Integrating theory and experimentation in the study of malaria

by
Nicole Mideo

A thesis submitted to the Department of Biology
in conformity with the requirements
for the degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
August 2009

Copyright © Nicole Mideo, 2009
ABSTRACT

Malaria poses a serious threat to much of the developing world and an enormous effort is under way to design vaccines and other novel interventions. Nevertheless, we understand very little about the ecology and evolution of malaria parasites. For instance, while scientists have had considerable success identifying factors involved in regulating parasite growth within hosts, it is extremely hard to disentangle the relative influences of host immunity and other within-host factors on infection dynamics. Many mathematical models have been directed at understanding the dynamics of malaria infections, and these have provided valuable insights. However, these models have also been criticized, most notably for lacking any statistical analysis of the goodness of fit of model predictions to data. Here, we develop a new modeling approach that improves on previous work, and apply it to a novel data set from a simplified rodent malaria system. We find that resource availability and competition are important drivers of dynamics, and we identify a number of parasite traits that may underlie differences in virulence between parasite strains. These include the number of progeny parasites produced per infected cell (burst size) and the invasion rates of target cells. We test these predictions with further experiments and find broad support for the role of burst sizes in determining virulence, but the role of invasion rates is less certain. We also find evidence of potential plasticity in these parasite traits.
in response to within-host environmental factors. These within-host interactions between parasites and hosts have effects that will scale up to between-host processes; we discuss the growing body of theory that seeks to combine these levels (‘embedded models’). Using between-host and embedded models, we test the plausibility of various hypotheses to explain why there are so few transmissible malaria parasite forms, yet vast numbers of host-damaging asexual forms are produced. We show that a specific form of density-dependent transmission-blocking immunity and the occurrence of multiple infections can each generate selection for this pattern. Overall, this thesis contributes to a better understanding of malaria parasites, while providing a framework for addressing unanswered questions in disease biology, and offering interesting paths for future empirical work.
ACKNOWLEDGEMENTS

Foremost, I thank my supervisor Troy Day for his patience, sense of humour, and encouragement (in equal and abundant measure). For reasons I still don’t understand, you showed confidence in me when mine was tapped out. Fearing the proof of your beliefs as baseless has always been hugely motivating. I also want to thank you for shaping this thesis into what it is and letting me believe that you think I somehow came up with it. You have been a better supervisor than I could ever have hoped for. Everyone should be so lucky.

To the co-captain of my thesis, Andrew Read, I owe a huge deal of thanks for welcoming me on board like I was his own student. Without you, this thesis would have been drastically different, I suspect less interesting, and I certainly can’t imagine I would’ve had as much fun. Please stay in the fan club! I am indebted also to Sarah Reece who averted an experimental catastrophe and was (still) brave enough to spearhead an empirical study executed in part by a theorist, and to Nick Savill for teaching me about model fitting.

Special thanks go to my supervisory committee: Bob Montgomerie, my occasional stand-in supervisor who always had my back while keeping me on my toes; Peter Taylor, my academic grandfather who preached that math is beauty and was there when I finally realized it; and Chris Eckert, my academic great-uncle (maybe?) and stats/R guru.

I might not have made it to this point without Samuel Alizon and Andy Gardner who
Acknowledgements

contributed in different ways, but shared an infinite patience for incessant questions. I thank the members of the Day, Taylor, Montgomerie, Read, and Reece labs for good science discussions and great (non-science) memories, especially Karen, Amy, Daniel, Vicki, Silvie, Will, Brian, and Laura.

I thank my family for always being supportive and proud even when it all seems like Greek to them, and Ken for more things than I can mention here, the most relevant of which is reminding me how to work hard. Thank you for everything.

Now, here is a joke that I have been waiting years to put in my thesis and which may have been the biggest motivator for actually writing this thing. I thought it was apt.

Q: What do you get when you cross a mosquito with a mountain climber?

A: You can’t cross a vector with a scalar!
CO-AUTHORSHIP

This thesis conforms to the “Manuscript Format” outlined in the Department of Biology Guide to Graduate Studies. Each chapter is co-authored by my supervisor, Dr. Troy Day, who contributed to project and model design, analysis and presentation of results, and financially supported me. Below is a list of additional contributions, as well as publications resulting from this thesis. I am first author on each of these publications.

- Chapter 3: Andrew Read (project design, interpretation and presentation of results), Nicholas Savill (statistical analyses), Victoria Barclay and Brian Chan (data collection). Published in *The American Naturalist* (2008) **172**: E214-E238.
- Chapter 4: Sarah Reece (experimental design, data analyses and interpretation, presentation of results), and Andrew Read (experimental design, interpretation and presentation of results).
# CONTENTS

**Abstract** ................................................................................................. ii

**Acknowledgements** ........................................................................ iv

**Co-authorship** .................................................................................... vi

**Table of Contents** ................................................................................ vii

**List of Figures** .................................................................................... xii

**List of Tables** ..................................................................................... xiv

**Chapter 1. General introduction** ..................................................... 1

**Chapter 2. Modeling malaria pathogenesis** ...................................... 7

2.1 Introduction ..................................................................................... 8

2.2 Modeling approaches to malaria pathogenesis ........................... 12

2.3 Insights ........................................................................................... 17

2.3.1 Effective immune targets .......................................................... 18

2.3.2 Protective immune control ....................................................... 18
<table>
<thead>
<tr>
<th>2.3.3 Clonal interactions</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.4 Gametocytogenesis</td>
<td>20</td>
</tr>
<tr>
<td>2.3.5 Consequences of RBC preference</td>
<td>20</td>
</tr>
<tr>
<td>2.3.6 Manipulation of erythropoiesis</td>
<td>22</td>
</tr>
<tr>
<td>2.4 Theoretical developments</td>
<td>22</td>
</tr>
<tr>
<td>2.4.1 Biological realism</td>
<td>23</td>
</tr>
<tr>
<td>2.4.2 Tying theory to data</td>
<td>24</td>
</tr>
<tr>
<td>2.5 Future directions</td>
<td>26</td>
</tr>
<tr>
<td>2.6 Conclusions</td>
<td>28</td>
</tr>
</tbody>
</table>

**Chapter 3. Strain-specific patterns of pathogenesis in malaria** | 30 |
| 3.1 Introduction | 31 |
| 3.2 Methods | 33 |
| 3.2.1 Model development and data fitting | 34 |
| 3.2.2 Statistical analysis | 42 |
| 3.2.3 Coinfection experiments | 43 |
| 3.3 Results | 45 |
| 3.3.1 Model development and data fitting | 45 |
| 3.3.2 Statistical Analysis | 49 |
| 3.3.3 Coinfection experiments | 51 |
| 3.4 Discussion | 53 |
Chapter 4.  Host exploitation strategies in two strains of rodent malaria  .... 60

4.1 Introduction  ... 61

4.2 Methods  ... 66

4.2.1 Parasites and hosts  ... 66

4.2.2 Manipulating immunity  ... 66

4.2.3 Manipulating RBC age structure  ... 67

4.2.4 Experimental setup and sampling  ... 69

4.2.5 Statistical analyses  ... 73

4.3 Results  ... 75

4.3.1 Burst sizes  ... 75

4.3.2 Invasion rates  ... 77

4.4 Discussion  ... 81

Chapter 5.  Linking within- and between-host disease dynamics  ... 87

5.1 Introduction  ... 88

5.2 What are nested dynamical models of pathogen evolution?  ... 89

5.3 What have nested models of infectious disease taught us?  ... 91

5.3.1 Interdependence of parameters  ... 91

5.3.2 Conflicting selection  ... 94

5.4 Future prospects and challenges  ... 98

5.5 Conclusion  ... 102
## Chapter 6. On the evolution of reproductive restraint in malaria

6.1 Introduction .................................................. 104
   6.1.1 Why might reproductive restraint be adaptive? .......... 106
6.2 Model and approach ........................................ 107
6.3 Results ..................................................... 109
   6.3.1 Mosquito survival ....................................... 109
   6.3.2 Density-dependent transmission-blocking immunity ....... 111
   6.3.3 Relative density-dependent transmission-blocking immunity . 112
   6.3.4 Multiple infection ....................................... 114
6.4 Discussion .................................................. 118

## Chapter 7. General discussion

7.1 Ongoing research and future studies .......................... 127
   7.1.1 The mathematics of immune responses to malaria .......... 127
   7.1.2 The growth versus reproduction trade-off in malaria ....... 127
   7.1.3 Using models to inform public health policy ............... 129
7.2 Epilogue .................................................. 130
7.3 Thesis summary ........................................... 132

References ..................................................... 133

Appendices ..................................................... 145

Chapter 3 ..................................................... 146
LIST OF FIGURES

2.1 Schematic of two recent models of malaria pathogenesis . . . . . . . . . 10
2.2 Model selection and validation . . . . . . . . . . . . . . . . . . . . . . . 25
3.1 Schematic of blood stage malaria infections and model framework . . . 36
3.2 Experimental data and best fit curves for RBC dynamics . . . . . . . . 47
3.3 Experimental data and simulations of parasite dynamics . . . . . . . . 48
3.4 Experimental data and predictions for RBC dynamics in competition . . 52
3.5 Experimental data and predictions for parasite dynamics in competition . 52
4.1 Erythrocytic stages in Plasmodium development . . . . . . . . . . . . . 62
4.2 Burst size predictions and data for CD4\(^+\) depleted mice . . . . . . . 64
4.3 Invasion rate predictions and data for CD4\(^+\) depleted mice . . . . . . 65
4.4 Parasite dynamics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 67
4.5 Red blood cell dynamics . . . . . . . . . . . . . . . . . . . . . . . . . . 68
4.6 Effect of PHZ on RBC age structure . . . . . . . . . . . . . . . . . . . 69
4.7 Burst size data for all treatments . . . . . . . . . . . . . . . . . . . . . 77
4.8 Invasion rate data for all treatments . . . . . . . . . . . . . . . . . . . . 79
5.1 Schematic of a nested model . . . . . . . . . . . . . . . . . . . . . . . . 90
5.2 Inessential and essential nested models ........................................ 97
5.3 Disease life history .................................................................. 100
6.1 Optimal gametocyte production from a continuum of trait pairs. . . . . 111
6.2 The effect of multiple infection on the optimal conversion rate. ....... 117
C.1 Expected maximum log-likelihood distributions ............................ 157
C.2 Estimated invasion rate parameters ............................................. 158
C.3 Estimated burst size parameters ............................................... 159
C.4 Estimated RBC production parameters ....................................... 160
C.5 Estimated effects of inoculum size and clone .............................. 161
D.1 Sampling protocol .................................................................. 162
LIST OF TABLES

2.1 Example mathematical descriptions of different underlying assumptions . 15
3.1 Model parameters, published estimates and test ranges . . . . . . . . 40
3.2 Median parameter estimates relative to AS $10^6$ values . . . . . . . 51
4.1 Analysis of burst sizes in CD4$^+$ depleted mice . . . . . . . . . . 76
4.2 Analysis of burst sizes across eight treatments . . . . . . . . . . . 78
4.3 Analysis of RBC invasion rates in CD4$^+$ depleted mice . . . . . . 78
4.4 Analysis of RBC invasion rates across eight treatments . . . . . . . 80
C.1 Maximum log-likelihoods for each model . . . . . . . . . . . . . . . 155
C.2 Estimated parameter values for best-fit model . . . . . . . . . . . 156
C.3 Medians of pooled estimated parameter distributions . . . . . . . 156
Chapter 1. GENERAL INTRODUCTION

It has been said that time can make us fools. A poignant example of this adage is provided by the former U.S. Surgeon General William H. Stewart, who in 1967 infamously claimed that “the war against infectious diseases has been won” (Morens et al. 2004). The sobering reality is that, currently, 1 in 6 deaths worldwide can be attributed to infectious diseases and in developing regions like Africa, this ratio jumps to 1 in 2 (WHO, 2008a). One of the main contributors to these numbers is malaria, a disease that every year claims the lives of over one million people and generates at least 300 million infections (WHO, 2008b). Such a scourge is malaria that the UN named it (along with HIV/AIDS) as an object of one of the UN’s Millennium Development Goals (UN Millennium Project: Task force on HIV/AIDS, Malaria, TB, and Access to Essential Medicines, Working Group on Malaria, 2005) and The Global Malaria Action Plan outlined specific targets for its control, including to reduce both malaria cases and malaria deaths “from 2000 levels by 50% in 2010” (Roll Back Malaria Partnership, 2008). Alas, we are a few short months from 2010 and the stated goals are desperately far from being within reach. Time and miscalculated optimism are conspiring again.

We are in our current predicament despite the fact that funding for malaria research has increased since the Millennium Development Goals were established (Roll Back Malaria
Chapter 1. General introduction

Partnership, 2008), and there has been growing interest in its control, with a number of upstart organizations aiming to bring to light the impact of the disease (see, for example, Breman et al., 2004). While the focus of much of this research energy and money is being placed squarely on intervention invention (and probably rightly so), what goes unheralded is the fact that the biology of malaria parasites remains poorly understood. It is clear that we need to move beyond traditional approaches for dealing with malaria, and I argue that an appeal for a better understanding of malaria biology is critical in this process.

Malaria is a complex disease and many questions about its ecology still lack satisfying answers. For example, what is it specifically about the interaction between hosts and parasites that results in disease symptoms? How do these interactions differ between malaria parasite strains and species? What are the important elements of the natural immune response to malaria, and what is the relative importance of each? When it comes to an understanding of the evolutionary biology of malaria parasites we are perhaps even worse off, as evidenced most strikingly by the repeated failure of new anti-malarial drugs in the face of evolving resistance, and the seeming capitulation of public health institutions to the continuation of this pattern (Read & Huijben, 2009). Little is known about what factors have shaped the evolution of various life-history traits of malaria parasites and what constraints will be at play when selective forces change (e.g. due to interventions).

Understanding the factors that regulate the within- and between-host dynamics and how these factors interact to shape the evolution of parasite traits (including those with clinical significance, e.g. virulence) is extremely important for designing, evaluating and
monitoring control strategies. In understanding the basic biology of malaria parasites and infected hosts, we may identify targets for intervention, e.g. susceptibilities in the parasite or potential enhancements to the human immune response. Developing this knowledge base demands a multi-faceted approach that includes clinical data, field studies, and controlled laboratory experiments with model systems. Communication between these branches is vitally important, so a key challenge is to find ways to bridge the chasms that exist between groups that use different approaches (e.g. Dietz et al., 2006). Mathematical modeling is one such bridge.

While the variation in experimental infections is necessarily restricted (they aim to give finer resolution to one or a number of constituent pieces of a biological system) and is near-unquantifiable in natural infections, models allow us to include elements of variation, in isolation and in combination, to examine their effects and make testable predictions. While the scope of experiments is often narrow (measuring within-host dynamics can be destructive, obviating the possibility of measuring between-host transmission) and intractable in natural infections, modeling can provide a unified framework for examining biology across scales. While experimental systems allow us the (ethical) freedom to explore biological questions about malaria infections, modeling can do this more cheaply and with the power to sift through a vast number of biological hypotheses (or, say, potential interventions) in short order.

Ultimately, and if we are careful in building the bridge, modeling can be an extremely powerful approach for both basic malaria research and malaria control. In this thesis, I
aim to demonstrate the utility of mathematical modeling for studying malaria, and other infectious diseases, particularly (though not exclusively) when tied closely to data. “The history of modeling in malaria is long, but so erratic that we have not yet passed the threshold at which modeling becomes a standard tool” (McKenzie, 2000). Nevertheless, we are nearing this point.

**Thesis outline**

This thesis is conceptually divided into two sections. The first focuses on the within-host biology of malaria infections and the second scales up to consider between-host processes for studying the evolutionary biology of both malaria and other infectious diseases more broadly.

In Chapter 2, I review previous mathematical models of the within-host dynamics of malaria infections. Over the past few decades, mathematical models have provided valuable knowledge of the mechanisms involved in malaria pathogenesis. These models have become increasingly sophisticated: they continue to be tied more closely to data and experimentation, and incorporate rigorous model selection and validation techniques. The important role that quantitative approaches to studying malaria will have in the future is also emphasized in this review.

In Chapter 3, I use mathematical models to understand the within-host dynamics of malaria infections in an experimental system (*Plasmodium chabaudi* in mice). Some of the complexity of this system is removed by using data from an experiment in which host immunity was manipulated. Through the processes of model development and selection,
this work makes substantial progress towards deciphering the relative importance of ecological factors in determining malaria infection dynamics, how they differ among parasite clones and might contribute to different levels of virulence.

In Chapter 4, I test some of the predictions that come out of the theory in Chapter 3 with further experiments. In particular, I aim to measure two key host exploitation traits of malaria parasites, which the model predicted varied between clones and could underlie their differential virulence. More broadly, this experimental work investigates the differences in host exploitation strategies between clones and in response to ‘ecological’ factors like host immunity and anemia.

In Chapter 5, I present a review of the growing body of infectious disease theory that attempts to link within- and between-host processes in a unified model framework (i.e. nested models). This review espouses the view that the greatest utility of nested models is that they provide a useful framework for studying all aspects of disease life history evolution, moving away from what has been a narrow focus on the evolution of virulence (itself, often narrowly defined). This more holistic approach will help to make realistic predictions about other measures of parasite-induced harm, to study the evolution of different exploitation strategies, and to gain a better understanding of why there is such diversity in parasite life histories.

In Chapter 6, I focus again on malaria, and use between-host and nested models of malaria dynamics to answer an outstanding question about malaria biology that has confounded evolutionary biologists for years: why do malaria parasites invest very little in
the production of transmissible forms and instead produce vast numbers of host-damaging asexual forms? From an evolutionary perspective, the generation of new infections is paramount and it seems that a parasite would benefit from increasing the rate of conversion to transmissible forms. This apparent reproductive restraint is an evolutionary puzzle, especially given its absence in related taxa. Here, I present work that is the first to analytically address the plausibility of various hypotheses for the evolution of reproductive restraint in malaria. Evolutionary invasion analyses are used to determine the conditions under which low conversion rates are selectively advantageous.

This thesis makes contributions to a number of distinct fields. It contributes to the understanding of the basic biology of malaria parasites, it motivates future empirical work, it challenges the rigor of common methodologies in theoretical biology, and presents some advances on these. Finally, it demonstrates the value of mathematical modeling over a range of biological questions, encompassing different biological scales, and employing a number of theoretical approaches.
Chapter 2. MODELING MALARIA PATHOGENESIS

ABSTRACT

Almost 20 years after the development of models of malaria pathogenesis began, we are beyond the ‘proof-of-concept’ phase and these models are no longer abstract mathematical exercises. They have refined our knowledge of within-host processes, and have brought insights that could not easily have been obtained from experimentation alone. There is much potential that remains to be realized, however, both in terms of informing the design of interventions and health policy, and in terms of addressing lingering questions about the basic biology of malaria. Recent research has begun to iterate theory and data in a much more comprehensive way, and the use of statistical techniques for model fitting and comparison offers a promising approach for providing a quantitative understanding of the pathogenesis of such a complex disease.

2.1 Introduction

Identifying factors that are involved in pathogenesis (defined to be the within-host mechanisms through which a parasite causes disease) is important, but it is really only the first step towards a complete understanding of infectious diseases. In the words of Ronald Ross: “To say that a disease depends upon certain factors is not to say much, until we can also form an estimate as to how largely each factor influences the whole result” (cited by McKenzie, 2000). Ross was driven by this philosophy to find accurate mathematical models of malaria transmission and his efforts led to new insights into the biology of the disease and strategies for control. In particular, his quantitative analysis demonstrated that mosquito populations need not be eradicated, but rather need only be driven below a particular threshold in order to eradicate malaria. Subsequent implementation of malaria control measures validated these predictions (McKenzie, 2000) and, since then, numerous epidemiological models have been proposed and directed towards understanding various processes in malaria transmission.

The interesting and important questions for the mathematical study of malaria are not, however, exclusively epidemiological, but span several levels of biological organization. Despite a large body of research on malaria pathogenesis, the relative significance of different factors in the development of disease is still debated (Miller et al., 2002). Research has focused mainly on two broad categories of factors: those that are resource-mediated (e.g. availability of red blood cells (RBCs), in which malaria parasites undergo asexual replication; Figure 2.1a), and those that are immune-mediated (Figure 2.1b). That we
can categorize factors like this, however, in no way indicates that they are understood. Indeed, our resolution of the relative importance of these factors and of any interactions between them remains quite inadequate. For example, it is known that several immune components, like macrophages and natural killer cells, are likely involved in the innate immune response to malaria, but the interactions of these factors with acquired responses remain speculative, and conflicting experimental data leaves the role of many components in question (Stevenson & Riley, 2004). At best, we know that some factors are necessary for a particular outcome (e.g. interferon-\(\gamma\) and natural killer cells limit peak parasitemia, Stevenson & Riley, 2004), but the relative contribution of each factor to this outcome is still unclear.

In this article we follow Ross’ philosophy, and suggest that, only when we can quantitatively predict the pattern of pathogenesis as a function of the underlying within-host regulatory factors can we legitimately claim to understand the processes at work. Although some might argue that the complexity of malaria biology puts this criterion beyond reach, we contend that such an objection is simply an acknowledgement of how little is known. As such, the development and testing of mathematical models of within-host processes is necessarily a vital component of research on malaria pathogenesis.

The value of having a sound, predictive, mathematical model of malaria pathogenesis cannot be overstated. Such models would allow in silico experiments of drug treatments and other interventions, thereby focusing efforts on targets that are likely to have the greatest impact. They would also allow for the evaluation of the too-often overlooked
Figure 2.1: Schematic of two recent models of malaria pathogenesis. (a) Modified from Mideo et al. (2008b, Chapter 3), this model tracks the densities of red blood cells (RBCs), merozoites and gametocytes. The main regulatory mechanism here is resource (i.e., RBC) abundance. (b) The model of Dietz et al. (2006) focuses on the effects of innate and acquired immune responses and tracks the density of infected RBCs only. The abundances and actions of different immune effectors are translated into probabilities of infected RBCs surviving their attack.
evolutionary consequences of treatments before they are made into policy. “No one dies of theoretical infections” (McKenzie, 2000), which are also cheaper and quicker than animal experiments. Equally important, however, is the value of the whole process of developing such mathematical models. By developing models of pathogenesis, we are forced to confess our ignorance of many of the biological details of infection, and this alone can be useful for highlighting important areas for future empirical research. Mathematical models also remove all ambiguity in potential explanations for patterns of pathogenesis (there is no where to hide ignorance), and they clearly delineate the logical conclusions that stem from various hypotheses. If, for example, experimental inhibition of erythropoiesis suppresses parasite recrudescence during an infection, one might be led to suggest that RBC availability determines such recrudescence. But unless we can accurately predict such experimental changes in pathogenesis using a mechanistic model, our understanding of the processes involved is still incomplete. A mutual feedback between model development/testing and empirical (ideally experimental) research is necessary to develop this level of understanding.

The same mathematical tools that have become (almost) ubiquitous in studies of malaria epidemiology have the capacity to be equally informative when applied to questions of malaria pathogenesis. Yet, modeling the within-host dynamics of malaria is a comparatively new practice, beginning just 20 years ago (e.g. Anderson et al., 1989). But unlike population dynamical models, which can rarely be fitted to truly replicated populations, within-host models can be very stringently challenged with data from numerous hosts. For
the brave, this should make for much more rapid progress than is possible in epidemiology alone. Indeed, it may be no accident that the advent and expansion of in-host malaria models coincided with the explosive interest in mathematical modeling of within-host dynamics of HIV, where interactions between theorists and empiricist yielded considerable insights (see examples in Perelson & Nelson, 1999).

### 2.2 Modeling approaches to malaria pathogenesis

The types of mathematical models of malaria pathogenesis we discuss in this review are based on a mechanistic description of the underlying biology of the system. This contrasts with purely statistical (i.e., curve-fitting models), although our exclusion of this class of models is not a value judgement; such statistical models have also provided important insights (see Paul et al., 2007, for a recent example). Also, note that we do not restrict our discussion to models of human malaria. Much of the current theory is aimed at explaining experimental malaria infections in model organisms and, in particular, mice from which there is a wealth of data, and therefore we include these models as well.

A basic model of the dynamics of malaria infection would track the changes in density (e.g. number per microlitre) of the kinds of cells and/or molecules thought to be important in pathogenesis. For example, it might track the density of asexual and sexual parasite forms, red blood cells, and various types of immune effectors. As factors are identified as being potentially important in regulating these variables, they are translated into a mathematical formulation that can be incorporated into the basic model structure. The ultimate aim when building models of malaria pathogenesis is to simplify the highly complex bi-
ological processes occurring during an infection into a comprehensible mathematical system from which inferences and predictions can be drawn. Critical in this process is the recognition that not all details of the biological system are relevant for understanding and predicting pathogenesis. Models need not (in fact should not) incorporate everything that we know about the biological system: we seek to understand the important components of reality, not to replicate the reality we do not understand. Indeed, the true power of a good model lies in its ability to expose the central agents responsible for the biological patterns under investigation by dispensing with the irrelevant details. Models help us to determine what is irrelevant.

To illustrate what such models typically look like and how, through the process of model development, the irrelevant biological details are uncovered, we present a simplified generic example based on Mideo et al. (2008b, Chapter 3). The model is in discrete time to account for the distinctly discrete life cycle of malaria parasites, and the densities are evaluated every day, corresponding to the 24-hour cycle of the rodent malaria system on which this model was based. The model predicts how the density of three quantities, free-living parasites (merozoites), $M$, transmission-stage parasites (gametocytes), $G$, and red blood cells, $R$, change from one day, $t$, to the next:

$$M_{t+1} = f(M_t, R_t),$$

$$G_{t+1} = h(M_t, R_t, G_t),$$

$$R_{t+1} = k(M_t, R_t).$$

The above equations capture the idea that the density of each quantity in the next day
(time $t + 1$) is some function of their densities on the present day (time $t$). Notice that two of these functions do not depend on the gametocyte density, $G_t$, reflecting an assumption that gametocytes play no role in determining the merozoite or RBC counts on the next day. Other assumptions about how various biological processes work (e.g. erythropoiesis, RBC infection, gametocytogenesis, etc.) are captured by the specific forms of the functions $f(M,R)$, $h(M,R,G)$, and $k(M,R)$ (see Table 2.1).

