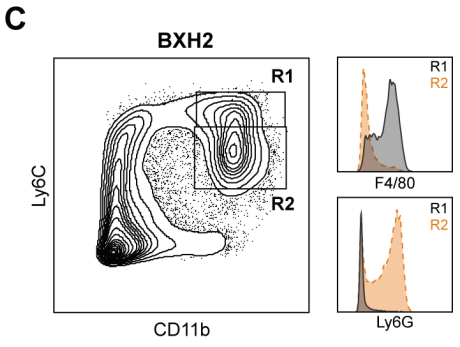

A
B6
PbA d7/d0
B
BXH2
PbA d7/d0
Scrib
Ank2
Fkbp5
Ppp5c
Na+,K+-ATPase
Plac8
Erk1/2
H2-Eb1
Il17f
Il18r
H2-Ab1
Tob1
Il7a
Ccl4
Cd74
Rab6
Ccl7
Hba/Hbb
Tia1
IgG2b
Igg2
Klf7
Fcer1g
Cxcl9
Hbb
Hbb-bh1
Arid1a
Oasl2
Pbrm1
hemoglobin
Alas2
Hist2h2be
Slc40a1
Cd47
Cenpa
C
D
Il27l2
Ifitm1
Irgm1
Ifit3
Oasl
Ifi47
Ifna induced
IRGs
Usp18
Ifitm3
Ifi27
Ifit2
Serpina3g
Stat1
Mx1
Irf7
Gbp2
Isg15
Ifna/b
Cxcl10
Oasl1
Socs3
Batf2
Jak
Ifn
Irf family
Eif2ak2
Oasl2
Osmr
Dgkz
Erk
Strpb
E
F
Camkk2
Igtp
Phldb1
Nfkb-RelA
Sema6
Arpc1b
Fcgr
Pea15
Akt
Fcgr4
Fcgr3
Nckap1
Ctsc
Fc receptor
Arp2/3
Syk
Emp1
Immunoglobulin
IgG
BCR complex
Atp6v0d1
IgG1
H2-Ab1
Rnd2
Ly6a
Ms4ad6
Rsad2
Igg3
C1qb
C1q
Serping1
-7 -6 -5 -4 -3 -2 2 3 4 5 6 7+
Down
Up
IRF8-bound
Members of family bound by IRF8

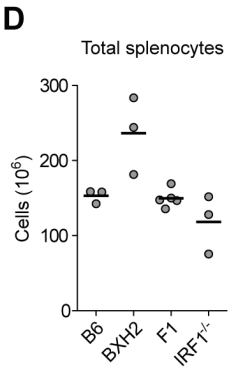
## Figure 3-9 (Supplementary S2). Gating strategy for immunophenotyping of spleen cells from PbA-infected mice.

The gating strategy is shown for representative specimens of B6 (A) and BXH2 (B) mice. An initial gate to isolate single cells was established based on the FSC-H and FSC-A. Cellular debris and dead cells were then excluded by SSC-A and FSC-A gating, and leukocytes selected according to the $CD45^+$ staining. For the myeloid compartment, granulocytes were identified as $CD11b^+Ly6G^+$ and monocytes/macrophages as $CD11b^+F4/80^+$. Myeloid dendritic cells were first gated as $Ly6C^-CD11b^+$ (excluding granulocytes and monocytes) and then further specified as $CD11c^+MHCII^+$. $CD8^+$ and $CD4^+$ T cells were identified as $CD8^+CD4^-$ and $CD4^+CD8^-$ staining. Every counter plot title denotes the gate within. (C) The characteristic BXH2 immature myeloid expansion, as shown in Figure 3-6A, is comprised of monocyte-like (Gate R1 cells are $F4/80^+$ and $Ly6G^-$) and granulocyte-like cells (Gate R1 cells are $F4/80^-$ and $Ly6G^+$). (D) Total number of spleen cells used to report cell type numbers in Figure 3-6 and 3-7.


A
B6
Singlets
Debris exclusion
CD45⁺ leukocytes
Granulocytes
Monocytes
Ly6C⁻CD11b⁺
Myeloid dendritic cells
T cells
FSC-H
FSC-A
SSC-A
CD45
Ly6G
F4/80
CD11b
Ly6C
CD11c
MHC II
CD8
CD4


B
BXH2
Singlets
Debris exclusion
CD45⁺ leukocytes
Granulocytes
Monocytes
Ly6C⁻CD11b⁺
Myeloid dendritic cells
T cells
FSC-H
FSC-A
SSC-A
CD45
Ly6G
F4/80
CD11b
Ly6C
CD11c
MHC II
CD8
CD4


C
BXH2
R1
R2
Ly6C
CD11b
F4/80
Ly6G


D
Total splenocytes
Cells (10⁶)
300
200
100
0
B6
BXH2
F1
IRF1-/-

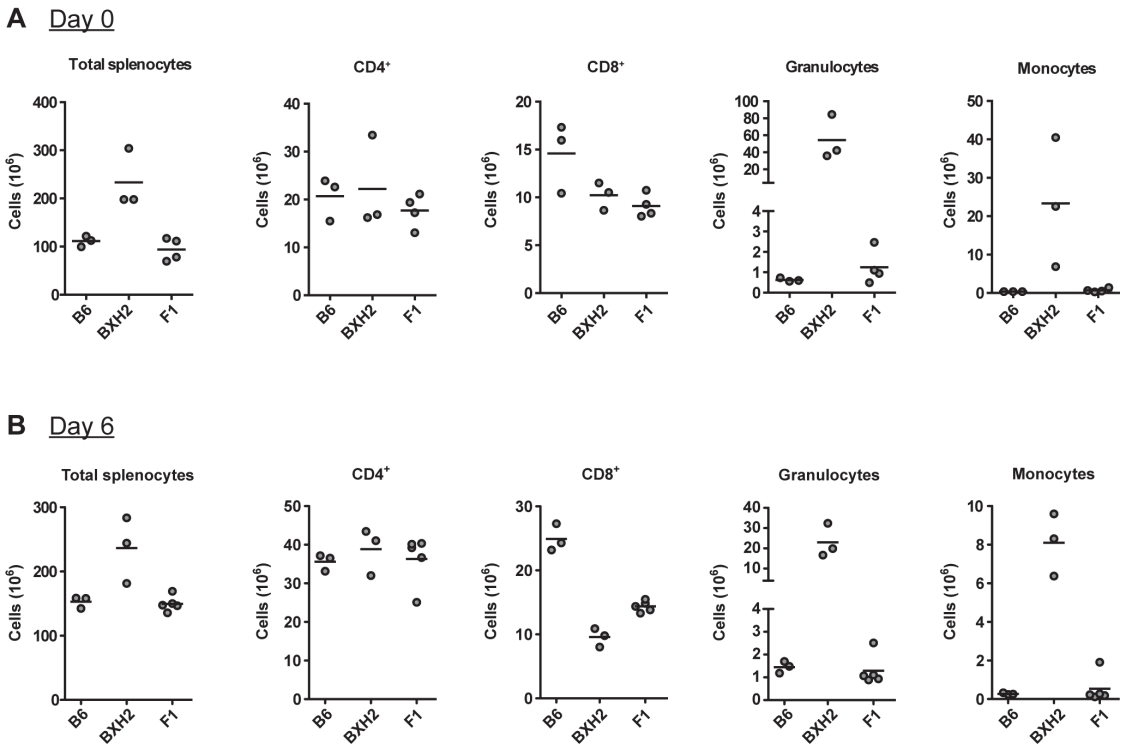



**Figure 3-10 (Supplementary S3). Total cell counts in spleen prior to and after PbA infection**

(A) Absolute number of splenocytes and leukocyte populations in B6, BXH2 and [BXH2×B6]F1 mice prior to infection (day 0). As shown in Figure 3-6A, BXH2 splenomegaly is associated with an immature myeloid cell hyperplasia. B6 and [BXH2×B6]F1 mice have highly similar cell numbers, with the exception of [BXH2×B6]F1 exhibiting a BXH2-like reduced number of $CD8^+$ T. (B) Total cell numbers in mice 6 days after PbA infection. This collection of graphs is reproduced from Figure S2, 3-6 and 3-7 for a basis of comparison with the day 0 values shown in (A).

**Table 3-1. Supplementary Table 1. Strain and infection specific regulation of gene expression in PbA infected B6 and BXH2 mice.**
Fold change of transcripts d7/d0 is shown, asterisks indicate that the fold change reported is the average of two or more probes (2-way ANOVA, 2>fold change cut-off, $p_{adj}$-interaction<0.05). Gene order is as seen in Figure 3C; and shaded rows indicate an IRF8 binding site within 20kb of the transcription start site.

| Group 1 genes | | | Group 2 genes | | | Group 3 genes | | |
|---|---|---|---|---|---|---|---|---|
| GeneID | B6 | BXH2 | GeneID | B6 | BXH2 | GeneID | B6 | BXH2 |
| Igtp | 17.29 | 6.56 | Xaf1 | 5.22 | 2.40 | H2afv | -1.44 | 1.06 |
| Cxcl10 | 14.90 | 3.18 | Ccl12* | 5.67 | 1.62 | EG381438 | -1.36 | 1.07 |
| Rsad2 | 12.79 | 3.94 | Psmb9 | 6.09 | 1.16 | Prkag2 | -1.39 | -1.06 |
| Fcgr4 | 11.50 | 2.33 | Cdkn1a* | 5.17 | -1.36 | Tmcc1 | -1.42 | 1.00 |
| Cd274 | 10.38 | 4.03 | Serpina3k | 4.71 | -1.11 | Atp2c1 | -1.40 | 1.01 |
| Gbp3* | 9.58 | 4.03 | Fosb | 4.19 | -1.34 | 1700123O20Rik | -1.35 | -1.00 |
| Irgm1 | 9.44 | 2.76 | Socs3 | 4.19 | 1.14 | Ccng2 | -1.33 | 1.00 |
| Isg15 | 7.61 | 3.99 | Ccl5 | 4.16 | 1.24 | Arl8b | -1.25 | 1.01 |
| Oasl2 | 8.31 | 3.25 | Ms4a6d | 4.07 | 1.21 | Slc4a4 | -1.39 | -1.14 |
| Plac8 | 7.41 | 2.49 | Emp1 | 3.81 | 1.10 | Emid2 | -1.57 | -1.15 |
| Cd74* | 6.66 | 2.97 | Ccl7* | 3.77 | 1.15 | Hes5 | -2.68 | -1.54 |
| Irf1* | 6.29 | 2.24 | C4b* | 4.54 | 1.32 | Gga3 | -2.24 | -1.41 |
| Cxcl9 | 8.25 | 1.57 | H2-Ab1* | 4.66 | 1.37 | D14Abb1e | -2.41 | -1.29 |
| Ccl4* | 7.20 | 1.39 | Tap1 | 4.63 | 1.63 | Flt1 | -2.89 | -1.29 |
| Plin4 | 7.11 | -2.20 | Cd52 | 4.19 | 1.57 | Hcn3 | -2.01 | -1.31 |
| | | | Chi3l4 | 4.15 | 1.71 | Akap1 | -2.14 | -1.25 |
| | | | Tap2 | 3.39 | 1.40 | Zbtb44 | -2.05 | -1.26 |
| | | | Lyz1 | 3.62 | 1.36 | Zfp523 | -2.07 | -1.21 |
| | | | Nlrc5 | 3.61 | 1.94 | E330009J07Rik | -2.45 | -1.14 |
| | | | Fkbp5 | 3.07 | -1.44 | Pak1 | -2.11 | -1.12 |
| | | | Map3k6* | 2.85 | -1.13 | Ccm2 | -2.07 | -1.12 |
| | | | Atp5k | 2.72 | -1.09 | E2f6 | -2.00 | -1.08 |
| | | | Angptl4 | 2.71 | -1.08 | 6330407J23Rik | -2.04 | -1.05 |
| | | | Txnip* | 2.60 | -1.09 | Fcrls | -2.15 | -1.03 |
| | | | AA467197 | 3.37 | 1.15 | Slco1c1 | -2.27 | -1.03 |
| | | | Slc10a6 | 3.23 | 1.01 | 2510022D24Rik | -1.90 | 1.01 |
| | | | Mt2 | 3.31 | -1.08 | Raf1 | -1.67 | 1.02 |
| | | | H2-L | 2.89 | 1.08 | Mettl17 | -2.27 | 1.08 |
| | | | Sult1a1 | 3.01 | 1.03 | Slc38a5 | -2.95 | 1.12 |
| | | | Slc15a3 | 3.04 | 1.15 | Tia1 | -3.07 | 1.03 |
| | | | Ch25h | 3.00 | 1.22 | | | |
| | | | Serping1* | 2.91 | 1.24 | | | |
| | | | Tagln2 | 2.81 | 1.28 | | | |
| | | | Fpr2 | 2.80 | 1.43 | | | |
| | | | Ifi47 | 3.01 | 1.60 | | | |
| | | | Oasl1 | 3.01 | 1.46 | | | |
| | | | Samhd1 | 3.10 | 1.50 | | | |
| | | | Icam1 | 3.10 | 1.36 | | | |

| | | | | |
|---|---|---|---|---|
| | Psmb8 | 2.40 | 1.64 | |
| | Batf2 | 2.26 | 1.50 | |
| | Osmr | 2.39 | 1.23 | |
| | Ifi205 | 2.47 | 1.34 | |
| | Ubd | 2.63 | 1.18 | |
| | Cyba | 2.29 | 1.12 | |
| | Saa3 | 2.26 | 1.07 | |
| | Bcl2a1b | 2.35 | 1.05 | |
| | H2-Eb1 | 2.41 | 1.05 | |
| | Bcl2a1d | 2.48 | 1.05 | |
| | 2410039M03Rik | 2.41 | -1.09 | |
| | Mobp | 2.37 | -1.10 | |
| | Pnpla2 | 2.28 | -1.13 | |
| | Gna13 | 2.28 | -1.06 | |
| | Cenpa | 2.20 | -1.32 | |
| | Synpo | 2.21 | -1.16 | |
| | Itgad* | 2.03 | -1.07 | |
| | Ugt1a6a | 2.05 | -1.02 | |
| | Adamts9 | 2.12 | 1.02 | |
| | Phyhd1 | 2.19 | -1.03 | |
| | Fcgr3 | 2.14 | 1.09 | |
| | Ctsc | 2.07 | 1.10 | |
| | Upp1* | 2.11 | 1.14 | |
| | Arpc1b | 2.04 | 1.05 | |

**Table 3-2. Supplementary Table 2. Genes regulated in B6 mouse brains during PbA infection, and comparison to infection-induced fold change to what has been reported in the lungs of B6 mice during *M. tuberculosis* infection.** The number of IRF8 ChIP-seq binding peaks within 20 kb of the TSS is also displayed where applicable. Genes bearing one or more IRF8 binding sites, and regulated in both cerebral malaria and pulmonary tuberculosis are highlighted in yellow.

| Gene ID | PbA d7/d0[a] | Mtb d30/d0 [b] | adj. p-value[c] | #IRF8 peaks[d] | Gene Name |
|---|---|---|---|---|---|
| Gh | 84.82 | | 1.1E-02 | 1 | Growth hormone |
| Gbp2 | 25.38 | 16.53 | 2.0E-03 | 1 | Guanylate binding protein 2 |
| Prl | 22.77 | | 1.6E-02 | | Prolactin |
| Igtp | 17.29 | 14.99 | 2.3E-03 | 3 | Interferon gamma induced GTPase |
| Cxcl10 | 14.90 | 92.96 | 1.1E-02 | 3 | Chemokine (C-X-C motif) ligand 10 |
| Ifit3 | 14.88 | 5.28 | 2.4E-03 | 3 | Interferon-induced protein with tetratricopeptide repeats 3 |
| Rsad2 | 12.79 | 3.93 | 5.3E-03 | 4 | Viperin |
| Fcgr4 | 11.50 | 13.19 | 3.1E-03 | 2 | Fc receptor, IgG, low affinity IV |
| Cd274 | 10.38 | 11.25 | 3.1E-03 | 2 | CD274 antigen |
| Irgm1 | 9.44 | 10.82 | 5.3E-03 | 1 | Immunity-related GTPase family M member 1 |
| Gbp3 | 9.40 | 12.61 | 2.0E-03 | 2 | Guanylate binding protein 3 |
| Cxcl9 | 8.25 | 413.81 | 8.8E-03 | 1 | Chemokine (C-X-C motif) ligand 9 |
| Ifi27l2a | 7.91 | 2.22 | 5.3E-03 | | Interferon stimulated gene 12 |
| Usp18 | 7.68 | 3.63 | 5.3E-03 | 3 | Ubiquitin specific protease 18 |
| Isg15 | 7.61 | | 2.0E-03 | 1 | ISG15 ubiquitin-like modifie |
| Plac8 | 7.41 | 4.20 | 7.6E-03 | | C15 protein |
| S3-12 | 7.11 | -4.02 | 1.9E-02 | | Perilipin |
| Cd74 | 6.51 | 2.25 | 2.5E-03 | 2 | CD74 antigen |
| Irf1 | 6.49 | 3.53 | 2.5E-03 | 2 | Interferon regulatory factor 1 |
| Rnf213 | 6.38 | 2.06 | 2.4E-03 | 2 | D11Ertd759e |
| Lgals3bp | 6.36 | 3.86 | 6.5E-03 | | Lectin, galactoside-binding, soluble, 3 binding protein |
| Psmb9 | 6.09 | 8.46 | 2.0E-03 | 5 | Proteasome (prosome, macropain) subunit, beta type 9 (large multifunctional peptidase 2) |
| Ccl12 | 6.06 | 10.68 | 1.0E-02 | 1 | Chemokine (C-C motif) ligand 12 |
| Irgm2 | 5.79 | 7.88 | 2.7E-02 | 2 | Immunity-related GTPase family M member 2 |
| Serpina3f | 5.63 | | 5.3E-03 | 1 | serine (or cysteine) peptidase inhibitor, clade A, member 3F |
| Oasl2 | 5.58 | 5.97 | 8.7E-03 | 1 | 2-5 oligoadenylate synthetase-like 2 |
| Ifitm3 | 5.57 | | 5.0E-03 | 5 | Interferon induced transmembrane protein 3 |

| | | | | | |
|---|---|---|---|---|---|
| Serpina3g | 5.49 | 24.45 | 2.5E-03 | 1 | serine (or cysteine) peptidase inhibitor, clade A, member 3G |
| Xaf1 | 5.22 | | 5.3E-03 | 1 | gene model 881 |
| Cdkn1a | 5.17 | | 4.5E-02 | | Cyclin-dependent kinase inhibitor 1A |
| Ifitm1 | 4.84 | | 1.2E-02 | 1 | Interferon induced transmembrane protein 1 |
| Parp14 | 4.74 | 2.78 | 8.8E-03 | 2 | Poly (ADP-ribose) polymerase family, member 14 |
| Ccl4 | 4.71 | 4.51 | 2.7E-02 | 4 | Strain SJL/J small inducible cytokine A4 |
| C4b | 4.64 | | 8.8E-03 | 1 | Complement component 4B |
| Tap1 | 4.63 | | 6.9E-03 | 5 | Transporter 1, ATP-binding cassette, sub-family B |
| Cd52 | 4.19 | 7.79 | 2.3E-02 | 3 | CD52 antigen |
| Socs3 | 4.19 | 3.06 | 4.5E-02 | 2 | Suppressor of cytokine signaling 3 |
| H2-K1 | 4.17 | 7.92 | 3.0E-03 | 1 | MRNA similar to histocompatibility 2, D region locus 1 |
| Ccl5 | 4.16 | 23.91 | 2.4E-02 | 4 | Chemokine (C-C motif) ligand 5 |
| Chi3l4 | 4.15 | | 5.6E-03 | 1 | Chitinase 3-like 4 |
| Ms4a6d | 4.07 | 18.49 | 1.5E-02 | | Membrane-spanning 4-domains, subfamily A, member 6D |
| Emp1 | 3.81 | | 3.2E-02 | | Epithelial membrane protein 1 |
| H2-Ab1 | 3.08 | 7.76 | 5.3E-03 | 1 | histocompatibility 2, class II antigen A, beta 1 |
| Mx2 | 3.75 | | 1.9E-02 | 1 | Myxovirus (influenza virus) resistance 2 |
| H2-D1 | 3.19 | 4.07 | 5.9E-03 | 4 | MHC class Ib antigen Qa-1 |
| Irf7 | 3.65 | 8.19 | 2.0E-03 | 1 | Interferon regulatory factor 7 |
| Lyz1 | 3.62 | | 1.0E-02 | 2 | Lysozyme 1 |
| Nlrc5 | 3.61 | 8.55 | 2.0E-03 | 1 | expressed sequence AI451557 |
| Tap2 | 3.39 | 4.21 | 5.6E-03 | 4 | Transporter 2, ATP-binding cassette, sub-family B |
| Txnip | 3.36 | | 4.1E-02 | 2 | Thioredoxin interacting protein |
| Chi3l3 | 3.36 | | 2.6E-03 | 1 | Chitinase 3-like 3 |
| Eif2ak2 | 3.32 | 2.53 | 4.6E-03 | | Eukaryotic translation initiation factor 2-alpha kinase 2 |
| Mt2 | 3.31 | | 4.5E-02 | | Metallothionein 2 |
| Samdh9l | 3.30 | | 1.0E-02 | 1 | sterile alpha motif domain containing 9-like |
| Stat1 | 3.25 | 17.30 | 2.1E-02 | | Signal transducer and activator of transcription 1 |
| Oas1g | 3.22 | 5.59 | 2.5E-03 | 1 | 2-5 oligoadenylate synthetase 1G |
| Ifit2 | 3.19 | 8.05 | 6.3E-03 | 1 | Interferon-induced protein with tetratricopeptide repeats 2 |
| Glipr2 | 3.13 | 5.28 | 3.5E-02 | 2 | GLI pathogenesis-related 2 |
| Tgm2 | 3.11 | | 3.0E-02 | 2 | Transglutaminase 2, C polypeptide |
| Icam1 | 3.10 | | 8.8E-03 | 2 | Intercellular adhesion molecule 1 |