Models like that above can be further refined as necessary, by including things such as RBC age structure and time-lags in erythropoiesis, or can be extended to include regulatory factors related to innate and adaptive immune responses. For example, one might introduce variables that represent the densities of different immune effectors. If we use $T_t$ to denote the density of specific T cells on day $t$, then the model might be extended as

\[
M_{t+1} = f(M_t, R_t, T_t), \\
G_{t+1} = h(M_t, R_t, G_t, T_t), \\
R_{t+1} = k(M_t, R_t, T_t), \\
T_{t+1} = l(M_t, R_t, T_t),
\]

where the functions $f(M,R,T)$, $h(M,R,G,T)$, $k(M,R,T)$ and $l(M,R,T)$ are specified to account for the relevant assumptions about how these processes work (Table 2.1). The predictions of each model obtained by employing a different set of assumptions can then be tested against data to determine its ability to explain known patterns of pathogenesis (Mideo et al., 2008b, Chapter 3).
<table>
<thead>
<tr>
<th>Equations/ variations</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Merozoite density</strong></td>
<td></td>
</tr>
<tr>
<td>$M_{t+1} = \omega a(M_t) R_t (1 - \epsilon)$</td>
<td></td>
</tr>
<tr>
<td>- A proportion of susceptible RBCs becomes infected; this proportion is described by a function, $a(M_t)$, which depends on the density of merozoites. Asexual replication occurs in a proportion $1 - \epsilon$ of all infected RBCs, each producing $\omega$ progeny merozoites.</td>
<td></td>
</tr>
<tr>
<td>$M_{t+1} = \omega a(M_t) R_t (1 - \epsilon) e^{-\lambda T_t}$</td>
<td></td>
</tr>
<tr>
<td>- As above but with an immune response as well. The probability of an infected RBC surviving immune attack is given by $e^{-\lambda T_t}$, and decreases with increasing immune cell density. $\lambda$, describes the susceptibility of infected RBCs to immune attack.</td>
<td></td>
</tr>
<tr>
<td><strong>Red blood cell density</strong></td>
<td></td>
</tr>
<tr>
<td>$R_{t+1} = \theta + R_t - a(M_t) R_t$</td>
<td></td>
</tr>
<tr>
<td>- A constant number, $\theta$, of RBCs (per $\mu l$) are produced daily. There is no natural death of RBCs, but RBCs are lost through infection.</td>
<td></td>
</tr>
<tr>
<td>$R_{t+1} = b(R_t) + R_t - a(M_t) R_t$</td>
<td></td>
</tr>
<tr>
<td>- Here, RBC production, $b$, is an arbitrary function of current RBC density.</td>
<td></td>
</tr>
<tr>
<td>$R_{t+1} = b(R_{t-\tau}) + R_t - a(M_t) R_t$</td>
<td></td>
</tr>
<tr>
<td>- As above, but now RBC production is time-lagged to account for the maturation time of RBC precursors (production is a function of RBC density $\tau$ days earlier).</td>
<td></td>
</tr>
<tr>
<td><strong>Gametocyte density</strong></td>
<td></td>
</tr>
<tr>
<td>$G_{t+1} = G_t + \epsilon a(M_t) R_t$</td>
<td></td>
</tr>
<tr>
<td>- A proportion, $\epsilon$, of all infected RBCs produce gametocytes.</td>
<td></td>
</tr>
<tr>
<td>$G_{t+1} = G_t + \epsilon a(M_t) R_t - \delta G_t$</td>
<td></td>
</tr>
<tr>
<td>- As above, but a proportion, $\delta$, of gametocytes die each day as well.</td>
<td></td>
</tr>
<tr>
<td>$G_{t+1} = G_t + \epsilon a(M_{t-\tau}) R_{t-\tau}$</td>
<td></td>
</tr>
<tr>
<td>- As in the first model, but gametocytes are sequestered for $\tau$ days before maturing and being released into the bloodstream.</td>
<td></td>
</tr>
<tr>
<td><strong>Immune cell density</strong></td>
<td></td>
</tr>
<tr>
<td>$T_{t+1} = T_t + \sigma T_t$</td>
<td></td>
</tr>
<tr>
<td>- Immune cell densities increase exponentially. Each immune cell activates $\sigma$ others. Immune cell activation does not require contact with infected RBCs or merozoites.</td>
<td></td>
</tr>
<tr>
<td>$T_{t+1} = T_t + \gamma a(M_t) R_t$</td>
<td></td>
</tr>
<tr>
<td>- Production of immune cells is proportional to infected RBC density.</td>
<td></td>
</tr>
</tbody>
</table>
Each model, with its associated assumptions about how the important regulatory factors come into play, represents a different biological hypothesis about what determines malaria pathogenesis (Johnson & Omland, 2004). It might well turn out that a given model does not accurately reflect the biology of the system, but if so, something new about the biology and the importance of various regulatory factors would thereby have been learned. The clarity that comes from having made such unambiguous assumptions often will also point towards new empirical questions that need to be answered.

Consider again, the recrudescences in parasite density that occur during some malaria infections. One explanation for these peaks is that they are the result of antigen switching and immune escape by the parasite (Brown & Brown, 1965; Phillips et al., 1997). That secondary peaks are antigenically distinct does not, however, mean that the peaks exist because they are antigenically distinct, and an alternative explanation is that the peaks simply track resource availability – parasite densities increase as red blood cell densities rebound from the destruction wrought by the initial wave of parasites. An experimental approach to testing these hypotheses might alter an immune component of a model organism, generate malaria infections, and then see if these peaks still occur. Deciding what immune component to alter, when to alter it, and by how much, however, makes this line of investigation nontrivial.

A modeling approach to this question can provide two valuable kinds of information. First, as is often done in physics, mathematical simulations of experimental procedures can help to inform researchers which manipulations, of the vast array of possibilities, are
likely to be the most informative. As with the case of RBC availability mentioned earlier, ultimately we cannot claim to understand the processes that are occurring unless we can accurately predict the outcome of such experimental manipulations.

Second, mathematical models can provide inferential power. By building models that include or exclude antigen switching mechanisms and immune responses, we can determine the conditions under which secondary parasite peaks are predicted to occur and if these regulatory factors are relevant for this particular pattern. In fact, models have revealed that both resource availability (e.g. Hetzel & Anderson, 1996; Mideo et al., 2008b, Chapter 3) and certain forms of antigen switching (e.g. Paget-McNicol et al., 2002) are plausible explanations for parasite recrudescences. A more thorough modeling approach that incorporates each of these processes, in isolation and in combination, will then help to resolve this issue, and inform future experiments. Later, we discuss how statistical model selection techniques are increasingly being used to compete multiple models and test alternative hypotheses.

2.3 Insights

Mathematical models of malaria pathogenesis have helped develop our knowledge in many ways. Due to an increasing interest in this approach, we cannot cite all important findings, and instead we therefore focus on a few developments that reflect the range of applications of these models.
2.3.1 Effective immune targets

One area that has received a lot of attention is host immune responses to malaria. Despite our incomplete knowledge of these processes (as described above), models have helped elucidate some of their general characteristics. Models of malaria have repeatedly demonstrated that immune responses are more effective if directed towards infected RBCs rather than free-living merozoites (Anderson et al., 1989; Hetzel & Anderson, 1996; Haydon et al., 2003). This is perhaps not surprising given the short lifespan of merozoites in the bloodstream (on the order of minutes), but tackling this question theoretically allows for the quantification of this difference in efficacy (Haydon et al., 2003). These results have obvious implications for vaccine design.

2.3.2 Protective immune control

Also, as we would expect, innate immune responses are predicted to be most important during initial parasite peaks and progressively less important throughout the course of infection (Molineaux et al., 2001). Here, Molineaux and colleagues have actually quantified the relative importance of three types of immune responses (innate, acquired variant-specific, and acquired non-variant-specific) over successive parasite peaks, finding, for example, that innate immune responses are almost entirely responsible for controlling the primary peak but are completely absent later on in infections.

The approach of Molineaux and colleagues is also unique in that they estimate model parameters for several individual host datasets, allowing variation in parasite traits be-
tween variants and between hosts. This results in some novel and testable predictions. In particular, to recover differences in patterns of infection among hosts, the rank order of the parasite variants’ baseline multiplication factors differed between hosts (i.e. the fastest growing variant in one host was not the fastest growing in all hosts, Molineaux et al., 2001). While the authors suggest this might not be a biologically justified conclusion, such a variant by host interactive effect on parasite growth rate could be tested experimentally with a model system.

### 2.3.3 Clonal interactions

Immune responses will also have effects on the competition between parasite clones within a host. Hellriegel (1992) showed that competitive suppression of a superior clone, via a common immune response, was predicted to occur when an inferior clone arrived first (i.e. several days earlier). This prediction has since been tested and validated with experimental infections in mice (de Roode et al., 2005a). The paradoxically low numbers of transmissible parasite forms in malaria infections (Taylor & Read, 1997) can also be explained by competition mediated by a shared immune response (McKenzie & Bossert, 1998). Instead of producing multiple daughter merozoites (each with the potential to infect another RBC), a small fraction of infected RBCs produce transmissible gametocytes. A common immune response that was elicited by, and targeted, merozoites would favour parasite clones that could quickly build up a ‘stock’ of merozoites during competition, i.e., those with low levels of conversion to gametocytes (McKenzie & Bossert, 1998). Compe-
tition for access to red blood cells may also be sufficient for generating selection for low levels of conversion to gametocytes (Mideo & Day, 2008, Chapter 6). In the timeframe of a single infection, there is experimental evidence that parasites alter gametocyte sex ratios in response to coinfection (Reece et al., 2008), and alter rates of conversion to gametocytes in response to drug treatment (Buckling et al., 1999); however, adjustment of conversion rates in response to coinfection has yet to be unequivocally demonstrated (Wargo et al., 2007a). Further investigation of a potentially plastic response is warranted, as is a comparison of levels of conversion in *Plasmodium* species that have been exposed to different amounts of coinfection over an evolutionary timeframe (Mideo & Day, 2008, Chapter 6).

### 2.3.4 Gametocytogenesis

Models of human malaria, fitted to data from malarial therapy patients, have shown that the level of conversion to gametocytes changes significantly during the course of infection (Diebner et al., 2000; Eichner et al., 2001). However, the pattern of the shift in conversion rates remains unexplained. Several factors have been identified as likely influencing gametocyte production (Dyer & Day, 2000; Talman et al., 2004), and incorporating these into models could help in understanding the dynamics of gametocytogenesis.

### 2.3.5 Consequences of RBC preference

In humans, parasite preference for certain ages of RBCs seems to relate to disease severity. The most severe disease is caused by *P. falciparum*, which infects RBCs of all ages. In contrast, less deadly species show preferences for infecting either younger (reticulocytes)
or older RBCs (Paul et al., 2003). These cell preferences have been incorporated into models of human malaria, where they have been shown to have a significant effect on infection dynamics, and also help to explain the differences in clinical observations between species (McQueen & McKenzie, 2004). Recent models have explored whether this sort of cellular tropism is an important determinant of the dynamics of different rodent malarias. For example, a strong preference of *P. berghei* for reticulocytes can explain prolonged low levels of circulating reticulocytes that tend to be thought of as the result of suppressed RBC production (Cromer et al., 2006).

In contrast, *P. chabaudi* is generally thought to indiscriminately infect RBCs of all ages, but parameter estimates from a recent model show that parasites invade older RBCs at a rate that is an order of magnitude higher than for reticulocytes (Mideo et al., 2008b, Chapter 3). Despite this, parasites were predicted to do better in reticulocytes, producing more daughter merozoites per infected cell. This effect, however, appears to be clone-specific which suggests it might be a mechanism that explains different levels of virulence between clones. Another model allows for *P. chabaudi* parasites to further discriminate between ages of RBCs and predicts that the size of the age range a particular clone can infect is correlated with its virulence (Antia et al., 2008). Each of these predictions can (and should) be tested empirically (see Chapter 4).
2.3.6 *Manipulation of erythropoiesis*

The idea that erythropoiesis is suppressed during malaria infection has been refuted by models of both human (Jakeman et al., 1999) and rodent (Mideo et al., 2008b, Chapter 3) malaria. In fact, during periods of disease-induced anemia, these models predict that hosts increase RBC production well above baseline rates. Intriguingly, models fitted to data from experimental *P. chabaudi* infections suggest that the host response to anemia varies depending on the genotype of the infecting parasites (Haydon et al., 2003; Mideo et al., 2008b, Chapter 3). Whether this effect can also help explain different levels of virulence between parasite clones remains unclear, as in one case the more virulent clone induced faster rates of RBC production (Haydon et al., 2003) and in the other, the opposite was true (Mideo et al., 2008b, Chapter 3). Unlike in Haydon et al. (2003), the data in Mideo et al. (2008b) came from CD4$^+$ T-cell depleted mice, indicating a possible effect of an interaction between parasite genotype and host immune status on RBC production. Regardless of the relationship between virulence and erythropoiesis, the prediction of a clone-specific effect on this host trait could not likely be made from looking at data alone.

2.4 *Theoretical developments*

In this section, we discuss criticisms of previous models of malaria pathogenesis and the improvements both realized in recent work and hoped for in the future. It should be noted that the issues we highlight are not malaria-specific; rather, they really represent challenges to all modelers of disease dynamics.
2.4.1 Biological realism

The best level of mathematical abstraction for any biological process should be determined by the nature of the question being addressed. For example, if one wants to identify mechanisms that can plausibly explain dynamics, schematic models are often sufficient. Many of the insights discussed above come from model approaches of this type. However, many of these same models have been criticized for their lack of realism (Molineaux & Dietz, 1999) and if one wants to use a model to make quantitative predictions about the dynamics of malaria pathogenesis, then a key goal is realistically capturing as much of the biology of the system as is necessary to explain data. In particular, early models of malaria pathogenesis allowed for no individual variation between hosts or parasites, and failed to account for the distinctly discrete life cycle of malaria parasites (Molineaux & Dietz, 1999). The field is maturing and recent work has addressed these concerns (e.g. Molineaux et al., 2001; Dietz et al., 2006; Mideo et al., 2008b, Chapter 3), hopefully broadening the appeal (and thus the impact) of theoretical approaches for studying malaria pathogenesis.

Clearly there is still some way to go regarding the mathematical capture of immunological processes. A major future challenge is to move beyond the mathematical caricatures of immune control (e.g. arbitrary functions, the grouping of distinct populations of immune effectors into ‘immune cells’) towards mathematical descriptions that more closely accord with the qualitative picture emerging from experimental immunology (e.g. Stevenson & Riley, 2004). Indeed, we expect mathematical models to play a critical role in identifying the key processes involved, and determining their relative importance.
2.4.2 Tying theory to data

Another significant criticism of many models is that, while they often include some indication of the qualitative agreement between model predictions and experimental or clinical data, they tend to lack an investigation into whether alternative models could do as good a job, and whether the model predictions provide a statistically good fit for data (Molineaux & Dietz, 1999). More rigorous approaches to model design, fitting, and selection can help to resolve conflicts between current results. As a starting point, many theoretical studies consider multiple biological hypotheses for what regulates malaria pathogenesis. Ideally, these should be translated into mathematical descriptions (i.e., models) of the hypotheses, and the best model should be chosen from this pool of potential descriptions. For example, a resource-based model may capture the details schematized in Figure 2.1a; a competing model, the immune regulation of Figure 2.1b. Statistical procedures exist to compare the fit of each model’s predictions to data using one of a number of different information-theoretic techniques (Hilborn & Mangel, 1997; Burnham & Anderson, 2002; Johnson & Omland, 2004); we give an illustrative example in Figure 2.2. Once a best model is chosen, determining whether it provides an acceptable quantitative description of observed dynamics requires further assessment. Such “goodness-of-fit” analyses remain relatively uncommon in mathematical studies of malaria, but recent research has begun to take this approach (Molineaux et al., 2001; Dietz et al., 2006; Mideo et al., 2008b, Chapter 3).

Finally, models ought to be validated by altering some model component, generating new predictions, and then empirically testing these with independent experiments. The
Figure 2.2: Model selection and validation. Data from a single CD4$^+$ T cell depleted mouse (dashed lines and dots) and model predictions (solid lines). Top panels, RBC densities; bottom panels, parasite densities. Predictions are from four models representing different hypotheses about what regulates the dynamics of pathogenesis: i. no RBC age structure or parasite cell age preference and constant erythropoietic response; ii. no RBC age structure or parasite cell age preference and variable erythropoietic response; iii. RBC age structure, possible parasite cell age preference and constant erythropoietic response; iv. RBC age structure, possible parasite cell age preference and variable erythropoietic response. Models iii and iv provide statistically significantly better fits to the RBC data than models i and ii. As the models were fit only to the RBC data, the parasite data provide a means of model validation. It is clear that model iv is better than iii at qualitatively capturing the parasite dynamics. Model iv is selected as the best model among those tested. See Mideo et al. (2008b, Chapter 3).

Experimental manipulations required to do this are only ethically feasible in model organisms. While there is still debate about the relevance of these experiments for human malaria (see Dietz et al., 2006), the fact that we can use them to test and refine theory highlights the utility of using model malaria systems. Further, if we cannot design theory that accurately captures the dynamics of malaria pathogenesis in a highly controlled setting (like clonal infections in inbred laboratory mice), there is no hope for a full understanding of more heterogeneous human malaria infections.
2.5 Future directions

An important goal of building mathematical models of malaria pathogenesis is to use them to evaluate interventions. By translating the mode of action of an intervention into changes in particular model parameters, the effects on disease progression within a host can be predicted. Recently, Dietz et al. (2006) took this approach to explore the effects of vaccination by systematically altering several model parameters corresponding to different (plausible) modes of vaccine action. They determined that the outcome of infection after vaccination is strongly host-dependent, and that certain types of vaccines are better at protecting against severe versus mild malaria. These results suggest that current methods of evaluating vaccines may be inappropriate (Dietz et al., 2006).

Changes to within-host parameters due to interventions will have influences on processes at higher levels of biological organization. Simply put, epidemiological processes like transmission are determined by parasite densities within hosts, which are regulated by all the factors we’ve discussed above and, likely, many more. Interventions like drug treatment and vaccination will alter within-host regulatory factors, with effects that will scale up from the within- to the between-host level, influencing, for example, disease prevalence.

One way to account for these interactions is by using a theoretical framework that nests levels of biological organization (Mideo et al., 2008a, Chapter 5). A few recent models have taken this approach and combined models of the within-host and transmission dynamics of malaria (McKenzie & Bossert, 2005; Mideo & Day, 2008, Chapter 6). The
combined model of McKenzie & Bossert (2005) leads to some novel predictions about how within-host details affect host population level processes. For example, they predict that increased antigenic diversity in a parasite population can lead to increased persistence of individual parasite genotypes. This is explained by the action of innate immune responses during superinfections, which allow recovery from secondary infection by a particular genotype before acquired responses (specific to that genotype) can build up, maintaining a pool of hosts susceptible to that genotype. Another finding of McKenzie & Bossert (2005) is that the number of hosts infected in their model is insensitive to changes in within-host gametocyte survivorship, since only below a certain threshold density does the actual number of gametocytes (versus presence) seem to have an effect on transmission success. Given these results, any intervention targeting this parasite stage “would need to be astonishingly effective” (McKenzie & Bossert, 2005) to be of any consequence.

Although the aim of McKenzie & Bossert (2005) was not to evaluate interventions, nested models can be directed at this purpose since we know that changes to within-host parameters (as a result of, say, vaccine action) will affect higher level processes. For example, nested models have demonstrated the potential for unintended, negative consequences in response to vaccination, including selection for increased virulence, depending on the specific vaccine target (e.g. Gandon et al., 2001b; Ganusov & Antia, 2006) or vaccine coverage (e.g. André & Gandon, 2006). However, these are highly generalized models with non-disease-specific mathematical descriptions of within-host processes. As we’ve argued above, if our aim is to make quantitative predictions about outcomes for a specific disease,
we ought to use strongly supported, biologically based models. Improving within-host models of malaria pathogenesis and combining these with epidemiological models will lead to better predictions about the effects of interventions.

Some of the potential evolutionary effects of interventions have been studied with model organisms and while their results are important, they are not always well understood. Why, for example, does passaging malaria parasites through immunized mice result in selection for more virulent parasites (Mackinnon & Read, 2004)? In particular, what is the mechanism of this increased virulence, i.e., on what trait is selection acting? These empirical results may have serious implications for human health policy, yet these questions remain unanswered. With a good model of mouse malaria (one that has been derived from, calibrated and validated with data) we can replicate this experiment \textit{in silico} with the aim of predicting what kinds of malaria parasites (e.g. those that undergo rapid antigen switching, replicate at higher rates or infect RBCs at faster rates) have an advantage in immunized hosts. Theory offers an easier and powerful approach for teasing apart mechanisms.

\section*{2.6 Conclusions}

The fundamental goal of any study of malaria pathogenesis is to bring new insights towards developing successful treatment and control measures for this disease. Given the lack of progress in the past, our best hope for tackling the problem of malaria is through a more comprehensive understanding of the mechanisms that determine its pathogenesis. Unless these processes can be translated into mathematical models that accurately capture
the dynamics of pathogenesis, this understanding will remain out of reach. In the process of striving for a mathematical account of malaria pathogenesis, we will likely discover the existence of new regulatory factors and that some regulatory factors are largely unimportant. By determining the relative importance of what does matter, and how those factors interact, we will be able to predict the likely consequences of clinical interventions, such as vaccines and chemotherapeutic agents targeted at particular parasite stages, and novel interventions aimed at host factors which determine disease (immunopathology). Achieving this promise requires the careful interactions between experimental biologists who appreciate that useful models need not include every last detail of every pathway, and biomathematicians who are prepared to tackle the jargon, the huge experimental literature, and the fuzzy uncertainties of real experimental data.
Chapter 3. UNDERSTANDING AND PREDICTING STRAIN-SPECIFIC PATTERNS OF PATHOGENESIS IN THE RODENT MALARIA, *Plasmodium chabaudi* 

**ABSTRACT**

Scientists have had considerable success elucidating important immunological and resource-based mechanisms that control the dynamics of infection in some diseases, yet little is known about how differences in these mechanisms results in strain differences in patterns of pathogenesis. Using a combination of data and theory we disentangle the role of ecological factors (e.g. resource abundance) in the dynamics of pathogenesis for the malaria species *Plasmodium chabaudi*, in CD4$^+$ T cell depleted mice. We build a series of nested models to systematically test a number of potential regulatory mechanisms and determine the ‘best’ model using statistical techniques. The best-fit model is further tested using an independent data set from mixed-clone competition experiments. Results reveal that parasites preferentially invade older red blood cells even when more fecund in younger reticulocytes, and that inoculum size has a strong effect on burst size in reticulocytes. Importantly, the results suggest that strain-specific differences in virulence arise from differences in red-blood-cell-age-specific invasion rates and burst sizes, since these

are lower for the less virulent strain, as well as differences in levels of erythropoiesis induced by each strain. These results highlight the importance of model selection and validation for revealing new biological insights.

3.1 Introduction

Malaria is one of the leading causes of death among infectious diseases in the world, killing over one million people every year (WHO, 2008b). Despite this enormous burden and the large proportion of the world’s population threatened by malaria, we know little about the relative importance of mechanisms that regulate malaria parasite growth within infected hosts, or how variation in these mechanisms gives rise to different patterns of morbidity and mortality between strains. A better understanding of these mechanisms is of interest in its own right, and it will likely yield important insights for improving control strategies, as well as for understanding the potential evolutionary consequences of these controls.

The rodent model system of malaria, *Plasmodium chabaudi*, provides an excellent opportunity to address this question because controlled and replicated experimental manipulations are possible. In this chapter we combine data from experimental manipulations with the development and testing of mathematical models. Our overriding goal is to provide a quantitative (mathematical) description of the main mechanisms governing the regulation of *P. chabaudi* growth within hosts, and how these differ among strains. If a mathematical description cannot be made for mouse models of malaria, where host and parasite genotype together with a vast number of environmental variables can be carefully controlled, then there is little hope for a quantitative understanding of human infections.
Many factors regulating malaria infection dynamics have been identified. Among these, immune responses play a large role in the control of parasite densities and may take many forms with respect to the trigger and target of the response (i.e., free-living merozoites or infected red blood cells), the level of clone-specificity of the response, and the significance of antigenic variation (for a review, see Stevenson & Riley, 2004). Other, ‘non-immunological’ mechanisms, however, are also likely to play a fundamental role. For example, simple resource abundance, i.e., the availability of red blood cells (RBCs), is likely a key regulator of parasite growth (Hellriegel, 1992; Hetzel & Anderson, 1996; Haydon et al., 2003) since it is primarily within these cells that asexual replication occurs. Also, the age-structure of available RBCs could be important (McQueen & McKenzie, 2004; Cromer et al., 2006) since many malaria species preferentially invade either young RBCs (reticulocytes) or fully mature normocytes (Paul et al., 2003). Further, the nature of the host’s erythropoietic response to malaria-induced anemia will affect both the abundance and age-structure of RBCs (Antia et al., 2008), as will temporal patterns of conversion to gametocytes during an infection (e.g. Eichner et al., 2001). While some previous theoretical studies have explored the role of these factors in determining within-host dynamics of malaria, our approach is unique in that we compare multiple model variations that systematically test different combinations of these regulatory mechanisms. Furthermore, we allow for individual variation by fitting models to data sets from individual hosts, and we then perform a rigorous statistical analysis of the fit to experimental data.

To dissect the relative importance of these factors, and how they differ among strains,
a combination of experimental manipulation and mathematical modeling is required. Experimental manipulations can be used to remove some potential regulatory mechanisms so that we can examine the dynamics of parasite growth when only a few factors are in operation. In this simplified setting, one can then build mathematical models to describe these dynamics. Rigorous tests of these models can then be conducted in this simplified setting, through a combination of statistical analyses and further experimental manipulations. Understanding, in a quantitative way, the strain-specific differences that occur in such ‘stripped down’ situations is an important step towards understanding the more complete picture in which all potential regulatory factors interact with one another.

The results presented here take this approach by focusing on *Plasmodium chabaudi* infections in mice with depleted CD4+ T cells. CD4+ T cells are essential for developing an effective immune response against malaria infections (Good & Doolan, 1999; Pombo et al., 2002). They play a crucial role both early in an infection by activating macrophages and initiating anti-parasitic cell-mediated immune responses, and in later stages of infection by helping B cells to produce antibody, and regulating the adaptive immune responses (Langhorne et al., 1990; Urban et al., 2005; Stephens & Langhorne, 2006). By removing this regulatory factor we can then examine the extent to which the above-mentioned ecological factors can explain strain-specific differences in infections.

### 3.2 Methods

Our approach and results have three main components. The first focuses on data from two clones of *P. chabaudi*, each singly infecting mice with depleted CD4+ T cells. We build
a suite of eight nested models of increasing complexity, representing eight potential descriptions of the within-host dynamics of infection. We then fit each of these eight models to data using maximum likelihood techniques. Second, we use model selection criteria to chose the ‘best’ model from the eight fitted models for each clone. After determining, statistically, that the best model does in fact provide a good fit, we then use this model to infer the main mechanisms underlying the clone-specific differences in infection dynamics. Third, we take our ‘best’ model for each clone and use them to derive predictions for the expected dynamics when both clones coinfect a single host. These predictions are then compared with experimental data on coinfection as a further, independent, means of validating the model.

3.2.1 Model development and data fitting

Previous experimental data

Experimental infections were generated using two genetically distinct Plasmodium chabaudi clones, denoted AS and DK, which were originally isolated from thicket rats, Thamnomys rutilans, in the Central African Republic (Beale et al., 1978). In mixed infections, DK is competitively suppressed by AS (Barclay et al., 2008) and previous studies have shown a strong relationship between competitive ability and virulence (de Roode et al., 2003, 2005a,b; Bell et al., 2006).

Fifteen inbred female C57BL/6JolaHsd mice aged 6-8 weeks (Harlan, UK) were depleted of their CD4+ T cells using a rat monoclonal antibody, GK1.5 (for details see Bar-
 Chapter 3. Strain-specific patterns of pathogenesis in malaria

clay et al., 2008). Mice were infected via intra-peritoneal injection. Five mice were inoculated with $10^5$ AS parasites, five with $10^6$ of this same clone and five with $10^6$ DK parasites. Mice were maintained as described previously (de Roode et al., 2004). Daily RBC densities were tracked using flow cytometry (Beckman Coulter) and parasite densities were measured using genotype-specific real-time quantitative PCR (qPCR). One mouse in each of the AS $10^5$ and DK $10^6$ treatments died prematurely and thus were excluded from our analysis.

**The model**

Our aim is to understand what factors determine the clone-specific differences in dynamics of infection in CD4$^+$ T cell depleted mice. We developed models that consider clone-specific effects on RBC age-structure and cellular tropism, gametocytogenesis, and erythropoiesis as possibilities. All of these factors were incorporated in a single, large model, and the suite of eight nested models that are the objects of our analysis are then obtained as certain special cases of this single model.

Malaria parasites in most infections display a distinctly discrete replication cycle during an infection, with synchronous bursting of infected RBCs occurring every 24 hours in *P. chabaudi* (Carter & Walliker, 1975). Our basic model therefore tracks the infection dynamics in discrete time, with one time step corresponding to a single day (Molineaux & Dietz, 1999). Immediately after bursting, merozoites begin infecting susceptible RBCs. This process occurs relatively quickly, with the majority of the 24 hours between burst-
ing events being made up of development within RBCs (Figure 3.1). For simplicity, in
the model, we census the populations of merozoites and RBCs immediately after burst-
ing but prior to the infection of new RBCs. The basic model thus tracks the following
events within a single day: i. Census, ii. RBC invasion by merozoites, iii. RBC turnover
(production and natural death), and iv. bursting of infected RBCs (Figure 3.1).