| | | | | | |
|---|---|---|---|---|---|
| Psmb8 | 3.10 | 7.99 | 6.2E-03 | 5 | Proteasome (prosome, macropain) subunit, beta type 8 (large multifunctional peptidase 7) |
| Samhd1 | 3.10 | 3.93 | 5.9E-03 | 3 | SAM domain and HD domain, 1 |
| Fkbp5 | 3.07 | | 3.0E-02 | 1 | FK506 binding protein 5 |
| B2m | 3.04 | | 5.3E-03 | 3 | Beta-2 microglobulin |
| Slc15a3 | 3.04 | 6.14 | 2.5E-02 | 1 | Solute carrier family 15, member 3 |
| Ifi47 | 3.01 | 8.99 | 6.2E-03 | 3 | Interferon gamma inducible protein 47 |
| Sult1a1 | 3.01 | | 1.0E-02 | | Sulfotransferase family 1A, phenol-preferring, member 1 |
| Ch25h | 3.00 | 5.86 | 1.2E-02 | | Cholesterol 25-hydroxylase |
| 8430408G 22Rik | 2.91 | | 3.5E-02 | | RIKEN cDNA 8430408G22 gene |
| Ccl7 | 2.91 | 5.76 | 2.6E-02 | 1 | Chemokine (C-C motif) ligand 7 |
| Trim21 | 2.90 | 2.87 | 2.5E-03 | 3 | Tripartite motif-containing 21 |
| Serping1 | 2.88 | | 1.9E-02 | | Serine (or cysteine) peptidase inhibitor, clade G, member 1 |
| H2-K2 | 2.86 | | 9.0E-03 | 1 | LOC56628 |
| H2-T22 | 2.85 | 4.82 | 7.6E-03 | 3 | Histocompatibility 2, T region locus 10 |
| Map3k6 | 2.85 | | 3.3E-02 | | Mitogen-activated protein kinase kinase kinase 6 |
| D14Ertd6 68e | 2.83 | | 6.0E-03 | 1 | DNA segment, Chr 14, ERATO Doi 668, expressed, |
| - | 2.81 | | 1.6E-02 | | RIKEN cDNA 1200016E24 gene |
| Tagln2 | 2.81 | | 2.0E-02 | | Transgelin |
| Fpr2 | 2.80 | 10.17 | 2.1E-02 | 2 | Formyl peptide receptor 2 |
| Irf9 | 2.80 | 2.15 | 1.3E-02 | 1 | Interferon regulatory factor 9 |
| Pglyrp1 | 2.80 | 3.50 | 1.8E-02 | | Peptidoglycan recognition protein 1 |
| H2-Q2 | 2.78 | | 1.4E-02 | 2 | Histocompatibility 2, Q region locus 2 |
| Angptl4 | 2.71 | | 2.9E-02 | 1 | Angiopoietin-like 4 |
| Fcer1g | 2.71 | 5.57 | 2.1E-02 | 1 | Fc receptor, IgE, high affinity I, gamma polypeptide |
| Ly6a | 2.68 | | 7.5E-03 | 1 | Lymphocyte antigen 6 complex, locus A |
| Upp1 | 2.65 | 4.24 | 2.2E-02 | 1 | Uridine phosphorylase 1 |
| Ube1l | 2.64 | 2.60 | 1.1E-02 | 1 | Ubiquitin-activating enzyme E1-like |
| Oasl1 | 2.63 | 3.93 | 2.2E-02 | | Oligoadenylate synthetase-like protein-2 |
| Ubd | 2.63 | 259.08 | 5.3E-03 | 1 | Ubiquitin D |
| H2-L | 2.59 | | 2.7E-02 | 1 | H2-L |
| Tspo | 2.54 | | 1.6E-02 | 1 | Translocator protein |
| H2-Q7 | 2.52 | | 8.8E-03 | 2 | Histocompatibility 2, Q region locus 7 |
| Trim25 | 2.51 | 2.40 | 2.2E-02 | 1 | tripartite motif-containing 25 |
| Anxa2 | 2.49 | | 3.3E-02 | 1 | Annexin A2 |
| Bcl2a1d | 2.48 | | 5.0E-03 | 1 | B-cell leukemia/lymphoma 2 related |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | protein A1d |
| Ifi205 | 2.47 | 8.65 | 2.5E-03 | 1 | Interferon activated gene 205 |
| 2410039 M03Rik | 2.41 | | 4.6E-02 | | 2410039M03Rik |
| H2-Eb1 | 2.41 | 4.78 | 2.5E-03 | | Histocompatibility 2, class II antigen E beta |
| Osmr | 2.39 | | 2.7E-02 | | Oncostatin M receptor |
| Mobp | 2.37 | | 4.0E-02 | 1 | Myelin-associated oligodendrocytic basic protein |
| Bcl2a1b | 2.35 | 2.63 | 1.3E-02 | 1 | B-cell leukemia/lymphoma 2 related protein A1a |
| Bst2 | 2.33 | | 2.0E-03 | 1 | Bone marrow stromal cell antigen 2 |
| Cyba | 2.29 | 4.38 | 3.5E-02 | 1 | Cytochrome b-245, alpha polypeptide |
| Pnpla2 | 2.28 | | 2.9E-02 | | Patatin-like phospholipase domain containing 2 |
| Batf2 | 2.26 | | 4.6E-03 | 1 | Basic leucine zipper transcription factor, ATF-like 2 |
| Saa3 | 2.26 | 155.32 | 4.1E-02 | | Serum amyloid A 3 |
| Cmpk2 | 2.24 | 3.44 | 3.1E-03 | 3 | Cytidine monophosphate (UMP-CMP) kinase 2, mitochondrial |
| Itgad | 2.23 | | 4.1E-02 | | integrin, alpha D |
| Cenpa | 2.20 | 4.28 | 1.1E-02 | | Centromere protein A |
| Phyhd1 | 2.19 | | 4.0E-02 | | Phytanoyl-CoA dioxygenase domain containing 1 |
| H2-T17 | 2.18 | | 1.4E-02 | 3 | H2-T17 |
| Gm12250 | 2.16 | | 3.0E-02 | 2 | LOC215405 |
| Xdh | 2.16 | | 1.3E-02 | 4 | Xanthine dehydrogenase |
| Lgals9 | 2.15 | | 2.4E-03 | | Lectin, galactose binding, soluble 9 |
| Fcgr3 | 2.14 | 7.42 | 5.4E-03 | | Fc gamma receptor III |
| Ier3 | 2.14 | | 4.8E-02 | 2 | Immediate early response 3 |
| Adamts9 | 2.12 | | 2.2E-02 | | Adamts9 |
| Ifi27l1 | 2.10 | | 6.2E-03 | | Interferon, alpha-inducible protein 27 like 1 |
| Ctsc | 2.07 | 5.59 | 8.8E-03 | 1 | Cathepsin C |
| Gvin1 | 2.06 | | 1.6E-02 | 1 | GTPase, very large interferon inducible 1 |
| Ugt1a6a | 2.05 | | 7.6E-03 | 2 | UDP glucuronosyltransferase 1 family, polypeptide A6B |
| Arpc1b | 2.04 | | 3.5E-02 | | Actin related protein 2/3 complex, subunit 1B |
| C1qb | 2.02 | 13.98 | 2.5E-02 | 2 | Complement component 1, q subcomponent, beta polypeptide |
| Atxn7l3b | -2.00 | | 2.3E-02 | | predicted gene, ENSMUSG00000074747 |
| E2f6 | -2.00 | | 3.4E-02 | | E2F transcription factor 6 |
| Nckap1 | -2.00 | | 6.2E-03 | 1 | NCK-associated protein 1 |
| Pcyt2 | -2.00 | | 2.5E-02 | | Phosphate cytidylyltransferase 2, |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | ethanolamine |
| Tob1 | -2.00 | | 3.1E-02 | | Transducer of ErbB-2.1 |
| Abcf2 | -2.01 | | 4.6E-02 | 1 | ATP-binding cassette, sub-family F (GCN20), member 2 |
| Hcn3 | -2.01 | | 7.4E-03 | | Hyperpolarization-activated, cyclic nucleotide-gated K+ 3 |
| Mus81 | -2.01 | | 1.8E-02 | | MUS81 endonuclease homolog (yeast) |
| Ncoa1 | -2.01 | | 1.6E-02 | | Nuclear receptor coactivator 1 |
| Ntrk3 | -2.01 | | 3.0E-02 | | Neurotrophic tyrosine kinase, receptor, type 3 |
| Palm | -2.01 | | 1.3E-02 | | Paralemmin |
| Pbrm1 | -2.01 | | 2.5E-02 | | polybromo 1 |
| Strbp | -2.01 | -2.17 | 3.3E-02 | | RIKEN cDNA 6430510M02 gene |
| Arl8b | -2.02 | 2.23 | 1.1E-02 | 1 | ADP-ribosylation factor-like 8B |
| D17Wsu92e | -2.02 | | 1.4E-02 | 1 | DNA segment, Chr 17, Wayne State University 92, expressed |
| Gria2 | -2.02 | | 1.5E-02 | | Glutamate receptor, ionotropic, AMPA2 (alpha 2) |
| Lonrf2 | -2.02 | | 2.9E-02 | | LON peptidase N-terminal domain and ring finger 2 |
| Pltp | -2.02 | -2.51 | 3.0E-02 | 2 | Phospholipid transfer protein |
| Myt1l | -2.03 | | 6.1E-03 | | Myelin transcription factor 1-like |
| Oxct1 | -2.03 | | 8.7E-03 | 1 | Scot mRNA for succinyl CoA transferase |
| Taok1 | -2.03 | | 2.8E-02 | | TAO kinase 1 |
| Zfp385b | -2.03 | | 4.6E-03 | | Zinc finger protein 385B |
| 6330407J23Rik | -2.04 | | 2.9E-03 | | RIKEN cDNA 6330407J23 gene |
| Apba2 | -2.04 | | 2.5E-03 | | X11 protein mRNA, 3 end |
| Fam126b | -2.04 | | 1.1E-02 | 1 | Family with sequence similarity 126, member B |
| Sdr39u1 | -2.04 | | 1.6E-02 | 1 | Short chain dehydrogenase/ reductase family 39U, member 1 |
| 1190002N15Rik | -2.05 | | 2.1E-02 | | RIKEN cDNA 1190002N15 gene |
| Atxn7l3 | -2.05 | | 2.3E-02 | | ataxin 7-like 3 |
| Cyp46a1 | -2.05 | | 2.0E-02 | | Cytochrome P450, family 46, subfamily a, polypeptide 1 |
| Gna11 | -2.05 | | 1.5E-02 | | Guanine nucleotide binding protein, alpha 11 |
| Ppp5c | -2.05 | | 1.9E-02 | | Protein phosphatase 5, catalytic subunit |
| Ptn | -2.05 | | 1.7E-02 | | Pleiotrophin |
| Sept3 | -2.05 | | 2.6E-02 | | Septin 3 |
| Thra | -2.05 | -2.12 | 8.8E-03 | | Thyroid hormone alpha |
| Zbtb44 | -2.05 | | 6.2E-03 | | BC038156 |

| | | | | | |
|---|---|---|---|---|---|
| Hist1h2bf | -2.06 | | 1.9E-02 | | Histone cluster 1 |
| Ppp1r35 | -2.06 | | 3.0E-02 | | RIKEN cDNA 2010007H12 gene |
| Ccm2 | -2.07 | | 1.6E-02 | 2 | Cerebral cavernous malformation 2 homolog (human) |
| Glg1 | -2.07 | | 1.3E-02 | | Golgi apparatus protein 1 |
| Tmco3 | -2.07 | | 6.3E-03 | 1 | Transmembrane and coiled-coil domains 3 |
| Zfp523 | -2.07 | | 8.7E-03 | 1 | Zinc finger protein 523 |
| - | -2.08 | | 3.0E-02 | | LOC385086 |
| Carm1 | -2.08 | | 1.3E-02 | | Coactivator-associated arginine methyltransferase 1 |
| Gtrgeo22 | -2.08 | | 3.5E-02 | 1 | Gene trap ROSA b-geo 22 |
| Pacsin1 | -2.08 | | 3.8E-03 | | Protein kinase C and casein kinase substrate in neurons 1 |
| Pde6d | -2.08 | | 7.6E-03 | | Phosphodiesterase 6D, cGMP-specific, rod, delta |
| Rnf167 | -2.08 | | 4.0E-02 | 1 | Ring finger protein 167 |
| Tmem63b | -2.08 | | 1.2E-02 | | Transmembrane protein 63b |
| Nfix | -2.09 | -2.02 | 1.3E-02 | | Nuclear factor I/X |
| Caln1 | -2.09 | | 3.1E-02 | | Calneuron 1 |
| Dgkb | -2.09 | | 8.7E-03 | | Diacylglycerol kinase, beta |
| Mbtps1 | -2.09 | | 2.9E-02 | | Membrane-bound transcription factor peptidase, site 1 |
| Rora | -2.09 | | 3.0E-02 | | RAR-related orphan receptor alpha |
| Scrib | -2.09 | | 1.7E-02 | | Scribbled homolog (Drosophila) |
| Shank3 | -2.09 | -3.49 | 3.8E-03 | | SH3/ankyrin domain gene 3 |
| Bcl11b | -2.10 | 5.31 | 8.8E-03 | | B-cell leukemia/lymphoma 11B |
| Cops7a | -2.10 | | 1.3E-02 | | COP9 (constitutive photomorphogenic) homolog, subunit 7a (Arabidopsis thaliana) |
| Fam178a | -2.10 | | 1.4E-02 | | family with sequence similarity 178, member A |
| Foxq1 | -2.10 | | 8.8E-03 | | Forkhead box Q1 |
| Rhobtb2 | -2.10 | | 1.6E-02 | 1 | Rho-related BTB domain containing 2 |
| Zfp612 | -2.10 | | 8.8E-03 | | Zinc finger protein 612 |
| Eif5a | -2.11 | | 1.6E-02 | 1 | Eukaryotic translation initiation factor 5A |
| Pak1 | -2.11 | 3.15 | 5.9E-03 | | P21 (CDKN1A)-activated kinase 1 |
| Tspan3 | -2.11 | | 3.6E-02 | | Tetraspanin 3 |
| 1110012J17Rik | -2.12 | -2.39 | 9.3E-03 | | RIKEN cDNA 1110012J17 gene |
| Dgkz | -2.12 | | 1.8E-02 | | Diacylglycerol kinase zeta |
| Gnao1 | -2.12 | | 2.8E-02 | | Guanine nucleotide binding protein, alpha O |
| Hsd3b2 | -2.12 | | 2.7E-02 | | Hydroxy-delta-5-steroid dehydrogenase, 3 beta- and steroid |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | delta-isomerase 2 |
| Ntm | -2.13 | | 1.5E-02 | | Neurotrimin |
| Slc38a9 | -2.13 | | 2.3E-02 | 1 | Solute carrier family 38, member 9 |
| Cd47 | -2.14 | | 1.6E-02 | 1 | CD47 antigen (Rh-related antigen, integrin-associated signal transducer) |
| Gats | -2.14 | | 2.2E-02 | | Opposite strand transcription unit to Stag3 |
| Rab5b | -2.14 | | 1.3E-02 | | RAB5B, member RAS oncogene family |
| Fam63b | -2.15 | | 6.5E-03 | 1 | MKIAA1164 protein |
| Fcrls | -2.15 | -2.29 | 7.4E-03 | | IFGP2 |
| Pea15a | -2.15 | | 3.9E-02 | 1 | Phosphoprotein enriched in astrocytes 15A |
| Skiv2l | -2.15 | | 2.9E-02 | | Superkiller viralicidic activity 2-like (S. cerevisiae) |
| Usf2 | -2.15 | | 4.0E-02 | 3 | Upstream transcription factor 2 |
| Jph4 | -2.16 | | 1.6E-02 | | Junctophilin 4 |
| Klhdc1 | -2.16 | | 7.6E-03 | | Kelch domain containing 1 |
| Ltbp4 | -2.16 | -4.99 | 1.5E-02 | | Latent transforming growth factor beta binding protein 4 long splice variant |
| Zbtb7a | -2.16 | | 2.5E-03 | | Zinc finger and BTB domain containing 7a |
| Akap1 | -2.17 | | 3.6E-02 | 1 | A kinase (PRKA) anchor protein 1 (Akap1), nuclear gene encoding mitochondrial protein |
| Gtf3a | -2.17 | | 7.6E-03 | | General transcription factor III A |
| Mau2 | -2.17 | | 1.6E-02 | | RIKEN cDNA 9130404D08 gene |
| Rasgef1a | -2.17 | -2.06 | 6.6E-03 | | RasGEF domain family, member 1A |
| Slc40a1 | -2.17 | | 1.6E-02 | | Solute carrier family 40 (iron-regulated transporter), member 1 |
| Klf7 | -2.18 | | 2.3E-02 | | Kruppel-like factor 7 |
| Camkk2 | -2.18 | | 2.6E-03 | 1 | Calcium/calmodulin-dependent protein kinase kinase 2, beta |
| Fam171b | -2.18 | | 1.1E-02 | | Family with sequence similarity 171, member B |
| Kalrn | -2.18 | | 6.3E-03 | | 2210407G14Rik |
| Scamp3 | -2.18 | | 7.6E-03 | | CDC-like kinase 2 |
| Mzt1 | -2.19 | | 3.0E-02 | | RIKEN cDNA 2410129H14 gene |
| Phldb1 | -2.19 | | 2.5E-03 | | Pleckstrin homology-like domain, family B, member 1 |
| Sgtb | -2.19 | | 2.0E-02 | | Small glutamine-rich tetratricopeptide repeat (TPR)-containing, beta |
| Rnd2 | -2.20 | | 2.7E-02 | 1 | Rho family GTPase 2 |
| Tcf4 | -2.20 | -2.12 | 7.6E-03 | | Transcription factor 4 |

| | | | | | |
|---|---|---|---|---|---|
| Celsr2 | -2.21 | | 2.0E-02 | | Cadherin, EGF LAG seven-pass G-type receptor 2 (flamingo homolog, Drosophila) |
| Ptprd | -2.21 | -2.54 | 1.4E-02 | | protein tyrosine phosphatase, receptor type, D |
| 3110047P20Rik | -2.22 | | 2.2E-02 | | RIKEN cDNA 3110047P20 gene |
| Dlgap1 | -2.22 | | 8.8E-03 | | Discs, large (Drosophila) homolog-associated protein 1 |
| Mgat4b | -2.22 | | 1.6E-02 | 3 | Mannoside acetylglucosaminyltransferase 4, isoenzyme B |
| Usp11 | -2.22 | | 3.1E-02 | | Ubiquitin specific peptidase 11 |
| Fbxo41 | -2.23 | | 9.8E-03 | | F-box protein 41 |
| B4galt3 | -2.24 | | 1.6E-02 | 1 | UDP-Gal:betaGlcNAc beta 1,4-galactosyltransferase, polypeptide 3 |
| Fjx1 | -2.24 | | 3.4E-02 | | Four jointed box 1 (Drosophila) |
| Gga3 | -2.24 | | 6.1E-03 | | Golgi associated, gamma adaptin ear containing, ARF binding protein 3 |
| Sema6d | -2.24 | -2.23 | 1.4E-02 | | Sema domain, transmembrane domain (TM), and cytoplasmic domain, (semaphorin) 6D |
| Smpd1 | -2.24 | | 2.9E-02 | | Sphingomyelin phosphodiesterase 1, acid lysosomal |
| Rbfox1 | -2.25 | | 1.3E-02 | 1 | Hexaribonucleotide binding protein 1 |
| Arid1a | -2.25 | | 7.6E-03 | | AT rich interactive domain 1A (SWI-like) |
| Bcan | -2.25 | | 4.6E-03 | | Brevican |
| Tmod1 | -2.25 | -2.04 | 3.1E-02 | | Tropomodulin 1 |
| Rab14 | -2.26 | | 1.0E-02 | 2 | RAB14, member RAS oncogene family |
| Brunol4 | -2.27 | | 1.1E-02 | | CUGBP, Elav-like family member 4 |
| Mettl17 | -2.27 | | 1.3E-02 | | D14Ertd209e |
| Slco1c1 | -2.27 | | 2.0E-02 | | Solute carrier organic anion transporter family |
| Tbc1d17 | -2.27 | | 5.3E-03 | | TBC1 domain family, member 17 |
| Acot7 | -2.28 | | 3.7E-02 | 1 | BACH mRNA for acyl-CoA hydrolase |
| Efna5 | -2.28 | | 1.5E-02 | | Ephrin A5 |
| Ank2 | -2.29 | | 2.3E-02 | | Ankyrin 2, brain |
| C1qtnf4 | -2.29 | | 5.3E-03 | | C1q and tumor necrosis factor related protein 4, |
| Orf61 | -2.29 | | 9.0E-03 | | open reading frame 61 |
| AI593442 | -2.30 | | 1.5E-02 | | expressed sequence AI593442 |
| Jhdm1d | -2.30 | 2.03 | 2.2E-02 | 1 | jumonji C domain-containing histone demethylase 1 homolog D (S. cerevisiae) |
| Spnb4 | -2.30 | | 2.0E-02 | 2 | BetaIV-spectrin sigma1 |

| | | | | | |
|---|---|---|---|---|---|
| Acvr2b | -2.32 | | 6.3E-03 | | Activin receptor IIB |
| Elavl3 | -2.39 | | 5.3E-03 | | RNA-binding protein mHuC-S |
| Rab6 | -2.34 | | 1.2E-02 | | RAB6, member RAS oncogene family |
| Pak3 | -2.35 | | 1.1E-02 | | P21-activated kinase 3 |
| Dusp8 | -2.38 | | 1.9E-02 | | Dual specificity phosphatase 8 |
| Tnrc6a | -2.38 | | 1.1E-02 | | Trinucleotide repeat containing 6a |
| Arl2bp | -2.39 | | 5.6E-03 | | ADP-ribosylation factor-like 2 binding protein (Arl2bp), transcript variant 1 |
| Mtss1l | -2.39 | | 2.9E-02 | 1 | Metastasis suppressor 1-like |
| Cacng3 | -2.40 | | 1.2E-02 | 1 | Calcium channel, voltage-dependent, gamma subunit 3 |
| Ppp1ca | -2.40 | | 2.3E-02 | | Protein phosphatase 1, catalytic subunit, alpha isoform |
| Rtn1 | -2.40 | | 8.8E-03 | | Reticulon 1 |
| Gabrb1 | -2.41 | | 1.6E-02 | | Gamma-aminobutyric acid (GABA-A) receptor, subunit beta 1 |
| D14Abb1e | -2.41 | | 9.7E-03 | | D14Abb1e |
| Rgs7bp | -2.42 | | 1.5E-02 | | Regulator of G-protein signalling 7 binding protein |
| Syt4 | -2.42 | | 6.6E-03 | | Synaptotagmin IV |
| Ugt8a | -2.44 | | 5.0E-02 | | UDP galactosyltransferase 8A |
| E330009J07Rik | -2.45 | | 3.6E-02 | | RIKEN cDNA E330009J07 gene |
| Gjc2 | -2.45 | | 2.5E-03 | 1 | Gap junction protein, gamma 2 |
| Kif3a | -2.46 | | 7.6E-03 | | Kinesin family member 3A |
| Nfib | -2.48 | -5.75 | 6.2E-03 | | Strain C57BL/6J nuclear factor I/B (Nfib) |
| Unc13c | -2.48 | | 3.5E-03 | | LOC235480 |
| Atp6v0d1 | -2.49 | | 2.7E-02 | | ATPase, H+ transporting, lysosomal V0 subunit D1 |
| Prrt1 | -2.50 | | 1.7E-02 | 1 | Proline-rich transmembrane protein 1 |
| 4930402H24Rik | -2.50 | -2.44 | 1.6E-02 | | RIKEN cDNA 4930402H24 gene |
| Hes5 | -2.53 | | 1.1E-02 | | Hairy and enhancer of split 5 (Drosophila) |
| Mark2 | -2.59 | | 1.3E-02 | | MAP/microtubule affinity-regulating kinase 2 (Mark2), transcript variant 1 |
| Msl1 | -2.59 | | 8.8E-03 | | Male-specific lethal 1 homolog (Drosophila) |
| Gdi1 | -2.60 | | 3.1E-02 | | Guanosine diphosphate (GDP) dissociation inhibitor 1 |
| Slc38a2 | -2.64 | | 2.1E-02 | 1 | Solute carrier family 38, member 2 |
| Psd2 | -2.68 | | 1.3E-02 | | Pleckstrin and Sec7 domain containing 2 |
| Ppap2b | -2.69 | -2.65 | 1.5E-02 | 1 | Phosphatidic acid phosphatase type |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | 2B |
| Grm4 | -2.75 | | 2.9E-02 | | glutamate receptor, metabotropic 4 |
| Epn2 | -2.79 | -2.06 | 2.1E-02 | | Epsin 2 |
| 2210018 M11Rik | -2.83 | | 1.0E-02 | | 2210018M11Rik |
| Lphn1 | -2.85 | -2.17 | 1.6E-02 | | latrophilin 1 |
| Flt1 | -2.89 | -2.08 | 5.6E-03 | | FMS-like tyrosine kinase 1 |
| Slc38a5 | -2.95 | -3.45 | 2.3E-02 | | Solute carrier family 38, member 5 |
| Alas2 | -3.04 | -2.56 | 2.9E-02 | | Aminolevulinic acid synthase 2 |
| Itm2a | -3.05 | | 5.9E-03 | | Integral membrane protein 2A |
| Tia1 | -3.07 | | 1.5E-02 | 1 | cytotoxic granule-associated RNA binding protein 1 |
| Prkcb | -3.21 | 2.70 | 2.0E-02 | 1 | Protein kinase C, beta |
| Cxcl12 | -4.18 | 2.58 | 1.4E-02 | | Chemokine (C-X-C motif) ligand 12 |
| Hbb-b1 | -23.47 | | 2.1E-02 | | Hemoglobin, beta adult minor chain |
| Hba-a1 | -24.73 | | 2.9E-02 | | Hemoglobin alpha, adult chain 1 |