<table>
<thead>
<tr>
<th>Day $i$</th>
<th>i. Census</th>
<th>ii. Invasion</th>
<th>iii. RBC turnover</th>
<th>iv. Bursting</th>
</tr>
</thead>
<tbody>
<tr>
<td>merozoites</td>
<td>$P_{i}$</td>
<td>$R_{1,i}$</td>
<td>$I_{1}(\infty)$</td>
<td>$P_{i+1}$</td>
</tr>
<tr>
<td>RBCs</td>
<td>$R_{2,i}$</td>
<td>$I_{2}(\infty)$</td>
<td>$R_{1,i+1}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$R_{3,i}$</td>
<td>$I_{3}(\infty)$</td>
<td>$R_{2,i+1}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$N_{i}$</td>
<td>$I_{N}(\infty)$</td>
<td>$R_{3,i+1}$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.1: Schematic of asexual bloodstage of malaria infections and our model framework. It
takes 24 hours for *Plasmodium chabaudi* merozoites to invade RBCs, replicate and
burst from infected cells. We model this process in discrete time, tracking: the number
of merozoites, $P_{i}$, reticulocytes of age $j$, $R_{j,i}$, and normocytes, $N_{i}$, on day $i$. RBC
invasion occurs quickly relative to this one-day cycle. We model this invasion step in
continuous-time and calculate the number of infected reticulocytes of age $j$, $I_{j}(\infty)$, and
infected normocytes, $I_{N}(\infty)$, at the end of this process, when all merozoites are either
infecting RBCs or dead.
The discrete-time nature of the model requires that we define distinct age classes of RBCs for exploring the importance of RBC age-structure to the within-host dynamics. In rats, reticulocytes take 30 to 75 hours to mature once released into the bloodstream, taking longer when the rat is anemic (Ganzoni et al., 1969; Wiczling & Krzyzanski, 2007). This timeframe is in agreement with the 2-3 day maturation time observed in C57 black mice (S. Reece, personal communication, January 27, 2006). Since the experimental mice have depressed RBC densities for a substantial duration of the experiment, we assume the reticulocytes will, on average, take 3 days to mature. Thus, we define four age classes of RBCs: reticulocytes that have newly been introduced to the bloodstream, $R_{1,i}$; reticulocytes in their second day of being in the bloodstream, $R_{2,i}$; reticulocytes in their third day of being in the bloodstream, $R_{3,i}$; and fully mature normocytes, $N_i$ (the subscript $i$ denotes a value on day $i$ post-inoculation). In addition there are four classes of infected RBCs corresponding to each type of blood cell ($I_{1,i}$, $I_{2,i}$, $I_{3,i}$ and $I_{N,i}$). Merozoites may infect reticulocytes and normocytes at different per capita rates (reflecting a particular cell age ‘preference’) and may produce different numbers of progeny merozoites per infected cell upon bursting. In the absence of evidence to the contrary, merozoites are assumed to respond to all reticulocytes the same way, regardless of how many days the RBC has been in the bloodstream. RBC production varies according to the density of RBCs (as compared to an ‘equilibrium’ density in the absence of infection, $K$, which is calculated from data), increasing production when an individual becomes anemic, as is well documented (e.g. Mackey, 1997).
Chapter 3. Strain-specific patterns of pathogenesis in malaria

Incorporating all of these biological details in a single model gives the discrete-time dynamics system (see Appendix A for model derivation and Table 3.1 for parameter descriptions):

\[ P_{i+1} = \left( \omega_R \left( \sum_{j=1}^{3} R_{j,i} \left( 1 - e^{-\left(\frac{P_i \beta_R}{R_1,i + R_2,i + R_3,i}\right)} \right) \right) + \omega_N N_i \left( 1 - e^{-\left(\frac{P_i \beta_R}{R_1,i + R_2,i + R_3,i}\right)} \right) \right) \left(1 - d\right) \left(1 - g\right), \]

\[ R_{1,i+1} = \theta(K - (R_{1,i-\tau} + R_{2,i-\tau} + R_{3,i-\tau} + N_{i-\tau})), \]

\[ R_{2,i+1} = (1 - d) R_{1,i} e^{-\left(\frac{P_i \beta_R}{R_1,i + R_2,i + R_3,i}\right)}, \]

\[ R_{3,i+1} = (1 - d) R_{2,i} e^{-\left(\frac{P_i \beta_R}{R_1,i + R_2,i + R_3,i}\right)}, \]

\[ N_{i+1} = (1 - d) \left( R_{3,i} e^{-\left(\frac{P_i \beta_R}{R_1,i + R_2,i + R_3,i}\right)} + N_i e^{-\left(\frac{P_i \beta_N}{R_1,i + R_2,i + R_3,i}\right)} \right). \]

We fit eight variations of this model to the data to determine which factors are important for explaining the dynamics of disease. These eight variations are all special cases of the above general model. Briefly, when there is no age structure, all RBCs (i.e. \( R_1, R_2, R_3 \) and \( N \)) are collapsed into a single class and there is no heterogeneity in burst size or invasion rate; constant recovery of RBC deficit means erythropoiesis is a linear function of RBC density (\( \theta_0 = \theta_A \)), while variable recovery assumes it is a piecewise linear function (\( \theta_0 \neq \theta_A \)), accounting for saturation as described earlier; gametocyte production, governed by the parameter \( g \), is either assumed to not occur (\( g=0 \)) or to occur at some constant rate throughout infection.

The model variations, i.e., the different biological hypotheses and resultant parameter
Chapter 3. Strain-specific patterns of pathogenesis in malaria

constraints, are as follows (see Table C.1 for the numbers of fitted parameters):

1. No age structure (AS), constant recovery of RBC deficit (RBC recovery), no gametocyte production (GP); \( \beta_R = \beta_N, \omega_R = \omega_N, \theta_0 = \theta_A, g = 0 \).
2. No AS, constant RBC recovery, constant GP; \( \beta_R = \beta_N, \omega_R = \omega_N, \theta_0 = \theta_A \).
3. No AS, variable RBC recovery, no GP; \( \beta_R = \beta_N, \omega_R = \omega_N, g = 0 \).
4. No AS, variable RBC recovery, constant GP; \( \beta_R = \beta_N, \omega_R = \omega_n \).
5. AS, constant RBC recovery, no GP; \( \theta_0 = \theta_A, g = 0 \).
6. AS, constant RBC recovery, constant GP; \( \theta_0 = \theta_A \).
7. AS, variable RBC recovery, no GP; \( g = 0 \).
8. AS, variable RBC recovery, constant GP.

Curve fitting

Initial conditions for individual mice were determined from experimental measurements of both RBC and parasite densities two days post-inoculation. Parameter estimates were obtained from the literature and our fitting routine searched parameter space slightly larger than the biologically reasonable ranges (see Table 3.1). We fit all parameters except for the merozoite and RBC death rates, which were fixed for computational practicality, and because they are well known (Table 3.1).

The models were fit to the data from individual mice using the maximum-likelihood method. Our approach was to lay a coarse grid over the entire parameter space and for each parameter set and data to calculate the likelihood of the parameters given the data. Once the likelihood was maximized over a particular grid, we adjusted the grid, using finer-scale steps over a smaller section of parameter space. To find our maximum-likelihood parameter estimates, this process was repeated until finer grids failed to achieve a maxi-
Table 3.1: Model parameters, published estimates and test ranges. (Unless otherwise noted, estimates are restricted to *Plasmodium chabaudi* and C57 black mice.)

<table>
<thead>
<tr>
<th>Parameter and definition</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_R$, $\omega_N$</td>
<td>Mean number of merozoites produced per infected reticulocyte, normocyte</td>
<td>4-10</td>
</tr>
<tr>
<td>$\beta_R$, $\beta_N$</td>
<td>Invasion rate of reticulocytes, normocytes ((cell/µl)$^{-1}$day$^{-1}$)</td>
<td>$10^{-6}$</td>
</tr>
<tr>
<td>$d$</td>
<td>Death rate of RBCs (day$^{-1}$)</td>
<td>0.02-0.025</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Death rate of merozoites (day$^{-1}$)</td>
<td>40-50</td>
</tr>
<tr>
<td>$g$</td>
<td>Proportion of infected RBCs that produce gametocytes</td>
<td>&lt;0.1</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Lag in RBC production (transit time, days)</td>
<td>2-3</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Proportion of RBC deficit that is made up each day</td>
<td>0.15-0.25</td>
</tr>
</tbody>
</table>

* $P. berghei$, † CBAxC57BL mice, ‡ C3H mice.

...mum log-likelihood greater than the previous grid, up to four decimal places. Since we fit the models to individual mice, we expect the only uncertainty in our data to be due to our measurement apparatus and sampling error. A prior experiment estimated the repeatability of RBC densities measured by flow cytometry and parasite densities measured by qPCR. From this we know that the error of the log$_{10}$-transformed RBC and parasite densities are normally distributed with standard deviations of 0.034 and 0.2, respectively.

Ideally, we would fit the models to both the RBC and the parasite data; however, we have less confidence in our parasite data because of the greater error associated with it.
Furthermore, given the discrete-time nature of our model we can make predictions about the density of merozoites either just after all infected RBCs have burst, or just after all merozoites have either died or invaded an RBC but before any asexual replication within RBCs has occurred. Our experimental measurements are unlikely to have captured either of these events exclusively, so it is unclear which predictions should be fit to data. Using only the RBCs to fit the models allows us to avoid this difficulty. Also, because we do not fit our model directly to the parasite data, the parasite density predictions generated from best-fit models provide another qualitative test of the appropriateness of this model. The parasite predictions we use to qualitatively compare with our data are those that calculate the density of parasites that have successfully invaded RBCs, prior to asexual replication.

With this error structure, the probability of observing $10^D$ RBCs per $\mu l$ of blood given the model predicts $10^M$ RBCs per $\mu l$ of blood is

$$\text{Prob}(D|M) = \frac{1}{0.034\sqrt{2\pi}} e^{-\frac{(D-M)^2}{2(0.034)^2}}. \quad (3.2)$$

The likelihood of a particular set of parameters given the data is proportional to the product of the probabilities of observing the data given the model predictions. For numerical accuracy, we take the natural logarithm of the likelihood, so that

$$L = \sum_{q=1}^{t_{\text{max}}} \ln \left( \frac{1}{0.034\sqrt{2\pi}} e^{-\frac{(D_q-M_q)^2}{2(0.034)^2}} \right) \quad (3.3)$$

where $D_q$ is the log$_{10}$ of the measured RBC density at time $q$, and $M_q$ is found by simulating the model for a particular data set and taking the log$_{10}$ of the total RBCs (reticulocytes and normocytes) at time $q$. The best-fit parameters for a given model maximize the log-likelihood, $L$. 


Chapter 3. Strain-specific patterns of pathogenesis in malaria

3.2.2 Statistical analysis

Model selection and Goodness-of-fit

Given the nested nature of our eight model variations, we use the likelihood ratio test to determine the best-fit model. If the calculated maximum log-likelihoods for models A and B are $L_A$ and $L_B$, and if model A has fewer fitted parameters then we define

$$R = 2(L_B - L_A). \quad (3.4)$$

If $R$ is greater than the chi-squared critical value, as determined from a $\chi^2$ distribution with degrees of freedom equal to the difference in fitted parameters between models A and B, then model B is a significantly better fit at the 5% level (e.g. Hilborn & Mangel, 1997; Grimshaw et al., 2001; Johnson & Omland, 2004). We do these pairwise model comparisons for each individual mouse to determine the best-fit model. In addition, we reject any models with best-fit parameters that are in strong disagreement with experimentally-determined parameter estimates (Table 3.1).

Once we choose a best-fit model for each mouse, we test the goodness-of-fit of the maximum-likelihood parameters of this model, knowing that the likelihood of these parameters given the data is proportional to $L_{max}$. To do this, we assume that the model and parameters are true and simulate 1000 artificial data sets by generating RBC predictions and incorporating the error structure described in section 3.2.1. Since we assume that the model is true, any of these artificial data sets could have been the measured one, so we can generate an expected distribution for $L_{max}$ by calculating the likelihoods of the parameters
given these artificial data sets. If the observed $L_{max}$ lies within the 95% highest density region of its expected distribution then the parameters are considered a good fit and are accepted at the $\alpha = 5\%$ level.

**Parameter significance and parameter error estimates**

To determine whether fitted parameters are significant, we set each parameter in turn to zero, recalculate the maximum likelihood and then compare this maximum-likelihood to the original using the likelihood-ratio test, as described above. If the likelihood score of the model including the parameter is significantly better than the model with the parameter set to zero, then the parameter is significant at the 5% level.

To approximate the uncertainty in our estimated parameters we generate a probability distribution for each. We again assume that the best-fit parameters and model are the true ones and generate 100 artificial data sets for each mouse. This gives us 100 datsets that each could have been the one that was measured and which would have resulted in a slightly different set of maximum-likelihood parameter estimates. With each of these synthetic data sets we redo the parameter-fitting routine, using the best-fit model, and use the resulting maximum-likelihood estimates to generate a probability distribution for each parameter.

**3.2.3 Coinfection experiments**

The above model fitting and statistical analysis provides a rigorous approach for model selection, but all of these techniques make use of the original data set for which the models
were constructed. Another powerful approach for further examining the validity of the model is to test it with an independent data set. For example, if the model is a valid description of the disease dynamics in CD4$^+$ T cell depleted mice, then it can be used to predict how the dynamics would be altered by different conditions. These predictions can then be tested with new experimental data.

We take this approach by using the best fit models obtained above to make predictions about the disease dynamics that are expected if both clones simultaneously infect a single host. To generate model predictions we extend the model to allow for two distinct clones (see Appendix B) and set appropriate parameter values. For clone-specific burst size and invasion rate parameters we use the median values from the two $10^6$ experiments. For the parameters governing RBC production, we use the mean of the medians from the two clones except for the time lag, which must be an integer, and thus was set at 2.

The predictions from our model are then compared with data from an experiment. Specifically, we obtained data from 4 mice that received the same anti-CD4$^+$ treatment as in the previous experiments, were contemporaneously inoculated with $10^6$ AS and $10^6$ DK parasites, and survived to day 19 post-inoculation. One mouse (mouse 3) died prematurely in this experiment and so was excluded from the analysis. Experimental details are given in Barclay et al. (2008).
3.3 Results

3.3.1 Model development and data fitting

Maximum log-likelihood values for each model and mouse are reported in Appendix C in Table C.1. Using the likelihood-ratio test, model 7 was chosen as the best-fit for eleven of the thirteen mice. The two that failed to choose model 7 as the best fit both picked model 8, but had maximum-likelihood estimates of gametocyte conversion rates that are not easily reconciled with published values (AS 10^5 mouse 1: g=0.307; AS 10^5 mouse 3: g=0.198).

Empirical estimates of gametocyte conversion rates are hard to obtain since only the end product of this process can be counted and the immune system may be quickly disposing of maturing gametocytes (Taylor & Read, 1997). Instead, some studies have measured the daily proportion of total parasites (gametocytes and merozoites) that are gametocytes (Buckling et al., 1999; Shutler et al., 2005). This serves as a reasonable proxy for conversion rate in the absence of evidence for a strong, gametocyte-specific immune response and considering the relatively long life-span of gametocytes in the bloodstream compared to merozoites (gametocyte half-life is estimated to be 14 hours; Reece et al., 2003). Under normal conditions, in experimental infections with *P. chabaudi* clone DK, gametocytes make up around 1-2% of all circulating parasites and, even at their maximum density, do not constitute more than 10% (Shutler et al., 2005). Under maximal stimulation it is possible that conversion rates reach values above 0.1, but there is no empirical evidence of rates this high being maintained for the duration of the acute phase of infection and so we exclude model 8 for these mice. The next best model for both of these mice is model 7,
so we took this as the best fit model for these two mice for the remainder of the analyses. None of our conclusions about inoculum size or clone effects qualitatively change by choosing model 8 instead.

The best-fit curves for all mice are in excellent qualitative agreement with measured RBC densities (see Figure 3.2). From the best-fit model and parameter estimates, we generated predictions for the parasite dynamics and these are shown, along with experimental data in Figure 3.3. Considering that we did not use this data for fitting, the model does a good job of qualitatively explaining the parasite dynamics, with a few notable exceptions. For mouse 4 from the AS 10\(^5\) experiment the model predicts unreasonably high parasite densities for the second peak. There is one potential outlier in the RBC measurements for this mouse occurring on day 10 post-inoculation. We omitted this data point, refit the model, and found much more reasonable predictions for the parasite dynamics (as depicted by the blue curves in Figures 3.2 and 3.3). For the remainder of the analyses, we use this modified data set with one outlier omitted for mouse 4. Also, for mouse 2 from the AS 10\(^6\) experiment the model fails to capture the timing of parasite peaks. However, in this dataset there were no obvious outliers in the RBC measurements as the RBC dynamics look very different from the other mice in this experiment. In particular, RBC density drops about two days later than in the other mice despite similar timing in the parasite peaks, and RBC density fails to show any real increase after the second parasite peak. This mouse was excluded from further statistical analysis.
Figure 3.2: Experimental data and best fit curves for RBC dynamics. Circles are observed values, solid lines are best fits. Infections are with either $10^5$ AS, $10^6$ AS, or $10^6$ DK parasites. Blue line represents best-fit curve for mouse 4 in the $10^5$ AS experiment with one outlier (day 10) removed, denoted mouse 4'. The fit of the model is very good for all mice. (Goodness-of-fit, AS $10^5$: mouse 1, 388; mouse 3, 134; mouse 4, 739; mouse 5, 134; AS $10^6$: mouse 1, 824; mouse 3, 581; mouse 4', 590; mouse 5, 642; DK $10^6$: mouse 1, 610; mouse 2, 866; mouse 3, 727; mouse 4, 368. Values above 50 represent good fits.).
Figure 3.3: Experimental data and simulations of parasite dynamics. Circles are observed values, solid lines are model predictions (using parameters fitted to RBC data only). Infections are with either $10^5$ AS, $10^6$ AS, or $10^6$ DK parasites. Blue line represents model predictions for mouse 4 in the $10^5$ AS experiment after one outlier (day 10) was removed.
3.3.2 Statistical Analysis

The goodness-of-fits of model 7 for all mice are given in the caption of Figure 3.2 and are shown graphically in Figure C.1 of Appendix C. The fits were very good for all mice and all parameters were significant for every mouse. Best-fit parameters for model 7 are shown for individual mice in Table C.2 and boxplots of estimated parameter distributions are shown in Figures C.2-C.4.

Some trends can be seen when comparing estimated values of certain parameters within individuals. In particular, RBC deficits tend to be made up faster under anemic conditions than when RBC densities are normal (i.e. $\theta_0 < \theta_A$, see Figure C.4). In addition, one aspect of the importance of RBC age structure bears itself out in these distributions: for every individual, the invasion rate of fully mature RBCs is higher than for reticulocytes (i.e. $\beta_R < \beta_N$, see Figure C.2), often by an order of magnitude. These invasion rates control for differences in availability, so truly represent a preference for mature RBCs. Despite this preference, for most individuals burst size is higher in reticulocytes than in normocytes (i.e. $\omega_R > \omega_N$, see Figure C.3).

Effect of parasite clone

By pooling the data from individuals we are able to compare the distributions of the parameter estimates to see what effect parasite clone has on the infection dynamics and, in particular, to try to identify the basis for differences in virulence between the DK (less virulent) and AS (more virulent) clones. We compared only those data from inoculations with
10^6 parasites to control for possible inoculum size effects. Table 3.2 contains estimated DK parameter values relative to the AS values and we assume, somewhat arbitrarily, that differences of greater than 10% of the AS value are evidence of a clone effect (see also Appendix C for raw median values and pooled distributions). The median invasion rates of reticulocytes are approximately equal for both clones, but the AS clone has a higher invasion rate of normocytes than DK. Median values for burst size are also higher for the AS than the DK parasites and this difference is even more pronounced in reticulocytes than in normocytes. The mice in the AS-infected mice have slower rates of RBC production, both at normal RBC densities and during anemia.

**Effect of inoculum size**

We also compared the distributions of the parameter estimates to see what effect inoculum size has on the infection dynamics. We compared only those data from inoculations with AS parasites to control for possible clone effects. Estimated parameter values for the 10^5 experiment relative to the 10^6 experiment are given in Table 3.2 (see also Appendix C). Median values of RBC invasion rates and normocyte burst size are approximately equal. However, there is a marked difference in the median reticulocyte burst sizes, with the low inoculum size obtaining more almost twice as many merozites per infected reticulocyte than the high inoculum size. Estimates of RBC production rates are higher in the 10^6 inoculum size than in 10^5, both at normal RBC densities and during anemia.
Table 3.2: Medians of pooled estimated parameter distributions relative to AS $10^6$ estimated values. We assume a difference of 10% or greater is evidence of an effect, indicated by a √.

<table>
<thead>
<tr>
<th>parameter</th>
<th>AS $10^5$</th>
<th>AS $10^6$</th>
<th>DK $10^6$</th>
<th>Inoculum size</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_0$</td>
<td>0.759</td>
<td>1</td>
<td>1.312</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\theta_A$</td>
<td>0.381</td>
<td>1</td>
<td>1.180</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\beta_R$</td>
<td>1.020</td>
<td>1</td>
<td>0.940</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_N$</td>
<td>0.906</td>
<td>1</td>
<td>0.791</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>$\omega_R$</td>
<td>1.943</td>
<td>1</td>
<td>0.765</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>$\omega_N$</td>
<td>1.002</td>
<td>1</td>
<td>0.877</td>
<td></td>
<td>✓</td>
</tr>
</tbody>
</table>

3.3.3 Coinfection experiments

Model predictions and experimental data from coinfection experiments are plotted in Figures 3.4 and 3.5. Despite not allowing any individual variation except for starting RBC and parasite densities, the model predictions provide a reasonable qualitative fit to the data, particularly for the early phase of the parasite dynamics. The predicted peaks and troughs in the RBC densities have similar amplitude to the data, but the timing is slightly different. This is especially clear in mice 1 and 2, where RBC densities are predicted to rebound faster than they actually do. This suggests that RBC production in these mice with mixed infections is slower than in the single-clone infected mice, since we used the single-clone parameter estimates to generate these predictions. Despite these differences, the parasite density predictions are very good, predicting both the correct timing of peaks and magnitude of the first wave of parasites. Model predictions and data start to diverge after the first peak with the model overestimating the level of competitive exclusion of the less virulent DK by the relatively more virulent AS.
Figure 3.4: Experimental data and model predictions for RBC dynamics in competition experiments. Circles are observed values, solid lines are model predictions. Infections are with $10^6$ AS and $10^6$ DK parasites. $\beta_R$, $\beta_N$, $\omega_R$, and $\omega_N$ are clone-specific and set to values listed in Table C.2 for the $10^6$ single-clone experiments. $\theta_0$ and $\theta_A$ are set to the mean of the values obtained from these experiments and $\tau$ is set to 2, as obtained from the AS $10^6$ data.

Figure 3.5: Experimental data and model predictions for parasite dynamics in competition experiments. Red is AS parasites, blue is DK. Circles are observed values, solid lines are model predictions. Parameters are the same as in Figure 3.4.
3.4 Discussion

We have used a comprehensive combination of mathematical modeling and experimental data to explore the relative importance of ecological factors to the within-host dynamics of *P. chabaudi* infections. From our best-fit model we conclude that RBC availability is an important regulator of parasite growth in CD4+ T cell depleted mice and our systematic approach to model fitting uncovered some of the complexities of this resource and its interaction with the parasites.

Consistent with previous work (Hetzel & Anderson, 1996), we find that the loss of RBCs due to infection is sufficient to down-regulate parasite densities after the initial peak and the subsequent increase in RBC production alone generates a second wave of parasites. Recrudescences are conventionally explained as outgrowths by antigenic variants able to escape protective immunity, and it is true that recrudescences are antigenically distinct from the primary wave of parasites (Brown & Brown, 1965; Phillips et al., 1997). However, since malaria parasites are generating antigenic variation through time, recrudescences would be antigenically distinct even if parasite densities were entirely controlled by RBC dynamics. The relative importance of immune evasion and resource limitation remains to be determined in animals with fully intact immune systems.

During periods of anemia, we find that the rate of RBC production is upregulated (similarly to Jakeman et al., 1999). Intriguingly, we find that mice infected with the AS clone of *P. chabaudi* have slower rates of RBC production than those inoculated with DK. A similar effect of clone differences on erythropoiesis has been estimated before (Haydon
et al., 2003). In our study, it remains unclear whether RBC production is suppressed by AS or enhanced by DK. If future empirical work can establish a baseline response to anemia, this question could be resolved. Either way, our finding that mice inoculated with DK suffer less anemia may help to explain why AS is relatively more virulent than DK. The details, ubiquity and adaptive value of this sort of resource manipulation offer many avenues for future investigation.

Recent theory has pointed to a role for RBC age-structure in determining relative virulence of *P. chabaudi* clones, with more virulent clones estimated to be able to invade a greater age range of RBCs than less virulent clones (Antia et al., 2008). Our results echo the importance of age structure, but our model takes a different approach and suggests different mechanisms. In particular, Antia et al. (2008) assume that no invasion is possible for RBCs outside a predicted age range and estimate the youngest RBCs each clone is able to infect (8 days for the more virulent and 12 days for the less virulent clone). In contrast, we allow for different invasion rates and burst sizes between types of RBCs and sought to define these types in a biologically meaningful way (i.e., reticulocytes and normocytes). We could then determine if malaria parasites interact differently with each type of cell and if these interactions differ between clones.

First, we find that the more virulent clone (AS) has an advantage early on in infections with a greater invasion rate of, and higher burst size in, normocytes than the less virulent clone (DK). Second, while both AS and DK merozoites have lower invasion rates of reticulocytes, AS gains a ‘fecundity’ benefit from these invasions, producing more daugh-
ter merozoites in younger cells than in fully mature ones. This likely gives AS a big advantage as anemia sets in, RBC production increases and the system is flushed with reticulocytes. While no biological mechanism for this higher burst size in reticulocytes is yet known, potential candidates include: the larger size of reticulocytes allows more room for merozoites, older cells may be more likely to rupture under stress, or, given that reticulocytes are newly introduced into the bloodstream, their expected circulation time before clearance by the spleen is longer (Loeffler et al., 1989), thus, the relative risk to the merozoites of undergoing an addition round of asexual multiplication is lower.

In intact mice, *P. chabaudi* merozoites infect the most abundant cell, showing a preference for normocytes early in infections then switching to reticulocytes when the mouse becomes anemic and the bloodstream flushes with new RBCs (Jarra & Brown, 1989; Taylor-Robinson & Phillips, 1994). In CD4+ T cell depleted mice, infections become chronic, RBC densities remain depressed and the proportion of circulating reticulocytes remains higher than under normal conditions, though they remain less abundant than mature RBCs (Taylor-Robinson & Phillips, 1994). In broad agreement with our ‘fecundity benefit’ hypothesis of reticulocyte invasion, other experimental studies have shown that in immune-depleted mice, AS parasites continue to show a preference for reticulocytes (Taylor-Robinson & Phillips, 1994), i.e., invade reticulocytes more often than would be expected given their relative density. In these mice, the cost of overcoming a lower invasion rate in reticulocytes may be less than the burst size benefit, while in the presence of an immune response the cost of not infecting the most easily accessible cell may be
prohibitively high. Predicting a facultative adjustment of a parasite life-history trait in response to the host environment may be novel in the case of RBC preference and host immunity, but it is not unprecedented. Rodent malaria parasites have been shown to alter either their investment in gametocytes or the gametocyte sex ratio in response to host stress (e.g. Buckling et al., 1999; Paul et al., 2003; Reece et al., 2005).

Given this empirical evidence of variable investment in gametocytes throughout infections and depending on the host environment, modeling conversion rate as a constant throughout the course of infection is inadequate. However, our results suggest that, on average, conversion to gametocytes has no significant effect on the infection dynamics. An alternative approach would be to allow this parameter to vary over time. While this would perhaps be more biologically realistic, it would also make the parameter-fitting process computationally impractical. In its simplest form, our model need not explicitly track gametocytogenesis, as this ‘loss’ of asexuals could be subsumed by the estimates for burst sizes or invasion rates.

Our results present some unexpected differences due to parasite inoculum size. First, inoculum size matters for RBC production, with higher inoculum sizes inducing higher production rates. These parameter estimates are likely due to the fact that in the experimental data, mice inoculated with $10^5$ AS parasites reached lower minimum RBC densities than those given $10^6$ parasites. Greater anemia with a lower inoculum size is not consistent with previous results on inoculum size effects (Timms et al., 2001), though that study was done in mice with intact immune systems and with different *P. chabaudi* clones. Second,
reticulocyte burst sizes are about two-fold higher in the $10^5$ inoculum size experiment than in the $10^6$ experiment. One potential explanation of this is the rigidity of the model framework. It could be that with the bigger inoculum size, as well as producing more RBCs, the RBCs are maturing faster so that what we are modeling as reticulocytes are functionally fully mature RBCs (and we would then expect the best-fit $\beta_R$ to be closer to $\beta_N$ for the $10^6$ experiment). A more intriguing possibility is that burst size is density-dependent so that the merozoites regulate their usage of RBCs.

We found no consistent estimate of the lag in RBC production from our individual mice. While published estimates of the transit time of blood cell precursors to the bloodstream tend to be around three days (Mary et al., 1980; Chang et al., 2004) this process has been estimated to take as little as 33 hours under maximal erythropoietic stimulation (Loeffler et al., 1989). Again, our model did not allow any flexibility in this parameter over time because the discretized description of RBC production meant it had to be an integer. The medians from the pooled distributions do match up with the predictions about RBC production rates – DK induced the highest production rate and had the shortest lag, while inoculation with a smaller number of merozoites induced the lowest production rate and has the longest lag. Though the medians from both high inoculum size experiments are different, the means were more similar (AS: 2.0 days, DK: 1.7 days) and the true transit time need not be an integer.