**Table 3-3. Supplementary Table 3. Transcriptional response to *P. berghei* in CM-susceptible B6 mice, sorted according to ontology category (AMIGO).** IRF8 targets (indicated by bold text) are significantly enriched in upregulated genes. Superscript labels refer to genes where the human ortholog has been identified in GWAS studies for psoriasis (P), rheumatoid arthritis (RH), celiac disease (C), Crohn's disease (CD), ulcerative colitis (UC), diabetes (D), multiple sclerosis (MS), systemic lupus erythmatous (SLE), irritable bowel disease (IBD), or where the human ortholog is found in the MHC (MHC).

| Ontology | Upregulated genes | Downregulated genes |
|---|---|---|
| **Innate immunity** | **C1qb, C4b$^{MHC}$, Chi3l3, Chi3l4, Cyba, Gvin1, Ifi205, Ifi47$^{CD}$, Ifit3**, Map3k6, **Nlrc5**, Oasl1, **Oasl2**, Pglyrp1, Saa3, **Samd9l**, Serping1, **Trim21** | C1qtnf4 |
| **Response to virus** | **Bcl2a1b, Bst2**, Eif2ak2, **Ier3$^{MHC}$**, Ifi27l2a, **Ifit2, Ifitm1, Ifitm3, Isg15**, Ly6a, **Mx2, Oas1g$^{D}$, Rsad2, Samhd1** | Skiv2l$^{MHC}$ |
| **Chemokines, cytokines, receptors** | **Ccl4, Ccl5, Ccl7$^{UC}$, Ccl12$^{IBD}$, Cxcl9, Cxcl10**, Osmr, **Socs3$^{MS,IBD,D}$** | Cxcl12 |
| **Response to stimulus, signal transduction** | **Angptl4, Ctsc, Fpr2**, Mt2, **Serpina3f**, | Acvr2b, Dgkb, Dgkz, Dlgap1, Grm4, Lphn1, **Mtss1l**, Pacsin1$^{MHC}$, **Prrt1$^{MHC, IBD}$**, Psd2, Rasgef1a, Rgs7bp, Tnrc6a, Unc13c |
| **Adaptive Immunity, antigen processing and presentation** | **B2m, Cd274, Cd52, Cd74, Fcer1g**, Fcgr3$^{SLE,IBD}$, **Fcgr4$^{SLE, IBD}$, H2-Ab1$^{C,MHC}$, H2-D1/L$^{MHC}$,** H2-Eb1$^{MHC, MS,RA}$, **H2-K1$^{MHC}$, H2-K2$^{MHC}$, H2-Q2$^{MHC}$, H2-Q7$^{MHC}$,** H2-Qa1$^{MHC}$, **H2-T22$^{MHC}$, Psmb8$^{MHC}$, Psmb9$^{MHC}$, Tap1$^{MHC}$, Tap2$^{MHC}$** | Fclrs |
| **Transcription factor, regulation of transcription** | **Batf2, D14Ertd668e, Irf1$^{UC}$, Irf7$^{SLE}$, Irf9**, Stat1$^{IBD}$, **Txnip** | 2210018M11Rik, Arid1a, Atxn7l3, Atxn7l3b, Bcl11b, Carm1, E2f6, Foxq1, Gtf3a, Hes5, Hist1h2bf, Hopx, **Jhdm1d**, Klf7, Msl1, Myt1l, Ncoa1, Nfix$^{C}$, Pbrm1, **Prkcb$^{IBD}$, Rbfox1**, Rora, Tcf4, Usf2, Zbtb44, Zbtb7a, **Zfp523** |

| | | |
|---|---|---|
| GTP signaling | **Gbp2**, **Gbp3**, **Igtp**, **Irgm1**[IBD], **Irgm2** | Gdi1, Gna11[MHC], Gnao1, **Rab14**, Rab5b, Rab6, **Rhobtb2**, **Rnd2**, Sept3, Tbc1d17 |
| Cell cycle and proliferation, cellular differentiation | Arpc1b, Cdkn1a, Cenpa, Emp1, **Gh**, Prl, Tagln2, **Xdh** | **Arl8b**, Efna5, Elavl3, **Gm16517**, Itm2a, Ltbp4, Mau2, Mzt1, **Nckap1**, Nfib, Ntrk3, Ptn, **Rnf167**, Scrib, Sema6d, Strbp, Thra, Tmod1, Tob1 |
| Adhesion | **Icam1**[P,MS,IBD], Itgad, Lgals3bp, Lgals9[CD] | Bcan, **Cd47**, Celsr2, Ntm |
| Apoptosis | **Bcl2a1d**, Ifi27l1, **Serpina3g**, **Tspo**, **Xaf1** | Ank2, **Tia1** |
| Protein kinase, phosphotase | **Cmpk2** | **Akap1**, **Camkk2**, Dusp8, Fjx1, Kalrn, Mark2, Pak3, Ppp1r35, Ppp5c, Ptprd, Taok1 |
| Ubiquitination | **Parp14**, **Rnf213**, **Trim25**, **Ubd**[MHC], **Ube1l**, **Usp18** | Fbxo41, Usp11 |
| RNA processing, translation | | Brunol4, **Eif5a**, Mettl17, Pabpn1, Rbfox1 |
| Transport | **Slc15a3** | **Abcf2**, Apba2, Arf5, **Cacng3**, Gabrb1, Gga3, Gria2, Hcn3, **Pea15a**, **Pltp**, Scamp3[7], Slc38a2, Slc38a5, Slc38a9, Slc40a1, Slco1c1, Syt4, **Tmco3**, Ugt8a |
| Blood cells and vessels | **Anxa2**, **Tgm2** | Alas2, **Ccm2**, Flt1, Hba-a1, Hbb-b1, Pak1, **Ppap2b** |
| Neuronal and junctions | **Mobp** | Epn2, **Gjc2**, Kif3a, Klc1, Palm, Shank3, **Spnb4** |
| Metabolic processes | Adamts9, Ch25h, Fkbp5[MHC], **Lyz1**, Phyhd1, Pnpla2, Sult1a1, **Ugt1a6a**, **Upp1** | 0610007P14Rik, 1190002N15Rik, **Acot7**, Atp6v0d1, **B4galt3**, Cyp46a1, Glg1, Hsd3b2, Mbtps1, **Mgat4b**, Mus81, **Oxct1**, Pcyt2, Pde6d, Phldb1, Ppp1ca, **Sdr39u1**, Smpd1 |
| Biological processes or unannotated | 2410039M03Rik, 8430408G22Rik, **Glipr2**, **Gm12250**, Ms4a6d, Plac8, Plin4 | 1110012J17Rik, 2600009P04Rik, 2900011O08Rik, 3110047P20Rik, 4930402H24Rik, 6330407J23Rik, AI593442, Caln1, Cops7a, D14Abble, **D17Wsu92e**, E330009J07Rik, **Fam126b**, Fam171b, Fam178a, **Fam63b**, Fndc5, Gats, Jph4, Klhdc1, Lonrf2, Orf61, Rtn1, Sgtb, Tmem63b, Tspan3, Zfp385b, Zfp612 |

## Preface to Chapter 4

In the previous chapters, we have demonstrated how mouse models can be used to both (1) uncover novel genetic underpinnings of resistance to malaria via QTL mapping, and (2) to dissect some of the mechanisms of pathology. The following chapter builds on work previously done in the lab by Gundula Min-Oo, who used the blood stage model of *P. chabaudi* infection to identify erythrocyte pyruvate kinase, *Pklr,* as the gene underlying *Char4*, a *P. chabaudi* resistance QTL.[102] Promisingly, in collaboration with Kevin Kain's research group, we were able to reproduce indicators of malaria resistance (decreased parasite replication, increased phagocytosis) in PKLR deficient blood using *P. falciparum* in vitro cultures.

To determine if the relationship between PKLR deficiency and malaria resistance has been relevant within malaria endemic population

# Chapter 4 : Genetic diversity in human erythrocyte pyruvate kinase

## 4.1 Abstract

Previously, we have shown that pyruvate kinase, liver and red cell isoform (PKLR) deficiency protects mice *in vivo* against blood-stage malaria, and observed that reduced PKLR function protects human erythrocytes against *Plasmodium falciparum* replication *ex vivo*. Here, we have sequenced the human *PKLR* gene in 387 individuals from malaria-endemic and other regions in order to assess genetic variability in different geographical regions and ethnic groups. Rich genetic diversity was detected in PKLR, including 59 single-nucleotide polymorphisms and several loss-of-function variants (frequency 1.5%). Haplotype distribution and allele frequency varied considerably with geography. Neutrality testing suggested positive selection of the gene in the sub-Saharan African and Pakistan populations. It is possible that such positive selection involves the malarial parasite.

## 4.2 Introduction:

Erythrocytes rely exclusively on glycolysis for production of adenosine triphosphate (ATP), and, therefore, loss of pyruvate kinase, liver and red cell isoform PKLR activity has major consequences for cell function and lifespan. PK deficiency is the second most common cause of hereditary non-spherocytic hemolytic anemia in humans[298]. The large mutational spectrum in *PKLR* is associated with a broad clinical spectrum, varying from mild anemia to a transfusion-dependent disorder[299, 300]. Clinical PK deficiency is usually associated with homozygosity or compound heterozygosity of mutant alleles[300].

PK-deficiency protects mice against blood-stage malaria (*Plasmodium chabaudi*) *in vivo*[102]; and studies with *P. falciparum*-infected erythrocytes demonstrate both homozygosity and heterozygosity for mutant *PKLR* alleles decrease parasite invasion and replication, and increase phagocytosis of infected cells[101, 104]. These findings raise the possibility that human *PKLR* variants may have a protective effect if retained in malaria-pressured areas. To address this possibility, we sequenced the *PKLR* gene from 387 individuals, primarily those of African and Asian origin living in malaria-endemic areas (Figure 4-1a).

## 4.3 Results and discussion:

Sequencing *PKLR* identified 59 polymorphic sites, including 30 previously unreported single nucleotide polymorphisms (SNPs; Figure 4-1b, Supplementary table 2 for details). Full nucleotide sequences can be found by accessing the CEPH database. Our data is almost certainly an underestimate as the sequencing technology failed to accurately read through two highly repetitive A-T rich regions of the gene [301]. The majority of SNPs were in non-coding regions, though 7 synonymous coding SNPs (A36A, Q57Q, L61L, R194R, L236L, L301L and R569R), and 7 non-synonymous mutations were detected (Figure 4-2). Many SNPs (21/59) were limited to a single heterozygote individual. The sub-Saharan population (SS) showed the highest diversity with 31 polymorphisms, including 13 unique sites (Figure 4-1b; Supplementary Table 3). The European (EUR) dataset had the lowest genetic diversity (eight SNPs), though only a small number of individuals were included (n=23).

All SNPs were in Hardy-Weinberg equilibrium within the populations in which they were detected with the exception of rs3762272 in Pakistan (PAK), which had an excess of minor allele homozygotes ($p<0.0002$). Most of these were Hazara, likely indicating population substructure in the PAK group. This SNP was also enriched in PAK (minor allele frequency, $MAF_{PAK}=11.6\%$) and Southeast Asia (SEA, $MAF_{SEA}=19.6\%$) while being nearly absent in Africa.

We observed linkage disequilibrium across the entire gene in all but the EUR population. To compare haplotype distribution between populations, we used the seven SNPs indicated by arrows in Figure 4-1b which assigned >90% of samples to a haplotype with a frequency >2% (Figure 4-1c). Two haplotypes dominated: 5'-TTGGCCT-3' (haplotype 1), which accounts for the majority of samples in North Africa (NA), PAK, SEA and EUR and matches the USCS reference genome, and 5'-CCCGATC-3' (haplotype 2), which accounts for the majority of SS and South African (SA) haplotypes. The ancestral haplotype 7 was represented at 10.4% frequency in SS, but below 5% in other populations. All populations had significantly different haplotype distributions ($X^2$ test, $p<0.05$).

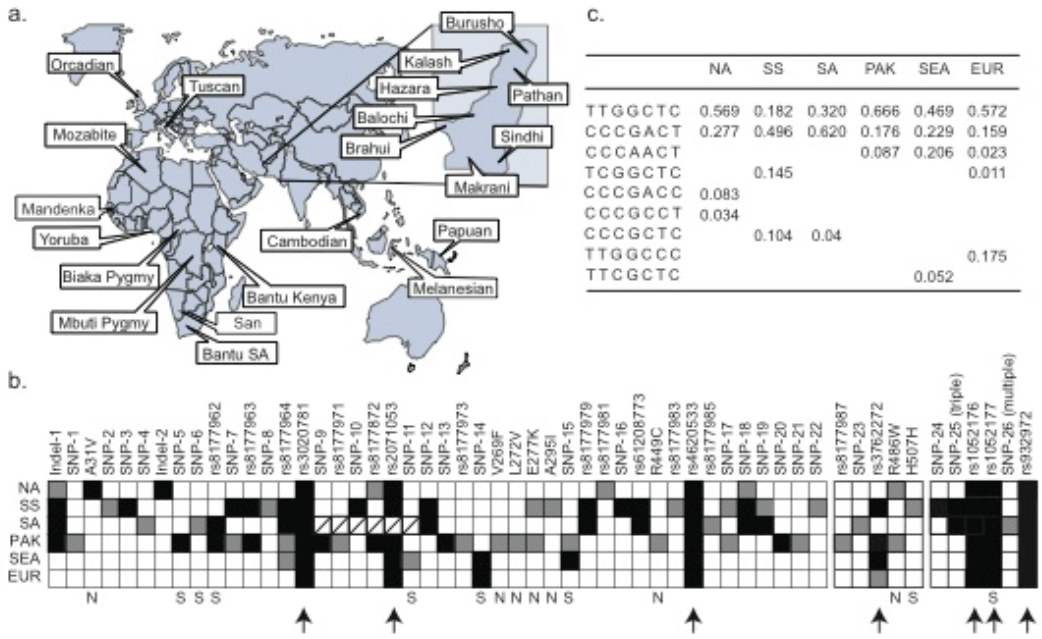
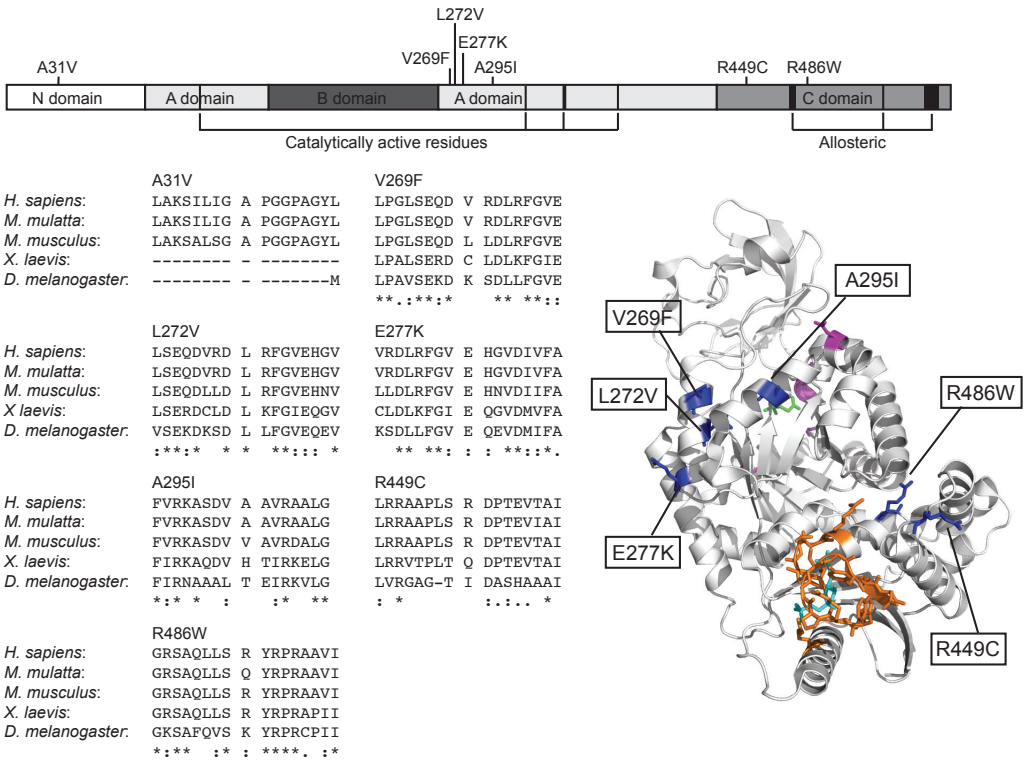



**Figure 4-1 Geographical origin of human DNA samples, along with common haplotypes and SNPs detected in *PKLR* for each population**

(a) 387 unrelated DNA samples were selected from the H952 subset[302] of the Human Genetic Diversity Panel provided by the Centre d'Etude du Polymorphisme Humain[303]. Ethnic groups were pooled into geographically based populations: "North African" (Mozabite n=29), "Sub-Saharan" (Biaka Pygmy n=23, Mbuti Pygmy n=13, Yoruba n=22, Mandenka n=22), "South African" (San n=6, Bantu n=8, Kenyan Bantu n=11), "Pakistan" (*inset*) (Brahui n=25, Balochi n=24, Hazara n=22, Makrani n=25, Sindhi n=24, Pathan n=24, Kalash n=23, Burusho n=25), "Southeast Asian" (Cambodian n=10, Melanesian, n=11, Papuan n=16) and European (Tuscan n=8, Orcadian n=15). All samples were originally collected following proper informed consent. (b) Coding regions of PKLR were PCR amplified and sequenced (see Supplementary Table 1 for primers). Polymorphic sites are identified as columns, listed from 5' to 3'. Each SNP is scored within the given population as having only the reference allele detected (white), alternative allele detected as a single heterozygous individual (gray), or alternative allele detected in more than one individual (black). Slashes indicate >75% missing data at this locus. Letters below the grid indicate synonymous (S) or non-synonymous (N) substitutions in the amino acid sequence of the protein. (c) Haplotype frequency distributions were determined in Haploview[304] according to Gabriel's method[305], selecting SNPs with a MAF of at least 10% that defined whole gene linkage disequilibrium (arrows in panel b).

The seven non-synonymous *PKLR* coding mutations suggest an overall mutant allele frequency of 1.5±0.9% (Figure 4-2). All mutations were found as heterozygotes, and no one had more than one mutation present. Pakistan had the highest mutation diversity with V269F (Hazara), L272V (Balochi), R449C (Makrani), E277K (Brahui) and R486W (Burusho) appearing as singletons. When the data were phased using BEAGLE software[306], all of the PAK SNPs except E277K appeared on haplotype 1, which was the most frequently observed haplotype in this population. R486W in NA appeared on the same haplotype. E277K appeared in a Brahui (PAK) individual on haplotype 7 and on a rare haplotype (5'-CCGGACT-3', <2%) in a Mandenka individual (SS), having likely arisen separately. The A295I mutation in a Biaka Pygmy appeared on the most common haplotype in SS (haplotype 2). All of these mutations are too rare to have influenced the underlying haplotype distribution when we examined the extended haplotype homozygosity using Sweep[307] (not shown). In the NA dataset, A31V appeared in four individuals ($MAF_{NA}$=6.9%), three times on haplotype 2 (27% frequency) and a fourth on haplotype 5 that differs by a single T/C polymorphism at the distal end of the haplotype block. This polymorphism may have appeared recently on a common core haplotype within the Mozabite lineage and been retained at a moderate frequency either by positive selection or by relatedness within the population (Supplementary Figure 1). Although the

mutated individuals are not first- or second-degree relatives[302], CEPH panel Mozabites have a high inbreeding coefficient[308] and historically small effective population size[309]. Recent common ancestry or ascertainment bias inherent to the small sample size may be confounding variables.




## Figure 4-2 Peptide position and conservation of non-synonymous substitutions in *PKLR*

a) Schematic representation of the PKLR protein, known structural and functional domains and location of non-synonymous amino acid substitutions. (b) Subset of multiple species protein alignments indicating conservation of mutated residues. “*” indicates 100% identity across all 13 species, “:” strong group conserved and “.” weak group conserved according to ClustalW. (c) Three-dimensional modeling of the non-synonymous substitutions in the PKLR protein based on the 2VG crystal structure[299].

To examine more closely the potential functional consequences of the mutations, we compared the PKLR protein across 13 species (*Homo sapiens* NP_000289.1, *Pan troglodytes* XP_524896.2, *Macaca mulatta*, XP_001112902.2, *Canis lupus familiaris* XP_547547.2, *Bos taurus* NP_001069644.1, *Mus musculus* NP_038659.2, *Rattus norvegicus* NP_036756.3, *Oryctolagus cuniculus* XP_002715467.1, *Monodelphis domestica* XP_001374169.1, *Gallus gallus* NP_990800.1, *Xenopus laevis* NP_001083514.1, *Danio rerio* NP_958446.1, *Drosophila melanogaster* AAF55979.3) using ClustalW. Overall, PKLR is well conserved with a minimum of 60% identity across the species examined. To quantify conservation, we used SIFT which assesses conservation and the physical properties of amino acids to assign scores where values <0.05 are predicted to affect protein function. A31V (median sequence conservation (MSC)=4.32; score=0.00), L272V (MSC=3.84; 0.00), E277K (MSC=3.84; 0.02), R449C (MSC=3.84; 0.03) and R486W (MSC=3.84; 0.00) were predicted to substantially affect function, while V269F (MSC=3.84, 0.21) and A295I (MSC=3.84, 0.57) may be tolerated.

Further analyses support that these substitutions are likely to reduce PK enzymatic activity. Notably, L272V, E277K and R486W have been reported in PK-deficient patients[299, 300] with R486W being the most common mutation associated with PK deficiency[310]. While A295I has not been previously

associated with PK deficiency, a similar mutation (A295V) does cause clinical symptoms[311]. Additionally, V269F, L272V, E277K and A295I are clustered within a 25 residue segment of the A domain of the enzyme, which contains active site residues, while R449C and R486W map near the allosteric regulatory site (C domain).

The large number and variety of PKLR mutations reported to cause PK deficiency suggest that the protein is highly sensitive to even small changes. Modest changes in protein folding/stability could have serious consequences for long-term erythrocyte function as mature cells lack the means to compensate by generating more protein.

A31V merits further mention as it appeared at high frequency (6.9%) in NA. This mutation lies in the erythrocyte-specific N-domain, affecting a residue highly conserved in mammals. This mutation is either fairly neutral, offers some advantage, or is in linkage disequilibrium with a polymorphism increasing fitness. Distinguishing these possibilities will require characterizing A31V enzymatic activity and testing its association with response to malaria in case-control studies.

Should all coding variants be loss-of-function alleles, the expected overall rate of PK deficiency would be ~1/4203 births. Stratified by geography, NA had the highest rate with an 8.6% mutant allele frequency, followed by PAK (1.3%)

and SS (1.25%). No mutants were detected in our SA, SEA or EUR datasets. Although our report is the first to determine mutant allele frequency by complete sequencing, other studies have estimated mutant allele frequencies between 0.4-6% by genotyping for common variants, with the highest mutation prevalence in Saudi Arabia (6%) and Hong Kong (3%)[310, 312-314].

We applied neutrality tests in DnaSP 5.10[315] to determine whether selection has altered the expected pattern of genetic variability in PKLR[316]. Neutrality tests are based on the principle that the mutation rate is constant and synonymous substitutions will behave in a neutral fashion, whereas non-synonymous changes will be either highly beneficial or deleterious to fitness, and thus targets of evolutionary selection[317]. As up to two thirds of statistically significant non-neutrality observations can be influenced by demographic history[316] we also generated 1000 coalescent simulations using the CoSi package (http://www.broad.mit.edu/~sfs/cosi) and Best-fit model described by Schaffner et al[318] to determine empirical p values (emp-p). This model takes into account migration history, bottleneck, agriculture history, recombination rate and mutation rate, which were optimized to fit well with empirical data of each population. Simulations were run on a sequence length of 1725bp corresponding to the coding region of PKLR, and population sizes were matched with our experimental data. In PAK, Tajima's test suggested non-neutrality, specifically

when examining the distribution of non-synonymous SNPs (D=-1.37292, emp-p=0.156, $D_{NS}$=-1.78894, emp-p=0.035). Fu and Li's D* and F* also rejected neutrality for SS (D*=-2.89519, F*=-2.46698, emp-p=0.001) and PAK (D*=-4.55934, F*=-4.07077, emp-p=0.001). Both populations were significantly non-neutral when compared to chimpanzee using the MacDonald-Kreitman test ($p<0.02$).

Together, these observations are in agreement with the proposal that the *PKLR* gene has been under selective pressure in some of the populations examined, possibly mediated by the malarial parasite[319, 320].

## 4.4 Acknowledgements:

This study was funded by a Canadian Institutes of Health Research (CIHR) Team Grant in Malaria (KCK, PG), CIHR MOP-13721 (KCK), Genome Canada through the Ontario Genomics Institute (KCK), CIHR Canada Research Chairs (KCK) and CIHR studentships (JB, SH).

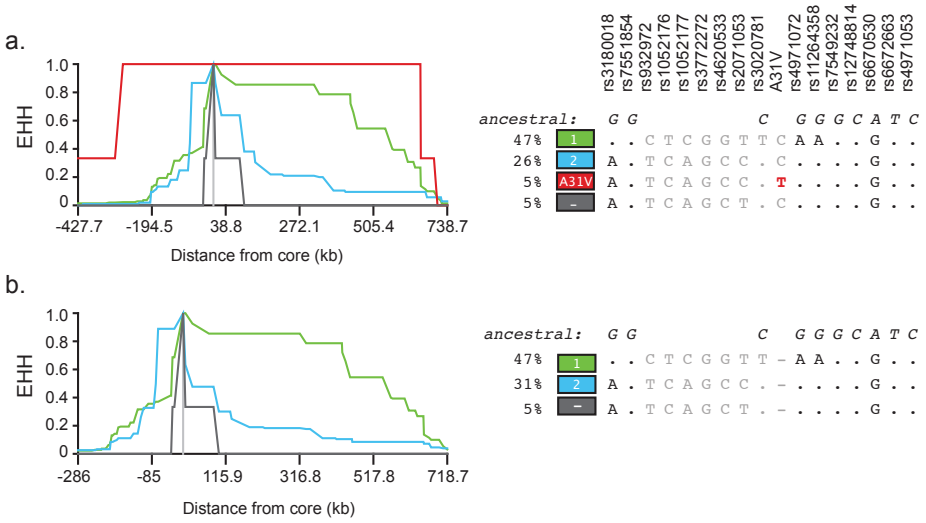
## 4.5 Supplementary Material

Figure S1 compares extended haplotype homozygosity within North African Mozabites, including those who carry the A31V mutation.

Table S1 lists the primers used in sequencing PKLR.

Table S2 lists each SNP that was identified along with its chromosomal position, minor allele frequency across the total dataset along with other identifying information and total population Hardy-Weinberg p-values.

Table S3 describes the population-specific allele diversity and frequency.




## Figure 4-3 Supplementary Figure 1. Extended haplotype homozygosity for *PKLR* in North African Mozabites

Extended haplotypes surrounding *PKLR* in the North African population. Genotype data extending from 151.9 kb to 154.3 kb on chromosome 1 (hg16) was downloaded from the CEPH website (ftp://ftp.ceph.fr) and combined with sequence data for the seven haplotype defining SNPs used in previous analyses, either (a) with A31V included to tag mutant chromosomes (red) or (b) without A31V. Data were phased in Beagle and analyzed using Sweep 1.1 software. A core haplotype was defined which included the entire *PKLR* gene (sequenced SNPs indicated with gray letters) and spanned 17 SNPs from *SCAMP3* through *ASH1L*. Mutant chromosomes had significant extended haplotype homozygosity in this region, likely indicating recent common ancestor.

## Table 4-1 Supplementary Table 1. Primers used for *PKLR* sequencing

| Fragment | Forward | Reverse |
|---|---|---|
| Exon1 | 5'- gaa aat ttc cct ggg ggt ag -3' | 5'- gcc tat agg gcc tgg aaa ag -3' |
| Exon2 | 5'- act ggg tga ttc tgg gtc tg -3' | 5'- ccc aag ggt agg gat ttt tg -3' |
| Exon3/4/5 | 5'- gtg agg cgt tct gag aat ggt -3' | 5'- gaa ggt gtg atc ggt ctg agg -3' |
| Exon6/7 | 5'- cat ggt tcc tga cct ctt gc -3' | 5'- gga tgt cac acc cat aag tg -3' |
| Exon8/9 | 5'- ttt cag ggg ttg tga ctg tga -3' | 5'- tcc tgt taa tcc tgc caa cc -3' |
| Exon10 | 5'- aag cac ata atg gac ttt ctc aaa -3' | 5'- tgg cag gga agg tct agg ta -3' |
| Exon11A | 5'- gcc ttg atg tgg tga aag gt -3' | 5'- ggt cgg ttt gac cct aat tt -3' |
| Exon11A_resequ | 5'- gca cga gat gcc cag att at -3' | 5'- gcc tgg tcc ctc aat gac ta -3' |
| Exon11B | 5'- cca aag ccc cac ctt taa gt -3' | 5'- agc cac atg aga gag ggaga -3' |

## Table 4-2 Supplementary Table 1: Minor allele frequencies of all polymorphic sites identified in *PKLR*, along with genetic location, minor allele frequency and other identifiers.

SNP locations are according to UCSC Feb 2009 build (GRCh37/hg19).

| SNP ID | Location | MAF | HWE p-value | Function | Flanking Sequence | % Genotype |
|---|---|---|---|---|---|---|
| Indel-1 | 155271689 | 0.014 | ns | 5' UTR | CATCTTGG[+]TTCAAAC | 96.0 |
| SNP-1 | 155271550 | 0.001 | ns | 5' UTR | GTGATCCC[T/C]TTTGTCTT | 94.1 |
| A31V | 155271095 | 0.006 | ns | **A31V** | GATTCCCC[C/T]TCCAGGAG | 95.0 |
| SNP-2 | 155270398 | 0.001 | ns | intron | CTGGGTTG[T/C]GCGGGGT | 92.5 |
| SNP-3 | 155270396 | 0.003 | ns | intron | GGGTTGTG[C/T]GGGGTCA | 94.1 |
| SNP-4 | 155270272 | 0.001 | ns | intron | TTTCTCAG[C/T]CTCTGTTT | 95.4 |
| Indel-2 | 155270190 | 0.003 | ns | intron | ATAATAAA[A/-]ATAATATA | 97.0 |
| SNP-5 | 155270064 | 0.006 | ns | **A36A** | GGGCCAGC[G/A]GGGGTAT | 95.2 |
| SNP-6 | 155270001 | 0.003 | ns | **Q57Q** | TTCTTCCA[G/A]CAGCAGCA | 95.2 |
| rs8177962 | 155269991 | 0.007 | ns | **L61L** | AGCAGCAG[C/T]TGCCAGCT | 95.2 |
| SNP-7 | 155269865 | 0.007 | ns | intron | CCCCTGCA[G/A]CCACACAG | 95.2 |
| rs8177963 | 155269831 | 0.006 | ns | intron | TGAGGTGC[T/A]TCTTCATC | 95.2 |
| SNP-8 | 155269825 | 0.001 | ns | intron | GCTTCTTC[A/G]TCTTTTGT | 95.2 |
| rs8177964 | 155269780 | 0.035 | 0.0125 | intron | CAGTCATT[C/T]ATTTAACA | 95.2 |
| rs3020781 | 155269776 | 0.417 | ns | intron | TCATTCATT[T/C]AACAAAAA | 95.2 |
| SNP-9 | 155265576 | 0.005 | ns | intron | GGGGGAAC[G/A]TTGTCTGA | 95.0 |
| rs8177971 | 155265446 | 0.002 | ns | intron | TGCGGGAC[G/T]GGCCGCC | 96.6 |
| SNP-10 | 155265214 | 0.002 | ns | intron | CAGTGGGG[C/A]TGGGACT | 96.6 |
| rs8177872 | 155265192 | 0.010 | ns | intron | GCCAGGGC[C/T]GGAAAGG | 96.6 |
| rs2071053 | 155265177 | 0.384* | $5.0 \times 10^{-4}$ | intron | GCGGCGCC[T/C]GTAGGGT | 96.6 |
| SNP-11 | 155265021 | 0.002 | ns | **R194R** | TCCGGACG[C/A]GGGGGAA | 96.3 |
| SNP-12 | 155264791 | 0.007 | 0.0279 | intron | GTTTTTGC[C/G]AGCCCGTC | 93.5 |
| SNP-13 | 155264686 | 0.011 | ns | intron | TGCCTCTT[C/T]CGGGCGAA | 95.3 |
| rs8177973 | 155264623 | 0.001 | ns | intron | CTTGCGGG[T/C]CCAAGCCC | 95.6 |
| SNP-14 | 155264530 | 0.005 | 0.0164 | **L236L** | GAGGGACT[G/C]GTGACCC | 95.6 |
| V269F | 155264433 | 0.001 | ns | **V269F** | AGCAGGAC[C/A]TCCGAGAC | 95.0 |
| L272V | 155264424 | 0.001 | ns | **L272V** | TCCGAGAC[G/T]TGCGCTTC | 95.0 |
| E277K | 155264409 | 0.003 | ns | **E277K** | TCGGGGTG[C/G]AGCATGG | 95.0 |
| A295I | 155264354/5 | 0.001 | ns | **A295I** | GCGACGTG[GC/AT]TGCCGT | 95.0 |
| SNP-15 | 155264330 | 0.001 | ns | **L301L** | GCTGCTCT[G/A]GGTCCGGA | 95.6 |
| rs8177979 | 155263430 | 0.015 | ns | intron | CAGGGTCC[C/T]CAGTCACA | 86.9 |
| rs8177981 | 155263384 | 0.001 | ns | **splice site** | ctgtcccc[c/t]agATGCTG | 89.6 |
| SNP-16 | 155263174 | 0.003 | ns | intron | TTCCATAC[C/T]CCAGTGCC | 95.2 |
| rs61208773 | 155263166 | 0.007 | ns | intron | CCCAGTGC[C/A]CCTTCCCA | 95.2 |
| R449C | 155263059 | 0.001 | ns | **R449C** | CACTAAGC[C/T]GTGATCCC | 90.1 |
| rs8177983 | 155262639 | 0.002 | ns | intron | TCTTTACT[C/-]AGAAATGT | 87.0 |
| rs4620533 | 155262613 | 0.389 | 0.0086 | intron | AGCCTGCT[G/C]TCTTTCCT | 85.8 |
| rs8177985 | 155262414 | 0.001 | ns | intron | GGCAGCTC[G/A]CTCTGTTG | 95.7 |

| | | | | | | |
|---|---|---|---|---|---|---|
| SNP-17 | 155262370 | 0.004 | ns | intron | TCTCAGCT[C/T]ACCACAAC | 95.5 |
| SNP-18 | 155262334 | 0.027 | 0.0054 | intron | GATTCTCC[T/C]GCCTCAGC | 95.7 |
| SNP-19 | 155262313 | 0.010 | ns | intron | CGAGTAGC[T/C]GGGATTAC | 95.7 |
| SNP-20 | 155262277 | 0.003 | ns | intron | CTGCATAT[T/G]TTTTGTATT | 99.0 |
| SNP-21 | 155262209 | 0.001 | ns | intron | TGACCTCA[G/C]GTGATCCA | 95.7 |
| SNP-22 | 155262087 | 0.001 | ns | intron | ATCTGTAT[T/C]ATCTCATA | 95.7 |
| polyT region -- could not score | | | | | | |
| rs8177987 | 155262014 | 0.001 | ns | intron | ATAGTGAG[A/T]CCCCATCT | 93.3 |
| SNP-23 | 155262002 | 0.001 | ns | intron | ACTATGTT[G/T]CCTAGGCT | 96.7 |
| rs3762272 | 155261777 | 0.078 | $5.03 \times 10^{-9}$ | intron | CTTTCCTC[G/A]TTCACCAC | 96.8 |
| R486W | 155261709 | 0.003 | ns | **R486W** | TTCTGTCT[C/T]GGTACCGA | 95.0 |
| H507H | 155261644 | 0.001 | ns | **H507H** | | 95.0 |
| repetitive region -- could not score | | | | | | |
| SNP-24 | 155260668 | 0.009 | ns | intron | CTCCCTAG[T/C]AGCTGGGA | |
| SNP-25 (triple) | 155260595-625 | 0.035 | ns | intron | TTTT[G/A]TATTTT[C/T]AGT…CCA[T/C]GTTG | 68.3 |
| rs1052176 | 155260383 | 0.444 | ns | **R569R** | ACATCATG[C/A]GGGTGCTA | 68.8 |
| rs1052177 | 155260350 | 0.444 | ns | 3' UTR | CCCCTCCC[T/C]CCTCTGGC | 68.8 |
| rs932972 | 155260096 | 0.381 | ns | 3' UTR | GTACATTC[C/T]CTGCATCC | 77.6 |
| SNP-26 (multiple) | 15526049 - 66 | 0.002 | ns | 3' UTR | CTCC[A/C]T[T/C]TCA[-/T]CAC[AG-/TTT]G[T/G]CCT[-/A]GAAA | 79.9 |
| rs8847 | 155259323 | 0.235 | ns | 3' UTR | AGGTAAAC[G/T]CGGTCCTT | 90.2 |
| rs41264939 | 155259470 | 0.002 | ns | 3' UTR | AGTGAGGG[T/G]GGTAAAG | 76.9 |

## Table 4-3. Supplementary 2. Population specific breakdown of genetic diversity statistics and polymorphic content

| | NorthAfrica | SubSaharan | SouthAfrica | Pakistan | Southeast Asia | Europe |
|---|---|---|---|---|---|---|
| #Individuals | 29 | 80 | 25 | 191 | 37 | 23 |
| #Polymorphic loci | 13 | 31 | 17 | 29 | 11 | 8 |
| #Singletons | 5 | 13 | 4 | 14 | 2 | 1 |
| Unique sites | 3 | 13 | 4 | 11 | 1 | 0 |
| Ave gene diversity | 0.020 +/- 0.016 | 0.030 +/- 0.021 | 0.041 +/- 0.027 | 0.035 +/- 0.022 | 0.033 +/- 0.031 | 0.021 +/- 0.016 |
| Fst | 0.08553 | 0.07967 | 0.766 | 0.08417 | 0.08808 | 0.08697 |
| | | | | | | |
| SNP ID | NAfrican MAF | SubSaharan MAF | SAfrican MAF | Pakistan MAF | SEA MAF | Europe MAF |
| Indel-1 | 0.018[a] | 0.019 | 0.040 | 0.008 | | |
| SNP-1 | | | | 0.003[ab] | | |
| A31V | 0.069[b] | | | | | |
| SNP-2 | | 0.007[ab] | | | | |
| SNP-3 | | 0.013[b] | | | | |
| SNP-4 | | | 0.021[ab] | | | |
| Indel-2 | 0.034[ab] | | | | | |
| SNP-5 | | | | 0.011 | | |
| SNP-6 | | | 0.021[ab] | | | |
| rs8177962 | | | 0.042 | 0.011 | | |
| SNP-7 | | 0.032 | | 0.003[a] | | |
| rs8177963 | | 0.019 | | 0.008 | | |
| SNP-8 | | 0.006[a] | | | | |
| rs8177964 | | 0.110 | 0.125 | 0.003[a] | 0.022[a] | |
| rs3020781 | 0.429 | 0.649 | 0.667 | 0.307 | 0.457 | 0.182 |
| SNP-9 | | | - | 0.008[b] | | |
| rs8177971 | | | - | 0.003[ab] | | |
| SNP-10 | | 0.013[b] | - | | | |
| rs8177872 | 0.019[a] | | - | 0.013 | | |
| rs2071053 | 0.423 | 0.833 | - | 0.305 | 0.389 | 0.227 |
| SNP-11 | | | - | | 0.014[ab] | |
| SNP-12 | | 0.021 | 0.042 | | | |
| SNP-13 | | | | 0.022[b] | | |
| rs8177973 | | | | 0.003[ab] | | |
| SNP-14 | | | | | 0.030 | 0.043 |
| V269F | | | | 0.003[ab] | | |
| L272V | | | | 0.003[ab] | | |
| E277K | | 0.006[a] | | 0.003[a] | | |
| A295I | | 0.007[ab] | | | | |
| SNP-15 | | | | 0.003[a] | 0.030 | |
| rs8177979 | | 0.080 | 0.045 | | | |
| rs8177981 | 0.017[ab] | | | | | |
| SNP-16 | | 0.014[b] | | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| **rs61208773** | | 0.014 | 0.068 | | | |
| **R449C** | | | | 0.003[ab] | | |
| **rs8177983** | | 0.009[ab] | | | | |
| **rs4620533** | 0.462 | 0.623 | 0.625 | 0.297 | 0.500 | 0.227 |
| **rs8177985** | | | 0.020[ab] | | | |
| **SNP-17** | 0.017 | 0.008[a] | | 0.003[a] | | |
| **SNP-18** | | 0.111 | 0.060 | | | |
| **SNP-19** | | 0.007[a] | 0.120 | | | |
| **SNP-20** | | | | 0.006[b] | | |
| **SNP-21** | | | | 0.003[ab] | | |
| **SNP-22** | | 0.007[ab] | | | | |
| **polyT region** | | | | | | |
| **rs8177987** | | | | 0.004[ab] | | |
| **SNP-23** | | | 0.025[ab] | | | |
| **rs3762272** | | 0.022 | | 0.116[c] | 0.196 | 0.022[a] |
| **R486W** | 0.017 | | | 0.003[a] | | |
| **H507H** | | 0.007[ab] | | | | |
| **repetitive region** | | | | | | |
| **SNP-24** | | 0.025[b] | | | | |
| **SNP-25 (triple)** | | 0.110 | 0.042 | | | |
| **rs1052177** | 0.444 | 0.500 | 0.646 | 0.315 | 0.469 | 0.409 |
| **rs1052176** | 0.389 | 0.500 | 0.646 | 0.349 | 0.469 | 0.227 |
| **SNP-26 (multiple)** | | | | 0.004[a] | | |
| **rs932972** | 0.382 | 0.493 | 0.646 | 0.287 | 0.469 | 0.227 |
| **rs8847** | - | | - | 0.235 | - | - |

Shading indicates linkage disequilibrium

[a]single heterozygote observation (singleton) in this population

[b]polymorphism unique to this population

[c]Hardy-Weinberg equilibrium, p=0.0004

# Chapter 5 Summary and Future Perspectives

## 5.1 Summary

It is indisputable that host genetic factors play a major role in determining resistance to severe malarial disease. It is also unquestionable that multiple genes and pathways, many of which remain unknown, can influence resistance to severe and cerebral pathologies that are a consequence of infection. Uncovering what these pathways are can suggest new avenues for disease management and treatment, which is especially necessary for cerebral malaria. At present, prognosis is poor for patients exhibiting cerebral symptoms due, in part, to the lack of specialized treatments beyond anti-parasitic drugs. With no clinical measures to combat cellular sequestration, inflammation or other pathological features, cerebral-malaria associated cascades can lead to lethality or cause persistent damage[321]. Being able to address the specific causes of some of the more damaging sequelae would be a boon to clinicians and a potential lifesaver to thousands or millions of people in endemic areas. In addition, better understanding of the key players in the host:parasite interactions can be used to design entirely new classes of drugs as a way to combat increasing drug resistance and/or may indicate promising vaccine targets. Further to this, we have also generated data that suggest discoveries made in the context of mouse models of cerebral malaria may be applicable to other inflammatory conditions as well.

Animal models are a valuable complement to human studies in malaria and have identified ten loci contributing to parasitemia burden (and/or survival) in response to *P. chabaudi* (*Char1-11*)[120, 180] and ten loci contributing to the onset of cerebral disease in response to *P. berghei* (*Berr1-8, cmsc,* unnamed locus on Chr19)[192, 193], which includes three new genetic loci described since

the work presented here. While candidate gene identification from these loci has been difficult, mutation identification and follow-up on the *Char4*/*Pklr*/*PKLR* locus proved that if the variant gene or pathway can be mapped, translation across species from mice into humans is possible. This translatability goes both ways as mouse models with targeted knockout alleles have also been shown to validate the involvement of multiple loci and pathways initially uncovered by human studies. Commonalities between mice and humans allow further exploration into shared responses and resistance mechanisms.

In Chapter 2 we identified the *Berr5* region on mouse chromosome 19 that strongly influences cerebral malaria susceptibility. In an F2 population of mice with mixed genetic background composed of C57BL/6 and BALB/c alleles, over 90% of the mice carrying the susceptible C57BL/6 haplotype at *Berr5* developed neurological symptoms compared to only 61% of mice harboring the resistant BALB/c derived haplotype[191]. Interestingly, this locus also overlapped with three other QTLs that had been mapped in the context of inflammation (*Eae19*, experimental autoimmune encephalitis, *Tsiq2*, IL-4 secretion of T-cells) or response to infection (*Trl4*, tuberculosis susceptibility). This suggests that one or several adjacent genes within this region may serve as important regulators of immune balance following a variety of insults. Differential transcript expression analysis in susceptible and resistant mouse brains implicated the *Ifit1-3* gene cluster or *Fas* as promising positional candidate genes for this effect. This was one of the first unbiased dissections of the genetic basis underlying differential cerebral malaria susceptibility in (non-wild derived) inbred strains, and the first to examine C57BL/6 versus BALB/c, though these two are often used as “susceptible” and “resistant” inbred models. One of the strengths of QTL mapping is the unbiased nature of gene discovery and it’s ability to suggest causative drivers of pathology without blurring into downstream hallmarks of

terminal processes already in progress. Finding these tipping points and critical nodes opens up the possibility of blocking upstream initiator events and preventing serious symptoms from ever occurring. Localizing *Berr5* provides potential positional candidates for these investigations.