Both empirical and theoretical studies have attempted to understand the mechanisms of competition in genetically diverse malaria infections. Experiments with *P. chabaudi* in
athymic mice revealed some evidence for immune-mediated apparent competition as the initial wave of parasites receded (Råberg et al., 2006), but comparisons of the CD4+ T cell depleted mice, used in the analyses reported here, with intact control mice failed to find evidence for CD4+ T cell-mediated competition (Barclay et al., 2008). In the presence of a weak immune response, theory predicts that competition for RBCs is a limiting factor for coinfecting parasites (Hellriegel, 1992). Our results confirm this: RBC availability alone can explain the first peak of parasites in CD4+ T cell depleted mice. The differences in burst sizes and invasion rates of the clones that were derived from our single-strain model generate the appropriate relative heights of the first AS and DK merozoite peaks in coinfections. Beyond the initial peak our model predictions and the data start to deviate. The model predicts much stronger competitive suppression of the less virulent DK by AS than is seen in the data. Our data is consistent with previous studies in immune-intact mice which showed that after the initial peaks, parasite densities were no different or higher in competition than they were in single infections (Bell et al., 2006). The authors of that study point to the immune response becoming more clone-specific as an explanation of why the competitively superior clone cannot completely suppress the other. Here, we propose that a similar, but CD4+ T cell independent specific immune response is responsible for the competitive release of the DK clone that our model could not account for. Further insights into the dynamics of coinfections could be found by fitting our competition model to the available data to see if differences in infection dynamics are solely due to differences in parameters or, alternatively, if other mechanisms (like our proposed immune response) are...
in operation.

While we believe that the ability to predict the outcome of the acute phase of competition experiments provides strong validation of our model, it could be further tested in a number of ways. One option is to empirically test our parameter estimates, though our results show that there will be substantial variation between individual mice. Another is to experiment with other clones to see if their relative virulence can be explained by the mechanisms we have proposed. It would also be interesting to see if there was more evidence of density-dependent regulation of RBC usage and burst sizes, perhaps through experimental manipulation of RBC availability.

The results presented here take a significant first step towards a comprehensive model of the within-host dynamics of malaria infections. We have shown ecological factors are crucial for explaining infection dynamics and with this foundation, we can build up the complexity towards a better understanding of the dynamics in the more biologically interesting setting of immune-intact hosts.
Chapter 4. HOST EXPLOITATION STRATEGIES IN TWO STRAINS OF THE RODENT MALARIA, *P. chabaudi*

ABSTRACT

Malaria parasites are known to differ in their strategies for exploiting host resources both between species and between genotypes of the same species. This variation results in different levels of disease severity (i.e., virulence) and is predicted to determine the outcome of competitive interactions between co-infecting parasites. Using a model malaria system (*P. chabaudi* infections in mice) we measure two important host exploitation traits of malaria parasites – the invasion rates of target red blood cells (RBCs) and the number of progeny parasites produced per infected cell (burst size) – in two different strains. We use our data to test predictions from our recent theoretical work on infections with these two strains in CD4+ depleted mice. We find that, as predicted, the more virulent strain has higher burst sizes, but the role of invasion rates in determining virulence is equivocal. We also investigate the effect of within-host environmental factors on these traits, including host immunity and anemia. Our data suggest that parasites are able to vary their host exploitation strategy, in response to changes within a host, in ways that may be adaptive.
4.1 Introduction

A defining characteristic of parasites is that they exploit host resources for their own survival, growth or reproduction. This has obvious implications for host health and, as such, theory has often treated virulence (e.g. reduction in host fitness as a result of infection) as synonymous with host exploitation (Day, 2002). In natural infections, subtle variation in exploitation strategies can indeed lead to qualitatively different disease outcomes. Take, for example, members of the *Plasmodium* genus (the infectious agents of malaria) that exploit host red blood cells (RBCs) during one stage of their complex life cycle. After invading a host RBC, a malaria parasite undergoes multiple rounds of asexual replication, eventually bursting the RBC and releasing several progeny parasites (merozoites), each with the capacity to invade another RBC (Figure 4.1). The four malaria species that infect humans differ in a number of host exploitation traits, including the number merozoites produced per infected cell (referred to here as ‘burst size’) and the ages of host RBCs they show a preference for invading. *P. malariae*, *P. vivax* and *P. ovale* each invade only a particular subset of RBCs, either young reticulocytes or fully mature normocytes (Paul et al., 2003). However *P. falciparum*, the species that is responsible for the overwhelming majority of clinical malaria cases (and subsequently, death and disease associated with malaria; WHO, 2008b), shows no strong preference for specific ages of RBCs and has the highest estimated range of burst sizes (Paul et al., 2003). These differences in host exploitation traits are likely a key component of virulence and are predicted to affect the outcome of within-host competitive interactions between species (McQueen & McKenzie, 2006).
While an emphasis seems often to be placed on differences between malaria species, there is also evidence of variation in life-histories within a species (Reece et al., 2009). For example, experiments with a number of genetically distinct clones of the rodent malaria \textit{P. chabaudi}, reveal substantially different infection dynamics and outcomes (i.e., virulence; Mackinnon & Read, 1999) across clones. The competitive ability of clones in this system is positively correlated with their virulence (Bell et al., 2006). Furthermore, competition in mixed infections is at least partially resource-driven (Barclay et al., 2008; Råberg et al., 2006), suggesting that some of the variation in life history (and ultimately, virulence) between clones can be attributed to differences in host exploitation strategies. Elucidating which traits underlie these strategies is not easily accomplished with an empirical approach alone; a more efficient approach is to integrate theory with data.

In Chapter 3, we used mathematical models to understand the factors that regulate within-host dynamics of infections with different pairs of \textit{P. chabaudi} clones. We predicted that burst sizes and RBC age-specific preferences (defined as invasion rates of different ages of RBCs) are key traits that underlie patterns of infection dynamics. These traits vary across clones and could therefore explain differences in virulence and competitive ability.
between them. Specifically, the more virulent clone, AS, is predicted to obtain a higher burst size in all ages of RBC than the less virulent clone, DK. AS is also predicted to gain a ‘fecundity’ advantage from infecting reticulocytes, producing more merozoites in these cells than in normocytes. Further, while both clones are predicted to ‘prefer’ normocytes (i.e., they invade them at higher rates than reticulocytes), the more virulent clone, AS, is predicted to gain a competitive advantage by invading normocytes at a higher rate than DK. Invasion rates of reticulocytes are predicted to be similar between the two clones.

Differences in infection dynamics between malaria parasite clones have previously been investigated, but few have attempted to quantify that variation at the level of host exploitation traits. Here, we use an experimental approach to study two traits of *P. chabaudi* parasites – RBC age-specific burst sizes and invasion rates. Our aims are twofold. First, we seek to test predictions from our previous theoretical work, as described above, and to verify whether these host exploitation traits can indeed explain differences in virulence between clones. Second, we seek to determine whether patterns of host exploitation between clones are mediated by host immunity or anemia, i.e., if there is evidence of parasites altering burst size and invasion rates in response to their within-host environment.

To test the model predictions, we repeated the same CD4$^+$ depletion as in Chapter 3. Since it is possible only to measure average, rather than RBC age-specific, burst sizes it was necessary to manipulate RBC age structure with phenylhydrazine (PHZ, described in detail below). This allowed us to compare burst sizes in mice with enhanced reticulocyte densities (due to PHZ) and burst sizes in control mice in which reticulocytes are scarce
(see Figure 4.2a for general model predictions). RBC age-specific invasion rates could be measured more directly in CD4$^+$ depleted mice (see Figure 4.3a for model predictions). To examine the effects of immunity and anemia on the two traits, we also measure burst sizes and invasion rates in immune-intact hosts, both PHZ-treated and PHZ-control.

**Figure 4.2:** Burst size predictions and data for CD4$^+$ depleted mice. (a) Our model of *P. chabaudi* infection in CD4$^+$ T cell depleted mice predicts that burst sizes are higher for AS than for DK. AS is also predicted to have a higher burst size in reticulocytes than in normocytes, while the opposite is predicted for DK. Since PHZ treatment should increase the proportion of infected RBCs that are reticulocytes, we expect to see average burst sizes move in these opposite directions in PHZ treated mice. The magnitude of these burst sizes cannot be easily predicted since the proportions of infected RBCs that are reticulocytes and normocytes are unknown, so we present only qualitative predictions. (b) AS obtained a higher burst size than DK in experimental infections, but PHZ resulted in an increase across both clones. Means are calculated from average burst sizes in four or five individual infections from each treatment (genotype x PHZ) in CD4$^+$ depleted mice. Error bars, $\pm$ 1 s.e.m. Parasite genotype is denoted by symbols (circle, DK; triangles, AS).
Figure 4.3: Invasion rate predictions and data for CD4$^+$ depleted mice. (a) Our model of *P. chabaudi* infection in CD4$^+$ T cell depleted mice predicts that the invasion rate of reticulocytes (‘Retic’) will be similar for AS and DK and, overall, much lower than the invasion rates of normocytes (‘Normo’). AS is predicted to have a higher invasion rate of normocytes than DK. (b) Invasion rates of both clones are similar in reticulocytes, but we did not observe the predicted increase of invasion rates in normocytes across clones as expected. Means of $\log_{10}$ transformed invasion rate x merozoite lifespan estimates are presented from four (AS) or five (DK) individual infections in CD4$^+$ depleted mice (with no PHZ). Error bars, ± 1 s.e.m. Parasite genotype is denoted by symbols (circle, DK; triangles, AS).
4.2 Methods

4.2.1 Parasites and hosts

Figure 4.1 illustrates the relevant blood stages in the life cycle of *P. chabaudi* within hosts. Development through these stages is synchronous between parasites and takes 24 hours to complete. We initiated experimental infections with one of two genetically distinct *Plasmodium chabaudi* clones denoted AS and DK (WHO Registry of Standard Malaria Parasites, University of Edinburgh, UK), originally isolated from thicket rats, *Thamnomys rutilans*, in the Central African Republic (Beale et al., 1978). All mice were inbred female C57BL/6JolaHsd mice aged 6-8 weeks old (Harlan-Olac, UK) and received an intra-peritoneal inoculation of $10^6$ parasitized red blood cells in a 0.1 ml dose consisting of 47.5% Ringers (27 mM KCl, 27 mM CaCl$_2$, 0.15M NaCl), 50% heat inactivated calf serum and 2.5% heparin (200 units mlK1). Mice were housed randomly in groups of five at 20°C with a 12 hour light/12 hour dark cycle and provided food (41B, Harlan-Teklad, UK) and water with 0.05% PABA (to enhance parasite growth) ad libitum.

4.2.2 Manipulating immunity

CD4$^+$ T cell depleted mice received intra-peritoneal injections of anti-CD4$^+$ antibody (GK1.5) and control mice received injections of a non-depleting antibody of the same isotype (IgG I4131, Sigma) dissolved in PBS, according to the following regimen: 500$\mu$g 5 days before infections, 250$\mu$g 4 days before infections, and 250$\mu$g 1 day before infection. Using FACS analysis (as described in Barclay et al., 2008) we verified that CD4$^+$ depletion
was successful. Mean parasite densities for each treatment over the initial peak of infection are shown in Figure 4.4; as expected, parasite densities are higher in depleted mice than in immune-intact (PHZ control) mice (Barclay et al., 2008).

Figure 4.4: Parasite dynamics. Plots show mean parasite density in different treatments (grouped by parasite genotype) over time. Error bars, ± 1 s.e.m. Colours indicated what treatments, in addition to infection, hosts received: black, none; red, phenylhydrazine (PHZ); blue, CD4+ depletion (CD4-); purple, PHZ and CD4+ depletion. Over the course of the initial peak, parasite densities were higher in CD4+ depleted hosts as compared to immune-intact (PHZ control) hosts.

4.2.3 Manipulating RBC age structure

Treated mice received intra-peritoneal injections of 40mg/kg of phenylhydrazine hydrochloride dissolved in phosphate-buffered saline (PBS), 2 days prior to infection. Control mice received equal volume injections of PBS alone. This dose of PHZ causes a rapid decrease in RBC densities and a subsequent influx of reticulocytes as the system ‘re-equilibrates’. RBC densities in our experiment do show a marked decrease in response to PHZ treat-
ment (Figure 4.5). The maximum proportion of reticulocytes was predicted to occur day 3 post-infection, according to previous experiments (Savill et al., in press). In our PHZ treated mice, the proportion of reticulocytes remained significantly higher than in control mice through days 6 and 7 ($F_{1,32} = 189.716, P < 0.001$; Welch, $F_{1,20.11} = 34.798, P < 0.001$; Figure 4.6a). Further, the proportion of infected RBCs that were reticulocytes was also significantly higher in PHZ treated mice on days 6 and 7 ($F_{1,31} = 22.803, P < 0.001$; Welch, $F_{1,19.452} = 17.576, P < 0.001$; Figure 4.6b).

**Figure 4.5:** Red blood cell (RBC) dynamics. Plots show mean RBC density in different treatments (grouped by parasite genotype) over time. Error bars, ± 1 s.e.m. Colours indicated what treatments, in addition to infection, hosts received: black, none; red, phenylhydrazine (PHZ); blue, CD4$^+$ depletion (CD4-); purple, PHZ and CD4$^+$ depletion. PHZ treatment (two day prior to infection, i.e., day -2) resulted in a decrease in RBC densities. This effect coincided with the initial stages of infection.
Figure 4.6: Effect of phenylhydrazine (PHZ) on red blood cell (RBC) age structure. (a) Mean proportion of all RBCs that are reticulocytes. (b) Mean proportion of all infected RBCs that are reticulocytes. Open squares show averages for mice that did not receive PHZ and filled squares show averages for mice that did receive PHZ. Error bars, ± 1 s.e.m. At the early stages of infection, mice treated with PHZ had a significantly higher proportion of reticulocytes, as expected, and subsequently, a significantly higher proportion of all infected cells were reticulocytes.

4.2.4 Experimental setup and sampling

We ran a fully cross-factored experiment with three treatments in which we manipulated the host immune response, anemia/RBC age structure and parasite clone. For both AS and DK infections, groups of five mice were treated with one of the following regimes:

a. control antibodies, PBS (PHZ control)
b. control antibodies, PBS+PHZ
c. anti-CD4⁺ antibodies, PBS
d. anti-CD4⁺ antibodies, PBS+PHZ
giving a total of 40 mice. Control groups were used to determine if patterns of host exploitation traits were affected by the within-host environment. In addition to the samples described in detail below (and outlined in Appendix D, Figure D.1), we took daily
measurements of the mass of individual mice, and daily blood samples from the tail to prepare Giemsa-stained thin blood smears, to estimate RBC density (using flow cytometry, Beckman Coulter), and to measure parasite density with genotype-specific real-time quantitative PCR (as described in Barclay et al., 2008).

**Measuring burst size**

It is not possible to sample mature parasites (schizonts) *in vivo*, because all asexual parasites leave the circulation during their development to sequester in host tissues and reach maturity. Therefore, to measure burst size without destructive sampling, we harvested parasites at trophozoite stage, before sequestration, and cultured them *in vitro* to mature schizont stage. In addition to allowing us to sample schizonts, this method enabled us to collect parasites from each host on multiple occasions. Specifically, we cultured 10µl parasitised blood from each mouse in 1ml complete culture medium (RPMI medium with NaHCO₃, Hepes and L-glutamine (Invitrogen), containing 25% heat inactivated foetal calf serum (Gibco), at pH7.25) in the presence of 10% O₂, 5% CO₂, 85% N₂, at 30°C. We sampled cultures at 18.5, 20 and 21 hours to enable us to capture the optimal time for assaying burst size (i.e., when mature schizonts are frequent but ruptured schizonts and invasion of RBC by the released merozoites is rare). We cultured parasites from each infection on both days 5 and 6 post-infection to ensure we captured the optimal percent parasitemia for schizont culture (~10%). The cultures from day 5 were more successful than day 6 (when several treatment groups exceeded 10% parasitemia) and 20 hours proved to be optimal
for capturing the highest proportion of mature schizonts.

To measure burst size we took thin blood smears of cultured blood and counted the number of merozoites in at least 25 mature schizonts for each individual mouse. The average burst size for each individual mouse was calculated and used in statistical analyses. The distribution of burst sizes within an individual mouse was used to estimate the density of parasites available on the following day (day 6) to begin invasion of RBCs ($\hat{P}_i$, see next section). For unknown reasons, two mice produced blood smears with no schizonts at all on day 5 and an absence of parasites was confirmed by PCR. These two mice (both CD4$^+$ T cell depleted, one infected with AS and the other treated with PHZ and infected with DK) were excluded from all of the analyses.

**Measuring RBC invasion rate**

From Chapter 3 Appendix A, we can find expressions for the invasion rates of different ages of RBCs by rearranging equations A.14 and A.15. The invasion rates of reticulocytes, $\beta_R$, and normocytes, $\beta_N$, are

$$\beta_R = -\frac{\mu \ln \left( \frac{R_i - I_{R,i}}{R_i} \right)}{\hat{P}_i + \hat{R}_i \ln \left( \frac{R_i - I_{R,i}}{R_i} \right) + \hat{N}_i \ln \left( \frac{N_i - I_{N,i}}{N_i} \right)},$$

$$\beta_N = -\frac{\mu \ln \left( \frac{N_i - I_{N,i}}{N_i} \right)}{\hat{P}_i + \hat{R}_i \ln \left( \frac{R_i - I_{R,i}}{R_i} \right) + \hat{N}_i \ln \left( \frac{N_i - I_{N,i}}{N_i} \right)},$$

where $1/\mu$ is the expected lifespan of free living merozoites in the blood stream, $P_i$ is the density of merozoites on day $i$, $R_i$ and $N_i$ are the densities of uninfected reticulocytes and mature RBCs on day $i$, and $I_{R,i}$ and $I_{N,i}$ are the densities of infected reticulocytes and
mature RBCs on day $i$, respectively. A hat (\(^\hat{\cdot}\)) indicates a density immediately before the invasion process begins and the absence of a hat indicates a density after the invasion process is finished, when all merozoites are assumed to have invaded an RBC or died. Therefore, we have $\hat{I}_{R,i} = \hat{I}_{N,i} = 0$ and $P_i = 0$. Our experimental procedures were directed at measuring $\hat{P}_i$, $\hat{R}_i$, $\hat{N}_i$, $I_{R,i}$ and $I_{N,i}$.

To estimate $\hat{P}_i$, two pieces of information are required. First, from thin blood smears prepared when parasites were at their ring-stage on day 5, we calculated the proportion of RBCs that were infected by estimating the total number of RBCs in at least 5 microscopic fields under 1000x magnification and counting at minimum 100 RBCs. This information gives us the density of infected RBCs when multiplied by the total RBC density, as measured by flow cytometry, on day 5. Second, from the burst size measurements, we obtain estimates of the distribution of burst sizes of those infected RBCs on day 5. Putting this together, we calculate the number of merozoites that are (theoretically) present on day 6 to start the invasion cycle, $\hat{P}_i$.

Measuring the other variables is more straightforward, but requires a few assumptions to be made. From thin blood smears taken at ring-stage on day 6, we calculate the densities of infected and uninfected reticulocytes and mature RBCs by counting at least 40 of each type of cell, or up to 20 microscopic fields, and multiplying the resultant proportions by the total RBC density measured by flow cytometry. We assume that no RBC death takes place before or during parasite invasion so that the total density of RBCs we observe in the blood smears, after invasion is complete, is equal to the density of RBCs the parasites
were exposed to at the start of the invasion cycle. Thus, $I_{R,6}$ is the density of infected reticulocytes on day 6 and $\hat{R}_6$ is estimated as the sum of the densities of uninfected and infected reticulocytes on day 6. A similar estimation is done for normocytes. Finally, since the exact value of the expected merozoite lifespan, $1/\mu$, is not known, we rearrange equations 4.1 and 4.2 and estimate the invasion rate multiplied by the merozoite lifespan, i.e., $\beta_R/\mu$ and $\beta_N/\mu$. Any conclusions about differences in invasion rates thus implicitly assume that merozoite lifespan is equal across treatments. The invasion rate estimates from two mice (both infected with DK, CD4$^+$ T cell depleted, and treated with PHZ) were negative, and these were therefore excluded from the analyses. We consider what these negative values mean and what could be their cause in the discussion.

4.2.5 Statistical analyses

All statistical analyses were carried out using R version 2.7.1 (The R Foundation for Statistical Computing, 2008, http://www.R-project.org). One-way analysis of variance tests were used to assess the effects of PHZ on RBC age structure, as reported above in the methods (stats package in R). Statistics are presented from linear models of burst size variation (stats package in R) and linear mixed effects models of invasion rate variation (nlme package in R). In the latter case, a random effect of mouse was included to account for dependence between the two measures of invasion rate (i.e., of reticulocytes and of normocytes) from each mouse.

We ran two different analyses for burst sizes: (1) To test whether burst sizes differ be-
between the clones and according to RBC age structure, as predicted by our model, we fitted a linear model to all of the data from CD4+ depleted mice and included genotype (AS or DK), PHZ (treated or control) and their interaction as fixed effects. (2) To determine if parasite burst sizes in different clones or in response to PHZ are mediated by host immune status, we fitted a linear model to the entire data set, with parasite genotype (DK or AS), CD4+ depletion (depleted or intact control) and PHZ (administered or not), as well as all two-way interactions and the three-way interaction as fixed effects. This latter analysis, although not directly relevant to testing our model predictions, nevertheless provides useful and interesting additional information on host exploitation strategies. In each case, parasite density, uninfected RBC density, and host mass were included as covariates. We evaluated the significance of fixed effects by contrasting models following stepwise deletion of terms with an incremental $F$ test. Non-significant interactions were removed first, followed by non-significant main effects and finally, non-significant covariates; however, none of our findings change when covariates are removed first. Maximal models were simplified until only significant terms remained in the model ($\alpha < 0.05$).

Estimates of invasion rates were log$_{10}$-transformed and similar analyses were conducted. (1) To test whether invasion rates differ between AS and DK and between reticulocytes and normocytes in CD4+ depleted mice, according to our model predictions, we fitted a linear model to the data from CD4+ depleted mice (without PHZ treatment only) and included parasite genotype, RBC age (reticulocyte or normocyte; identifies which age of RBC the invasion rate estimate is for) and their interaction as fixed effects. (2) To inves-
tigate patterns of invasion rates as mediated by host immune status and RBC age structure and abundance, we fitted a linear model to the entire data set, with the same fixed effects structure as described for the second analysis of burst sizes above, but with the addition fixed effect of RBC age (and all its relevant interactions). In each of the invasion rate analyses, the only covariate fitted was host mass since both parasite density and RBC density were used in the calculation of the invasion rates. Maximum-likelihood techniques were used to generate minimal models.

4.3 Results

4.3.1 Burst sizes

We predicted that in CD4\(^+\) depleted mice, burst sizes in both PHZ-treated and PHZ-control groups would be higher for the more virulent clone, AS, than for DK, and that the average burst size of AS would be higher in PHZ-treated than PHZ-control hosts while the opposite would be true for DK (Figure 4.2a). Our results from CD4\(^+\) depleted mice were partially in agreement with these patterns (Figure 4.2b). AS parasites obtained significantly higher burst sizes than DK (\(F_{1,15} = 11.021, P = 0.005\); see Table 4.1) and PHZ resulted in an increase of average burst sizes across both clones (\(F_{1,15} = 4.067, P = 0.062\)). Whilst the effect of PHZ is marginal we include it in the minimal model since it is a main effect and since the minimal model explains 10% more variation than the model that excludes PHZ (adjusted \(R^2 = 0.465\) versus 0.363). Given our model predictions, we expected to find a significant interaction between PHZ and genotype (since we expected AS burst sizes to
increase, but DK burst sizes to decrease in PHZ treated mice). However, the interaction between genotype and PHZ was not significant ($F_{1,13} = 1.738, P = 0.210$), though the main effect of PHZ is larger for AS than for DK (Figure 4.2b). All of the covariates fitted in the maximal model were non-significant (see Table 4.1 for statistics).

Table 4.1: Analysis of burst sizes in CD4+ depleted mice

<table>
<thead>
<tr>
<th>Minimal model</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype</td>
<td>$F_{1,15} = 11.021$</td>
<td>0.005</td>
</tr>
<tr>
<td>Phenylhydrazine</td>
<td>$F_{1,15} = 4.067$</td>
<td>0.062</td>
</tr>
</tbody>
</table>

Non-significant terms deleted from maximal model

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass of mouse</td>
<td>$F_{1,14} = 0.893$</td>
<td>0.361</td>
</tr>
<tr>
<td>Parasite density</td>
<td>$F_{1,13} = 0.181$</td>
<td>0.677</td>
</tr>
<tr>
<td>Uninfected red blood cell density</td>
<td>$F_{1,12} = 0.317$</td>
<td>0.584</td>
</tr>
<tr>
<td>Genotype:Phenylhydrazine</td>
<td>$F_{1,11} = 0.975$</td>
<td>0.345</td>
</tr>
</tbody>
</table>

We then tested whether the effects of parasite genotype and/or PHZ on burst sizes was mediated by host immune status. The average burst size data across all of the different treatment groups is shown in Figure 4.7 and it can be seen that the pattern in immune intact mice is different than in CD4+ depleted mice. Only PHZ ($F_{1,33} = 15.614, P < 0.001$; see Table 4.2) and an interaction between genotype and CD4+ depletion ($F_{1,33} = 5.678, P = 0.023$) remained in the minimal model. PHZ treatment resulted in higher burst sizes, irrespective of parasite genotype or host immune status. The effect of CD4+ depletion varied by clone and resulted in lower burst sizes for DK and higher burst sizes for AS (see Figure 4.7). There is some indication that the burst sizes of different clones in different ages of cells is mediated by immunity, given the marginally significant three-way interac-
tion ($F_{1,27} = 3.390, P = 0.059$). Across treatment groups, AS tends to have a greater burst size than DK, but this trend is reversed in immune-intact, PHZ-treated mice.

![Figure 4.7](image-url)

**Figure 4.7:** Average burst sizes across treatment groups (circles, clone DK; triangles, clone AS). Error bars, ± 1 s.e.m. For consistency, treatments in addition to infection, are coloured as before. Burst sizes were significantly higher in PHZ treated infections and CD4⁺ depletion reduced the burst size of DK parasites, but increased the burst size of AS.

### 4.3.2 Invasion rates

We predicted that in CD4⁺ depleted mice, invasion rates of normocytes would be substantially higher than invasion rates of reticulocytes across clones, and that the more virulent clone, AS, would have a slightly higher invasion rate of normocytes than DK (Figure 4.3a). We found little support for the predicted patterns of invasion rates in CD4⁺ depleted mice (Figure 4.3b) as there was no general increase in invasion rates of normocytes as compared to reticulocytes across clones. There was a significant interaction between parasite genotype and RBC age ($\chi^2 = 8.243, P = 0.004$; see Table 4.3); the invasion rate of normocytes
Chapter 4. Host exploitation strategies in two strains of rodent malaria  

78

Table 4.2: Analysis of burst sizes across eight treatments

<table>
<thead>
<tr>
<th>Minimal model</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genotype</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>CD4⁺ depletion</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Phenylhydrazine</td>
<td>F₁,₃₃ = 15.164</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Genotype:CD4⁺ depletion</td>
<td>F₁,₃₃ = 5.678</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Non-significant terms deleted from maximal model

<table>
<thead>
<tr>
<th></th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass of mouse</td>
<td>F₁,₃₂ = 2.470</td>
<td>0.126</td>
</tr>
<tr>
<td>Parasite density</td>
<td>F₁,₃₁ &lt; 0.001</td>
<td>0.988</td>
</tr>
<tr>
<td>Uninfected red blood cell density</td>
<td>F₁,₃₁ = 0.034</td>
<td>0.856</td>
</tr>
<tr>
<td>Phenylhydrazine:CD4⁺ depletion</td>
<td>F₁,₂₉ = 1.310</td>
<td>0.262</td>
</tr>
<tr>
<td>Genotype:Phenylhydrazine</td>
<td>F₁,₂₈ = 0.425</td>
<td>0.520</td>
</tr>
<tr>
<td>Genotype:Phenylhydrazine:CD4⁺ depletion</td>
<td>F₁,₂₇ = 3.390</td>
<td>0.059</td>
</tr>
</tbody>
</table>

was higher than the invasion rate of reticulocytes for AS, as predicted, but the opposite was true for DK.