Advances in genetic research, resources and technologies allow us to re-examine the *Berr5* interval to improve the candidate gene prediction or propose additional strategies for cloning the locus successfully. *Ifit1*$^{-/-}$ mice were obtained and infected with *P. berghei* as part of the work presented in Chapter 3, where they were found to be as susceptible as B6 (Figure 3-5)[322]. This argues against variation within the *Ifit1* gene, and potentially other members of the Ifit cluster, being the root cause of differential susceptibility. *Fas* remains a good candidate, Cd274 – with it's role in CD8+ T-cell proliferation and immune modulation[323] – may be worth further consideration, and hypotheses can be built off of several other genes within the span.

To differentiate between myriad possibilities and provide evidence for one requires that several steps be taken. First, some form of sequence variation between the two strains must exist and be identified. Second, that variant (or some surrogate, such as a gene knockout) must be isolated and still able to recapitulate a portion of the observed variance in the phenotype. Alternatively, or thirdly, some biological mechanism based on known gene functions of the positional candidates can be hypothesized and tested. The first goal was made far easier in 2012 when a group from the Wellcome Trust sequenced and publicly hosted the complete genome for 17 inbred mouse strains including B6 (the existing reference genome) and BALB/cByJ[324]. Even more recently, the Mouse ENCODE Consortium has published a functional catalog of elements genome-wide, including transcription factor binding sites, open and closed chromatin and sites of epigenetic modification in 100 distinct tissues[325].

Combining these bioinformatic resources allows large-scale identification of potentially functional variation without laborious Sanger sequencing of candidate regions. Variant isolation and testing can also now be improved over the long and expensive process of congenic breeding (5-10 generations, maintenance of multiple sub-lines) or embryonic stem cell targeting and screening to generate a knockout or specific knock-in allele. Now, endonuclease-mediated gene editing (Zinc finger nucleases, TALENs and CRISPR/Cas9)[326] can be used to generate mice carrying an interesting, purportedly functional variant or even set of variants within a single generation. The CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats with CRISPR-associated protein 9) system is particularly revolutionary for its versatility, ease of use and speed[327]. Prioritized variants, individually or as a complete modified transcript, from the BALB/c genome can be introduced specifically and directly into the genome of B6 mice, who can then be phenotyped to see if the incidence of pathology is reduced.

The overlap of *Berr5* with three other infection and inflammatory QTLs pointed again towards what others have observed that the active host inflammatory response is a major driver in experimental cerebral malaria. *Berr5* mirrored what was seen in *Tsiq2* with C57BL/6 alleles across Chr19 leading to a more pro-inflammatory profile than BALB/c and indicating a role for this region in T-cell activation specifically. In addition, it drew pathway, and possibly mechanistic, parallels between cerebral malaria pathogenesis and other diseases including *M. tuberculosis* and experimental autoimmune encephalitis (a model for multiple sclerosis). The overlap between the genetics involved in mouse models of cerebral malaria and the genetic loci implicated in multiple sclerosis was further expanded in the next chapter.

Further linking experimental cerebral malaria and inflammation, we explored a pathogenic role for the transcription factor *Irf8* in Chapter 3. We determined that IRF8 is critical for the development of cerebral malaria, and that IRF8 mediated pathways are central within the characteristic lethal neuroinflammatory response. Cellular profiling indicated that both myeloid cellular development and activation play significant roles as while the parental strains do differ, the resistance observed in heterozygous F1 mice was not tied to a strong pre-existing cell composition phenotype. The major activation pathways highlighted by examination of differential transcript expression were interferon associated responses (both Type 1 and Type 2), innate immunity (including cytokine signaling and cell recruitment), Fc-receptors, and antigen presentation. Many of the same pathways had previously been noted as upregulated during tuberculosis infections, and comparing the expression profiles of mouse brains during *P. berghei*-induced cerebral malaria with mouse lungs during *M. tuberculosis* response, there was a highly significant overlap. We further identified a network of 53 genes that: (1) possessed IRF8 binding sites and (2) were upregulated at least two-fold in both disease models which may represent a common core for these two pathogen responses.

We were also able to draw parallels to other human diseases of pathological inflammation including multiple sclerosis, systemic lupus erythematosus and inflammatory bowel disease which have all implicated SNPs at *Irf8* in pathogenesis at the level of genome wide significance[273]. In each of these conditions, immune system processes become dysregulated, leading to a chronic inflammatory state. Understanding the activation and amplification of IRF8 and its downstream responses may provide relevant information for some of these diseases too.

Chapter 4 returned the focus to malaria as we explored population level variation in the human *PKLR* gene (responsible for the *Char4* QTL effect in mice). Previous work by us and others has shown that (1) PKLR mutations are protective against blood stage malaria in humans[104] and animal models[102] and (2) PKLR deficiency is a very common disorder[310]. Whether the two findings are causally linked or merely coincident remains to be proven, however as a first step we sequenced the PKLR coding region in DNA from 387 humans across 21 global ethnic populations including those under current or historical malarial pressure. What we found was rich genetic diversity with pronounced haplotype structuring by geography and a 1.5% mutant allele frequency (seven separate mutations in twelve independent individuals). Putatively loss-of-function mutations were most commonly observed in Pakistani and African (North African and sub-Saharan African) populations. This study lays the groundwork required in order to pursue a larger scale association study with case and control cohorts in African populations. Analysis of SNPs with high local frequencies allow the identification of haplotype-defining “tag” SNPs, and elucidation of locally segregating non-synonymous alleles is also extremely valuable for designing a genotyping, as opposed to a sequencing, study.

Together, these three stories demonstrate some of the ways in which mouse models of infection can generate new insights into the genetic basis of malaria resistance or translate to other immunopathological conditions. While the limitations of mouse models should be kept in the back of one’s mind, many biological response pathways are well conserved and observations made in mice can be used to inform human studies as well as vice versa. Identifying and exploiting malarial protective effects due to host driven mechanisms can circumvent increasing drug resistance seen in *Plasmodium* parasites and focus

on preventing or reversing the most severe pathologies, which are dependent on specific interactions between the host and parasite.

## 5.2 Future Directions

### 5.2.1 Ongoing malaria-influencing gene discovery in mice

As was touched upon in the introduction, one of the drawbacks of using mice to uncover new genetic hypotheses is the limited genetic diversity within established laboratory strains of mice. Classical inbred strains exhibit a substantial degree of identity by descent as a result of limited initial founders and breeding history[328]. While the reproducibility of an inbred strain is a significant advantage, this lack of variation does not fully cover the range that can be observed in humans or other natural populations. For a trait such as malaria resistance with a very complex etiology and progression, focusing exclusively on this handful of genetically identical (within-strain), homozygous lines may inadvertently over-emphasize or neglect certain important aspects of the host response. While they are still extremely valuable for dissecting the complex responses or validating targets, using targeted knockout lines as may also be akin to attacking the genome with a hammer and causing multiple pleiotropic phenotypes, potentially obscuring the roles for genes that affect viability and basic functions. These also require *a priori* hypotheses, so truly novel genes are unlikely to be uncovered without a systematic large-scale effort.

To address these issues and introduce new variability into mouse models available for study can be accomplished by either using a random mutagenesis screen on an existing inbred line, or by exploring novel populations and allele combinations: perhaps with wild derived or outbred mice. *Berr1* and *Berr2* were,

in fact, identified by Campino, Bagot and colleagues[187, 188, 195] using the wild-derived WLA strain, which is resistant to experimental cerebral malaria following infection with *P. berghei*. In this series of papers, the authors demonstrated that the added wild variation was sufficient to introduce enough novelty that two QTLs could be identified, and an unprecedented phenotype of parasite clearance and increased immunity to subsequent experimental re-infection was observed in a subset of F2 progeny (further mapped and explored as *Berr3*)[188].

The WLA strain originated from mice trapped in the Toulouse region of France[329]. Though they are genetically distinct from the founder stocks of the commonly used inbred lines (and associated “fancy mouse” breeding bottlenecks), they do still cluster with the common inbred strains[330] in genetic analysis due to both carrying predominantly European *M. musculus domesticus* ancestry, and both having presumably been exposed to (or exempt from) certain similar evolutionary pressures throughout their population history. It would be quite interesting to perform a similar study using a wild derived population originating from Thailand (such as the CAST/Ei, CASA/RkJ, or CAROLI strains)[331], or by examining rodents native to central Africa where *P. berghei* and other *Plasmodium spp*. are natively distributed, allowing for the disease to progress in a host to which it is has naturally adapted (and can also produce a fatal infection)[332, 333], and allowing the host response to employ whatever defense mechanisms may have been evolutionarily honed by infectious pressure. The risk of this, of course, is that moving away from well-established laboratory resources into wild populations means that many of the features and analysis tools can no longer be taken advantage of (reproducibility, manageable heterogeneity, sequenced and annotated genome, validated markers, etc). Prioritization and identification of important variation may also be intractable in

these populations as the sheer number of changes – both sequence and structural variants – can be overwhelming.

An alternative strategy then, is to introduce manageable novel variation into laboratory mice, which can be done with chemical mutagens, radiation, or insertional mutagenesis. Up until the last five years, while any of these mutagenesis strategies might successfully be able to introduce variants affecting mouse susceptibility to a given phenotype, the road to variant discovery and characterization once an interesting pedigree was identified was both time and resource prohibitive. Researchers would be obligated to set up informative, multigenerational crosses for meiotic mapping, then, if a positional interval could be determined, narrow the interval as in a dihybrid cross between two inbred strains. As evinced by the as yet unsuccessful search for a causative variant(s) underlying the *Berr5* QTL in Chapter 2, this can be an arduous task with multiple positional candidates capable of influencing such a complex trait in some way, along with a wealth of accessory variation that needs to be filtered away. Searching for *de novo* mutations is more straightforward as variants present in either parent can be excluded, but the use of two parental strains required for informative mapping markers can also introduce modifiers which combined with the time required makes this unsuitable as a large scale screen. In fact, *Berr7* and *Berr8* (formerly *Berr6*) are an example of this occurring where cerebral malaria susceptible C57BL/6J mice were crossed to the susceptible 129S1/SvImJ strain[192]. Rather than merely providing informative markers, an epistatic two-locus interaction generated a subset of F2 progeny exhibiting resistance above the expected background levels. In the pedigree used to initially map this effect, resistance was observed at a rate of 20-30%, which was consistent across approximately ten other unrelated pedigrees. The observation of multiple pedigrees segregating resistance at matching genetic loci on

Chromosome 4 confirmed that the effect was due to the background interaction rather than a pedigree-specific spontaneous or induced mutation[192]. Positional candidates have been proposed for these QTL effects (as others), but no variants have been validated.

To avoid meiotic mapping and the challenges it brings requires technological advancement. In the last five years, whole exome sequencing technology has advanced to a point where it is rapid, accurate and economical enough to use for phenotypic mutation discovery without the need for linkage analysis. With ENU mutagenesis, individual gametes from treated male G1 founders are expected to carry ~3000 mutations, with a proportionate ~30-45 mutations appearing in coding regions[334, 335]. While coding regions represent only 1-1.5% of the human (or mouse) genome, functional impact and interpretation are more straightforward for missense, nonsense or splice site mutations, particularly in traits with Mendelian inheritance. At present, the overwhelming majority of ENU-induced phenotypes that have been molecularly characterized have been attributed to protein coding or splicing changes, with only rare reports of phenotypes being attributed to mutations in gene regulatory regions[336] or non-coding expressed sequences such as miRNA[337]. While sequencing exonic regions alone does miss these functional non-coding portions it is very effective at capturing the most impactive variants. These coding substitution variants are the basis for 85% of monogenic disorders where the genetic basis is known [338]. Additionally, even amongst complex traits with multiple common alleles contributing small biological effects at the molecular and phenotypic level, a strongly pathogenic or complete null allele within a key gene can exaggerate the gene-disease link or lead to an atypical pseudo-Mendelian disease state[339]. Plenty of such “low hanging fruit” remains to be discovered in experimental or clinical contexts for both rare and common disorders. Introducing

*de novo* mutations, using an appropriate phenotypic screen, and exome sequencing of either founders or specific progeny exhibiting a phenotypic difference is an effective way to mimic these conditions and add a wealth of new gene-phenotype and gene-disease connections.

### 5.2.2 Exome sequencing

Exome sequencing technology roared onto the stage in 2009 with a proof-of-concept paper by Ng and colleagues studying Freeman-Sheldon syndrome (FSS), a rare, dominant Mendelian disease with characteristic facial and joint abnormalities, known to be caused by mutations in myosin heavy chain 3 (*MYH3*)[338, 340]. Four unrelated FSS patient DNAs were used for sequencing, and compared with both eight control DNAs processed in parallel, and the set of variants deposited in dbSNP. They were able to achieve 50-fold coverage across 95% of the coding bases of 78% of all coding genes, with at least 8-fold coverage across 99.7% of targeted bases. After variant calling, they found that each individual carried an average of 17,272 coding single nucleotide polymorphisms (cSNP), with ~10,000 heterozygous segregating sites, ~6500 non-reference homozygous variants and 166 coding indels. *MYH3* was the only gene where (a) at least one non-synonymous or other pathogenic cSNP was observed in all four patient DNAs and (b) none of the observed cSNPs also appeared in dbSNP, HapMap or the control exomes[340]. Successfully identifying the known causal variant in such a small sample size with few steps is certainly exciting on it's own, but what this additionally demonstrated is the high coverage that exome sequencing technology makes possible, and how that translates into high confidence, with low false negative and low false positive rates in called variants. This proof-of-concept where researchers successfully

chased a known gene was rapidly followed by identification of the first unknown genes in rare recessive (*DHODH* in Miller syndrome)[341] and dominant (*SETBP1* in Schinzel-Giedion syndrome)[342] human diseases. The technology was soon adapted for mouse, allowing comprehensive analysis of the variation within inbred strains and mutant discovery in spontaneous or induced phenotypic alleles [334, 343].

Several different platforms have been developed for exome sequencing that rely on similar technical principles[338]. The first step requires non-specific PCR amplification of the DNA sample, ligating in common 5' and 3' linkers to each segment of sheared DNA in order to generate a library. Second, magnetic beads or some other pull-down substrate are added, with each carrying a complementary DNA sequence label specifically designed to target and capture the regions of interest, here, the exome. The design of the exome capture reagents is critical, as platform developers need to define their targets in an accurate and informative way, accounting for exons involved in alternative splicing and attempting to minimize the impact of common variation on efficiency[344]. Simply locating the exons remains a moving target with estimates for mouse ranging from 191,946[345] to 213,500[346] loci, with the more recent estimate on the lower end. From there, highly repetitive regions, variable regions between strains, or regions that remain poorly sequenced or assembled in the reference can all impact on successful exome capture.

Exon containing fragments that are captured undergo a second round of PCR amplification in order to enrich the exome from the original library. These are hybridized to an array and then undergo rounds of massively parallel sequencing to generate millions of short reads (100-300 bp). If multiple samples are hybridized simultaneously, DNA barcodes are then read in order to assign specific reads to the appropriate sample. These reads are aligned to the

reference genome, which, in mice, is derived from the C57BL/6 strain. Variants are called, and then appropriate bioinformatics filters are applied to sift the important variants from the background noise.

This technology is particularly useful for discovery of *de novo* and rare, high impact mutations, which makes it an extremely valuable accompaniment to mutagenesis screens, or mutant discovery in spontaneous heritable phenotypes. The use of inbred mice allows background variation to be filtered away and functional prediction algorithms to determine pathogenicity (conserved missense, nonsense, splice site disruption, etc), and segregation or pedigree analysis to confirm the variant-phenotype relationship. The combination of introduced mutations with exome sequencing has been exploited to uncover new gene-phenotype associations in hundreds of new pedigrees, including screens for hematological profile variation, organogenesis and development, and a variety of other observable characteristics [334, 335, 347].

### 5.2.3 Translation of malaria resistance loci from mice to humans

Identification of genes and pathways involved in the host response to infection is half the story. There are several steps required in order for gene discoveries made in mice to be considered relevant to human global health. First and foremost, the gene itself needs to be confidently associated with the observed resistance phenotype. A candidate gene can be confirmed by using targeted gene knock-out mice, allelic complementation (for recessive phenotypes where two potentially loss-of-function mutants are available), or by reproduction of the phenotype using pharmacological regulation of an enzyme or pathway (antibody knockdown of proteins or cells, small molecule activators or inhibitors).

Second, human homologs (or homologous pathways) need to be identified and a role for them must be established within human malarial disease. Sticking to a primarily genetic viewpoint – this can be done using an association or case control study in endemic populations. If SNPs within the human gene can be associated with variable response to infection (higher prevalence in cases versus controls or vice versa), or geographic distribution of a SNP mirrors malaria endemicity, it adds support to the hypothesis that the gene has evolved under malaria pressure and may be critically involved in the host response.

For PKLR, which is an erythrocyte protein, this was accomplished by obtaining erythrocytes from PKLR-deficient patients undergoing transfusion and (1) infecting them with *Plasmodium falciparum* in culture to monitor parasite growth and maturation rates as a marker for how readily an infection could progress within a PK-deficient patient and (2) comparing phagocytosis of *P. falciparum*-infected PK-deficient erythrocytes with infected normal erythrocytes[104]. In this way, it was found that parasites have difficulty invading human PK-deficient cells and clearance was markedly higher. Going forward, it would be interesting to take the lessons learned so far about the nature and prevalence of non-synonymous SNPs as well as African haplotypes to correlate genetics with malaria infection, progression and outcome in a genotyping or sequencing study in malaria-exposed human populations. Examining the extended haplotype structure surrounding PKLR-mutant individuals can also ascertain whether these alleles have been specifically retained as conferring a fitness advantage despite the drawbacks of complete deficiency[307]. Positive selection signatures can reveal data about our past as a species and how humanity has adapted to counter survival challenges in diverse ways throughout pre-history.

In the case of the transcription factor IRF8, the approach to validate the relevance of the mouse findings in humans would have to be a little bit different. Complete loss-of-function mutations in human IRF8 are rare and highly pathological[266], however immune cells from IRF8-deficient patients or transfected monocytes could be introduced into an *in vitro* assay to compare phagocytic ability, responsiveness to parasite antigens (by cytokine release) and migration behavior in response to chemoattractants.

Pure phagocytosis assays using *Irf8*$^{-/-}$ cells may be of scientific interest, but are unlikely to be particularly fruitful for modeling cerebral malaria as symptoms are not strictly correlated with parasite burden and circulating parasitemia is an unreliable predictive biomarker[348], suggesting that adhesion or other complex host-mediated factors are more central to pathology[349]. Additionally, both *P. chabaudi*[124] and *P. berghei* infections in *Irf8*-deficient mice both indicate increased permissiveness to blood stage replication, likely as a result of an inadequate immune response.

As IRF8 is a transcriptional regulator, a second strategy could be to examine transcriptional profiles of patients suffering from cerebral malaria for signatures of IRF8 involvement by searching for altered expression of multiple IRF8-target genes. A recent study characterizing whole blood transcriptomes of children with low and high *P. falciparum* parasitemia indicated several significant genome-by-infection interactions[118]. Analyses revealed a set of genes whose transcript abundance was regulated by the presence of *P. falciparum* infection, a cohort of genes whose expression was associated with differential parasite burden (high versus low) and those whose responsiveness to infection were genetically controlled (i.e. expressed in a SNP dependent fashion). None of the SNPs in *IRF8* reached genome-wide significance, however regulation of several genes in an IRF8 and infection associated manner was observed (Idaghdour,

personal communication) which could be very interesting to follow up on further. While this focused on blood stage infection including severe anemia, a similar approach comparing profiles of children with or without cerebral involvement could uncover cerebral malaria specific pathways which could be interrogated for IRF8 targets or other patterns.

If a role for the human ortholog can be established in human malaria infections, the second step is determining if the gene or pathway can be used clinically as either a diagnostic/predictive biomarker or therapeutic target. Given how little we know about cerebral malaria at this point, both applications can be of substantial value in a clinical context. Targets that are directly druggable present the most attractive candidates for therapeutic development, as their potential is immediately obvious and testable. Similarly, genetic findings that point towards a druggable pathway can present up or downstream candidates to enact the outcome if a mechanism can be established. Currently, neither PKLR nor IRF8 present an obvious means to direct therapeutic manipulation for malaria treatment, though the IRF8 story does lend support to investigating immunosuppression as a strategy when cerebral malaria is confirmed.

However, there is still room for these genetic findings to become clinically useful outside of immediate therapeutics. As field diagnostic tools continue to improve for use in developing countries, a spike in some specific RNA production, appearance of a biomarker protein in serum, or even predicted predisposition to severe disease based on genotype can all be used diagnose or prioritize resources and treatment options where needed. Rapid tests based on ELISA are currently available for HIV diagnosis and serve as practical and economical point-of-care tools in resource poor settings[350]. Similar diagnostics exist for visceral leishmaniasis or malaria infection *per se*[351]. RNA-based tests are also under study, or used, to rapidly diagnose infections including HIV, Hepatitis C

and visceral leishmaniasis [351-354]. At this point, all of these dipstick rapid diagnostics are based exclusively on detection of pathogen markers and antigens. However, if a reliable early host marker can be found, there is the potential for a similar test to be designed in order to differentiate between those progressing to severe disease versus those who will not. Human biomarkers in disease have been developed with varying success for multiple cancer types as diagnoses and prognostic indicators, which can serve as a starting platform and proof of concept[355].

### 5.2.4 Relevance to other conditions

One of the observations from both Chapter 2 (*Berr5*) and Chapter 3 (*Irf8*) was how much cerebral malaria resistance loci overlapped with resistance to other infectious or immunopathological conditions. The *Berr5* locus overlapped with a mouse tuberculosis resistance QTL (*Trl4*), an experimental autoimmune encephalitis QTL (*Eae19*), and an immune balancing QTL (*Tsiq2*)[191]. Though the range is too large to draw any substantive conclusions, the syntenic regions in humans (chr9:2-6.5Mb; chr10:50.3-52.7Mb; and chr10:87.5-100.2Mb) have been implicated in GWAS studies for circulating VEGF levels (chr9p24.2, rs10738760, $p=1.96x10^{-34}$, author candidate: *VLDLR*)[356] and Crohn's disease, inflammatory bowel disease (IBD) and ulcerative colitis (chr9.p24.1, rs10758669, $1x10^{-6}>p>8x10^{-45}$, author candidate: upstream JAK2)[268, 269, 286, 357, 358]. As discussed in Chapter 3, IRF8 and binding partners or downstream targets have been similarly associated with tuberculosis and a range of inflammatory disorders including Crohn's, IBD, colitis and psoriasis as well[271, 322]. What this suggests is that the genetic findings that emerge from the murine experimental cerebral malaria model may have broader interest and applicability.

Immunopathologies like IBD affect over 1.1 million people in the United States alone, with currently no cure, but requiring a lifetime of management[359]. Multiple sclerosis affects 1.1-2.5 million people and rheumatoid arthritis an additional 1.5 million US adults. The striking overlap in the genetic etiology of these autoimmune pathologies with each other has been well noted[360]. The mouse model of experimental cerebral malaria seems to share that overlapping genetic etiology and genetic dissection of the trait in a mouse model demonstrated to share pathway similarity is more cost effective, more amenable to hypothesis testing, and overall more controllable than human association or intervention studies. New targets that are truly causative to the pathology, or elucidation of the critical turning points and mechanisms can assist with the search for new understanding and new approaches to therapeutics in these conditions. Experimental cerebral malaria provides a rapid, sensitive and easy to score model for tackling some of these issues. It also is certainly not impossible to imagine that the pathways going awry in autoimmune pathologies are the same as those unbalanced during a host response to pressure from parasitic infections, making this overlap more than coincidental. Proponents of the hygiene hypothesis have theorized that our comparative freedom from parasites and childhood infectious diseases may be contributing to the rise of chronic inflammatory conditions, as the same genetics are placed in new environmental context[361].