Table 4.3: Analysis of RBC invasion rates in CD4⁺ depleted mice

<table>
<thead>
<tr>
<th>Minimal model</th>
<th>LRT (χ²)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBC age</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Genotype</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Genotype:RBC age</td>
<td>χ₁² = 8.234</td>
<td>0.004</td>
</tr>
</tbody>
</table>

Non-significant terms deleted from maximal model

|                          | χ₁² = 0.380 | 0.538 |

Assuming our model provides an accurate description of the RBC invasion dynamics in immune-intact and PHZ-treated mice, we estimated invasion rates multiplied by merozoite lifespan, β/µ, of both reticulocytes and normocytes for all treatment groups (Figure 4.8). Overall, CD4⁺ depletion resulted in significantly higher invasion rates of both normocytes
and reticulocytes ($\chi^2_1 = 7.184$, $P = 0.007$; see Table 4.4), which is most likely attributable to an increasing merozoite lifespan in the bloodstream in the absence of a main arm of immunity. Again, there was a significant interaction between parasite genotype and RBC age ($\chi^2_1 = 6.913$, $P = 0.009$) with AS, but not DK, showing an overall increase in invasion rates of normocytes compared to reticulocytes. Finally, there was a significant interaction between PHZ and RBC age ($\chi^2_1 = 29.982$, $P < 0.001$), with PHZ resulting in lower invasion rates of reticulocytes, but higher invasion rates of normocytes.

Figure 4.8: Means of the log$_{10}$-transformed invasion rate x merozoite lifespan measurements across the treatments are shown for DK (circles) and AS (triangles). Open symbols are estimates for reticulocyte (‘retic’) invasion and closed symbols are normocyte (‘normo’) invasion. Error bars, ± 1 s.e.m. For consistency, treatments in addition to infection, are coloured as before. CD4$^+$ depletion resulted in a significant increase in all invasion rates. DK and AS had similar invasion rates of reticulocytes, but DK has a lower invasion rate of normocytes. Finally, PHZ treatment resulted in decreased invasion rates of reticulocytes, but increased invasion rates of normocytes.
Table 4.4: Analysis of RBC invasion rates across eight treatments

<table>
<thead>
<tr>
<th>Minimal model</th>
<th>LRT ($\chi^2$)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBC age</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Genotype</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Phenylhydrazine</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>CD4$^+$ depletion</td>
<td>$\chi^2 = 7.184$</td>
<td>0.007</td>
</tr>
<tr>
<td>Phenylhydrazine:RBC age</td>
<td>$\chi^2 = 29.982$</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Genotype:RBC age</td>
<td>$\chi^2 = 6.913$</td>
<td>0.009</td>
</tr>
</tbody>
</table>

Non-significant terms deleted from maximal model

<table>
<thead>
<tr>
<th></th>
<th>$\chi^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass of mouse</td>
<td>$\chi^2 = 0.123$</td>
<td>0.726</td>
</tr>
<tr>
<td>Genotype:Phenylhydrazine</td>
<td>$\chi^2 = 0.533$</td>
<td>0.465</td>
</tr>
<tr>
<td>Genotype:CD4$^+$ depletion</td>
<td>$\chi^2 = 0.036$</td>
<td>0.849</td>
</tr>
<tr>
<td>Phenylhydrazine:CD4$^+$ depletion</td>
<td>$\chi^2 = 0.258$</td>
<td>0.612</td>
</tr>
<tr>
<td>CD4$^+$ depletion:RBC age</td>
<td>$\chi^2 = 1.507$</td>
<td>0.220</td>
</tr>
<tr>
<td>Genotype:Phenylhydrazine:CD4$^+$ depletion</td>
<td>$\chi^2 = 2.175$</td>
<td>0.140</td>
</tr>
<tr>
<td>Genotype:Phenylhydrazine:RBC age</td>
<td>$\chi^2 = 1.523$</td>
<td>0.217</td>
</tr>
<tr>
<td>Genotype:CD4$^+$ depletion:RBC age</td>
<td>$\chi^2 = 0.022$</td>
<td>0.881</td>
</tr>
<tr>
<td>Phenylhydrazine:CD4$^+$ depletion:RBC age</td>
<td>$\chi^2 = 0.178$</td>
<td>0.673</td>
</tr>
<tr>
<td>Genotype:Phenylhydrazine:CD4$^+$ depletion:RBC age</td>
<td>$\chi^2 = 0.409$</td>
<td>0.523</td>
</tr>
</tbody>
</table>
4.4 Discussion

In Chapter 3, we presented a model of *P. chabaudi* infection in CD4+ depleted mice, and made predictions about how differences in host exploitation strategies between clones lead to their different levels of virulence. We found experimental evidence to support some of those predictions. In particular, the number of merozoites produced per infected RBC was higher for the more virulent clone, AS, and our results suggest that a ‘fecundity’ advantage might be obtained from infecting reticulocytes. While our model predicted that this benefit would arise for AS only, we observed increases in burst sizes in response to PHZ treatment in both clones, although the effect was less pronounced in DK. On the other hand, our predictions for invasion rates were not generally supported, although the more virulent clone, AS, tended to have higher invasion rates in the relatively more abundant normocytes. This result, and the overall higher burst size of AS, could help to explain both its competitive advantage and higher virulence.

There are at least two potential (and non-exclusive) explanations for the overall increasing effect of PHZ on burst sizes. The first is that burst sizes are truly higher in reticulocytes, and the higher proportion of infected reticulocytes in PHZ-treated mice drives the average up. In Chapter 3, we offer several explanations for why burst sizes could be higher in reticulocytes. First, reticulocytes are larger and therefore might allow more room for merozoites. Second, reticulocytes have a longer expected circulation time in the bloodstream (Loeffler et al., 1989), so the relative risk to merozoites of undergoing an additional round of asexual multiplication is decreased. We measured burst sizes at the same time
of day and therefore would have expected that parasites in all infected RBCs had roughly the same amount of time to replicate; however, it is possible that replication could begin earlier in reticulocytes, if their residual intracellular components allow for more efficient remodeling of the RBC (which is required for parasite growth; Haldar & Mohandas, 2007).

The above explanations assume that the age of the RBC a parasite infects imposes some constraint on parasite replication. In other words, parasites employ a fixed strategy, but reticulocytes provide a better quality habitat and infection of these RBCs results in more progeny merozoites. However, the effect of PHZ we observed could also be the result of parasites altering their strategy in response to within-host conditions; for example, burst sizes in reticulocytes and/or normocytes may increase facultatively in response to decreasing RBC densities. If parasites use PHZ and subsequent anemia as a signal for imminent immune attack, a beneficial plastic response could be to increase the number of merozoites produced per infected cell so that, when they are released, some are likely to survive immune attack through sheer force of numbers. Malaria parasites are known to use correlates of anemia, thought also to be correlates of developing fertilization-blocking immune responses, to alter the sex ratios of transmission stages and counter these responses (Paul et al., 2000; Reece et al., 2005). Here, anemia could serve as a cue to invest in quantity of merozoites per schizont in order to mitigate the effects of immune responses that target this parasite stage. If this indirect effect of PHZ produced a sufficiently strong plastic response in parasite burst sizes, it could help to explain why (1) we observed an increase in burst sizes in DK in CD4⁺ depleted, PHZ-treated mice when our model did not
predict this, and (2) the overall difference in burst sizes between clones was not significant in immune intact hosts, which appears to be mostly driven by the PHZ treatment. These results are also suggestive of DK being better able to measure host immune responses, as this clone exhibits a much larger increase in burst sizes in response to PHZ in immune-intact mice versus CD4$^+$ depleted. The change in burst size in response to PHZ treatment need not be adaptive, however. The effects of PHZ, in general, are not well known, and the impact that it might have directly on parasite replication is unclear. With improved methods for differentiating burst sizes between reticulocytes and normocytes, future work could look for variation in response to different levels of anemia to test these imposed constraints and parasite plasticity hypotheses.

Our invasion rate results reveal that parasites have a harder time infecting RBCs in immune-intact mice, likely because of higher merozoite death rates in these hosts. This highlights the putative benefit of increasing burst sizes in response to (cues for) immunity. Comparing the invasion rates in CD4$^+$ depleted mice with the expectations from our model, we find little support for our predictions; however, treating hosts with PHZ resulted in increasing normocyte, and decreasing reticulocyte, invasion rates, producing patterns in both PHZ-treated groups that are more similar to what was expected. It could be argued that invasion rates in these treatments are better resolved, since, with a greater number of circulating reticulocytes, our estimates of the proportion of reticulocytes and of infected reticulocytes (and thus, our calculated invasion rates) might be more accurate. If the effect of PHZ on invasion rates is not due to our resolution on the measurements, it could other-
wise be the result of RBCs entering the blood stream earlier to make up for the drastic loss in densities. An explanation for younger cells being harder to invade in either case is that RBC deformability increases with age (Leblond et al., 1971), and conformational changes to an RBC are key for parasite penetration (Dvorak et al., 1975). All of this remains speculation, however, as our model in Chapter 3 was not designed with PHZ treatment in mind, and the mathematical description of the invasion process might therefore be inaccurate for a PHZ-treated host (or even an immune-intact host). For example, our assumption of no RBC death during the course of infection may be ungrounded for PHZ-treated mice because this treatment results in abnormally high RBC death rates for several days (Savill et al., submitted).

Such deviations from model assumptions might also explain the two mice with negative invasion rate estimates (both infected with DK, CD4+ T cell depleted, and treated with PHZ). A negative invasion rate implies that, for a given parasite density, more RBCs were invaded in the experimental infection than was expected from the model. Loss of RBCs due to natural death could increase our experimental measure of infected proportions of RBCs and this was not accounted for in the model. Our model also allows parasites to invade already infected RBCs, and if there is some biological limitation on these multiple infections, then we might observe more infected RBCs than predicted. Alternatively, assuming the model is a valid description of the real biological processes occurring in all of our different treatment groups, negative invasion rate estimates, and deviations from model expectations in general, could be explained by the strong dependency of calculated inva-
sion rates on our burst size measurements. If these are not accurately captured, then our estimates of invasion rates will be correspondingly in error. We believe our estimates of burst sizes are quite accurate, however, because our culturing and sampling methods were aimed at finding the optimal time for counting merozoites in mature schizonts. A final explanation for discrepancies between our data and model predictions (and a cautionary note about these results in general) is that our experiment provides merely a snapshot of the host exploitation traits of *P. chabaudi* parasites during an infection. These traits may vary over the course of a single infection, as do other life history traits of malaria parasites (Reece et al., 2009). Our model does not allow for within-infection variation, but makes predictions about average host exploitation strategies over the first three weeks of infection. Ideally, we would like point estimates of different traits over the course of an infection from both models and data, thereby allowing one to explore the importance of temporal changes in these traits, but such estimates would make for extremely challenging, and quite possibly impracticable, experiments and model fitting procedures.

Our results, in general, support a main hypothesis of our model; namely, that host exploitation traits are important determinants of virulence. We have also provided some evidence of these traits varying in response to the within-host environment, which is a possible indication that parasites employ plastic host exploitation strategies and can respond adaptively to changes within a given infection. A better understanding of the interactions between parasites, host resources and the within-host environment, and how these vary between clones, is important for studying the evolutionary trajectory of disease outcomes.
of interest (e.g. virulence). In particular, this knowledge would aid in evaluating the potential evolutionary consequences of interventions that are aimed at altering the within-host environment (like drugs or vaccination). Key to these evaluations is a better understanding of whether parasite traits are shaped more by biological constraints or plasticity (Mideo, *in press*).
Chapter 5. LINKING WITHIN- AND BETWEEN-HOST DYNAMICS IN THE EVOLUTIONARY EPIDEMIOLOGY OF INFECTIOUS DISEASES

ABSTRACT

Nested models (also called embedded models) explicitly link dynamical processes that occur at different scales. Recently, there has been considerable interest in linking within- and between-host levels of disease dynamics in the study of pathogen evolution. Here we review the extent to which these nested models have increased our understanding of pathogen evolution. We suggest that, although such models have been useful for determining the nature of tradeoffs between epidemiological parameters and for evaluating the consequences of conflicting selection pressures at different scales, the vast majority of previous results could likely have been obtained without the use of nested models per se. Nevertheless, these models have proven very useful through their highlighting of the impact of within-host disease dynamics on pathogen evolution.

5.1 Introduction

Biological processes occur at several nested levels of organization. Biological communities are comprised of interacting species, and these species are comprised of interacting individuals with differing phenotypes. The phenotype of each of these individuals is affected by the dynamics of their component physiological systems, and these systems are affected by smaller scale biological processes from the cellular down to the molecular level. One of the major challenges in evolutionary biology is to understand the evolutionary causes and consequences of these levels of organization (Maynard-Smith & Szathmary, 1995).

The importance of nested levels of biological organization is particularly apparent in the evolutionary epidemiology of infectious disease. Pathogens can have demographically significant effects at the level of the host population, as a result of their transmission among individuals (see Grenfell & Harwood, 1997, for a review of the importance of metapopulation structure in disease dynamics). At the same time, pathogen transmission is intimately tied to the expression of disease in infected individuals, and this disease expression arises from the within-host dynamics of pathogen replication and its interaction with host defense mechanisms. One important goal of evolutionary epidemiology is to understand how these nested processes affect the epidemiological and evolutionary dynamics of host-parasite interactions.

In this review, we discuss recent theoretical research that uses so-called nested dynamical models to study the evolutionary epidemiology of infectious diseases. Our main objective is to evaluate whether or not the nesting of models per se has been important
in developing our understanding of pathogen evolution. As will be seen, our answer is a qualified ‘yes’. We conclude by pointing to some interesting open questions for which this modeling approach is likely to be useful.

5.2 What are nested dynamical models of pathogen evolution?

The focus of this review is on linking between- and within-host dynamics in the evolutionary biology of infectious diseases. Between-host models of infectious diseases follow the dynamics of disease spread at the level of the host population by tracking the number of susceptible and infected individuals of different types (Anderson & May, 1991; Hethcote, 2000). These models have been widely used to study the evolution of pathogens by allowing for multiple pathogen strains to circulate in the population, and then determining which strains are best able to persist.

Models for the evolutionary dynamics of infection at the within-host level have also been developed, and these typically track the dynamics of the pathogen density of different strains, as well as the state of host defense mechanisms (e.g. density of lymphocytes) over the course of an infection. These within-host models of pathogen evolution have been used, in particular, to study rapidly mutating viruses such as HIV or hepatitis C, and to understand how within-host dynamics affect antigenic escape (Sasaki, 1994), viral diversity (Nowak et al., 1990; De Boer & Boerlijst, 1994) and fitness evolution (Iwasa et al., 2005).

More recently, efforts have been directed towards nesting models of within-host dynamics into models of the between-host dynamics when studying pathogen evolution (Fig-
These are what we refer to as nested models in this review. Notice that this definition excludes models of multilevel selection for pathogen evolution in which there is no explicit dynamic model at the within-host level (for example, models based on the superinfection or coinfection frameworks, e.g. Nowak & May, 1994; May & Nowak, 1995; van Baalen & Sabelis, 1995). We recognize that this distinction is somewhat ambiguous (e.g. models of superinfection implicitly rely on some form of within-host model) but it is the explicit nature of the within-host dynamics that sets recent work apart from other approaches.

Figure 5.1: Schematic of a nested model. A basic nested model is constructed by defining a between-host model, in this case a simple susceptible-infected (SI) epidemiological framework, and linking parameters of that model to within-host dynamics. Some of the potentially important elements and interactions of a within-host model are highlighted.
It is also useful to distinguish two ways in which nested dynamical models might contribute to our understanding of pathogen evolution. The first, and most significant, would be if such models provided insights that were not possible without the explicit nesting of within- and between-host processes. In particular, perhaps some aspects of pathogen evolution cannot be understood without allowing for a reciprocal feedback between processes occurring at different levels of biological organization. The second, and more modest, possibility would be if nested models simply refined our understanding of pathogen evolution by providing an explicit link between the within-host dynamics of a pathogen’s replication and its spread at the between-host population level. In this case, the explicit nesting of models is, in a sense, inessential as there is no reciprocal feedback between the within- and between-host dynamics (Figure 5.2a).

5.3 What have nested models of infectious disease taught us?

The importance of nesting models of within-host dynamics into epidemiological frameworks for between-host processes was recognized over a decade ago (Anderson, 1994). Since then, much has been done in this area and many valuable insights have been gained.

5.3.1 Interdependence of parameters

One important insight gained from nested models is that epidemiological parameters such as disease-induced mortality (virulence), recovery rate, and transmission rate are often interrelated in subtle and complex ways. The majority of current theory for virulence evolution assumes that there exists a tradeoff between virulence and transmission rate across
pathogen genotypes, and builds such constraints into models for the between-host dynamics in a phenomenological way (e.g. Ewald, 1983; Sasaki & Iwasa, 1991; van Baalen & Sabelis, 1995; Frank, 1996).

Recent work has developed an explicit description of the within-host dynamics of pathogen replication and the host’s defensive response, and nested this into the traditional between-host models of pathogen evolution (see Appendix E for a simple example of how this is accomplished). This approach provides an explicit link between the mechanism of pathogen replication (e.g. the host cells or resources used by the pathogen to replicate, and mechanisms by which the host attempts to rid itself of infection) and the virulence and transmission rate of the pathogen at the level of the host population. Thus, the nature of any tradeoffs between virulence, recovery rate, and/or transmission rate can be directly linked to mechanistic hypotheses of within-host dynamics. Indeed, it has been demonstrated that such tradeoffs can emerge through a variety of interactions between pathogens and both host immune responses (Ganusov et al., 2002; Gilchrist & Sasaki, 2002; André et al., 2003; Alizon & van Baalen, 2005) and host resources (Gilchrist & Coombs, 2006), leading to the evolution of intermediate levels of virulence (e.g. Sasaki & Iwasa, 1991; Antia et al., 1994; Ganusov et al., 2002).

While the above examples have certainly deepened our understanding of pathogen evolution, the nesting of models per se in these examples is inessential. The within-host dynamical model provides an explicit description of how the mechanisms of pathogen replication and host defense gives rise to transmission rates, recovery rates, and virulence,
and these feed into the between-host dynamical model so as to affect the epidemiological dynamics of disease at the level of the host population. There is not, however, any reciprocal feedback from the between-host dynamics back down to the within-host dynamics in these studies (Figure 5.2a). In this sense the nesting of models is inessential – the within-host model could be analyzed in isolation from the between-host model, and any important insights drawn from it then simply incorporated into the between-host model in a phenomenological way.

That none of the above-mentioned studies have nested models in an essential way is not meant to be a major criticism, however, since much has still been learned from these. For example, they have demonstrated that the mechanism through which a pathogen causes harm to its host, and how it is transmitted from one host to another, are important determinants of virulence (Ganusov & Antia, 2003; Gilchrist & Coombs, 2006; Frank & Schmid-Hempel, 2008; Alizon & van Baalen, 2008b). Studies have also shown that the optimal virulence is strongly dependent on within-host parameter values, potentially providing an explanation for why transmission-virulence trade-offs have been difficult to demonstrate empirically (Alizon & van Baalen, 2005).

Nested models have also demonstrated that the common practice of treating recovery rate as an independent parameter in between-host models of pathogen evolution is typically not justifiable. A primary focus of between-host models has been on determining how the recovery rate affects the evolution of virulence and transmission (Anderson & May, 1982; Frank, 1996; van Baalen, 1998; Gandon & Michalakis, 2000; Gandon et al.,
A similar analysis of recovery rate has also been used to study the effects of vaccination on virulence evolution (enhanced recovery rate is one way in which vaccines might act; Gandon et al., 2001a). A vaccine is unlikely to simply increase recovery rate, however, and instead it might decrease the time until a significant immune response occurs, or increase the abundance of relevant immune effector molecules. While this might increase recovery rate, it will also change the within-host dynamics in a way that can simultaneously alter transmission rate, virulence, and even any tradeoff between the two (Ganusov & Antia, 2006; André & Gandon, 2006). Therefore, even though the nested nature of these recent studies is inessential, they have nevertheless refined our understanding of how differences in the mode of vaccine action can result in differences in pathogen evolution.

5.3.2 Conflicting selection

Pathogen fitness is determined by processes that occur both within and between hosts. Within a host, pathogens might compete directly for host resources, as well as indirectly due to the presence of a common immune response. The strain that is most successful at this within-host competition, however, need not be the one that is best able to transmit among hosts. The ultimate evolutionary outcome will depend on how these potentially conflicting patterns of selection at the within- and between-host levels is resolved (Levin & Bull, 1994; van Baalen & Sabelis, 1995; Frank, 1996).

The idea of conflicting natural selection at the within- and between-host levels is
broadly supported by a substantial body of experimental work. Serial passage of parasites often results in increased parasite virulence, a phenomenon that may be explained by artificial transmission regimes largely removing the constraints imposed on virulence at the between-host level (see Ebert, 1998, for a review). Similarly, in diverse *Plasmodium chabaudi* (rodent malaria) infections in mice, more virulent clones (i.e., those that induce greater red blood cell or weight loss) will competitively suppress less virulent competitors (de Roode et al., 2005b), and make up a greater proportion of the total parasite burden within a host (Bell et al., 2006).

Nested dynamical models can be useful for elucidating the consequences of these different levels of selection. Such models have been used to show that the shape of the tradeoff between virulence and transmission, mediated by within-host mechanisms, determines whether selection acting on pathogen growth rate pushes this trait to different optima at the within- and between-host levels (i.e., whether selection is in conflict between levels; Gilchrist & Coombs, 2006). Extending nested models to the case of multiple infections has also allowed for the exploration of how selection in response to both within-host competition and between-host transmission is resolved (Coombs et al., 2007; Mideo & Day, 2008) and the conditions under which one level of selection dominates (Boldin & Diekmann, 2008). Further, these models have shown that multiple infections can lead to evolutionary diversification in the pathogen population, with one strain specializing on infecting susceptible hosts and another, more virulent, strain specializing on co-infecting already infected hosts (Boldin & Diekmann, 2008; Alizon & van Baalen, 2008a).
As before, virtually all of the above examples represent cases of inessential nesting of models. Although studying the effects of selection at different scales requires some implicit model of dynamics at each scale, this need not be accomplished through an explicit mathematical tracking of within-host dynamics. For instance, implicit within-host modeling is sufficient to show that adding multiple infections can favour more virulent strains (Nowak & May, 1994; May & Nowak, 1995; van Baalen & Sabelis, 1995), even without a trade-off between transmission and virulence (Bonhoeffer & Nowak, 1994). As far as we are aware, there are only two studies for which nesting is essential: one uses a nested model to explore conflicting levels of selection (Coombs et al., 2007) and the other investigates multiple infections more generally (Alizon & van Baalen, 2008a). In these studies, the within-host dynamics of pathogen replication influence the between-host epidemiological dynamics, just as is in many of the other nested models mentioned earlier. In addition, however, because the between-host epidemiological dynamics affect the relative frequency of different pathogen strains circulating in the host population, they also affect the composition of the inoculum entering newly infected hosts (Coombs et al., 2007) or the order of arrival of co-infecting strains (Alizon & van Baalen, 2008a). These, in turn, influence the within-host dynamics, resulting in a reciprocal feedback between levels of biological organization that cannot be captured without an explicitly nested model (Figure 5.2b).

The reciprocal feedback in Coombs et al. (2007) leads to the interesting new insight that the balance between conflicting selection at the within- versus between-host level can
Figure 5.2: Schematics of inessential and essential nested models. (a) Inessential. Within-host dynamics influence between-host processes but not vice versa. For example, suppose that parasite replication within a host determines the rate of transmission to new hosts as well as the rate of disease-induced mortality. Also, suppose that the availability of target cells and the immune cell dynamics determine the rate of host recovery. In this case, the interaction between levels of biological organization in this nested model occurs only in one direction – from the within-host to the between-host level. Using a nested model in this case allows for the relationships between within- and between-host processes to be defined mechanistically. Nevertheless, from the standpoint of the between-host dynamics, these relationships could just as well have been captured phenomenologically by choosing appropriate mathematical descriptions for transmission, virulence and recovery rate as a function of infection age, without any explicit reference to within-host dynamics. (b) Essential. A model in which nesting is essential is one wherein there is reciprocal feedback between levels of organization. Consider a disease that is transmitted via environmental spores. The within-host dynamics will determine transmission, recovery and mortality rates, and thereby will affect the between-host dynamics, including the number of infected individuals. This, in turn, will affect the number of infective spores in the environment, which might then alter inoculum size. If inoculum size influences the progression of disease within a host (Schmid-Hempel & Frank, 2007), this would then lead to a feedback from the between-host dynamics back down to the within-host dynamics. Nesting of models is therefore essential to capture this reciprocal, dynamic feedback.
be shifted depending upon the timing of infection. In particular, strains that are ultimately destined to be outcompeted within a host might nevertheless reach a very high frequency within hosts early in an infection when target cells are abundant (see Ball et al., 2007). As a result, if there is a premium placed on early transmission during an infection then the effects of within-host selection will be mitigated and a strain that is less competitive within hosts can dominate these early transmission events. Early transmission will be important when future transmission opportunities are likely to be limited by host death, or when there is an abundance of susceptible hosts in the population. Of course the depletion of susceptible hosts at the population level will eventually feed back, making the longer-term competitive success of strains within an infection more evolutionarily important. Interestingly, there is some empirical evidence suggesting that competitive abilities of different strains do, in fact, change during the course of an infection (Bell et al., 2006).

5.4 Future prospects and challenges

We have suggested that, out of all previous studies that use nested models for studying pathogen evolution, in only two of these has this nesting per se been essential. What does this imply about the utility of nested models, and how should such models be used in future research? While we cannot claim to have the definitive answers to these questions, we close by offering a few possibilities.

Although the insights gained from most existing nested models might just as well have been obtained without such explicit nesting, we nevertheless believe that these models have been useful because they have emphasized the importance of the timing of events
during an infection (i.e., a disease’s life history; Figure 5.3) on the epidemiological and evolutionary dynamics of pathogens. Indeed, these studies have highlighted an important, and still unanswered, question in evolutionary epidemiology: what factors govern the evolution of disease life histories? A great deal of theory has been devoted to understanding virulence evolution, but surprisingly little has been done to examine the evolution of other important aspects of disease. Parasite transmission rate, parasite-induced mortality (virulence), and infection recovery rate are the three most important epidemiological quantities related to disease (Anderson & May, 1991). Each of these typically changes quite dramatically over the course of an infection, yet very little theory has been devoted to understanding the evolution of the temporal patterns of these traits. Why, for example, do some diseases, like measles, have substantial overlap in the timing of mortality and transmission during infection, whereas in other diseases, like HIV, these are temporally separated (Figure 5.3)? Furthermore, what factors are most important in driving evolutionary change in these patterns within any given pathogen species?

Nested models can be used to address such questions about the evolution of disease life history. Strain-specific differences in the replication rate of a pathogen, as well as in interactions with host defense mechanisms and resources, will give rise to strain-specific differences in the temporal patterns of transmission, virulence, and recovery during an infection. These effects are explicitly captured by within-host models which, when nested in a between-host model, will allow for predictions to be made about the evolution of disease life history at the host population level.
Figure 5.3: An illustration of disease life history. The term ‘life history’ refers to the age-specific pattern of survival and reproduction of an organism. If we view an infection as the ‘organism’ in question, then a disease’s life history is its age-of-infection-specific pattern of transmission, virulence, and recovery (Day, 2003) Parasites with different disease life histories therefore have very different clinical profiles. Solid lines show the within-host pathogen dynamics and the boxes indicate the approximate timing and duration of transmission and disease-associated mortality (recovery rate is ignored for simplicity). (a) For a disease such as measles, transmission begins with the onset of symptoms and, while the host’s immune response is usually sufficient to rapidly clear the infection, complications can lead to parasite-induced mortality. (b) In contrast, with a disease such as HIV, the timing of transmission and mortality is largely non-overlapping. Transmission can occur throughout the infection, with the highest risk of transmission associated with high viral loads (e.g. during the initial acute phase shortly after infection). Mortality associated with HIV generally occurs much later in the infection. From a host’s perspective, HIV is a chronic infection, characterized by a long asymptomatic period followed by rapid progression towards AIDS.

Such an approach will provide a more holistic theory of parasite evolution, with virulence being treated simply as one of several important aspects of disease life history. Furthermore, it opens the door to making realistic predictions about the evolution of other measures of parasite-induced harm, such as morbidity and case mortality, which typically result from a complex interaction between parasite-induced harm and host defense mechanisms (Gilchrist & Sasaki, 2002; Day, 2002; Day & Burns, 2003; Alizon & van Baalen, 2005; Graham et al., 2005; Day et al., 2007; Brown et al., 2008). Nested models also al-
low for predictions about whether a disease is expected to evolve a chronic or an acute life history, rather than assuming one of the two as in most current theory.

There are also other interesting open questions in evolutionary epidemiology for which nested models will be useful. For example, the nesting of models will be essential for studying the evolution of diseases in which the inoculum size affects the progression of the disease (Schmid-Hempel & Frank, 2007), since this will likely lead to a reciprocal feedback between within- and between-host dynamics (Figure 5.2b). In such cases, either the prevalence of the disease at the host population level, or the specific within-host details of the ‘donor’ at the time of transmission (e.g. pathogen density), will affect the dynamics of the ‘recipient’ host.

Previous results have focused on the fact that the composition of an inoculum can alter the pattern of disease progression as well (Coombs et al., 2007). This might be particularly relevant for diseases such as HIV for which the viral diversity in the early stages of infection has been shown to affect the speed of disease progression (Sagar et al., 2003). Genetic diversity in an infection is determined by a repeated cycle of within-host (e.g. immune pressure, replication and diversification) and between-host processes (e.g. transmission bottlenecks), with the outcome of selection at each level feeding into the other. Capturing all of these processes and their interactions in a nested model could help to answer important open questions about the dominance and pattern of disease progression of particular HIV subtypes (see Arien et al., 2007).

Finally, questions about public health policy might also benefit from the use of nested
models. An important aim of models of disease dynamics is to study the evolutionary responses of pathogens to treatments (Gandon et al., 2001a; Medley, 2002; van Baalen, 2002). If the prevalence of a disease in the host population affects, for example, either the probability of an individual receiving treatment or the efficacy of this treatment, a nested model might be useful for capturing the feedback that would result between the within- and between-host dynamics.

5.5 Conclusion

Linking different scales of biological organization with nested models is an approach that has long been used in ecology. The use of nested models for linking levels of disease dynamics has recently become more common, focusing on nesting a model of within-host dynamics into a model of the between-host epidemiological dynamics. Almost all of the studies that have used nested models of disease dynamics to date have not actually required nesting to accomplish their goals. Despite this, these models have provided important insights about pathogen evolution. Further, when there is reciprocal feedback between these two levels of biological organization, the nesting of models is the only way to capture this effect. The nesting of models will therefore continue to prove useful for tackling open questions in the evolutionary epidemiology of infectious diseases.
Chapter 6. ON THE EVOLUTION OF REPRODUCTIVE RESTRAINT IN MALARIA

ABSTRACT

Malaria is one of the leading causes of death among infectious diseases in the world, claiming over one million lives every year. By these standards, this highly complex parasite is extremely successful at generating new infections. Somewhat surprisingly, however, many malaria species seem to invest relatively little in gametocytes, converting only a small percentage of circulating asexual parasite forms into this transmissible form. In this article we use mathematical models to explore three of the hypotheses that have been proposed to explain this apparent ‘reproductive restraint’, and we develop a novel, fourth, hypothesis. We find that only one of the previous three hypotheses we explore can explain such low gametocyte conversion rates, and this hypothesis involves a very specific form of density-dependent transmission-blocking immunity. Our fourth hypothesis also provides a potential explanation and is based on the occurrence of multiple infection, and the resultant within-host competition between malaria parasite strains that this entails. Further experimental work is needed to determine which of these two hypotheses provides the most likely explanation.

6.1 Introduction

Malaria is responsible for a vast amount of mortality and morbidity in the developing world. The damage done to a human host infected with malaria is caused by merozoites, the asexual bloodstage forms of the parasite. These merozoites infect red blood cells (RBCs), replicate and then rupture the RBC, releasing several progeny merozoites each with the potential to invade another RBC. Occasionally, instead of following this pathway, an infected RBC will produce a single sexually-differentiated form of the parasite (gametocyte). While apparently benign inside a host, gametocytes are responsible for transmitting infections to mosquitoes through bloodmeals. Inside the vector, the parasites undergo the sexual phase of their life cycle, producing infectious sporozoites that can be transmitted to another human host in a subsequent bloodmeal. Thus, gametocytes offer the only opportunities for transmission of malaria infections to new hosts, while asexual forms of the parasite allow for the maintenance of a given infection within a host.

Were it not for the conversion of some infected RBCs to gametocytes, individual parasite ‘lines’ would be destined to die within a single host. Further, transmission of malaria to mosquitoes is generally positively related to gametocyte density within a host (Robert et al., 1996; Drakeley et al., 1999; Collins & Jeffery, 2003); but despite their crucial role in transmission, gametocytes are vastly outnumbered by asexuals (reviewed by Taylor & Read, 1997). It has been estimated, in vitro, that between 0 - 20% of infected RBCs go on to produce gametocytes (Carter & Miller, 1979), and data collected from 113 human infections suggests a geometric mean of only 0.64% (Eichner et al., 2001).
Considering that malaria parasites would seem to benefit, evolutionarily, from high levels of gametocytes relative to the more harmful (to the host) asexuals, Taylor & Read (1997) posed the question: why is the gametocyte conversion rate so low? The first step towards an answer is recognizing the tradeoff between replication and transmission (as described for parasites with two life-cycle stages by Koella & Antia, 1995). Converting to gametocytes at a high rate reduces the number of asexuals available for the next bout of RBC invasion, likely reducing the total number of RBCs infected over the course of an infection, and limiting total gametocyte production. On the other hand, while investing a lot in asexual proliferation early on provides a large pool of potential gametocytes later in an infection, the benefit of continuing this investment would be mitigated once a host’s immune system was waging its attack on the parasites; the optimal strategy when clearance of the infection is imminent would surely be to convert everything to gametocytes. Animal models provide evidence of increasing conversion to gametocytes under stress (e.g., Buckling et al., 1999), but even the maximum estimates suggest only about 10% of infected cells produce gametocytes. It would seem that malaria parasites could produce more transmission stages with modest increases in conversion rates.

Assuming available methods of detecting gametocytes are reliable, one is left to wonder why conversion rates and gametocyte densities are so low while the densities of harmful asexuals are relatively high. Taylor & Read (1997) offer two broad explanations: either conversion rates are high and large numbers of gametocytes are produced but quickly removed by some host immune response, or evolution has favoured low conversion rates and
Chapter 6. On the evolution of reproductive restraint in malaria

the apparent reproductive restraint of malaria. We leave for the empiricists the possibility of an elusive arm of immunity disposing of large numbers of gametocytes, and instead focus on the question of reproductive restraint.

6.1.1 Why might reproductive restraint be adaptive?

Taylor & Read (1997) put forth three reasons why it could be selectively advantageous to have low gametocyte conversion rates: (i) higher densities of gametocytes may ensure transmission to a mosquito through a bloodmeal, but may subsequently decrease the vector’s survival, (ii) gametocyte densities remain low to avoid eliciting specific transmission-blocking immune responses, and (iii) gametocyte densities remain low relative to asexual densities so that transmissible forms of the parasite are masked from non-specific immune responses by asexual forms. McKenzie & Bossert (1998) have also modeled another possibility, namely that apparent competition between parasites in a multiple infection, via the immune responses that they elicit, can select for low conversion rates. Our purpose here is to formulate the first three of these hypotheses precisely, in terms of mathematical models, and to determine the conditions under which each of them predicts low gametocyte conversion rates. We also develop a fourth hypothesis, related to that of McKenzie & Bossert (1998), but that focusses on the competition for access to red blood cells that occurs in a multiple infection. We formulate all four hypotheses within a common framework to clearly illustrate the connections among them, and we show that only two of the four hypotheses provide plausible evolutionary explanations for low conversion rates: (i)
some forms of relative density-dependent transmission-blocking immunity, and (ii) multiple infection, either through the process of apparent competition presented by McKenzie & Bossert (1998), or through competition for red blood cells as modeled here. These findings should help to sharpen the focus of empirical research on this important issue.

### 6.2 Model and approach

All of the results presented here are based on a modified version of the Ross-Macdonald model for malaria transmission (Macdonald, 1957; Koella, 1991) which assumes no acquired host immunity, but allows for host mortality. We suppose that the host population is homogeneous with respect to contact rates with mosquitoes and with respect to the probability of becoming infected when bitten by an infected mosquito. Biting rates of infected and uninfected mosquitoes are equal and the size of bloodmeals is constant. Finally, we make the simplifying assumption that the epidemiological parameters reflecting transmission and clearance rate in infected humans are constant during an infection. This is clearly not the case for malaria, but it greatly simplifies our analysis without sacrificing the important qualitative features under investigation. As such our analysis should be viewed as an examination of the validity of each of the four proposed hypotheses rather than an attempt to make quantitative predictions about gametocyte conversion rates.

We use an invasion analysis approach for modeling malaria evolution. Specifically, we suppose that a single strain of malaria is present in the population and examine the conditions under which a mutant strain with different properties can invade (Otto & Day, 2007). Of particular interest are strains that, once present, can exclude all other possible
mutants (i.e., evolutionarily stable strains). For the Ross-Macdonald model it can be shown (Appendix G) that the evolutionarily stable strain is the one that maximizes the following basic reproductive number,

$$R_0 = \frac{ma^2bpe^{-\mu T}}{(r + d)\mu}$$

(6.1)

where $m$ is mosquito density per human, $a$ is the biting rate, $b$ is the transmission rate from an infected host to a susceptible mosquito, $p$ is the transmission rate from an infected mosquito to a susceptible host, $\mu$ is the death rate of mosquitoes, $T$ is the incubation period of malaria within a mosquito, $r$ is the recovery rate of infected hosts, and $d$ is the death rate of humans.

Each of the hypotheses proposed for the evolution of reproductive restraint assumes different relationships between model parameters and the parasite traits of interest. To formalize these relationships we denote the total number of RBCs infected during an infection by $A(\epsilon, \phi)$, where this is written as a function of both the proportion of infected cells that are converted to gametocytes, $\epsilon$, and the within-host asexual growth factor, $\phi$. Formally, $\epsilon$ is a proportion but it is often incorrectly referred to as a ‘rate’ in the empirical literature, and we retain this convention here for consistency. The growth factor, $\phi$ represents the per parasite rate of invasion of RBCs multiplied by the burst size of an infected RBC and the expected lifespan of a merozoite in the bloodstream. For any value of $\epsilon$, increasing $\phi$ results in an increased number of RBCs infected during the course of infection. The total number of gametocytes produced is $G = \epsilon A(\epsilon, \phi)$. Each of the hypotheses can then be formalized by specifying the assumed relationships between the parameters in equation...
6.1 and $A$ and $G$. Malaria strains will be assumed to differ in their values of $\epsilon$ and $\phi$, which produces different values of $A$ and $G$, and therefore different values of $R_0$ in equation 6.1. Since evolution maximizes the basic reproductive number, we can determine the direction that selection acts on $\epsilon$ and $\phi$ by differentiating equation 6.1 with respect to these variables (Otto & Day, 2007). Our aim is to find the conditions under which evolution favours low rates of conversion to gametocytes.

### 6.3 Results

All four hypotheses assume that transmission to mosquitoes, $b$, is an increasing function of $G$, as without a benefit to increasing gametocyte production it would make no sense to ask why there are so few. This assumption is therefore implicit in all analyses below.

#### 6.3.1 Mosquito survival

The first hypothesis is that low numbers of gametocytes may be selectively advantageous if ingesting gametocytes leads to an increased risk of death in the mosquito vector, i.e., if $\mu$ is an increasing function of $G$. Without defining explicit forms of the functions $b(G)$ and $\mu(G)$, we can rewrite equation 6.1 to show these functional relationships as

$$R_0 = \frac{ma^2 b(G) p e^{-\mu(G)T}}{(r + d) \mu(G)}.$$

(6.2)

Evolution is expected to produce the values of $\epsilon$ and $\phi$ that maximize equation 6.2. The dependence of equation 6.2 on the traits $\epsilon$ and $\phi$ enters only through the dependence of $G$ on these traits, and therefore we obtain the same condition for evolutionary stability for
Chapter 6. On the evolution of reproductive restraint in malaria

Taking the derivative of $R_0$ with respect to each trait and setting it equal to zero, we see that the optimal values $\epsilon$ and $\phi$ must produce a gametocyte level, $G^*$, that satisfies the equation

$$
\mu(G^*) \left. \frac{db}{dG} \right|_{G=G^*} - b(G^*) \left. \frac{d\mu}{dG} \right|_{G=G^*} (\mu(G^*)T + 1) = 0.
$$

(6.3)

The actual value of this optimal gametocyte production level, $G^*$, will depend on the forms of the functions $b(G)$ and $\mu(G)$. Thus, gametocyte levels can be high or low depending on these functions. Regardless of what this optimum is, however, from the standpoint of conversion rates there is typically a continuum of values of $\epsilon$ and $\phi$ that give rise to the same level of gametocytes, $G^*$. In particular, for each growth factor there are typically two conversion rates that obtain the optimal $G^*$, one high and one low. The reason is that, for a fixed value of $\phi$, the same total number of gametocytes can be produced by having a low conversion rate, which results in a high total amount of parasite replication, or vice versa (Appendix H). Consequently, selection does not unequivocally favour reproductive restraint. Rather, an entire set of paired conversion rates and growth factors should be selectively neutral (Figure 6.1). This suggests that, barring other constraints, we would expect to see different conversion rates in different populations of the same species of malaria. It should be noted, however, that we have neglected the fact that strategies resulting in high levels of asexuals (i.e., those with low conversion rates) will likely cause their hosts to suffer the highest mortality rate. This suggests that, if any directional evolution on conversion rate is expected under this hypothesis, it would be towards high conversion rates (and low growth factors).
Chapter 6. On the evolution of reproductive restraint in malaria

(a) Total gametocyte production as predicted by the within-host model described in Appendix H at steady state (with $\theta=1$, $\eta=0.01$, $\zeta=2$). The total number of gametocytes, $G$, is the product of conversion rate, $\epsilon$ and total RBCs infected, $A(\epsilon, \phi)$. As conversion rate increases, fewer merozoites are available to infect RBCs, so $A$ decreases with $\epsilon$, resulting in the ‘hump’ shape of $G$. The total number of RBCs infected increases with $\phi$. In response to selection pressures, evolution favours a certain optimal level of gametocyte production, $G^*$, denoted by the gray line. For each $\phi$ we see two values of $\epsilon$ that obtain $G^*$ gametocytes. Here, $\phi_1 > \phi_2 > \phi_3$. (b) Extending these results we find that a continuum of trait pairs give rise to $G^*$ gametocytes.

6.3.2 Density-dependent transmission-blocking immunity

Low gametocyte production may be selectively advantageous if specific immune responses are elicited by gametocytes in a density-dependent manner. While transmission to mosquitoes should directly increase with gametocytes, high densities may inhibit successful transmission by eliciting a strong, specific immune response. To formalize this idea, we suppose that ‘immune cell activity’ against gametocytes is an increasing function of total gametocyte production, $G$, and that the transmission rate, $b$, is a function of both $G$ and immune cell activity, $I(G)$. For completeness, we keep mosquito mortality as an increasing function of $G$ as well, but the same qualitative results are obtained if we leave out this dependency.
With these functional relationships, equation 6.1 is written as

\[ R_0 = \frac{ma^2 b(G, I(G)) p e^{-\mu(G)T}}{(r + d)\mu(G)}. \]  

(6.4)

Again, the dependence of equation 6.4 on \( \epsilon \) and \( \phi \) enters only through the dependence of \( G \) on these traits, and therefore we again obtain a single condition for evolutionary stability of both traits:

\[ \frac{\partial b}{\partial G} \mu - b \frac{d\mu}{dG} (\mu T + 1) + \mu \frac{\partial b}{\partial I} \frac{dI}{dG} = 0. \]  

(6.5)

As before, there will be an optimal \( G^* \) that satisfies this equation, but since \( \frac{\partial b}{\partial I} < 0 \) and \( \frac{dI}{dG} > 0 \), density-dependent immune responses select for lower optimal gametocyte numbers than did mosquito survival alone. Nevertheless, these results again show that there will be both high and low conversion rates strategies that give rise to the same value \( G^* \), meaning reproductive restraint isn’t unequivocally favoured under this hypothesis either.

### 6.3.3 Relative density-dependent transmission-blocking immunity

Alternatively, low gametocyte production may be selectively favourable if it is the relative density of gametocytes to asexuals that determines success of transmission-blocking immune responses. In other words, non-specific immune responses with the ability to target both asexuals and gametocytes may be ‘distracted’ by asexuals if they are relatively more abundant, allowing gametocytes to be transmitted. As the ratio of gametocytes to asexuals increases, transmission will decrease. In this case, immune cell activity against gametocytes, \( I \), is an increasing function of \( \epsilon \) because \( I \) is now an increasing function of
Chapter 6. On the evolution of reproductive restraint in malaria

\[ G/A = \epsilon A/A = \epsilon. \] To show this relationship we write the basic reproductive number as

\[ R_0 = \frac{ma^2b(G, I(\epsilon)) pe^{-\mu(G)T}}{(r + d) \mu(G)}. \] (6.6)

Differentiating equation 6.6 with respect to \( \epsilon \) and \( \phi \) now produces two separate equations that must be satisfied by these traits at evolutionary equilibrium, because \( \epsilon \) now appears independently of \( G \) in this equation. The derivative with respect to \( \phi \) gives the condition

\[ \mu \frac{\partial b}{\partial G} - b \frac{d\mu}{dG} (\mu T + 1) = 0 \] (6.7)

which is what we found in the mosquito survival case. The derivative with respect to \( \epsilon \) gives the condition

\[ \frac{\partial G}{\partial \epsilon} \left( \mu \frac{\partial b}{\partial G} - b \frac{d\mu}{dG} (\mu T + 1) \right) + \mu \frac{\partial b}{\partial I} \frac{dI}{d\epsilon} = 0. \] (6.8)

This reveals that, if we were at the evolutionary equilibrium of the mosquito survival case, then selection acts on \( \epsilon \) in a direction given by the sign of \( \mu \frac{\partial b}{\partial I} \frac{dI}{d\epsilon} \). Thus, selection acts to decrease the conversion rate, as postulated by Taylor & Read (1997). In addition, as a consequence of this selection, optimal growth factors will be higher, explaining the high asexual densities that are common in infections.

A criticism of the above formulation of this hypothesis might be that while immune attack of gametocytes will increase with their relative density, the strength of the immune response elicited (in particular, the number of ‘immune cells’ recruited) will increase with the total number of infected cells as well. In other words, it may be more accurate to consider \( I \) as an increasing function of both \( \epsilon \) and \( A(\epsilon, \phi) \). In this case selection no longer
necessarily favors low conversion rates because such conversion rates require a high number of asexuals to reach the optimal level of gametocyte production, and such strategies are selected against by the increased immune response that they elicit. Therefore, to determine whether or not relative density-dependent transmission-blocking immune responses can unequivocally select for reproductive restraint through low conversion rates, we would need to know the precise details of the action of the immune mechanism.

6.3.4 Multiple infection

To this point, our analysis has focused on infections generated by a single ‘strain’ of a particular malaria species. If, however, infections are composed of multiple strains (or multiple species) of malaria, these strains may be in direct competition for both host resources (i.e., RBCs) and access to mosquitoes. Many studies have documented the high prevalence of people infected with multiple *Plasmodium* species, (e.g. Snounou et al., 1993; Bruce et al., 2000; Ebrahimzadeh et al., 2007) and with multiple strains of the same species, (e.g. Paul et al., 1995; Babiker et al., 1999; Konaté et al., 1999; Bruce et al., 2000; Engelbrecht et al., 2000; Magesa et al., 2002; Cole-Tobian et al., 2005). In stable endemic regions, where malaria transmission is consistently high, such as Nigeria (Engelbrecht et al., 2000), Tanzania (Babiker et al., 1999), and some parts of Senegal (Konaté et al., 1999) it has been estimated that between 70 and 90% of all infections harbour multiple strains.
McKenzie & Bossert (1998) examined the potential for multiple infections to select for low gametocyte conversion rates, under the assumption that different strains interact indirectly through a shared immune response. They found that low conversion rates were optimal, provided that the immune response resulted in a form of apparent competition between coinfecting strains. Here we explore an even simpler hypothesis, that multiple infection results in direct competition between strains for access to red blood cells, and this drives the evolution of lower conversion rates.

There are two main theoretical approaches for exploring multiple infections. A coinfection framework allows multiple strains to coexist within a host (May & Nowak, 1995) whereas a superinfection framework (Nowak & May, 1994) assumes that multiple infections always result in one strain immediately excluding the other from a host, owing to competition between the two. While coinfection is certainly closer to reality for malaria, an assumption of superinfection still accounts for the main qualitatively important evolutionary consequences of multiple infection with respect to conversion rates. Superinfection also results in much more transparent mathematical analyses, and since our main purpose is to use the models to draw out important conceptual aspects of multiple infection, it makes the most sense to employ this assumption here (qualitatively similar findings result from a model of coinfection; Mideo and Day, unpublished results).

The hypothesis is that competition for red blood cells alone can explain the evolution of low gametocyte conversion rates. To this end, we ignore any immune response of the host, and suppose that within-host dynamics of malaria are solely determined by the exploitation
of red blood cells. In this case, the strain with the largest within-host reproduction number will ultimately exclude all others from the infection (see Appendix H). Using $\omega$ to denote the number of merozoites produced by a single infected RBC, $\beta$ to denote the invasion rate of RBCs, and $\delta$ to denote the death rate of free-living merozoites, the strain that wins the within-host competition is that with the largest value of

$$\kappa_i = \phi_i (1 - \epsilon_i) \quad (6.9)$$

where a strain’s within-host growth factor is given by

$$\phi = \frac{\omega \beta}{\delta} \quad (6.10)$$

(see Appendix H). From equation 6.9 it can be seen that within-host competition always favours strains with high growth factors and low conversion rates.

This within-host competition model can be readily embedded in the theoretical framework developed for the previous hypotheses using the assumption of superinfection. Specifically, we suppose that the likelihood of a strain being replaced upon multiple infection depends on the difference between the new strain’s and the existing strain’s ‘competitiveness’. Thus, the likelihood of strain $i$ replacing strain $j$ is given by $\sigma(\kappa_i - \kappa_j)$, where $\sigma$ is an increasing function of its argument.

As shown in Appendix I, when multiple infections occur, the fitness of a mutant (strain 2) in a resident population (strain 1) is

$$R_0(G_2, G_1) = \frac{ma^2b_2(G_2)p e^{-\mu_2(G_2)t}}{\mu_2(G_2)} \frac{a^* + y_1^* \sigma(\kappa_2 - \kappa_1)}{r_2 + d + h_1^* \sigma(\kappa_1 - \kappa_2)} \quad (6.11)$$
where, in the absence of superinfection, \( x^* \) and \( y_1^* \) are the equilibrium densities of susceptible hosts and individuals infected with strain 1, and \( h_1^* \) is the equilibrium inoculation rate of strain 1. The first factor in the above fitness expression is equivalent to expression 6.2 from the mosquito survival hypothesis, up to a multiplicative constant. Therefore, the effect of multiple infections is given by the second factor. Since \( \sigma \) is an increasing function, we can see that multiple infection always favours mutant strains with large values of \( \kappa \), meaning low conversion rates and high growth factors (Figure 6.2).

![Figure 6.2](image)

**Figure 6.2:** The effect of multiple infection on the optimal conversion rate. Evolutionarily stable conversion rate strategies as predicted by the model of superinfection in Appendix I assuming steady state gametocyte production as described in Appendix H. (Here, \( \mu(G) = \mu_0 + \mu_1 G \), \( b(G) = \frac{G}{\alpha + G} \), \( \sigma(\kappa_i - \kappa_j) = \sigma_1 (1 + \tanh(h(\kappa_i - \kappa_j))) / 2 \), \( m=1000 \), \( a=1000 \), \( p=1 \), \( \mu_0 = 10^{-4} \), \( \lambda=1 \), \( \alpha=1 \), \( \tau=10 \), \( r=0.01 \), \( d=0.001 \), \( \theta=1 \), \( \eta=0.01 \), \( \zeta=2 \), \( \sigma_1=1 \).) Under these conditions, selection always acts to increase \( \phi \). Therefore, the equilibrium value of \( \phi \) will always be \( \phi_{\text{max}} = 10 \). The parameter \( h \) describes the relationship between the the likelihood of superinfection and the outcome of within-host competition. When \( h=0 \), superinfection occurs at a constant rate irrespective of the within-host growth factors of the competing strains. Superinfection in this case does not result in selection at the within-host level. As \( h \) increases, the likelihood of superinfection becomes increasingly dependent on the competing strains’ within-host growth factors, generating selection at this level. The equilibrium conversion rate therefore decreases as \( h \) increases. Also, as the cost of gametocytes goes up (by increasing mosquito mortality \( \mu_1 \)), lower conversion rates are favoured by selection.
This makes intuitive sense since it is only the asexual forms that play any role in competition for access to red blood cells. Thus, although a continuum of pairs of conversion rates and growth factors, \((\epsilon, \phi)\), can yield an optimal level of gametocyte production in singly infected hosts, if such a host should ever become multiply infected, it is those strains from this continuum that have a lower conversion rate and a higher growth factor that will do best. In the context of our assumption of superinfection, this is because the strain with a lower conversion rate will most often immediately exclude the other before it has any chance of further transmission. Under an assumption of coinfection a similar process occurs, but in this case the strain with the lower conversion rate does best, not by immediately excluding further transmission by the other strain, but by simply reducing the output of the other strain via competition for red blood cells during the remainder of the infection (Mideo & Day, unpublished results). Therefore, quite generally, the occurrence of multiple infections can drive the evolution of conversion rate to low values, solely through its effects on competition for access to red blood cells.

### 6.4 Discussion

Our results show that evolution can favour reproductive restraint in malaria. While all of the mechanisms offered by Taylor & Read (1997) to explain reproductive restraint can generate selection for low numbers of gametocytes, in only one case does this come about via selection for low conversion rates. Of the hypotheses we explore, multiple infection is the most likely explanation for the evolution of reproductive restraint in malaria.

Our results suggest that an increased risk of mosquito mortality associated with game-
tocytes does not clearly favour low conversion rates within a host. The empirical evidence for the mosquito survival hypothesis has itself been equivocal. While malaria has been shown to cause damage to its vector (reviewed by Ferguson & Read, 2002b), studies have found mosquito mortality to be unrelated to gametocyte density (Robert et al., 1990; Ferguson & Read, 2002a). How gametocyte density in a bloodmeal relates to mosquito mortality at later stages of the parasite’s sexual cycle remains unclear. Mosquitoes harbouring sporozoites suffer greater mortality associated with feeding than uninfected mosquitoes (Anderson et al., 2000), but only a limited number of gametocytes can be converted to sporozoites and early on in development parasite numbers are checked by apoptosis (Al-Olayan et al., 2002). Regardless of the shape of the relationship between gametocyte densities in a bloodmeal and sporozoite burden later in an infection, our results suggest that mosquito mortality could result in low optimal numbers of gametocytes, but is not sufficient for generating selection to achieve these low levels with a low conversion rate.

Similarly, a recent study estimated the relationship between gametocyte density and infection rates of mosquitoes and found an upper threshold above which infection rates leveled off (Paul et al., 2007). The authors conclude that the lack of further benefits to increased gametocyte density would favour reproductive restraint. While this might again explain low optimal numbers of gametocytes, this could be generated by a continuum of trait pairs, $\epsilon$ and $\phi$, including relatively higher conversion rates and lower growth factors.

Transmission-blocking immune responses to specific gametocyte antigens have been described (e.g. Healer et al., 1999) and could impose a strong selective force against high
Chapter 6. On the evolution of reproductive restraint in malaria

gametocyte production. Our results show that indeed, selection in response to this sort of immunity leads to lower optimal production of gametocytes but, again, not necessarily to low conversion rates. Transmission-blocking immune responses are generally considered to be mechanisms operating within the vector, for example by blocking fertilization. In our model, we assume that it is specifically the transmission of gametocytes to mosquitoes that is blocked by these immune responses though the precise timing of their action should not qualitatively change our results.