## 5.3 Final Conclusions

Malaria continues to place a steady pressure on children and adults living in most parts of the developing world. Current understanding of cerebral

pathologies and existing strategies to treat patients are clearly insufficient, and genetic approaches to dissect the complexity can provide valuable new insights. This work has presented new players in the host response to infection with the *Berr5* locus and *Irf8*, as well as drawn some tantalizing parallels between the pathways dysregulated during experimental cerebral malaria and in human autoimmune inflammatory conditions. Synthesizing the data from these two lines of research is ripe for future studies, and both common and unique pathways can be used to improve global health. We also were able to expand the translation of the *Pklr*/*PKLR* story out of experimental models and conditions, paving the way for hypothesis testing PKLR mutations in malaria resistance at the population, rather than just the individual level. These studies open new doors with diagnostic and therapeutic potential as well as provide an opportunity to further understand the conditions, pressures and evolutionary history of early humanity.

# Chapter 6 : Original Contributions to Knowledge

1. Identified a locus on mouse Chromosome 19 (*Berr5*) that differentiates the host response to experimental cerebral malaria between susceptible C57BL/6 and resistant BALB/c inbred strains. Positional candidate genes include the *Ifit1-3* gene family, *Fas* and 13 others with strain-dependent differential gene expression during infection.
2. Identification of interferon regulatory factor 8 (*Irf8*) as a critical factor for experimental cerebral malaria pathological processes to occur. Cerebral malaria resistance in hypomorphic $Irf8^{R294C}$ mutants is inherited semi-dominantly and associated with IRF8 activation signaling activity, as heterozygote $Irf8^{R294C/+}$ mice have normal cellular profiles but significantly improved survival versus $Irf8^{+/+}$.
3. Delineation of a gene expression core shared by mice with experimental cerebral malaria pathology and *Mycobacterium tuberculosis* pathology, where IRF8 transcription factor binding sites are extremely highly represented.
4. Characterization of the genetic diversity in human PKLR using a panel of malaria endemic and non-endemic ethnic DNA.
5. Identification of 7 putative loss-of-function PKLR alleles including novel variants in DNA samples derived from malaria-endemic populations. These population characteristics were suggestive of positive selection for non-synonymous variant retention by neutrality testing.

# Chapter 7 : References:


1. World Health Organization., *World malaria report: 2010*. 2010, Geneva: World Health Organization. 205 p.
2. Cibulskis, R.E., et al., *Worldwide incidence of malaria in 2009: estimates, time trends, and a critique of methods.* PLoS Med, 2011. **8**(12): p. e1001142.
3. Carlton, J.M., B.J. Sina, and J.H. Adams, *Why is Plasmodium vivax a neglected tropical disease?* PLoS Negl Trop Dis, 2011. **5**(6): p. e1160.
4. Farrow, R.E., et al., *The mechanism of erythrocyte invasion by the malarial parasite, Plasmodium falciparum.* Semin Cell Dev Biol, 2011. **22**(9): p. 953-60.
5. Bigoga, J.D., et al., *Malaria vectors and transmission dynamics in coastal south-western Cameroon.* Malar J, 2007. **6**: p. 5.
6. Dery, D.B., et al., *Patterns and seasonality of malaria transmission in the forest-savannah transitional zones of Ghana.* Malar J, 2010. **9**: p. 314.
7. Okello, P.E., et al., *Variation in malaria transmission intensity in seven sites throughout Uganda.* Am J Trop Med Hyg, 2006. **75**(2): p. 219-25.
8. Kelly-Hope, L.A. and F.E. McKenzie, *The multiplicity of malaria transmission: a review of entomological inoculation rate measurements and methods across sub-Saharan Africa.* Malar J, 2009. **8**: p. 19.
9. Beier, J.C., G.F. Killeen, and J.I. Githure, *Short report: entomologic inoculation rates and Plasmodium falciparum malaria prevalence in Africa.* Am J Trop Med Hyg, 1999. **61**(1): p. 109-13.
10. Akachi, Y. and R. Atun, *Effect of investment in malaria control on child mortality in sub-Saharan Africa in 2002-2008.* PLoS One, 2011. **6**(6): p. e21309.
11. Aponte, J.J., et al., *Efficacy and safety of intermittent preventive treatment with sulfadoxine-pyrimethamine for malaria in African infants: a pooled analysis of six randomised, placebo-controlled trials.* Lancet, 2009. **374**(9700): p. 1533-42.
12. Roca-Feltrer, A., et al., *The age patterns of severe malaria syndromes in sub-Saharan Africa across a range of transmission intensities and seasonality settings.* Malar J, 2010. **9**: p. 282.
13. Kaatano, G.M., et al., *Patterns of malaria related mortality based on verbal autopsy in Muleba District, north-western Tanzania.* Tanzan J Health Res, 2009. **11**(4): p. 210-8.
14. Carneiro, I., et al., *Age-patterns of malaria vary with severity, transmission intensity and seasonality in sub-Saharan Africa: a systematic review and pooled analysis.* PLoS One, 2010. **5**(2): p. e8988.
15. Haldar, K. and N. Mohandas, *Malaria, erythrocytic infection, and anemia.* Hematology Am Soc Hematol Educ Program, 2009: p. 87-93.
16. Sarkar, P.K., et al., *Critical care aspects of malaria.* J Intensive Care Med, 2010. **25**(2): p. 93-103.

17. Genton, B., et al., *Indicators of fatal outcome in paediatric cerebral malaria: a study of 134 comatose Papua New Guinean children.* Int J Epidemiol, 1997. **26**(3): p. 670-6.
18. Ikome, L.E., K.J. Ndamukong, and H. Kimbi, *Prevalence and case-control study of cerebral malaria in Limbe of the South-West Cameroon.* Afr J Health Sci, 2002. **9**(1-2): p. 61-7.
19. Kumar, A., et al., *Burden of malaria in India: retrospective and prospective view.* Am J Trop Med Hyg, 2007. **77**(6 Suppl): p. 69-78.
20. World Health Organization, R.O.f.S.-E.A., *Regional guidelines for the management of severe falciparum malaria in large hospitals*. 2006: World Health Organization.
21. Hulden, L., *Activation of the hypnozoite: a part of Plasmodium vivax life cycle and survival.* Malar J, 2011. **10**: p. 90.
22. Greenwood, B.M., et al., *Malaria: progress, perils, and prospects for eradication.* J Clin Invest, 2008. **118**(4): p. 1266-76.
23. Salmon, B.L., A. Oksman, and D.E. Goldberg, *Malaria parasite exit from the host erythrocyte: a two-step process requiring extraerythrocytic proteolysis.* Proc Natl Acad Sci U S A, 2001. **98**(1): p. 271-6.
24. Bagnaresi, P., et al., *Unlike the synchronous Plasmodium falciparum and P. chabaudi infection, the P. berghei and P. yoelii asynchronous infections are not affected by melatonin.* Int J Gen Med, 2009. **2**: p. 47-55.
25. Kwiatkowski, D. and M. Nowak, *Periodic and chaotic host-parasite interactions in human malaria.* Proc Natl Acad Sci U S A, 1991. **88**(12): p. 5111-3.
26. Garcia, C.R., R.P. Markus, and L. Madeira, *Tertian and quartan fevers: temporal regulation in malarial infection.* J Biol Rhythms, 2001. **16**(5): p. 436-43.
27. King, C.L., et al., *Fy(a)/Fy(b) antigen polymorphism in human erythrocyte Duffy antigen affects susceptibility to Plasmodium vivax malaria.* Proc Natl Acad Sci U S A, 2011. **108**(50): p. 20113-8.
28. Batchelor, J.D., J.A. Zahm, and N.H. Tolia, *Dimerization of Plasmodium vivax DBP is induced upon receptor binding and drives recognition of DARC.* Nat Struct Mol Biol, 2011. **18**(8): p. 908-14.
29. Hadley, T.J. and S.C. Peiper, *From malaria to chemokine receptor: the emerging physiologic role of the Duffy blood group antigen.* Blood, 1997. **89**(9): p. 3077-91.
30. Ord, R.L., et al., *Targeting sialic acid dependent and independent pathways of invasion in Plasmodium falciparum.* PLoS One, 2012. **7**(1): p. e30251.
31. Mayer, D.C., et al., *The glycophorin C N-linked glycan is a critical component of the ligand for the Plasmodium falciparum erythrocyte receptor BAEBL.* Proc Natl Acad Sci U S A, 2006. **103**(7): p. 2358-62.
32. Mayer, D.C., et al., *Glycophorin B is the erythrocyte receptor of Plasmodium falciparum erythrocyte-binding ligand, EBL-1.* Proc Natl Acad Sci U S A, 2009. **106**(13): p. 5348-52.
33. Orlandi, P.A., F.W. Klotz, and J.D. Haynes, *A malaria invasion receptor, the 175-kilodalton erythrocyte binding antigen of Plasmodium falciparum recognizes*

*the terminal Neu5Ac(alpha 2-3)Gal- sequences of glycophorin A.* J Cell Biol, 1992. **116**(4): p. 901-9.
34. Maier, A.G., et al., *Polymorphisms in erythrocyte binding antigens 140 and 181 affect function and binding but not receptor specificity in Plasmodium falciparum.* Infect Immun, 2009. **77**(4): p. 1689-99.
35. Tham, W.H., et al., *Complement receptor 1 is the host erythrocyte receptor for Plasmodium falciparum PfRh4 invasion ligand.* Proc Natl Acad Sci U S A, 2010. **107**(40): p. 17327-32.
36. Crosnier, C., et al., *Basigin is a receptor essential for erythrocyte invasion by Plasmodium falciparum.* Nature, 2011. **480**(7378): p. 534-7.
37. Kats, L.M., et al., *Plasmodium rhoptries: how things went pear-shaped.* Trends Parasitol, 2006. **22**(6): p. 269-76.
38. Lingelbach, K. and K.A. Joiner, *The parasitophorous vacuole membrane surrounding Plasmodium and Toxoplasma: an unusual compartment in infected cells.* J Cell Sci, 1998. **111 ( Pt 11)**: p. 1467-75.
39. Talman, A.M., et al., *Gametocytogenesis: the puberty of Plasmodium falciparum.* Malar J, 2004. **3**: p. 24.
40. Lawaly, Y.R., et al., *Heritability of the human infectious reservoir of malaria parasites.* PLoS One, 2010. **5**(6): p. e11358.
41. Baker, D.A., *Malaria gametocytogenesis.* Mol Biochem Parasitol, 2010. **172**(2): p. 57-65.
42. Babiker, H.A., P. Schneider, and S.E. Reece, *Gametocytes: insights gained during a decade of molecular monitoring.* Trends Parasitol, 2008. **24**(11): p. 525-30.
43. Billker, O., et al., *Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito.* Nature, 1998. **392**(6673): p. 289-92.
44. Vlachou, D., et al., *The developmental migration of Plasmodium in mosquitoes.* Curr Opin Genet Dev, 2006. **16**(4): p. 384-91.
45. Muhlberger, N., et al., *Epidemiology and clinical features of vivax malaria imported to Europe: sentinel surveillance data from TropNetEurop.* Malar J, 2004. **3**: p. 5.
46. Perkins, D.J., et al., *Severe malarial anemia: innate immunity and pathogenesis.* Int J Biol Sci, 2011. **7**(9): p. 1427-42.
47. Winter, G. and M. Wahlgren, *Severe anemia in malaria: defense gone wrong?* Blood, 2005. **106**(10): p. 3337-3338.
48. Kai, O.K. and D.J. Roberts, *The pathophysiology of malarial anaemia: where have all the red cells gone?* BMC Med, 2008. **6**: p. 24.
49. Nuchsongsin, F., et al., *Effects of malaria heme products on red blood cell deformability.* Am J Trop Med Hyg, 2007. **77**(4): p. 617-22.
50. Dondorp, A.M., et al., *Red blood cell deformability as a predictor of anemia in severe falciparum malaria.* Am J Trop Med Hyg, 1999. **60**(5): p. 733-7.
51. Layez, C., et al., *Plasmodium falciparum rhoptry protein RSP2 triggers destruction of the erythroid lineage.* Blood, 2005. **106**(10): p. 3632-8.
52. Chang, K.H., M. Tam, and M.M. Stevenson, *Inappropriately low reticulocytosis in severe malarial anemia correlates with suppression in the development of late erythroid precursors.* Blood, 2004. **103**(10): p. 3727-35.

53. Snow, R.W., et al., *Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population.* Bull World Health Organ, 1999. **77**(8): p. 624-40.
54. Postels, D.G. and G.L. Birbeck, *Children with retinopathy-negative cerebral malaria: a pathophysiologic puzzle.* Pediatr Infect Dis J, 2011. **30**(11): p. 953-6.
55. Renia, L., et al., *Cerebral malaria: Mysteries at the blood-brain barrier.* Virulence, 2012. **3**(2): p. 193-201.
56. Buffet, P.A., et al., *The pathogenesis of Plasmodium falciparum malaria in humans: insights from splenic physiology.* Blood, 2011. **117**(2): p. 381-92.
57. Ward, C.P., et al., *Analysis of Plasmodium falciparum PfEMP-1/var genes suggests that recombination rearranges constrained sequences.* Mol Biochem Parasitol, 1999. **102**(1): p. 167-77.
58. Ochola, L.B., et al., *Specific receptor usage in Plasmodium falciparum cytoadherence is associated with disease outcome.* PLoS One, 2011. **6**(3): p. e14741.
59. Janes, J.H., et al., *Investigating the host binding signature on the Plasmodium falciparum PfEMP1 protein family.* PLoS Pathog, 2011. **7**(5): p. e1002032.
60. Beare, N.A., et al., *Perfusion abnormalities in children with cerebral malaria and malarial retinopathy.* J Infect Dis, 2009. **199**(2): p. 263-71.
61. Serghides, L., et al., *Inhaled nitric oxide reduces endothelial activation and parasite accumulation in the brain, and enhances survival in experimental cerebral malaria.* PLoS One, 2011. **6**(11): p. e27714.
62. Milner, D.A., Jr., *Rethinking cerebral malaria pathology.* Curr Opin Infect Dis, 2010. **23**(5): p. 456-63.
63. Navas, A., et al., *Experimental validation of Haldane's hypothesis on the role of infection as an evolutionary force for Metazoans.* Proc Natl Acad Sci U S A, 2007. **104**(34): p. 13728-31.
64. Baeza, A., et al., *Climate forcing and desert malaria: the effect of irrigation.* Malar J, 2011. **10**: p. 190.
65. Kwiatkowski, D.P., *How malaria has affected the human genome and what human genetics can teach us about malaria.* Am J Hum Genet, 2005. **77**(2): p. 171-92.
66. Neghina, R., et al., *Malaria, a journey in time: in search of the lost myths and forgotten stories.* Am J Med Sci, 2010. **340**(6): p. 492-8.
67. Hawass, Z., et al., *Ancestry and pathology in King Tutankhamun's family.* JAMA, 2010. **303**(7): p. 638-47.
68. El-Hazmi, M.A., A.M. Al-Hazmi, and A.S. Warsy, *Sickle cell disease in Middle East Arab countries.* Indian J Med Res, 2011. **134**(5): p. 597-610.
69. Roseff, S.D., *Sickle cell disease: a review.* Immunohematology, 2009. **25**(2): p. 67-74.
70. Piel, F.B., et al., *Global distribution of the sickle cell gene and geographical confirmation of the malaria hypothesis.* Nat Commun, 2010. **1**: p. 104.
71. Allison, A.C., *Genetic control of resistance to human malaria.* Curr Opin Immunol, 2009. **21**(5): p. 499-505.

72. Hill, A.V., et al., *Common west African HLA antigens are associated with protection from severe malaria.* Nature, 1991. **352**(6336): p. 595-600.
73. May, J., et al., *Hemoglobin variants and disease manifestations in severe falciparum malaria.* JAMA, 2007. **297**(20): p. 2220-6.
74. Jallow, M., et al., *Genome-wide and fine-resolution association analysis of malaria in West Africa.* Nat Genet, 2009.
75. Modiano, D., et al., *Haemoglobin C protects against clinical Plasmodium falciparum malaria.* Nature, 2001. **414**(6861): p. 305-8.
76. Williams, T.N., et al., *Sickle cell trait and the risk of Plasmodium falciparum malaria and other childhood diseases.* J Infect Dis, 2005. **192**(1): p. 178-86.
77. Aidoo, M., et al., *Protective effects of the sickle cell gene against malaria morbidity and mortality.* Lancet, 2002. **359**(9314): p. 1311-2.
78. Taylor, S.M., C.M. Parobek, and R.M. Fairhurst, *Haemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis.* Lancet Infect Dis, 2012. **12**(6): p. 457-68.
79. Gong, L., et al., *Evidence for both innate and acquired mechanisms of protection from Plasmodium falciparum in children with sickle cell trait.* Blood, 2012. **119**(16): p. 3808-14.
80. Fairhurst, R.M., et al., *Artemisinin-Resistant Malaria: Research Challenges, Opportunities, and Public Health Implications.* Am J Trop Med Hyg, 2012. **87**(2): p. 231-241.
81. Piper, K.P., et al., *Malaria transmission and naturally acquired immunity to PfEMP-1.* Infect Immun, 1999. **67**(12): p. 6369-74.
82. Bongfen, S.E., et al., *Genetic and genomic analyses of host-pathogen interactions in malaria.* Trends Parasitol, 2009. **25**(9): p. 417-22.
83. Hutagalung, R., et al., *Influence of hemoglobin E trait on the severity of Falciparum malaria.* J Infect Dis, 1999. **179**(1): p. 283-6.
84. Chotivanich, K., et al., *Hemoglobin E: a balanced polymorphism protective against high parasitemias and thus severe P falciparum malaria.* Blood, 2002. **100**(4): p. 1172-6.
85. Agarwal, A., et al., *Hemoglobin C associated with protection from severe malaria in the Dogon of Mali, a West African population with a low prevalence of hemoglobin S.* Blood, 2000. **96**(7): p. 2358-63.
86. Moradkhani, K., et al., *Mutations in the paralogous human alpha-globin genes yielding identical hemoglobin variants.* Ann Hematol, 2009. **88**(6): p. 535-43.
87. Berghout, J. and P. Gros, *Susceptibility to Malaria, Genetics of*, in *eLS*. 2001, John Wiley & Sons, Ltd.
88. Higgs, D.R. and D.J. Weatherall, *The alpha thalassaemias.* Cell Mol Life Sci, 2009. **66**(7): p. 1154-62.
89. Kalleas, C., et al., *Phenotype and genotype frequency of beta-thalassemia and sickle cell disease carriers in Halkidiki, Northern Greece.* Hemoglobin, 2012. **36**(1): p. 64-72.
90. Higgs, D.R., J.D. Engel, and G. Stamatoyannopoulos, *Thalassaemia.* Lancet, 2012. **379**(9813): p. 373-83.
91. Min-Oo, G. and P. Gros, *Erythrocyte variants and the nature of their malaria protective effect.* Cell Microbiol, 2005. **7**(6): p. 753-63.

92. Danquah, I. and F.P. Mockenhaupt, *Alpha(+)-thalassaemia and malarial anaemia.* Trends Parasitol, 2008. **24**(11): p. 479-81.
93. Williams, T.N., *Red blood cell defects and malaria.* Mol Biochem Parasitol, 2006. **149**(2): p. 121-7.
94. Weatherall, D.J., et al., *The population genetics and dynamics of the thalassemias.* Hematol Oncol Clin North Am, 2010. **24**(6): p. 1021-31.
95. Howes, R.E., et al., *The global distribution of the Duffy blood group.* Nat Commun, 2011. **2**: p. 266.
96. Fortin, A., M.M. Stevenson, and P. Gros, *Susceptibility to malaria as a complex trait: big pressure from a tiny creature.* Hum Mol Genet, 2002. **11**(20): p. 2469-78.
97. Oliveira, T.Y., et al., *Molecular evolution of a malaria resistance gene (DARC) in primates.* Immunogenetics, 2012. **64**(7): p. 497-505.
98. Cortes, A., et al., *Ability of Plasmodium falciparum to invade Southeast Asian ovalocytes varies between parasite lines.* Blood, 2004. **104**(9): p. 2961-6.
99. Allen, S.J., et al., *Prevention of cerebral malaria in children in Papua New Guinea by southeast Asian ovalocytosis band 3.* Am J Trop Med Hyg, 1999. **60**(6): p. 1056-60.
100. Williams, T.N., *Human red blood cell polymorphisms and malaria.* Curr Opin Microbiol, 2006. **9**(4): p. 388-94.
101. Ayi, K., et al., *Adenosine triphosphate depletion of erythrocytes simulates the phenotype associated with pyruvate kinase deficiency and confers protection against Plasmodium falciparum in vitro.* J Infect Dis, 2009. **200**(8): p. 1289-99.
102. Min-Oo, G., et al., *Pyruvate kinase deficiency in mice protects against malaria.* Nat Genet, 2003. **35**(4): p. 357-62.
103. Min-Oo, G., et al., *Pyruvate kinase deficiency: correlation between enzyme activity, extent of hemolytic anemia and protection against malaria in independent mouse mutants.* Blood Cells Mol Dis, 2007. **39**(1): p. 63-9.
104. Ayi, K., et al., *Pyruvate kinase deficiency and malaria.* N Engl J Med, 2008. **358**(17): p. 1805-10.
105. Berghout, J., et al., *Genetic diversity in human erythrocyte pyruvate kinase.* Genes Immun, 2012. **13**(1): p. 98-102.
106. Traherne, J.A., *Human MHC architecture and evolution: implications for disease association studies.* Int J Immunogenet, 2008. **35**(3): p. 179-92.
107. McGuire, W., et al., *Variation in the TNF-alpha promoter region associated with susceptibility to cerebral malaria.* Nature, 1994. **371**(6497): p. 508-10.
108. Randall, L.M. and C.R. Engwerda, *TNF family members and malaria: old observations, new insights and future directions.* Exp Parasitol, 2010. **126**(3): p. 326-31.
109. Turner, G.D., et al., *Systemic endothelial activation occurs in both mild and severe malaria. Correlating dermal microvascular endothelial cell phenotype and soluble cell adhesion molecules with disease severity.* The American journal of pathology, 1998. **152**(6): p. 1477-87.
110. Turner, G.D., et al., *An immunohistochemical study of the pathology of fatal malaria. Evidence for widespread endothelial activation and a potential role*

*for intercellular adhesion molecule-1 in cerebral sequestration.* The American journal of pathology, 1994. **145**(5): p. 1057-69.
111. Cserti-Gazdewich, C.M., et al., *Cytoadherence in paediatric malaria: ABO blood group, CD36, and ICAM1 expression and severe Plasmodium falciparum infection.* British journal of haematology, 2012. **159**(2): p. 223-36.
112. Park, G.S., et al., *Evidence of Endothelial Activation in Asymptomatic Parasitemia and Effect of Blood Group on Levels of von Willebrand Factor in Malaria.* Journal of the Pediatric Infectious Diseases Society, 2012. **1**(1): p. 16-25.
113. Viebig, N.K., et al., *Direct activation of human endothelial cells by Plasmodium falciparum-infected erythrocytes.* Infection and immunity, 2005. **73**(6): p. 3271-7.
114. Newbold, C., et al., *Receptor-specific adhesion and clinical disease in Plasmodium falciparum.* The American journal of tropical medicine and hygiene, 1997. **57**(4): p. 389-98.
115. Fernandez-Reyes, D., et al., *A high frequency African coding polymorphism in the N-terminal domain of ICAM-1 predisposing to cerebral malaria in Kenya.* Human molecular genetics, 1997. **6**(8): p. 1357-60.
116. Jenkins, N.E., et al., *A polymorphism of intercellular adhesion molecule-1 is associated with a reduced incidence of nonmalarial febrile illness in Kenyan children.* Clinical infectious diseases : an official publication of the Infectious Diseases Society of America, 2005. **41**(12): p. 1817-9.
117. Paganotti, G.M., et al., *Genetic complexity and gametocyte production of Plasmodium falciparum in Fulani and Mossi communities in Burkina Faso.* Parasitology, 2006. **132**(Pt 5): p. 607-14.
118. Idaghdour, Y., et al., *Feature Article: Evidence for additive and interaction effects of host genotype and infection in malaria.* Proc Natl Acad Sci U S A, 2012.
119. Jackson-Laboratory, *Biology of the Laboratory Mouse*, E.L. Green, Editor. 2008.
120. Longley, R., et al., *Host resistance to malaria: using mouse models to explore the host response.* Mamm Genome, 2011. **22**(1-2): p. 32-42.
121. Liew, K.J., et al., *Defining species specific genome differences in malaria parasites.* BMC Genomics, 2010. **11**: p. 128.
122. Prugnolle, F., et al., *Selection shapes malaria genomes and drives divergence between pathogens infecting hominids versus rodents.* BMC Evol Biol, 2008. **8**: p. 223.
123. Silva, J.C., et al., *Genome sequences reveal divergence times of malaria parasite lineages.* Parasitology, 2011. **138**(13): p. 1737-49.
124. Turcotte, K., et al., *Icsbp1/IRF-8 is required for innate and adaptive immune responses against intracellular pathogens.* J Immunol, 2007. **179**(4): p. 2467-76.
125. Hernandez-Valladares, M., et al., *Pathology of Tnf-deficient mice infected with Plasmodium chabaudi adami 408XZ.* Exp Parasitol, 2006. **114**(4): p. 271-8.
126. Finney, C.A., et al., *Divergent roles of IRAK4-mediated innate immune responses in two experimental models of severe malaria.* Am J Trop Med Hyg, 2010. **83**(1): p. 69-74.