We have shown that relative density-dependent transmission-blocking immunity may generate selection for low conversion rates, but that this conclusion is dependent on the precise details of the immune response. In particular, it depends on the relative importance of immune stimulation versus immune evasion. Taylor & Read (1997) offer relative density-dependence as a slight variation on the immune pressure hypothesis and in an argument against its plausibility, cite Sinden (1991) who showed that infectiousness to mosquitoes actually decreases with increasing asexuals. Other studies, however have failed to find any relationship between asexual density and transmission success (Drakeley et al., 1999; Paul et al., 2007). In any case, the relationship between relative gametocyte density and transmission may be slightly more subtle. As transmission success depends on gametocyte maturity (Hallett et al., 2006), one of the biggest constraints on transmission might be the ability of gametocytes to successfully develop inside a host. Maturation of gametocytes takes around 10 days to complete during which time the dynamics of infection and immune activity are ongoing. Piper et al. (1999) have suggested that cross-stage
immune responses against PfEMP-1 may help explain the low gametocyte numbers seen in malaria infections. PfEMP-1 is an antigen that is expressed both on the surface of infected RBCs in which asexuals are developing and on the surface of infected RBCs which harbour early stages of gametocytes. In later stages of development, gametocytes lose this antigen and so are protected from immune responses directed towards it (Hayward et al., 1999). While low gametocyte numbers relative to asexuals may do little to prevent activating this anti-PfEMP-1 response (considering the vast numbers of asexuals stimulating immunity) their relative scarcity would provide early stage gametocytes some shelter from these immune responses. With transmission success depending on maturity, and gametocytes being more likely to survive the early stages of their development with many asexuals masking them from immune responses, the conditions would seem to be right for selection to favour low conversion rates.

Multiple infection has been suggested previously as an explanation for the apparent reproductive restraint of malaria and McKenzie & Bossert (1998) use numerical simulations to show that low conversion rates are favoured in multi-strain infections. In their model, all immune activity is targeted against asexual parasite forms and is fully cross-reactive between strains. A strain that invests more in asexuals will be primarily responsible for eliciting the immune response, but this response will affect all strains. Thus, a strain with a low conversion rate that produces many asexuals can reduce the success of a competitor with fewer. Our results are similar in that multiple infection can select for lower optimal conversion rates, but suggest that in addition to ‘apparent competition’ mediated through
a common immune response, competition for resources (i.e., RBCs) alone can provide the selective force. In particular, evolution favours strains with high growth factors and low conversion rates because these are best able to exploit the supply of RBCs within the host. This finding adds considerable further weight to the idea that multiple infections are responsible for low conversion rates since this mechanism does not even require specific assumptions about the nature of the immune response (e.g. target of response, level of cross-reactivity, etc).

Both of these competition theories suggest that in high-transmission areas, where a considerable fraction of infections are composed of multiple strains, we should be able to measure lower conversion rates in infected hosts. We are unaware of any studies that have specifically compared this trait in high- and low-transmission settings. However, a recent review (Drakeley et al., 2006) cites evidence of high-transmission areas harboring fewer gametocyte carriers among infectious individuals and a lower gametocyte to asexual parasite ratio within hosts. Further, there is more seasonal variation in both of these measures in low transmission areas (see Drakeley et al., 2006, and references therein), suggesting that something is constraining the flexibility in conversion rates in high transmission areas.

The two hypotheses that each provide a plausible explanation for low conversion rates posit that there is some additional benefit to high densities of asexuals beyond future production of gametocytes. In the relative density-dependent transmission-blocking immunity case, asexuals mask gametocytes from immune responses and in the multiple infection case, higher asexual densities are equated with a greater ability to exploit host resources.
and, consequently, a within-host competitive advantage. It should be noted that there are other plausible benefits for high asexual densities which could be incorporated into future theoretical frameworks. For example, greater asexual production will lead to greater host anaemia and other measures of parasite-induced harm. If these ‘symptoms’ of high asexual densities translate to reduced anti-vector behaviour, overall transmission success could be enhanced (Ewald, 1983). Alternatively, in the presence of specific immune responses there could be substantial selection pressure towards generating immune-escaping antigenic variants (Brown & Brown, 1965; Phillips et al., 1997) and increased asexual production could facilitate this.

We have claimed that selection for low optimal numbers of gametocytes does not necessarily lead to selection for low conversion rates, as the total number of gametocytes produced is the product of two traits that could equally be modified (i.e., growth factor and conversion rate). Our conclusions are contingent on this assumption and it is possible that we have overlooked some physiological constraint at play. It may be the case that to produce an optimal number of gametocytes a minimum number of asexuals must be produced, for example if parasites are unavoidably programmed to switch to gametocytogenesis after a certain number of rounds of asexual multiplication. If this threshold number of asexuals were high enough, and selection on asexual investment thus constrained, selection for low gametocyte production would result in low conversion rates. What determines the timing of gametocytogenesis is not completely clear, though putative roles have been established for factors such as host immune responses, host hormones and antimalarial
drugs (reviewed by Dyer & Day, 2000; Talman et al., 2004). In addition, gametocytogenesis is enhanced in vitro in cultures enriched with young RBCs (Trager & Gill, 1992; Trager et al., 1999). These reticulocytes make up 1% of normal blood, but this proportion can rise significantly with anemia as the body flushes the system with new cells to compensate. Anemia was able to predict the presence of gametocytes in patients in Thailand (Price et al., 1999), though the study’s authors suggest that the length of infection may have been a confounding factor. Still, there is some evidence of a critical asexual threshold required to produce conditions favourable to gametocytogenesis. Exactly how high that threshold might be is unclear so we don’t know how much (or if) the evolution of conversion rates is constrained, but indeed such constraints may exist.

Part of the motivation for resolving the case of reproductive restraint in malaria is the potential for new insights on control strategies (Taylor & Read, 1997). One intriguing potential aim of interventions is reducing the occurrence of multiple infections. Experimental multi-strain malaria infections in mice have shown that both immune- and resource-mediated competition may be generating selection for increased virulence in multiple infections (Råberg et al., 2006). Thus by limiting these types of infections, a beneficial evolutionary consequence would be selection for less virulent parasites (Bell et al., 2006). Consistent with this, our results suggest that limiting multiple infection would reduce selection for high growth factors and would have the added benefit of eliminating selection against high conversion rates. While this might seem to suggest an increased risk of transmission, selection on the optimal gametocyte production level, $G^*$, in theory should not
change, so the total transmission potential of an infection could remain the same. The benefit comes in the form of those infections with higher conversion rates destroying fewer RBCs of the infected host, while still achieving the same total gametocyte production.

An interesting line of future investigation would be to see if different conversion rates in other taxa could be explained by varying rates of occurrence of multiple infection. For example, gametocytes in lizard malaria, *P. mexicanum*, can make up the majority of parasites within a host (Bromwich & Schall, 1986). Other Apicomplexans closely related to *Plasmodium*, including *Hematocystis* and *Haemoproteus* spp., have life cycles that do not involve a period of asexual expansion within RBCs at all (Smith et al., 2002). Our results suggest that within-host competition is the main selective force favouring low conversion rates, so the occurrence of multiple infections with these related parasites should be lower than with human malaria. To our knowledge, empirical estimates of clonal diversity in these related infections do not yet exist.
Chapter 7. GENERAL DISCUSSION

“The utility of a malaria model depends not so much on how well a mathematical job has been accomplished as on how well a particular biological question has been translated, how thoroughly each assumption and its consequences have been tested, how carefully the range of relevance has been bounded, how closely descriptions and predictions fit data and the broader purpose, and how much its development has suggested explanations and deepened biological understanding.” – McKenzie (2000)

The research presented in this thesis has provided insights on the biology of malaria as well as on theoretical approaches for studying disease dynamics. We have captured the biological complexity of an experimental system with a simple mathematical model, identified virulence determinants of malaria parasites, observed evidence of phenotypic plasticity in parasite traits within infections, and found a plausible answer to the long-standing evolutionary puzzle of reproductive restraint in malaria. The value of this work reaches beyond these insights, however. Its true test is its ability to motivate novel experiments and provide frameworks for answering new questions. Below I discuss some potential extensions of this work, as well as ongoing research that has already stemmed from it.
7.1 Ongoing research and future studies

7.1.1 The mathematics of immune responses to malaria

A major theoretical challenge left unanswered in this thesis is to explain the dynamics of malaria infections in mice with intact immune systems. Unresolved are the importance of immune responses in acute infections and the best mathematical description of the dynamics of such responses. We are currently working on this challenge, using a similar combination of data and modeling as was presented in Chapter 3. Preliminary results suggest that the simple mathematical descriptions of immune responses that are often used in within-host models are not sufficient for capturing the dynamics of this system (Mideo et al., unpublished results). This project is an important step towards deciphering the relative roles of immune-mediated and resource-mediated control of malaria parasite growth and infection dynamics.

7.1.2 The growth versus reproduction trade-off in malaria

Malaria parasites offer a unique opportunity to explore the ubiquitous evolutionary trade-off between growth and reproduction, since resources can be easily identified as having been allocated either to producing asexual (growth) or sexual (reproduction) parasite stages. In Chapter 6 we showed that malaria parasites may invest very little in reproduction as compared to growth, as a result of selection in response to competition for host resources. This hypothesis is currently being tested with experimental infections of *P. chabaudi* in mice and, so far, the data appears to be supportive – parasites invest less in
transmission stages in multiple infections (Politt et al., unpublished results). Several other questions remain unanswered and could benefit from further theoretical or empirical study (or both).

1. How do host factors affect the resolution of the growth versus reproduction trade-off for malaria? There is evidence that parasites can facultatively adjust their conversion rate, as investment in gametocytes can be altered by administration of drugs (Buckling et al., 1997) and the hormone erythropoietin (Reece et al., 2005). Given this, it is important to understand how host heterogeneity will affect the optimal strategy for gametocyte production. An experimental approach (like the one presented in Chapter 4) could study how gametocyte production is affected by host immunity and anemia. Where malaria is endemic, a large proportion of the hosts that parasites encounter are likely to be anemic, due to repeated infection. The implications of this for transmission are not fully understood, and research on this question could provide valuable insight.

2. What is the optimal timing of investment in gametocytes and what factors determine the temporal dynamics of gametocyte production? While Chapter 6 took a ‘whole infection’ view of gametocyte production, it is clear that the pattern of production varies over the course of an infection (Diebner et al., 2000; Eichner et al., 2001). Within-host models that incorporate a dynamic gametocyte conversion rate could be used to identify optimal patterns of investment throughout an infection (e.g. Koella & Antia, 1995). Further, a combination of models and data could be used to test the role of different putative cues for conversion (e.g. red blood cell density, host hormones, and immune cell densities).
7.1.3 Using models to inform public health policy

A key goal of building mathematical models of infection dynamics is to direct them towards evaluating both the efficacy and the effects of disease control measures. The potential for unintended, negative evolutionary effects has already been realized in the case of drug resistance in many pathogens, but concerns about effects on virulence evolution also abound. Mathematical models are valuable tools for both predicting these effects and identifying promising approaches for avoiding negative outcomes, and can be used to direct empirical research in this area. Here, I describe a relatively straightforward, yet significant, extension of the work in this thesis, which was motivated in part by previous experimental results from *P. chabaudi* infections in mice.

Wargo et al. (2007b) showed that drug-resistant parasites are suppressed by competitively superior drug-sensitive clones in the absence of drugs. Curative drug treatment leads to ‘competitive release’ of the resistant clone (i.e., its growth is enhanced, achieving higher densities than in single infections, and potentially risking greater transmission of the drug-resistant clone). A subcurative drug treatment led to less substantial competitive release, raising the question of whether the spread of drug resistance could be slowed by not completely eliminating competitively superior drug-sensitive parasites. In collaboration with Dr. Andrew Read and Silvie Huijben, I have adapted the two-strain within-host model presented in Chapter 3 to study the outcome of competition between drug-sensitive and drug-resistant *P. chabaudi* clones, in the presence of various drug treatment regimes. With only a very simple modification to that model (i.e., drug action), we generated predictions
for a huge number of different drug doses and treatment protocols; a subset of those have subsequently been validated with experiments. These results suggest that there exist strategies for treating malaria, distinct from a WHO-style policy, which result in equal gains in terms of patient health and transmission, but an improved outlook for the spread of resistance (Huijben et al., unpublished results). Given the experimental finding that virulence and drug sensitivity are negatively correlated in *P. chabaudi* (Schneider et al., 2008), a similar theoretical approach could be used to determine the optimal treatment regime for limiting the evolution of increased virulence.

### 7.2 Epilogue

The research described above and presented in this thesis spans the spectrum from basic to more applied biology. Mathematical models can be useful tools for answering questions at either end (McKenzie, 2000), while the precise mathematical approaches used can vary. To me, one of the most interesting aspects of this thesis is that both simple, schematic models and more specific, detailed models generated important insights and stimulated further empirical research. This points to the value of choosing an appropriate approach based on the question at hand, rather than dogmatically adhering to a given technique. Perhaps such blind adherence has generated the chasm between malaria researchers, discussed in Chapter 1. I suspect a more fair evaluation of the cause of this chasm, though, is the tendency to generalize and extend results beyond their reach. Neither a mouse nor a petri dish is equivalent to a human, and neither will produce the exact same infection dynamics as in clinically relevant infections. Intuitively, one recognizes these differences
and suspends disbelief while extrapolating to human malaria. But can we do better? It is my belief that it is the role of mathematics to bring some clarity here, by providing a common framework that can span systems (i.e., malaria and host species), and which requires clearly defining assumptions about biological processes. Math can’t stop malaria transmission, but by bridging the disjoint groups all working towards the same endgame, it can offer a better understanding of this complex parasite in hopes of improving the outlook for hundreds of millions of people living at risk of malaria infections.
7.3 Thesis summary

1. We develop a new modeling approach that represents a significant improvement to previous work, and apply it to a novel data set from a simplified rodent malaria system in which a main immune component is removed. In doing so, we make substantial progress towards deciphering the relative role of ecological factors in determining malaria infection dynamics, and how they differ among parasite clones.

2. We discover that resource availability and resource competition are important determinants of infection dynamics, and these sorts of mechanisms can explain phenomena often attributed to immune-mediated processes.

3. Using a theoretical approach, we identify malaria parasite traits that potentially underlie the observed differential virulence of clones, and verify these with an experimental approach. We find that more virulent clones produce more progeny parasites per infected cell and tend to have faster rates of RBC invasion.

4. We analytically address the plausibility of various hypotheses to explain how natural selection could favour the production of very few transmissible parasite forms in malaria infections. In contrast to the conventional wisdom, we discover that a specific form of transmission-blocking immunity and the occurrence of multiple infections could each generate the patterns observed.

5. We show that verbal hypotheses pervading the literature often do not stand up to detailed, quantitative scrutiny.
REFERENCES


REFERENCES


REFERENCES 136


REFERENCES


REFERENCES


APPENDICES

Chapter 3 Appendices

A Model derivation

The basic structure of the discrete-time model in tracking merozoite density and RBC density is derived by considering the series of events occurring during a single day. Suppose the number of merozoites and susceptible RBCs of each age class per microlitre of blood at the census point on day \( i \) are given by \( P_i, R_{1,i}, R_{2,i}, R_{3,i} \) and \( N_i \). The next event in the replication cycle is the invasion of susceptible RBCs. We model this as a continuous-time process that occurs quickly, relative to the 24 hour discrete-time burst cycle. Free-living merozoites invade reticulocytes and normocytes at rates \( \beta_R \) and \( \beta_N \), respectively, and experience a natural mortality rate, \( \mu \) while in the bloodstream. We assume that merozoites can infect both susceptible and already infected RBCs, but for simplicity (and similar to Hetzel & Anderson, 1996) we assume that secondary invasions are lost and do not change the behaviour of the primary invasion.

Given these assumptions, we can write differential equations to describe the invasion phase. For \( j=1,2,3 \) (i.e., day post-release of reticulocyte in bloodstream):
\[
\frac{dP}{dt} = -P \left( \beta_R \sum_{j=1}^{3} (R_j + I_j) + \beta_N (N + I_N) \right) - \mu P, \tag{A.1}
\]
\[
\frac{dR_j}{dt} = -\beta_R PR_j, \tag{A.2}
\]
\[
\frac{dN}{dt} = -\beta_N PN, \tag{A.3}
\]
\[
\frac{dI_j}{dt} = \beta_R PR_j, \tag{A.4}
\]
\[
\frac{dI_N}{dt} = \beta_N PN. \tag{A.5}
\]

Pay special attention to the notation in equations (A.1-A.5). The variables in these equations are missing the subscript \(i\) referring to the day in question, because these are, in fact, different variables than the \(P_i, R_{j,i}, N_i, I_{j,i},\) and \(I_{N,i}\) introduced above. The variables in equations (A.1-A.5) above track the dynamics of the different kinds of cells during the RBC invasion phase only, which occurs as one of the events during each day. The variables subscripted with an \(i\), however, track these values from one day to the next. Thus, the initial conditions for the above system of differential equations on day \(i\) are \(P(0) = P_i, R_j(0) = R_{j,i}, N(0) = N_i, I_j(0) = 0\) and \(I_N(0) = 0\). The solution to this system is readily found (see below), allowing us to calculate the number of each type of cell after this invasion phase is complete; i.e., \(R_j(\infty), N(\infty), I_j(\infty)\) and \(I_N(\infty)\) (note that \(P(\infty) = 0\)).

The next event is RBC turnover through death and erythropoiesis. We suppose that a fraction, \(d\), of all RBCs die, then RBC aging occurs, and finally newly produced RBCs enter the age-1 class at a rate that depends on the RBC density \(\tau\) days earlier. Specifically,
we model RBC production in a density-dependent fashion as in Haydon et al. (2003), but
with a time-lag of $\tau$ days between the onset of anemia and the body’s response. This
represents the time it takes for new blood cells to develop in the bone marrow before being
released into the bloodstream. Thus, the number of susceptible age-1 RBCs produced on
day $i$ is $\theta(K - (R_{1,i-\tau} + R_{2,i-\tau} + R_{3,i-\tau} + N_{i-\tau}))$, where $K$ represents the equilibrium
density of RBCs in the absence of both merozoites and natural death and $\theta$ is the proportion
of the RBC deficit that is made up in one day. Other forms of RBC production such as a
saturating Hill function have been suggested (Mackey, 1997), but we use this linear form
of density dependence for simplicity, and we account for a type of saturation of production
by having two separate values of $\theta$; one for when the RBC level is above 50% of its normal
density ($\theta_A$), and another for when it is below 50% ($\theta_0$). We have done all of the analyses
in the main text with a Hill function as well, and this tended to result in qualitatively
reasonable, but statistically poorer fits to the data (N. Mideo, unpublished results). Lastly,
with the above assumptions, the number of susceptible RBCs of age 2 and age 3 on day $i$
becomes $(R_{1,i} - I_1(\infty))(1 - d)$ and $(R_{2,i} - I_2(\infty))(1 - d)$ respectively, and the number
of normocytes becomes $(R_{3,i} - I_3(\infty) + N_i - I_N(\infty))(1 - d)$.

The final event is the bursting of RBCs, and the resulting release of merozoites, at
which point a new census occurs. Using $\omega_R$ and $\omega_N$ as the burst size of infected retouched
cytes and normocytes respectively, the total number of merozoites produced on day $i$ is
then $$(\omega_R \sum_{j=1}^3 I_j(\infty) + \omega_N I_N(\infty))(1 - d)(1 - g),$$ where $g$ is the proportion of infected
RBCs that produce gametocytes rather than merozoites (referred to as the conversion rate).
Thus, the discrete-time system of equations from one day to the next is

\[ P_{i+1} = \left( \omega_R \sum_{j=1}^{3} I_j(\infty) + \omega_N I_N(\infty) \right) (1 - d) (1 - g), \quad (A.6) \]

\[ R_{1,i+1} = \theta(K - (R_{1,i-\tau} + R_{2,i-\tau} + R_{3,i-\tau} + N_{i-\tau})), \quad (A.7) \]

\[ R_{2,i+1} = (R_{1,i} - I_1(\infty))(1 - d), \quad (A.8) \]

\[ R_{3,i+1} = (R_{2,i} - I_2(\infty))(1 - d), \quad (A.9) \]

\[ N_{i+1} = (R_{3,i} - I_3(\infty) + N_i - I_N(\infty))(1 - d). \quad (A.10) \]

Finally, we derive explicit expressions for the number of each type of RBC invaded in a particular cycle by finding general solutions to the differential equations (A.1-A.5). Since the number of total RBCs (uninfected plus infected) remains constant during a bout of invasion (since RBC production and natural death occur in the next step of the discrete cycle) the parasite dynamics do not depend on the dynamics of either. An expression for \( P(t) \) is first found by solving equation A.1. This gives us

\[ P(t) = P_i e^{-t((R_{1,i}+R_{2,i}+R_{3,i})\beta_R+N_i\beta_N+\mu)}. \quad (A.11) \]

Substituting this into equations A.4 and A.5 we find

\[ I_j(t) = R_{j,i} \left( 1 - e^{-t\left(R_{j,i}+R_{2,i}+R_{3,i}\right)\beta_R+N_i\beta_N+\mu}\right) \quad (A.12) \]

for \( j=1,2,3 \), and

\[ I_N(t) = N_i \left( 1 - e^{-t\left(R_{1,i}+R_{2,i}+R_{3,i}\right)\beta_R+N_i\beta_N+\mu}\right). \quad (A.13) \]
Further simplification of these expressions is possible since, given our parameters, within a single day the number of parasites declines to zero. This makes biological sense since free-living parasites have only a matter of minutes to infect a RBC before they will die naturally. Thus, it is a reasonable approximation to assume that the numbers of infected cells just prior to bursting are given by the limits of equations A.12 and A.13 as $t$ goes to infinity. Thus,

$$I_j(\infty) = R_{j,i} \left( 1 - e^{-\frac{P_i \beta_R}{(R_{1,i} + R_{2,i} + R_{3,i}) \beta_R + N_i \beta_N + \mu}} \right)$$ (A.14)

$$I_N(\infty) = N_i \left( 1 - e^{-\frac{P_i \beta_N}{(R_{1,i} + R_{2,i} + R_{3,i}) \beta_R + N_i \beta_N + \mu}} \right).$$ (A.15)

Substituting the solutions for $I_j(\infty)$ and $I_N(\infty)$ into the above discrete-time system (A.6-A.10) gives the complete model (system 3.1) of the main text in Chapter 3.

**B  Competition model with two clones**

We can extend our original discrete-time framework (equations A.6-A.10) to allow for two different clones of parasite, $P_A$ and $P_B$. Densities of reticulocytes (on their $j$th day in the bloodstream) and normocytes infected with parasite $A$ are given by $I_{jA}$ and $I_{NA}$. Reticulocytes and normocytes infected with parasite $B$ are given by $I_{jB}$ and $I_{NB}$. As before we assume that only the first parasite to infect an RBC matters. Subsequent invasions are essentially “lost” so we only have to consider singly-infected RBCs.

The basic model tracking parasite and RBC densities is given by:
\[ P_{A,i+1} = \left( \omega_{RA} \sum_{j=1}^{3} I_{jA} (P_{A,i}, R_{j,i}) + \omega_{N} I_{NA} (P_{A,i}, N_{i}) \right) (1 - d) (1 - g) \] (B.1)

\[ P_{B,i+1} = \left( \omega_{RB} \sum_{j=1}^{3} I_{jB} (P_{B,i}, R_{j,i}) + \omega_{N} I_{NB} (P_{B,i}, N_{i}) \right) (1 - d) (1 - g) \] (B.2)

\[ R_{1,i+1} = \theta (K - (R_{1,i-\tau} + R_{2,i-\tau} + R_{3,i-\tau} + N_{i-\tau})) \] (B.3)

\[ R_{2,i+1} = (R_{1,i} - I_{1A} (P_{A,i}, R_{1,i}) - I_{1B} (P_{B,i}, R_{1,i})) (1 - d) \] (B.4)

\[ R_{3,i+1} = (R_{2,i} - I_{2A} (P_{A,i}, R_{2,i}) - I_{2B} (P_{B,i}, R_{2,i})) (1 - d) \] (B.5)

\[ N_{i+1} = (R_{3,i} - I_{3A} (P_{A,i}, R_{3,i}) - I_{3B} (P_{B,i}, R_{3,i})) (1 - d) + \]
\[ + (N_{i} - I_{NA} (P_{A,i}, N_{i}) - I_{NB} (P_{B,i}, N_{i})) (1 - d) . \] (B.6)

The dynamics of the invasion phase are now described by the following set of differential equations.

\[ \frac{dP_A}{dt} = -P_A \left( \beta_{RA} \sum_{j=1}^{3} (R_{j} + I_{jA} + I_{jB}) + \beta_{NA} (N + I_{NA} + I_{NB}) \right) - \mu P_A \] (B.8)

\[ \frac{dP_B}{dt} = -P_B \left( \beta_{RB} \sum_{j=1}^{3} (R_{j} + I_{jA} + I_{jB}) + \beta_{NB} (N + I_{NA} + I_{NB}) \right) - \mu P_B \] (B.9)

\[ \frac{dR_j}{dt} = -\beta_{RA} P_A R_j - \beta_{RB} P_B R_j \] (B.10)

\[ \frac{dN}{dt} = -\beta_{NA} P_A N - \beta_{NB} P_B N \] (B.11)

\[ \frac{dI_{jA}}{dt} = \beta_{RA} P_A R_j \] (B.12)

\[ \frac{dI_{NA}}{dt} = \beta_{NA} P_A N \] (B.13)
\[
\frac{dI_jB}{dt} = \beta_{RB} P_B R_j \quad (B.14)
\]
\[
\frac{dI_{NB}}{dt} = \beta_{NB} P_B N \quad (B.15)
\]

As before we can find expressions for \( P_A(t) \) and \( P_B(t) \). These expressions are equivalent to A.11 except with strain-specific invasion rates. Given the initial conditions \( R_j(0)=R_{j,i} \), \( N(0)=N_i \), \( P_A(0)=P_{A,i} \), \( P_B(0)=P_{B,i} \), \( I_{jA}(0)=0 \), \( I_{jB}(0)=0 \), \( I_{NA}(0)=0 \) and \( I_{NB}(0)=0 \).

\[
P_A(t) = P_{A,i} e^{-t((R_{1,i}+R_{2,i}+R_{3,i})\beta_{RA} + N_i\beta_{NA} + \mu)} \quad (B.16)
\]
\[
P_B(t) = P_{B,i} e^{-t((R_{1,i}+R_{2,i}+R_{3,i})\beta_{RB} + N_i\beta_{NB} + \mu)} \quad (B.17)
\]

Substituting these results into equations B.10 and B.11 we can find expressions for the \( R_j(t) \)’s and \( N(t) \). Letting \( Y_i=R_{1,i}+R_{2,i}+R_{3,i} \),

\[
R_j(t) = R_{j,i} e^{-\frac{(1-e^{-t(N_i\beta_{NA}+Y_i\beta_{RA}+\mu)) P_{A,i} \beta_{RA}}}{N_i \beta_{NA} + Y_i \beta_{RA} + \mu}} - \frac{(1-e^{-t(N_i\beta_{NB}+Y_i\beta_{RB}+\mu)) P_{B,i} \beta_{RB}}}{N_i \beta_{NB} + Y_i \beta_{RB} + \mu} \quad (B.18)
\]
\[
N(t) = N_i e^{-\frac{(1-e^{-t(N_i\beta_{NA}+Y_i\beta_{RA}+\mu)) P_{A,i} \beta_{NA}}}{N_i \beta_{NA} + Y_i \beta_{RA} + \mu}} - \frac{(1-e^{-t(N_i\beta_{NB}+Y_i\beta_{RB}+\mu)) P_{B,i} \beta_{NB}}}{N_i \beta_{NB} + Y_i \beta_{RB} + \mu} \quad (B.19)
\]

To a good approximation

\[
e^{-t(N_i\beta_{NA}+Y_i\beta_{RA}+\mu)} \approx e^{-t(N_i\beta_{NB}+Y_i\beta_{RB}+\mu)}, \quad (B.20)
\]

which allows us to solve for the numbers of different kinds of infected cells just prior to bursting. As before, we take the limits as \( t \) goes to infinity. Thus,

\[
I_{jA}(\infty) = \frac{P_{A,i} R_{j,i} \beta_{RA} (N_i \beta_{NB} + Y_i \beta_{RB} + \mu)}{N_i (P_{A,i} \beta_{RA} N_i \beta_{NB} + P_{B,i} \beta_{RB} \beta_{NA}) + P_{B,i} \beta_{RB} (Y_i \beta_{RA} + \mu) + P_{A,i} \beta_{RA} (Y_i \beta_{RB} + \mu)} \quad (B.21)
\]

152
\[ I_{NA}(\infty) = \frac{P_{A,i}N_i\beta_{NA} (N_i\beta_{NB} + Y_i\beta_{RB} + \mu)}{N_i (P_{A,i} + P_{B,i}) \beta_{NA}\beta_{NB} + P_{B,i} (Y_i\beta_{RA} + \mu) + P_{A,i}\beta_{NA} (Y_i\beta_{RB} + \mu)} \]

\[ I_{jB}(\infty) = \frac{P_{B,i}R_{j,i}\beta_{RB} (N_i\beta_{NA} + Y_i\beta_{RA} + \mu)}{N_i (P_{A,i}\beta_{RA} + P_{B,i}\beta_{RB}\beta_{NA}) + P_{B,i}\beta_{RB} (Y_i\beta_{RA} + \mu) + P_{A,i}\beta_{RA} (Y_i\beta_{RB} + \mu)} \]

\[ I_{NB}(\infty) = \frac{P_{B,i}N_i\beta_{NB} (N_i\beta_{NA} + Y_i\beta_{RA} + \mu)}{N_i (P_{A,i} + P_{B,i}) \beta_{NA}\beta_{NB} + P_{B,i} (Y_i\beta_{RA} + \mu) + P_{A,i}\beta_{NA} (Y_i\beta_{RB} + \mu)} \]

\[ (B.22) \]

\[ (B.23) \]

\[ (B.24) \]

\section*{C Supplementary results}

We use the maximum-likelihood method of fitting our model variations to the individual mouse data. Details of the methodology are presented in the main text of Chapter 3. The best-fit parameters for a given model maximize the log-likelihood, \( L \), and the maximum-likelihood values are presented in Table C.1. We use the likelihood ratio test to compare the fit of our nested model variations.