127. Ing, R., P. Gros, and M.M. Stevenson, *Interleukin-15 enhances innate and adaptive immune responses to blood-stage malaria infection in mice.* Infect Immun, 2005. **73**(5): p. 3172-7.
128. Elias, R.M., et al., *Role of CD28 in polyclonal and specific T and B cell responses required for protection against blood stage malaria.* J Immunol, 2005. **174**(2): p. 790-9.
129. van der Heyde, H.C., et al., *The time course of selected malarial infections in cytokine-deficient mice.* Exp Parasitol, 1997. **85**(2): p. 206-13.
130. White, N.J., et al., *The murine cerebral malaria phenomenon.* Trends Parasitol, 2010. **26**(1): p. 11-5.
131. Stevenson, M.M., et al., *Cerebral malaria: human versus mouse studies.* Trends in parasitology, 2010. **26**(6): p. 274-5.
132. Rank, G., et al., *Novel roles for erythroid Ankyrin-1 revealed through an ENU-induced null mouse mutant.* Blood, 2009. **113**(14): p. 3352-62.
133. Shear, H.L., et al., *Resistance to malaria in ankyrin and spectrin deficient mice.* Br J Haematol, 1991. **78**(4): p. 555-60.
134. Reimer, T., et al., *Experimental cerebral malaria progresses independently of the Nlrp3 inflammasome.* Eur J Immunol, 2010. **40**(3): p. 764-9.
135. Griffith, J.W., et al., *Toll-like receptor modulation of murine cerebral malaria is dependent on the genetic background of the host.* J Infect Dis, 2007. **196**(10): p. 1553-64.
136. Kordes, M., K. Matuschewski, and J.C. Hafalla, *Caspase-1 activation of interleukin-1beta (IL-1beta) and IL-18 is dispensable for induction of experimental cerebral malaria.* Infect Immun, 2011. **79**(9): p. 3633-41.
137. Belnoue, E., et al., *CCR5 deficiency decreases susceptibility to experimental cerebral malaria.* Blood, 2003. **101**(11): p. 4253-9.
138. Hansen, D.S., et al., *Regulation of murine cerebral malaria pathogenesis by CD1d-restricted NKT cells and the natural killer complex.* Immunity, 2003. **18**(3): p. 391-402.
139. Oakley, M.S., et al., *Pathogenic roles of CD14, galectin-3, and OX40 during experimental cerebral malaria in mice.* PLoS One, 2009. **4**(8): p. e6793.
140. Piguet, P.F., et al., *Role of CD40-CVD40L in mouse severe malaria.* Am J Pathol, 2001. **159**(2): p. 733-42.
141. Campanella, G.S., et al., *Chemokine receptor CXCR3 and its ligands CXCL9 and CXCL10 are required for the development of murine cerebral malaria.* Proc Natl Acad Sci U S A, 2008. **105**(12): p. 4814-9.
142. Akimitsu, N., et al., *Duffy antigen is important for the lethal effect of the lethal strain of Plasmodium yoelii 17XL.* Parasitol Res, 2004. **93**(6): p. 499-503.
143. Beattie, L., et al., *CD8+ T lymphocyte-mediated loss of marginal metallophilic macrophages following infection with Plasmodium chabaudi chabaudi AS.* J Immunol, 2006. **177**(4): p. 2518-26.
144. Ohno, T., F. Kobayashi, and M. Nishimura, *Fas has a role in cerebral malaria, but not in proliferation or exclusion of the murine parasite in mice.* Immunogenetics, 2005. **57**(3-4): p. 293-6.

145. Waisberg, M., et al., *Genetic susceptibility to systemic lupus erythematosus protects against cerebral malaria in mice.* Proc Natl Acad Sci U S A, 2011. **108**(3): p. 1122-7.
146. Pamplona, A., I.A. Clark, and M.M. Mota, *Severe malaria increases the list of heme oxygenase-1-protected diseases.* Future Microbiol, 2007. **2**: p. 361-3.
147. Seixas, E., et al., *Heme oxygenase-1 affords protection against noncerebral forms of severe malaria.* Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15837-42.
148. Hunt, N.H., C. Driussi, and L. Sai-Kiang, *Haptoglobin and malaria.* Redox Rep, 2001. **6**(6): p. 389-92.
149. Li, J., et al., *Intercellular adhesion molecule 1 is important for the development of severe experimental malaria but is not required for leukocyte adhesion in the brain.* J Investig Med, 2003. **51**(3): p. 128-40.
150. Voisine, C., et al., *Classical CD11c+ dendritic cells, not plasmacytoid dendritic cells, induce T cell responses to Plasmodium chabaudi malaria.* Int J Parasitol, 2010. **40**(6): p. 711-9.
151. Amani, V., et al., *Involvement of IFN-gamma receptor-medicated signaling in pathology and anti-malarial immunity induced by Plasmodium berghei infection.* Eur J Immunol, 2000. **30**(6): p. 1646-55.
152. Su, Z. and M.M. Stevenson, *Central role of endogenous gamma interferon in protective immunity against blood-stage Plasmodium chabaudi AS infection.* Infect Immun, 2000. **68**(8): p. 4399-406.
153. Shio, M.T., et al., *Malarial hemozoin activates the NLRP3 inflammasome through Lyn and Syk kinases.* PLoS Pathog, 2009. **5**(8): p. e1000559.
154. Fauconnier, M., et al., *IL-12Rbeta2 Is Essential for the Development of Experimental Cerebral Malaria.* J Immunol, 2012. **188**(4): p. 1905-14.
155. Senaldi, G., et al., *Protection against the mortality associated with disease models mediated by TNF and IFN-gamma in mice lacking IFN regulatory factor-1.* J Immunol, 1999. **163**(12): p. 6820-6.
156. Engwerda, C.R., et al., *Locally up-regulated lymphotoxin alpha, not systemic tumor necrosis factor alpha, is the principle mediator of murine cerebral malaria.* J Exp Med, 2002. **195**(10): p. 1371-7.
157. Togbe, D., et al., *Both functional LTbeta receptor and TNF receptor 2 are required for the development of experimental cerebral malaria.* PLoS One, 2008. **3**(7): p. e2608.
158. Clark, K., et al., *Lymphotoxin alpha and tumour necrosis factor are not required for control of parasite growth, but differentially regulate cytokine production during Plasmodium chabaudi chabaudi AS infection.* Parasite Immunol, 2007. **29**(3): p. 153-8.
159. Krucken, J., et al., *Deletion of LTbetaR augments male susceptibility to Plasmodium chabaudi.* Parasite Immunol, 2005. **27**(6): p. 205-12.
160. Coban, C., et al., *Pathological role of Toll-like receptor signaling in cerebral malaria.* Int Immunol, 2007. **19**(1): p. 67-79.
161. Togbe, D., et al., *Murine cerebral malaria development is independent of toll-like receptor signaling.* Am J Pathol, 2007. **170**(5): p. 1640-8.
162. Franklin, B.S., et al., *MyD88-dependent activation of dendritic cells and CD4(+) T lymphocytes mediates symptoms, but is not required for the immunological*

*control of parasites during rodent malaria.* Microbes Infect, 2007. **9**(7): p. 881-90.
163. Garnica, M.R., J.S. Silva, and H.F. de Andrade Junior, *Stromal cell-derived factor-1 production by spleen cells is affected by nitric oxide in protective immunity against blood-stage Plasmodium chabaudi CR in C57BL/6j mice.* Immunol Lett, 2003. **89**(2-3): p. 133-42.
164. Gramaglia, I., et al., *Low nitric oxide bioavailability contributes to the genesis of experimental cerebral malaria.* Nat Med, 2006. **12**(12): p. 1417-22.
165. van der Heyde, H.C., et al., *Nitric oxide is neither necessary nor sufficient for resolution of Plasmodium chabaudi malaria in mice.* J Immunol, 2000. **165**(6): p. 3317-23.
166. Srivastava, K., et al., *Platelet factor 4 mediates inflammation in experimental cerebral malaria.* Cell Host Microbe, 2008. **4**(2): p. 179-87.
167. Claser, C., et al., *CD8+ T cells and IFN-gamma mediate the time-dependent accumulation of infected red blood cells in deep organs during experimental cerebral malaria.* PLoS One, 2011. **6**(4): p. e18720.
168. Combes, V., et al., *Pathogenic role of P-selectin in experimental cerebral malaria: importance of the endothelial compartment.* Am J Pathol, 2004. **164**(3): p. 781-6.
169. Bullen, D.V., et al., *The lack of suppressor of cytokine signalling-1 (SOCS1) protects mice from the development of cerebral malaria caused by Plasmodium berghei ANKA.* Parasite Immunol, 2003. **25**(3): p. 113-8.
170. Thawani, N., M. Tam, and M.M. Stevenson, *STAT6-mediated suppression of erythropoiesis in an experimental model of malarial anemia.* Haematologica, 2009. **94**(2): p. 195-204.
171. Piguet, P.F., C.D. Kan, and C. Vesin, *Role of the tumor necrosis factor receptor 2 (TNFR2) in cerebral malaria in mice.* Lab Invest, 2002. **82**(9): p. 1155-66.
172. Li, C. and J. Langhorne, *Tumor necrosis factor alpha p55 receptor is important for development of memory responses to blood-stage malaria infection.* Infect Immun, 2000. **68**(10): p. 5724-30.
173. Min-Oo, G., et al., *Complex genetic control of susceptibility to malaria: positional cloning of the Char9 locus.* J Exp Med, 2007. **204**(3): p. 511-24.
174. Franke-Fayard, B., et al., *Sequestration and tissue accumulation of human malaria parasites: can we learn anything from rodent models of malaria?* PLoS Pathog, 2010. **6**(9): p. e1001032.
175. de Souza, J.B., et al., *Cerebral malaria: why experimental murine models are required to understand the pathogenesis of disease.* Parasitology, 2010. **137**(5): p. 755-72.
176. McQuillan, J.A., et al., *Coincident parasite and CD8 T cell sequestration is required for development of experimental cerebral malaria.* Int J Parasitol, 2011. **41**(2): p. 155-63.
177. Hunt, N.H., et al., *Immunopathogenesis of cerebral malaria.* Int J Parasitol, 2006. **36**(5): p. 569-82.
178. Medana, I.M. and G.D. Turner, *Human cerebral malaria and the blood-brain barrier.* Int J Parasitol, 2006. **36**(5): p. 555-68.

179. Amante, F.H., et al., *Immune-mediated mechanisms of parasite tissue sequestration during experimental cerebral malaria.* J Immunol, 2010. **185**(6): p. 3632-42.
180. Laroque, A., et al., *Genetic control of susceptibility to infection with Plasmodium chabaudi chabaudi AS in inbred mouse strains.* Genes Immun, 2012. **13**(2): p. 155-63.
181. Fortin, A., et al., *Genetic control of blood parasitaemia in mouse malaria maps to chromosome 8.* Nat Genet, 1997. **17**(4): p. 382-3.
182. Foote, S.J., et al., *Mouse loci for malaria-induced mortality and the control of parasitaemia.* Nat Genet, 1997. **17**(4): p. 380-1.
183. Burt, R.A., et al., *Temporal expression of an H2-linked locus in host response to mouse malaria.* Immunogenetics, 1999. **50**(5-6): p. 278-85.
184. Hernandez-Valladares, M., et al., *Confirmation and dissection of QTL controlling resistance to malaria in mice.* Mamm Genome, 2004. **15**(5): p. 390-8.
185. Hernandez-Valladares, M., et al., *Mapping of a new quantitative trait locus for resistance to malaria in mice by a comparative mapping approach with human Chromosome 5q31-q33.* Immunogenetics, 2004. **56**(2): p. 115-7.
186. Min-Oo, G., et al., *Mapping of Char10, a novel malaria susceptibility locus on mouse chromosome 9.* Genes Immun, 2010. **11**(2): p. 113-23.
187. Bagot, S., et al., *Identification of two cerebral malaria resistance loci using an inbred wild-derived mouse strain.* Proc Natl Acad Sci U S A, 2002. **99**(15): p. 9919-23.
188. Campino, S., et al., *Genetic control of parasite clearance leads to resistance to Plasmodium berghei ANKA infection and confers immunity.* Genes Immun, 2005. **6**(5): p. 416-21.
189. Ohno, T. and M. Nishimura, *Detection of a new cerebral malaria susceptibility locus, using CBA mice.* Immunogenetics, 2004. **56**(9): p. 675-8.
190. Nagayasu, E., et al., *Association of a determinant on mouse chromosome 18 with experimental severe Plasmodium berghei malaria.* Infect Immun, 2002. **70**(2): p. 512-6.
191. Berghout, J., et al., *Identification of a novel cerebral malaria susceptibility locus (Berr5) on mouse chromosome 19.* Genes Immun, 2010. **11**(4): p. 310-8.
192. Torre, S., et al., *Susceptibility to lethal cerebral malaria is regulated by epistatic interaction between chromosome 4 (Berr6) and chromosome 1 (Berr7) loci in mice.* Genes and immunity, 2013. **14**(4): p. 249-57.
193. Bopp, S.E., et al., *Identification of the Plasmodium berghei resistance locus 9 linked to survival on chromosome 9.* Malaria journal, 2013. **12**: p. 316.
194. Patel, S.N., et al., *C5 deficiency and C5a or C5aR blockade protects against cerebral malaria.* J Exp Med, 2008. **205**(5): p. 1133-43.
195. Bagot, S., et al., *Susceptibility to experimental cerebral malaria induced by Plasmodium berghei ANKA in inbred mouse strains recently derived from wild stock.* Infect Immun, 2002. **70**(4): p. 2049-56.
196. Sepulveda, N., et al., *Allelic penetrance approach as a tool to model two-locus interaction in complex binary traits.* Heredity, 2007. **99**(2): p. 173-84.

197. Nguyen, N., et al., *Random mutagenesis of the mouse genome: a strategy for discovering gene function and the molecular basis of disease.* Am J Physiol Gastrointest Liver Physiol, 2011. **300**(1): p. G1-11.
198. Gondo, Y., et al., *ENU-based gene-driven mutagenesis in the mouse: a next-generation gene-targeting system.* Exp Anim, 2010. **59**(5): p. 537-48.
199. Bongfen, S.E., et al., *An N-Ethyl-N-Nitrosourea (ENU)-Induced Dominant Negative Mutation in the JAK3 Kinase Protects against Cerebral Malaria.* PLoS One, 2012. **7**(2): p. e31012.
200. Coltel, N., et al., *Cerebral malaria -- a neurovascular pathology with many riddles still to be solved.* Curr Neurovasc Res, 2004. **1**(2): p. 91-110.
201. Ferreira, A., et al., *Sickle hemoglobin confers tolerance to Plasmodium infection.* Cell, 2011. **145**(3): p. 398-409.
202. Pamplona, A., et al., *Heme oxygenase-1 and carbon monoxide suppress the pathogenesis of experimental cerebral malaria.* Nat Med, 2007. **13**(6): p. 703-10.
203. Wu, Y., et al., *Amplification of P. falciparum Cytoadherence through induction of a pro-adhesive state in host endothelium.* PLoS One, 2011. **6**(10): p. e24784.
204. Dunstan, S.J., et al., *Variation in human genes encoding adhesion and proinflammatory molecules are associated with severe malaria in the Vietnamese.* Genes Immun, 2012.
205. Sun, G., et al., *Inhibition of platelet adherence to brain microvasculature protects against severe Plasmodium berghei malaria.* Infect Immun, 2003. **71**(11): p. 6553-61.
206. Talisuna, A.O., et al., *Mitigating the threat of artemisinin resistance in Africa: improvement of drug-resistance surveillance and response systems.* Lancet Infect Dis, 2012. **12**(11): p. 888-96.
207. WHO, *Malaria.* World Health Organization, 2007.
208. Lamb, T.J., et al., *Insights into the immunopathogenesis of malaria using mouse models.* Expert Rev Mol Med, 2006. **8**(6): p. 1-22.
209. Nyangoto, E.O., *Cell-mediated effector molecules and complicated malaria.* Int Arch Allergy Immunol, 2005. **137**(4): p. 326-42.
210. Mazier, D., J. Nitcheu, and M. Idrissa-Boubou, *Cerebral malaria and immunogenetics.* Parasite Immunol, 2000. **22**(12): p. 613-23.
211. Sexton, A.C., et al., *Transcriptional profiling reveals suppressed erythropoiesis, up-regulated glycolysis, and interferon-associated responses in murine malaria.* J Infect Dis, 2004. **189**(7): p. 1245-56.
212. Delahaye, N.F., et al., *Gene-expression profiling discriminates between cerebral malaria (CM)-susceptible mice and CM-resistant mice.* J Infect Dis, 2006. **193**(2): p. 312-21.
213. Lovegrove, F.E., et al., *Expression microarray analysis implicates apoptosis and interferon-responsive mechanisms in susceptibility to experimental cerebral malaria.* Am J Pathol, 2007. **171**(6): p. 1894-903.
214. Hansen, A.M., et al., *Increased expression of indoleamine 2,3-dioxygenase in murine malaria infection is predominantly localised to the vascular endothelium.* Int J Parasitol, 2004. **34**(12): p. 1309-19.

215. Rogerson, S.J., G.E. Grau, and N.H. Hunt, *The microcirculation in severe malaria.* Microcirculation, 2004. **11**(7): p. 559-76.
216. Pais, T.F. and S. Chatterjee, *Brain macrophage activation in murine cerebral malaria precedes accumulation of leukocytes and CD8+ T cell proliferation.* J Neuroimmunol, 2005. **163**(1-2): p. 73-83.
217. Renia, L., et al., *Pathogenic T cells in cerebral malaria.* Int J Parasitol, 2006. **36**(5): p. 547-54.
218. Fortin, A., et al., *Recombinant congenic strains derived from A/J and C57BL/6J: a tool for genetic dissection of complex traits.* Genomics, 2001. **74**(1): p. 21-35.
219. Boyartchuk, V.L., et al., *Multigenic control of Listeria monocytogenes susceptibility in mice.* Nat Genet, 2001. **27**(3): p. 259-60.
220. Broman, K.W., *Mapping quantitative trait loci in the case of a spike in the phenotype distribution.* Genetics, 2003. **163**(3): p. 1169-75.
221. Marquis, J.F., et al., *Fibrotic response as a distinguishing feature of resistance and susceptibility to pulmonary infection with Mycobacterium tuberculosis in mice.* Infect Immun, 2008. **76**(1): p. 78-88.
222. Penet, M.F., et al., *Imaging experimental cerebral malaria in vivo: significant role of ischemic brain edema.* J Neurosci, 2005. **25**(32): p. 7352-8.
223. Choi, P., et al., *Linkage analysis of the genetic determinants of T-cell IL-4 secretion, and identification of Flj20274 as a putative candidate gene.* Genes Immun, 2005. **6**(4): p. 290-7.
224. Butterfield, R.J., et al., *Identification of genetic loci controlling the characteristics and severity of brain and spinal cord lesions in experimental allergic encephalomyelitis.* Am J Pathol, 2000. **157**(2): p. 637-45.
225. Mitsos, L.M., et al., *Susceptibility to tuberculosis: a locus on mouse chromosome 19 (Trl-4) regulates Mycobacterium tuberculosis replication in the lungs.* Proc Natl Acad Sci U S A, 2003. **100**(11): p. 6610-5.
226. Haldar, K., et al., *Malaria: mechanisms of erythrocytic infection and pathological correlates of severe disease.* Annu Rev Pathol, 2007. **2**: p. 217-49.
227. Hunt, N.H. and G.E. Grau, *Cytokines: accelerators and brakes in the pathogenesis of cerebral malaria.* Trends Immunol, 2003. **24**(9): p. 491-9.
228. Clark, T.G., et al., *Tumor necrosis factor and lymphotoxin-alpha polymorphisms and severe malaria in African populations.* J Infect Dis, 2009. **199**(4): p. 569-75.
229. Bluyssen, H.A., et al., *Structure, chromosome localization, and regulation of expression of the interferon-regulated mouse Ifi54/Ifi56 gene family.* Genomics, 1994. **24**(1): p. 137-48.
230. Guo, J., K.L. Peters, and G.C. Sen, *Induction of the human protein P56 by interferon, double-stranded RNA, or virus infection.* Virology, 2000. **267**(2): p. 209-19.
231. Fensterl, V., et al., *Novel characteristics of the function and induction of murine p56 family proteins.* J Virol, 2008. **82**(22): p. 11045-53.
232. Berchtold, S., et al., *Forced IFIT-2 expression represses LPS induced TNF-alpha expression at posttranscriptional levels.* BMC Immunol, 2008. **9**: p. 75.
233. Strasser, A., P.J. Jost, and S. Nagata, *The many roles of FAS receptor signaling in the immune system.* Immunity, 2009. **30**(2): p. 180-92.

234. Villa-Morales, M., J. Santos, and J. Fernandez-Piqueras, *Functional Fas (Cd95/Apo-1) promoter polymorphisms in inbred mouse strains exhibiting different susceptibility to gamma-radiation-induced thymic lymphoma.* Oncogene, 2006. **25**(14): p. 2022-9.
235. Yanez, D.M., et al., *Participation of lymphocyte subpopulations in the pathogenesis of experimental murine cerebral malaria.* J Immunol, 1996. **157**(4): p. 1620-4.
236. Yoshimoto, T., et al., *A pathogenic role of IL-12 in blood-stage murine malaria lethal strain Plasmodium berghei NK65 infection.* J Immunol, 1998. **160**(11): p. 5500-5.
237. Tamura, T., et al., *The IRF family transcription factors in immunity and oncogenesis.* Annu.Rev Immunol, 2008. **26**: p. 535-584.
238. Wang, H. and H.C. Morse, 3rd, *IRF8 regulates myeloid and B lymphoid lineage diversification.* Immunol Res, 2009. **43**(1-3): p. 109-17.
239. Marquis, J.F., et al., *Interferon regulatory factor 8 regulates pathways for antigen presentation in myeloid cells and during tuberculosis.* PLoS Genet, 2011. **7**(6): p. e1002097.
240. Shi, L., et al., *Genome-wide analysis of interferon regulatory factor I binding in primary human monocytes.* Gene, 2011. **487**(1): p. 21-8.
241. Shin, D.M., C.H. Lee, and H.C. Morse, 3rd, *IRF8 governs expression of genes involved in innate and adaptive immunity in human and mouse germinal center B cells.* PLoS One, 2011. **6**(11): p. e27384.
242. Tamura, T., et al., *ICSBP directs bipotential myeloid progenitor cells to differentiate into mature macrophages.* Immunity, 2000. **13**(2): p. 155-65.
243. Holtschke, T., et al., *Immunodeficiency and chronic myelogenous leukemia-like syndrome in mice with a targeted mutation of the ICSBP gene.* Cell, 1996. **87**(2): p. 307-17.
244. Turcotte, K., et al., *Genetic control of myeloproliferation in BXH-2 mice.* Blood, 2004. **103**(6): p. 2343-50.
245. Turcotte, K., et al., *A mutation in the Icsbp1 gene causes susceptibility to infection and a chronic myeloid leukemia-like syndrome in BXH-2 mice.* J Exp Med, 2005. **201**(6): p. 881-90.
246. Gabriele, L., et al., *Regulation of apoptosis in myeloid cells by interferon consensus sequence-binding protein.* J Exp Med, 1999. **190**(3): p. 411-21.
247. Burchert, A., et al., *Interferon consensus sequence binding protein (ICSBP; IRF-8) antagonizes BCR/ABL and down-regulates bcl-2.* Blood, 2004. **103**(9): p. 3480-9.
248. Yang, J., et al., *Cutting edge: IRF8 regulates Bax transcription in vivo in primary myeloid cells.* J Immunol, 2011. **187**(9): p. 4426-30.
249. Tsujimura, H., T. Tamura, and K. Ozato, *Cutting edge: IFN consensus sequence binding protein/IFN regulatory factor 8 drives the development of type I IFN-producing plasmacytoid dendritic cells.* J Immunol, 2003. **170**(3): p. 1131-5.
250. Aliberti, J., et al., *Essential role for ICSBP in the in vivo development of murine CD8alpha + dendritic cells.* Blood, 2003. **101**(1): p. 305-10.

251. Schiavoni, G., et al., *ICSBP is essential for the development of mouse type I interferon-producing cells and for the generation and activation of CD8alpha(+) dendritic cells.* J Exp Med, 2002. **196**(11): p. 1415-25.
252. Wang, H., et al., *IRF8 regulates B-cell lineage specification, commitment, and differentiation.* Blood, 2008. **112**(10): p. 4028-38.
253. Wang, I.M., et al., *An IFN-gamma-inducible transcription factor, IFN consensus sequence binding protein (ICSBP), stimulates IL-12 p40 expression in macrophages.* J Immunol, 2000. **165**(1): p. 271-9.
254. Giese, N.A., et al., *Interferon (IFN) consensus sequence-binding protein, a transcription factor of the IFN regulatory factor family, regulates immune responses in vivo through control of interleukin 12 expression.* J Exp Med, 1997. **186**(9): p. 1535-46.
255. Zhu, C., et al., *Activation of the murine interleukin-12 p40 promoter by functional interactions between NFAT and ICSBP.* J Biol Chem, 2003. **278**(41): p. 39372-82.
256. Liu, J., et al., *Synergistic activation of interleukin-12 p35 gene transcription by interferon regulatory factor-1 and interferon consensus sequence-binding protein.* J Biol Chem, 2004. **279**(53): p. 55609-17.
257. Kim, Y.M., et al., *Roles of IFN consensus sequence binding protein and PU.1 in regulating IL-18 gene expression.* J Immunol, 1999. **163**(4): p. 2000-7.
258. Marquis, J.F., et al., *Disseminated and rapidly fatal tuberculosis in mice bearing a defective allele at IFN regulatory factor 8.* J Immunol, 2009. **182**(5): p. 3008-15.
259. Ouyang, X., et al., *Transcription factor IRF8 directs a silencing programme for TH17 cell differentiation.* Nat Commun, 2011. **2**: p. 314.
260. Scharton-Kersten, T., et al., *Interferon consensus sequence binding protein-deficient mice display impaired resistance to intracellular infection due to a primary defect in interleukin 12 p40 induction.* J Exp Med, 1997. **186**(9): p. 1523-34.
261. Fehr, T., et al., *Crucial role of interferon consensus sequence binding protein, but neither of interferon regulatory factor 1 nor of nitric oxide synthesis for protection against murine listeriosis.* J Exp Med, 1997. **185**(5): p. 921-31.
262. Hein, J., et al., *Interferon consensus sequence binding protein confers resistance against Yersinia enterocolitica.* Infect Immun, 2000. **68**(3): p. 1408-17.
263. Alter-Koltunoff, M., et al., *Innate immunity to intraphagosomal pathogens is mediated by interferon regulatory factor 8 (IRF-8) that stimulates the expression of macrophage-specific Nramp1 through antagonizing repression by c-Myc.* J Biol Chem, 2008. **283**(5): p. 2724-33.
264. Fortier, A., et al., *Restriction of Legionella pneumophila replication in macrophages requires concerted action of the transcriptional regulators Irf1 and Irf8 and nod-like receptors Naip5 and Nlrc4.* Infect Immun, 2009. **77**(11): p. 4794-805.
265. Ko, J., A. Gendron-Fitzpatrick, and G.A. Splitter, *Susceptibility of IFN regulatory factor-1 and IFN consensus sequence binding protein-deficient mice to brucellosis.* J Immunol, 2002. **168**(5): p. 2433-40.

266. Hambleton, S., et al., *IRF8 mutations and human dendritic-cell immunodeficiency.* N Engl J Med, 2011. **365**(2): p. 127-38.
267. Cunninghame Graham, D.S., et al., *Association of NCF2, IKZF1, IRF8, IFIH1, and TYK2 with systemic lupus erythematosus.* PLoS Genet, 2011. **7**(10): p. e1002341.
268. Anderson, C.A., et al., *Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47.* Nat Genet, 2011. **43**(3): p. 246-52.
269. Barrett, J.C., et al., *Genome-wide association defines more than 30 distinct susceptibility loci for Crohn's disease.* Nat Genet, 2008. **40**(8): p. 955-62.
270. Elding, H., et al., *Dissecting the genetics of complex inheritance: linkage disequilibrium mapping provides insight into Crohn disease.* Am J Hum Genet, 2011. **89**(6): p. 798-805.
271. De Jager, P.L., et al., *Meta-analysis of genome scans and replication identify CD6, IRF8 and TNFRSF1A as new multiple sclerosis susceptibility loci.* Nat Genet, 2009. **41**(7): p. 776-82.
272. Sawcer, S., et al., *Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis.* Nature, 2011. **476**(7359): p. 214-9.
273. *The genetic association of variants in CD6, TNFRSF1A and IRF8 to multiple sclerosis: a multicenter case-control study.* PLoS One, 2011. **6**(4): p. e18813.
274. Pino, P.A. and A.E. Cardona, *Isolation of brain and spinal cord mononuclear cells using percoll gradients.* J Vis Exp, 2011(48).
275. Du, P., W.A. Kibbe, and S.M. Lin, *lumi: a pipeline for processing Illumina microarray.* Bioinformatics, 2008. **24**(13): p. 1547-8.
276. Saeed, A.I., et al., *TM4 microarray software suite.* Methods Enzymol, 2006. **411**: p. 134-93.
277. Langlais, D., et al., *A pituitary-specific enhancer of the POMC gene with preferential activity in corticotrope cells.* Mol Endocrinol, 2011. **25**(2): p. 348-59.
278. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.
279. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.
280. Dorovini-Zis, K., et al., *The neuropathology of fatal cerebral malaria in malawian children.* Am J Pathol, 2011. **178**(5): p. 2146-58.
281. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat Genet, 2000. **25**(1): p. 25-9.
282. Sanni, L.A., et al., *Dramatic changes in oxidative tryptophan metabolism along the kynurenine pathway in experimental cerebral and noncerebral malaria.* Am J Pathol, 1998. **152**(2): p. 611-9.
283. Tan, R.S., et al., *Altered immune response of interferon regulatory factor 1-deficient mice against Plasmodium berghei blood-stage malaria infection.* Infect Immun, 1999. **67**(5): p. 2277-83.
284. Gramaglia, I., et al., *Cell- Rather Than Antibody-Mediated Immunity Leads to the Development of Profound Thrombocytopenia during Experimental*

*Plasmodium berghei Malaria.* The Journal of Immunology, 2005. **175**(11): p. 7699-7707.
285. Matsuyama, T., et al., *Targeted disruption of IRF-1 or IRF-2 results in abnormal type I IFN gene induction and aberrant lymphocyte development.* Cell, 1993. **75**(1): p. 83-97.
286. Jostins, L., et al., *Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease.* Nature, 2012. **491**(7422): p. 119-24.
287. Feng, J., et al., *IFN regulatory factor 8 restricts the size of the marginal zone and follicular B cell pools.* J Immunol, 2011. **186**(3): p. 1458-66.
288. Chen, L., Z.H. Zhang, and F. Sendo, *Neutrophils play a critical role in the pathogenesis of experimental cerebral malaria.* Clinical & Experimental Immunology, 2000. **120**(1): p. 125-133.
289. Porcherie, A., et al., *Critical role of the neutrophil-associated high-affinity receptor for IgE in the pathogenesis of experimental cerebral malaria.* The Journal of Experimental Medicine, 2011. **208**(11): p. 2225-2236.
290. Claser, C., et al., *CD8<sup>+</sup> T Cells and IFN-γ Mediate the Time-Dependent Accumulation of Infected Red Blood Cells in Deep Organs during Experimental Cerebral Malaria.* PLoS ONE, 2011. **6**(4): p. e18720.
291. Roebuck, K.A. and A. Finnegan, *Regulation of intercellular adhesion molecule-1 (CD54) gene expression.* J Leukoc Biol, 1999. **66**(6): p. 876-88.
292. Tamura, T., et al., *Prevention of experimental cerebral malaria by Flt3 ligand during infection with Plasmodium berghei ANKA.* Infect Immun, 2011. **79**(10): p. 3947-56.
293. Lundie, R.J., et al., *Blood-stage Plasmodium infection induces CD8+ T lymphocytes to parasite-expressed antigens, largely regulated by CD8alpha+ dendritic cells.* Proc Natl Acad Sci U S A, 2008. **105**(38): p. 14509-14.
294. Yamada, H., S. Mizuno, and I. Sugawara, *Interferon regulatory factor 1 in mycobacterial infection.* Microbiol Immunol, 2002. **46**(11): p. 751-60.
295. MacMicking, J.D., G.A. Taylor, and J.D. McKinney, *Immune control of tuberculosis by IFN-gamma-inducible LRG-47.* Science, 2003. **302**(5645): p. 654-9.
296. Sousa, A.O., et al., *Relative contributions of distinct MHC class I-dependent cell populations in protection to tuberculosis infection in mice.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 4204-8.
297. Gateva, V., et al., *A large-scale replication study identifies TNIP1, PRDM1, JAZF1, UHRF1BP1 and IL10 as risk loci for systemic lupus erythematosus.* Nat Genet, 2009. **41**(11): p. 1228-33.
298. Center, U.M., *PKLR Mutation Database*. 2007, Laboratory for Red Blood Cell Research.
299. Valentini, G., et al., *Structure and function of human erythrocyte pyruvate kinase. Molecular basis of nonspherocytic hemolytic anemia.* J Biol Chem, 2002. **277**(26): p. 23807-14.
300. van Wijk, R., et al., *Fifteen novel mutations in PKLR associated with pyruvate kinase (PK) deficiency: structural implications of amino acid substitutions in PK.* Hum Mutat, 2009. **30**(3): p. 446-53.

301. Machado, P., et al., *Malaria: looking for selection signatures in the human PKLR gene region.* Br J Haematol, 2010. **149**(5): p. 775-84.
302. Rosenberg, N.A., *Standardized subsets of the HGDP-CEPH Human Genome Diversity Cell Line Panel, accounting for atypical and duplicated samples and pairs of close relatives.* Ann Hum Genet, 2006. **70**(Pt 6): p. 841-7.
303. Cann, H.M., et al., *A human genome diversity cell line panel.* Science, 2002. **296**(5566): p. 261-2.
304. Barrett, J.C., et al., *Haploview: analysis and visualization of LD and haplotype maps.* Bioinformatics, 2005. **21**(2): p. 263-5.
305. Gabriel, S.B., et al., *The structure of haplotype blocks in the human genome.* Science, 2002. **296**(5576): p. 2225-9.
306. Browning, S.R. and B.L. Browning, *Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering.* Am J Hum Genet, 2007. **81**(5): p. 1084-97.
307. Sabeti, P.C., et al., *Detecting recent positive selection in the human genome from haplotype structure.* Nature, 2002. **419**(6909): p. 832-7.
308. Leutenegger, A.L., et al., *Consanguinity around the world: what do the genomic data of the HGDP-CEPH diversity panel tell us?* Eur J Hum Genet, 2011. **19**(5): p. 583-7.
309. Shi, W., et al., *A worldwide survey of human male demographic history based on Y-SNP and Y-STR data from the HGDP-CEPH populations.* Mol Biol Evol, 2010. **27**(2): p. 385-93.
310. Beutler, E. and T. Gelbart, *Estimating the prevalence of pyruvate kinase deficiency from the gene frequency in the general white population.* Blood, 2000. **95**(11): p. 3585-8.
311. Demina, A., et al., *Six previously undescribed pyruvate kinase mutations causing enzyme deficiency.* Blood, 1998. **92**(2): p. 647-52.
312. Abu-Melha, A.M., et al., *Erythrocyte pyruvate kinase deficiency in newborns of eastern Saudi Arabia.* Acta Haematol, 1991. **85**(4): p. 192-4.
313. Feng, C.S., S.S. Tsang, and Y.T. Mak, *Prevalence of pyruvate kinase deficiency among the Chinese: determination by the quantitative assay.* Am J Hematol, 1993. **43**(4): p. 271-3.
314. Yavarian, M., et al., *Prevalence of pyruvate kinase deficiency among the south Iranian population: quantitative assay and molecular analysis.* Blood Cells Mol Dis, 2008. **40**(3): p. 308-11.
315. Librado, P. and J. Rozas, *DnaSP v5: a software for comprehensive analysis of DNA polymorphism data.* Bioinformatics, 2009. **25**(11): p. 1451-2.
316. Akey, J.M., et al., *Population history and natural selection shape patterns of genetic variation in 132 genes.* PLoS Biol, 2004. **2**(10): p. e286.
317. Wayne, M.L., Simonsen K. L., *Statistical tests of neutrality in the age of weak selection.* Trends in Ecology and Evolution, 1998. **13**(6): p. 236-240.
318. Schaffner, S.F., et al., *Calibrating a coalescent simulation of human genome sequence variation.* Genome Res, 2005. **15**(11): p. 1576-83.
319. Zhang, L., et al., *Population genetics of duplicated disease-defense genes, hm1 and hm2, in maize (Zea mays ssp. mays L.) and its wild ancestor (Zea mays ssp. parviglumis).* Genetics, 2002. **162**(2): p. 851-60.

320. Mirabello, L. and J.E. Conn, *Molecular population genetics of the malaria vector Anopheles darlingi in Central and South America.* Heredity, 2006. **96**(4): p. 311-21.
321. Oluwayemi, I.O., et al., *Neurological sequelae in survivors of cerebral malaria.* The Pan African medical journal, 2013. **15**: p. 88.
322. Berghout, J., et al., *Irf8-regulated genomic responses drive pathological inflammation during cerebral malaria.* PLoS pathogens, 2013. **9**(7): p. e1003491.
323. Dong, H., et al., *B7-H1 determines accumulation and deletion of intrahepatic CD8(+) T lymphocytes.* Immunity, 2004. **20**(3): p. 327-36.
324. Yalcin, B., et al., *Next-generation sequencing of experimental mouse strains.* Mammalian genome : official journal of the International Mammalian Genome Society, 2012. **23**(9-10): p. 490-8.
325. Yue, F., et al., *A comparative encyclopedia of DNA elements in the mouse genome.* Nature, 2014. **515**(7527): p. 355-64.
326. Chen, L., et al., *Advances in genome editing technology and its promising application in evolutionary and ecological studies.* GigaScience, 2014. **3**: p. 24.
327. Singh, P., J.C. Schimenti, and E. Bolcun-Filas, *A Mouse Geneticist's Practical Guide to CRISPR Applications.* Genetics, 2014.
328. Frazer, K.A., et al., *A sequence-based variation map of 8.27 million SNPs in inbred mouse strains.* Nature, 2007. **448**(7157): p. 1050-3.
329. *Mouse Genome Database (MGD)*. Mouse Genome Informatics 2014 [cited 2014 July 9, 2014]; 5.18:[Available from: http://www.informatics.jax.org.
330. Trachtulec, Z., et al., *Fine haplotype structure of a chromosome 17 region in the laboratory and wild mouse.* Genetics, 2008. **178**(3): p. 1777-84.
331. Laboratory, T.J. *List of Wild-Derived Mouse Strains*. [webpage] 2012 [cited 2014 July 11, 2014]; Available from: http://jaxmice.jax.org/list/cat481389.html.
332. Adler, S., M. Yoeli, and A. Zuckerman, *Behaviour of Plasmodium berghei in some rodents.* Nature, 1950. **166**(4222): p. 571.
333. Killick-Kendrick, R., *Malaria parasites of Thamnomys rutilans (Rodentia, Muridae) in Nigeria.* Bulletin of the World Health Organization, 1968. **38**(5): p. 822-4.
334. Andrews, T.D., et al., *Massively parallel sequencing of the mouse exome to accurately identify rare, induced mutations: an immediate source for thousands of new mouse models.* Open biology, 2012. **2**(5): p. 120061.
335. Arnold, C.N., et al., *ENU-induced phenovariance in mice: inferences from 587 mutations.* BMC research notes, 2012. **5**: p. 577.
336. Masuya, H., et al., *A series of ENU-induced single-base substitutions in a long-range cis-element altering Sonic hedgehog expression in the developing mouse limb bud.* Genomics, 2007. **89**(2): p. 207-14.
337. Lewis, M.A., et al., *An ENU-induced mutation of miR-96 associated with progressive hearing loss in mice.* Nature genetics, 2009. **41**(5): p. 614-8.
338. Samuels, D.C., et al., *Finding the lost treasures in exome sequencing data.* Trends in genetics : TIG, 2013. **29**(10): p. 593-9.

339. Zhao, Y., et al., *A novel mutation in leptin gene is associated with severe obesity in Chinese individuals.* BioMed research international, 2014. **2014**: p. 912052.
340. Ng, S.B., et al., *Targeted capture and massively parallel sequencing of 12 human exomes.* Nature, 2009. **461**(7261): p. 272-6.
341. Ng, S.B., et al., *Exome sequencing identifies the cause of a mendelian disorder.* Nature genetics, 2010. **42**(1): p. 30-5.
342. Hoischen, A., et al., *De novo mutations of SETBP1 cause Schinzel-Giedion syndrome.* Nature genetics, 2010. **42**(6): p. 483-5.
343. Fairfield, H., et al., *Mutation discovery in mice by whole exome sequencing.* Genome biology, 2011. **12**(9): p. R86.
344. Asan, et al., *Comprehensive comparison of three commercial human whole-exome capture platforms.* Genome biology, 2011. **12**(9): p. R95.
345. Sharov, A.A., D.B. Dudekula, and M.S. Ko, *Genome-wide assembly and analysis of alternative transcripts in mouse.* Genome research, 2005. **15**(5): p. 748-54.
346. Waterston, R.H., et al., *Initial sequencing and comparative analysis of the mouse genome.* Nature, 2002. **420**(6915): p. 520-62.
347. Kildey, K., et al., *Characterization of ENU-induced Mutations in Red Blood Cell Structural Proteins.* Computational and structural biotechnology journal, 2013. **6**: p. e201303012.
348. Giha, H.A., et al., *Cerebral malaria is frequently associated with latent parasitemia among the semi-immune population of eastern Sudan.* Microbes Infect, 2005. **7**(11-12): p. 1196-203.
349. Armah, H.B., et al., *Cerebrospinal fluid and serum biomarkers of cerebral malaria mortality in Ghanaian children.* Malar J, 2007. **6**: p. 147.
350. Girardi, S.B., et al., *Evaluation of rapid tests for human immunodeficiency virus as a tool to detect recent seroconversion.* The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases, 2012. **16**(5): p. 452-6.
351. Boelaert, M., et al., *Rapid tests for the diagnosis of visceral leishmaniasis in patients with suspected disease.* The Cochrane database of systematic reviews, 2014. **6**: p. CD009135.
352. Jani, I.V., et al., *Accurate Early Infant HIV Diagnosis in Primary Health Clinics Using a Point-Of-Care Nucleic Acid Test.* Journal of acquired immune deficiency syndromes, 2014.
353. Saludes, V., et al., *Tools for the diagnosis of hepatitis C virus infection and hepatic fibrosis staging.* World journal of gastroenterology : WJG, 2014. **20**(13): p. 3431-42.
354. McCollum, E.D., et al., *Clinical versus rapid molecular HIV diagnosis in hospitalized African infants: a randomized controlled trial simulating point-of-care infant testing.* Journal of acquired immune deficiency syndromes, 2014. **66**(1): p. e23-30.
355. Soper, S.A., et al., *Point-of-care biosensor systems for cancer diagnostics/prognostics.* Biosensors & bioelectronics, 2006. **21**(10): p. 1932-42.

356. Debette, S., et al., *Identification of cis- and trans-acting genetic variants explaining up to half the variation in circulating vascular endothelial growth factor levels.* Circulation research, 2011. **109**(5): p. 554-63.
357. Franke, A., et al., *Genome-wide meta-analysis increases to 71 the number of confirmed Crohn's disease susceptibility loci.* Nature genetics, 2010. **42**(12): p. 1118-25.
358. McGovern, D.P., et al., *Genome-wide association identifies multiple ulcerative colitis susceptibility loci.* Nature genetics, 2010. **42**(4): p. 332-7.
359. Kappelman, M.D., et al., *Recent trends in the prevalence of Crohn's disease and ulcerative colitis in a commercially insured US population.* Digestive diseases and sciences, 2013. **58**(2): p. 519-25.
360. Beecham, A.H., et al., *Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis.* Nature genetics, 2013. **45**(11): p. 1353-60.
361. Zuckerman, M.K., et al., *The evolution of disease: anthropological perspectives on epidemiologic transitions.* Global health action, 2014. **7**: p. 23303.