As described in the main text, we reject model 8 as the best-fit for the two mouse data sets that chose this model, based on the estimated gametocyte conversion rate, \( g \). Thus, the best-fit model for each mouse is 7 and the best-fit parameters for each individual are given in Table C.2.

We test the goodness-of-fit of the maximum-likelihood parameters of model 7 to the data sets of each individual mouse. The likelihood of these parameters given the data is proportional to \( L_{max} \). For this test we assume that the model and parameters are true
and simulate 1000 artificial data sets by generating RBC predictions and incorporating the error structure described in section 3.2.1. Each of these data sets represent one that could have been the measured one, and we can generate an expected distribution for $L_{max}$ by calculating the likelihoods of the parameters given these artificial data sets. If the observed $L_{max}$ lies within the 95% highest density region (HDR) of its expected distribution then the parameters are considered a good fit and are accepted at the $\alpha = 5\%$ level. In Figure C.1 we plot these distributions and highlight the location of the observed $L_{max}$ within the 95% HDR. The fit is good for each mouse.

We can use the approach for generating artificial data sets described above to approximate the uncertainty in our estimated parameters. We generate 100 artificial data sets for each mouse, which again, could have been the measured one and would have resulted in a slightly different set of maximum-likelihood parameter estimates. With each of these synthetic data sets we redo the parameter-fitting routine, using the best-fit model, and use these maximum-likelihood estimates to generate a probability distribution for each parameter. These distributions are shown in Figures C.2, C.3 and C.4.

By pooling the individual probability distributions for each parameter according to experiment (i.e., AS $10^5$, AS $10^6$, DK $10^6$) we attempt to identify any effects of parasite clone or inoculum size. Median parameter values, according to treatment, are presented in Table C.3 and are shown graphically, along with distributions, in Figure C.5.
Table C.1: Maximum log-likelihoods for each model. A “*” indicates the best-fit model, as chosen by the likelihood-ratio test, for each mouse. (Models that share the same number of fitted parameters, given in parentheses, can be compared directly, i.e. a model with a larger $L_{\text{max}}$ is significantly better. For models that differ by one fitted parameter, twice the difference in $L_{\text{max}}$ values must be greater than 3.84 for a more complex model to provide a significantly better fit. See Chapter 3 for details.)

<table>
<thead>
<tr>
<th>(# fitted parameters)</th>
<th>model 1</th>
<th>model 2</th>
<th>model 3</th>
<th>model 4</th>
<th>model 5</th>
<th>model 6</th>
<th>model 7</th>
<th>model 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>treatment mouse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AS$10^5$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-252.0354</td>
<td>-254.2259</td>
<td>-198.3124</td>
<td>-188.9497</td>
<td>-71.9943</td>
<td>-60.4726</td>
<td>-29.2523</td>
<td>-16.5551*</td>
</tr>
<tr>
<td>4</td>
<td>-41.778</td>
<td>-41.3139</td>
<td>-34.0503</td>
<td>-33.1802</td>
<td>-32.7567</td>
<td>-31.4321</td>
<td>-15.2876*</td>
<td>-17.1742</td>
</tr>
<tr>
<td>AS$10^6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-95.3124</td>
<td>-95.3124</td>
<td>2.5347</td>
<td>3.3676</td>
<td>8.1647</td>
<td>8.1713</td>
<td>21.1703*</td>
<td>20.1022</td>
</tr>
<tr>
<td>5</td>
<td>-87.6311</td>
<td>-88.0733</td>
<td>-33.816</td>
<td>-33.816</td>
<td>-6.2623</td>
<td>-6.2629</td>
<td>0.7489*</td>
<td>0.276</td>
</tr>
<tr>
<td>DK$10^6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-260.4084</td>
<td>-259.8688</td>
<td>-130.6609</td>
<td>-128.4231</td>
<td>-133.3927</td>
<td>-120.12560</td>
<td>-86.8485*</td>
<td>-87.2826</td>
</tr>
</tbody>
</table>
### Table C.2: Estimated parameter values for best-fit model

<table>
<thead>
<tr>
<th>treatment</th>
<th>mouse</th>
<th>$\theta_0$</th>
<th>$\theta_A$</th>
<th>$\beta_R \times 10^6$</th>
<th>$\beta_N \times 10^6$</th>
<th>$\omega_R$</th>
<th>$\omega_N$</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AS10$^5$</td>
<td>1</td>
<td>0.06901</td>
<td>0.18889</td>
<td>1.0889</td>
<td>7.08999</td>
<td>13.8193</td>
<td>9.1826</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.05124</td>
<td>0.128</td>
<td>0.3918</td>
<td>15.06</td>
<td>21</td>
<td>6.072</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.83897</td>
<td>0.5099</td>
<td>9.997</td>
<td>0.9889</td>
<td>11.995</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.05181</td>
<td>0.1309</td>
<td>0.7839</td>
<td>12.0610</td>
<td>19.890</td>
<td>6.030</td>
<td>2</td>
</tr>
<tr>
<td>AS10$^6$</td>
<td>1</td>
<td>0.04942</td>
<td>0.41099</td>
<td>1.0009</td>
<td>16.28</td>
<td>4.649</td>
<td>6.001</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.01877</td>
<td>0.5</td>
<td>0.589</td>
<td>1.489</td>
<td>2</td>
<td>13.889</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.07664</td>
<td>0.5</td>
<td>0.2263</td>
<td>12.206</td>
<td>10.999</td>
<td>7.732</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.5</td>
<td>0.3715</td>
<td>0.8309</td>
<td>3.5002</td>
<td>19</td>
<td>11.021</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>0.0875</td>
<td>0.4491</td>
<td>0.843</td>
<td>11.749</td>
<td>15.89</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>DK10$^6$</td>
<td>1</td>
<td>0.5</td>
<td>0.4389</td>
<td>0.8508</td>
<td>4.104</td>
<td>18</td>
<td>7.111</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.04139</td>
<td>0.5</td>
<td>0.2759</td>
<td>19.825</td>
<td>11.799</td>
<td>5.991</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.1739</td>
<td>0.5</td>
<td>0.725</td>
<td>7.499</td>
<td>9.46</td>
<td>6.014</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>0.04983</td>
<td>0.3889</td>
<td>0.889</td>
<td>10.3891</td>
<td>2.98</td>
<td>8.078</td>
<td>3</td>
</tr>
</tbody>
</table>

### Table C.3: Medians of pooled estimated parameter distributions

<table>
<thead>
<tr>
<th>parameter</th>
<th>AS 10$^5$</th>
<th>AS 10$^6$</th>
<th>DK 10$^6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_0$</td>
<td>0.0759</td>
<td>0.10</td>
<td>0.1312</td>
</tr>
<tr>
<td>$\theta_A$</td>
<td>0.1999</td>
<td>0.52509</td>
<td>0.6196</td>
</tr>
<tr>
<td>$\beta_R$</td>
<td>1.086x10$^{-6}$</td>
<td>1.065x10$^{-6}$</td>
<td>1.001x10$^{-6}$</td>
</tr>
<tr>
<td>$\beta_N$</td>
<td>10.14x10$^{-6}$</td>
<td>11.189x10$^{-6}$</td>
<td>8.861x10$^{-6}$</td>
</tr>
<tr>
<td>$\omega_R$</td>
<td>15.778</td>
<td>8.12</td>
<td>6.21</td>
</tr>
<tr>
<td>$\omega_N$</td>
<td>7.998</td>
<td>7.979</td>
<td>7.0</td>
</tr>
<tr>
<td>$\tau$</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure C.1: Expected distributions of the maximum log-likelihood of the best-fit model given the data, $L_{max}$, for each individual mouse. These distributions are generated by calculating the log-likelihood of the best fit model and parameters given each of 1000 artificial data sets. Blue lines represent the observed $L_{max}$ (log-likelihood of best fit model and parameters given the measured data). In each case, the blue line is within the 95% highest density region of its expected distribution, as indicated by the dashed red lines, therefore the model is a good fit for each data set.
Figure C.2: Boxplots of the estimated distributions of invasion rate for reticulocytes (above) and normocytes (below). Each boxplot contains 100 parameter estimates obtained from refitting simulated data sets (100 data sets per mouse). Horizontal line is the median, the box contains 50% of the values and the circles represent outliers.
Figure C.3: Boxplots of the estimated distributions of burst size for reticulocytes (above) and normocytes (below). Each boxplot contains 100 parameter estimates obtained from refitting simulated data sets (100 data sets per mouse). Horizontal line is the median, the box contains 50% of the values and the circles represent outliers.
Figure C.4: Boxplots of the estimated distributions of RBC production parameters: proportion of RBC deficit made up in one day under normal conditions (above), proportion of RBC deficit made up in one day under anemic conditions (middle), time lag due to maturation of RBC (below). Each boxplot contains 100 parameter estimates obtained from refitting simulated data sets (100 data sets per mouse). Horizontal line is the median, the box contains 50% of the values and the circles represent outliers.
Figure C.5: Boxplots of the combined estimated parameter distributions for the AS $10^5$ (white), AS$10^6$ (gray) and DK$10^6$ (blue) experiments. Each boxplot contains 400 parameter estimates (100 data sets per mouse x 4 mice per dose). Horizontal line is the median, the box contains 50% of the values and the circles represent outliers.
Chapter 4 Appendix

D Sampling protocol

The main samples required to measure burst sizes and invasion rates are described in Figure D.1. We repeated these procedures on two different days, but for reasons outlined in the main text, we analyze only burst size data from day 5 and invasion rates from day 6.

<table>
<thead>
<tr>
<th>Day 5 (and 6)</th>
<th>Time</th>
<th>Sample</th>
<th>Measurements</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>10am</td>
<td>Ring-stage smears</td>
<td>• RBC density (all ages)</td>
<td>1. RBC production and natural death can be ignored.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Density of infected RBCs (all ages)</td>
<td>2. All merozoites are dead or infecting RBCs.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3. All parasitized cells are the product of a single invasion.</td>
<td></td>
</tr>
<tr>
<td>2pm</td>
<td>10μl blood for cultures</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Day 6 (and 7)</th>
<th>Time</th>
<th>Sample</th>
<th>Measurements</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>10am</td>
<td>Schizont stage smears</td>
<td>• # of merozoites per schizont (i.e. burst size)</td>
<td>4. All parasitized cells are the product of a single invasion.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>5. Asexual replication is complete, but infected cells have not burst yet.</td>
<td></td>
</tr>
<tr>
<td>10am</td>
<td>Ring-stage smears</td>
<td>• Normocyte density</td>
<td>Same as above (1-3).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Reticulocyte density</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Infected normocyte density</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Infected reticulocyte density</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: 
- Normocyte
- Reticulocyte
- Infected normocyte
- Infected reticulocyte
- Mature schizont

Figure D.1: Sampling protocol.
Chapter 5 Appendix

E Nested models

Below is a simplified example of how many studies construct nested models of within- and between-host dynamics. To account for changes in transmission, virulence, and recovery rates during an infection, the between-host model must first explicitly track the age of each infection. If \( S(t) \) is the density of susceptible hosts at time \( t \) and \( I(a, t) \) is the density of hosts at time \( t \) that were infected at time \( t - a \) (i.e., \( a \) is the ‘age of infection’), a simple Susceptible-Infected (SI) model is (Anderson & May, 1991; Diekmann & Heesterbeek, 2000; Day, 2003):

\[
\begin{align*}
\frac{dS(t)}{dt} &= \theta - \mu S(t) - S(t) \int_0^\infty \beta(a)I(a,t)da \\
\frac{\partial I(a,t)}{\partial t} &= -\frac{\partial I(a,t)}{\partial t} - v(a)I(a,t) \\
I(0,t) &= S(t) \int_0^\infty \beta(a)I(a,t)da,
\end{align*}
\]

where \( \theta \) is a constant input rate of susceptibles, \( \beta \) is the transmission rate, \( \mu \) is the constant host mortality rate and \( v \) is the loss rate of infected individuals through natural death, parasite-induced death (i.e., virulence), and recovery. Transmission, virulence and recovery are all assumed to vary with the age of the infection \( a \).

In addition to the between-host model (E.1), a within-host model can be constructed to follow the changes in parasite density \( x(a) \), the state of the immune response \( y(a) \), and any resource consumed by parasites \( r(a) \) (e.g. red blood cells) as a function of the ‘age of...
the infection, \( a \). In the simplest cases, this gives a dynamic model of the form

\[
\frac{dx}{da} = f(x, y, r, a) \\
\frac{dy}{da} = g(x, y, r, a) \\
\frac{dr}{da} = h(x, y, r, a).
\] (E.2)

The within-host model (E.2) is then embedded or nested into the between-host model (E.1) by specifying a functional relationship between the transmission rate and mortality rate at infection age \( a \) (i.e., \( \beta(a) \) and \( v(a) \)) and the within-host variables \( x, y, \) and \( r \) (Sasaki & Iwasa, 1991; Antia et al., 1994; Ganusov et al., 2002; Gilchrist & Sasaki, 2002; Ganusov & Antia, 2003). This model can allow for evolution by introducing multiple pathogen strains by, for example, keeping track of the number of hosts infected with each strain (i.e., \( I_j(a, t) \), where \( j \) denotes a particular strain) and by extending the within-host model (E.2) to keep track of different strains as well (Coombs et al., 2007; Boldin & Diekmann, 2008; Alizon & van Baalen, 2008a). In a similar fashion, other complexities such as differences in inoculum size or immune system states among hosts can, in principle, be incorporated by expanding the number of kinds of hosts in model (E.1), each with their own form of within-host dynamics in model (E.2).

Finally, some analyses rewrite system (E.1) in a different notation, expressing the dynamics in terms of mean values of \( \beta(a) \) and \( v(a) \) over the stable infection age distribution (André et al., 2003; André & Gandon, 2006), but the underlying model is the same (Day et al., 2007). Some treatments further simplify the analysis, however, by assuming that the within-host dynamics (E.2) are fast relative to the between-host dynamics (E.1). As such,
the within-host model is then assumed to remain at equilibrium, allowing the between-host parameters $\beta(a)$ and $v(a)$ to be treated as constant during an infection (Alizon & van Baalen, 2005; Gilchrist & Coombs, 2006).
Chapter 6 Appendices

G Evolution maximizes $R_0$

The dynamics of the frequencies of mosquitoes which are infected but not yet infectious, $v$, and inefectious, $w$, are given by

$$\frac{dv}{dt} = aby(1 - v - w) - ab\hat{y}(1 - \hat{v} - \hat{w}) e^{-\mu T} - \mu v \quad (G.1)$$

$$\frac{dw}{dt} = ab\hat{y}(1 - \hat{v} - \hat{w}) e^{-\mu T} - \mu w \quad (G.2)$$

where $a$ is the biting rate, $b$ is the transmission rate from infected humans to susceptible mosquitoes, $\mu$ is the death rate of mosquitoes, $T$ is the incubation period of malaria parasites in the vector, and $y$ is the frequency of infected human hosts. Variables with a hat, $\hat{y}$, $\hat{v}$, and $\hat{w}$, denote proportions at time $t-T$.

The dynamics of the frequencies of susceptible humans, $x$, and infected humans, $y$, are given by

$$\frac{dx}{dt} = d - dx - hx + ry \quad (G.3)$$

$$\frac{dy}{dt} = hx - ry - dy \quad (G.4)$$

where $r$ is the recovery rate from infection, $h$ is the inoculation rate, $d$ represents the rate of host death, and deaths are balanced by births into the susceptible class (i.e., a constant population size is assumed). The inoculation rate is given by

$$h = ma p w \quad (G.5)$$
where \( m \) is the per human mosquito density and \( p \) is the transmission rate from infected mosquitoes to susceptible human hosts. Assuming the mosquito population is at equilibrium with respect to changes in human dynamics, the quasi-equilibrium density of infected mosquitoes, \( w^* \) is

\[
\begin{align*}
w^* &= \frac{abe^{-\mu T}y}{aby + \mu}.
\end{align*}
\]  

(G.6)

The inoculation rate from equation G.5 can then be rewritten as

\[
\begin{align*}
h &= \frac{ma^2bpe^{-\mu T}y}{aby + \mu}.
\end{align*}
\]  

(G.7)

Substituting equation G.7 into equations G.3 and G.4, and solving for equilibrium proportions of infected humans, \( y^* \), and susceptible humans \( x^* \), we find the only non-zero equilibrium occurs where

\[
\begin{align*}
y^* &= \frac{ma^2bpe^{-\mu T} - \mu (r + d)}{ma^2bpe^{-\mu T} + ab(r + d)} \\
x^* &= 1 - y^*.
\end{align*}
\]  

(G.8)

(G.9)

It can be shown that this equilibrium is stable when the ‘infection growth’ rate outweighs the ‘infection loss’ rate, or

\[
ma^2bpe^{-\mu T} > \mu (r + d).
\]  

(G.10)

This condition must be true in order for infections to persist in the population, so for an endemic disease we can assume this condition is satisfied.

We can now imagine adding a mutant parasite, denoted with the subscript \( i \), which differs in its transmission rates and causes different rates of mosquito mortality, host mortality and host recovery. System (G.1-G.2) is altered and the dynamics in the mosquito
population are now described by

\[ \frac{dv}{dt} = aby(1 - v - v_i - w - w_i) - ab\hat{y}(1 - \hat{v} - \hat{v}_i - \hat{w} - \hat{w}_i) e^{-\mu T} - \mu v \]  \hspace{1cm} (G.11)

\[ \frac{dv_i}{dt} = ab_i y_i (1 - v - v_i - w - w_i) - ab_i \hat{y}_i (1 - \hat{v} - \hat{v}_i - \hat{w} - \hat{w}_i) e^{-\mu_i T} - \mu_i v_i \]  \hspace{1cm} (G.12)

\[ \frac{dw}{dt} = ab\hat{y}(1 - \hat{v} - \hat{v}_i - \hat{w} - \hat{w}_i) e^{-\mu T} - \mu w \]  \hspace{1cm} (G.13)

\[ \frac{dw_i}{dt} = ab_i \hat{y}_i (1 - \hat{v} - \hat{v}_i - \hat{w} - \hat{w}_i) e^{-\mu T} - \mu_i w_i. \]  \hspace{1cm} (G.14)

The equilibrium densities of mosquitoes infected with the resident strain, \( w^* \), and with the mutant strain, \( w^*_i \) are

\[ w^* = \frac{abe^{-\mu T} y \mu_i}{aby \mu_i + ab_i y_i \mu + \mu \mu_i} \]  \hspace{1cm} (G.15)

\[ w^*_i = \frac{ab_i e^{-\mu_i T} y_i \mu_i}{aby \mu_i + ab_i y_i \mu + \mu \mu_i}. \]  \hspace{1cm} (G.16)

The human population is now described by the following set of equations

\[ \frac{dx}{dt} = d - dx - hx - h_i x + ry + r_i y_i \]  \hspace{1cm} (G.17)

\[ \frac{dy}{dt} = hx - (r + d)y \]  \hspace{1cm} (G.18)

\[ \frac{dy_i}{dt} = h_i x - (r_i + d)y_i \]  \hspace{1cm} (G.19)

\[ \frac{h}{dt} = \frac{ma^2 bpe^{-\mu T} y \mu_i}{aby \mu_i + ab_i y_i \mu + \mu \mu_i} \]  \hspace{1cm} (G.20)

\[ \frac{h_i}{dt} = \frac{ma^2 b_i pe^{-\mu_i T} y_i \mu_i}{aby \mu_i + ab_i y_i \mu + \mu \mu_i}. \]  \hspace{1cm} (G.21)

To determine if the mutant can invade, we look at the stability matrix of the equilibrium with the mutant absent (i.e., \( y_i^*=0 \)) described by G.8 and G.9. A mutant will be able to
invade a population when this equilibrium is unstable (when at least one of the eigenvalues of the stability matrix is positive). This analysis results in a stability matrix of the following form

$$\begin{pmatrix} J_{\text{res}} & v \\ 0 & J_{\text{mut}} \end{pmatrix}$$  \hspace{1cm} \text{(G.22)}$$

where $J_{\text{res}}$ describes the stability of the equilibrium in the absence of the mutant strategy and

$$J_{\text{mut}} = \frac{ma^2b_ip_i e^{-\mu_i T} x^s \mu}{(aby^s + \mu) \mu_i} - (r_i + d).$$  \hspace{1cm} \text{(G.23)}$$

We have already shown that the eigenvalues of $J_{\text{res}}$ are negative because of inequality G.10. So, a mutant will be able to invade when $J_{\text{mut}}>0$. Substituting in the equilibrium values $y^*$ and $x^*$ (from G.8 and G.9) and simplifying we find that a mutant can invade when

$$\frac{ma^2b_ip_i e^{-\mu_i T} x^s \mu}{(r_i + d) \mu_i} > \frac{ma^2bpe^{-\mu T}}{(r + d) \mu}.$$  \hspace{1cm} \text{(G.24)}$$

Since

$$R_0 = \frac{ma^2bpe^{-\mu T}}{(r + d) \mu},$$  \hspace{1cm} \text{(G.25)}$$
a mutant can invade only if a single host infected with the mutant, in a wholly susceptible population, leads to more secondary infections than a host infected with the non-mutant.

**H  Within-host model of infection**

We present a simplified model of the dynamics of malaria infection within a host, assuming there is no immunity. Assuming a mass action infection rate of RBCs, the dynamics of RBCs, $S$, and two coinfecting parasite populations (tracking their merozoite, $M$, and
gametocyte, \( G \), densities) in the bloodstream is

\[
\frac{dS}{dt} = \theta - \beta_1 S M_1 - \beta_2 S M_2 - \eta S \quad \text{(H.1)}
\]

\[
\frac{dM_1}{dt} = \omega (1 - \epsilon_1) \beta_1 S M_1 - \delta_1 M_1 \quad \text{(H.2)}
\]

\[
\frac{dM_2}{dt} = \omega (1 - \epsilon_2) \beta_2 S M_2 - \delta_2 M_2 \quad \text{(H.3)}
\]

\[
\frac{dG_1}{dt} = \epsilon_1 \beta_1 S M_1 - \zeta_1 G_1 \quad \text{(H.4)}
\]

\[
\frac{dG_2}{dt} = \epsilon_2 \beta_2 S M_2 - \zeta_2 G_2 \quad \text{(H.5)}
\]

where \( \theta \) is a constant influx of new RBCs, \( \beta \) is the strain-specific invasion rate of RBCs, \( \eta \) is the natural death rate of RBCs, \( \omega \) is the number of merozoites produced by an infected RBC, and \( \delta \) and \( \zeta \) are the strain-specific merozoite and gametocyte death rates in the bloodstream. This model is different from previous ones incorporating gametocytogenesis (e.g. Hellriegel, 1992) since we assume that infected RBCs burst immediately (so we do not need to track densities of infected RBCS). This is clearly not the case in real malaria infections, but we are interested in the number of RBCs that are infected over the entire course of infection and the number of gametocytes that are produced, rather than the details of transient within-host dynamics. Our simplification should not qualitatively affect our results as long as different parasite strains take the same amount of time to mature within RBCs.

At steady state it is easy to show that the strain that can reduce the RBC number, \( S \), to the lowest value will competitively exclude the other. The winning strain thus has the highest value of \( \omega \beta (1 - \epsilon) / \delta \).
At steady state, a strain produces
\[ \epsilon \left( \frac{\phi(1 - \epsilon)\theta - \eta}{\phi(1 - \epsilon)} \right) \] (H.6)
gametocytes, where \( \phi = \frac{\omega_\beta}{\delta} \). This expression is used to generate Figure 6.1 and is incorporated into the model of superinfection (Appendix I) to generate Figure 6.2.

Similar qualitative results are obtained with a model displaying nonequilibrium within-host dynamics, as is typical of real malaria infections. In particular, malaria infections are characterized by multiple waves of parasitemia, with the first peak of parasitemia occurring before any significant immune responses have built up against the parasite. We can model this by assuming there is some fixed pool of RBCs available for invasion, and their abundance governs the dynamics of infection at this early stage. In this case, we can use a variant of the above within-host model, with \( \theta = 0 \) and \( \eta = 0 \). We can then ask how the total number of gametocytes produced during this wave of parasitemia changes with our traits of interest, and we obtain results qualitatively identical to those in Figure 6.1. This model also allows one to explore the evolutionary consequences of coinfection (as opposed to superinfection) as well. Again, the qualitative results are similar to those presented in Section 6.3.4.

I Ross-Macdonald and superinfection

Under superinfection, the mosquito equations are unchanged from the two strain model presented in Appendix G (except instead of having a resident and a mutant, which was denoted by the subscript \( i \), we have two strains denoted by the subscripts 1 and 2). Equi-
librium values of \( w_1 \) and \( w_2 \) are equivalent to those in equations G.15 and G.16,

\[
\begin{align*}
  w_1^* &= \frac{a b_1 e^{-\mu_1 T \tau} y_1 \mu_2}{a b_1 y_1 \mu_2 + a b_2 y_2 \mu_1 + \mu_1 \mu_2} \quad (I.1) \\
  w_2^* &= \frac{a b_2 e^{-\mu_2 T \tau} y_2 \mu_1}{a b_1 y_1 \mu_2 + a b_2 y_2 \mu_1 + \mu_1 \mu_2}. \quad (I.2)
\end{align*}
\]

The human population is described by the following set of equations

\[
\begin{align*}
  \frac{dx}{dt} &= d - dx - h_1 x - h_2 x + r_1 y_1 + r_2 y_2 \quad (I.3) \\
  \frac{dy_1}{dt} &= h_1 x + h_1 \sigma (R_1 - R_2) y_2 - (r_1 + d + h_2 \sigma (R_2 - R_2)) y_1 \quad (I.4) \\
  \frac{dy_2}{dt} &= h_2 x + h_2 \sigma (R_2 - R_1) y_1 - (r_2 + d + h_1 \sigma (R_1 - R_2)) y_2 \quad (I.5)
\end{align*}
\]

where \( h_j = \text{map} w_j^* \) and \( \sigma \) is a function that represents the ‘dominance’ (Bonhoeffer & Nowak, 1994) of an introduced strain over one that is established within a host, and this function depends on the difference in the values of \( \kappa \) for the two strains.

In the absence of superinfection, this model reduces to the one presented in equations G.3-G.4. To arrive at an expression for \( R_0 \) in this system we perform an invasion analysis similar to that in Appendix G. Here,

\[
J_{\text{mut}} = \frac{ma^2 b_2 p e^{-\mu_2 T \tau} y_1 \sigma (\kappa_2 - \kappa_1) \mu_1}{(a b_1 y^* + \mu_1) \mu_2} + \frac{ma^2 b_2 p e^{-\mu_2 T \tau} y_1^* \sigma (\kappa_2 - \kappa_1) \mu_1}{(a b_1 y^* + \mu_1) \mu_2} - (h_1^* \sigma (\kappa_1 - \kappa_2) + r_2 + d) \quad (I.6)
\]

The mutant can increase in frequency if \( J_{\text{mut}} > 0 \), which can be rearranged to give the mutant fitness expression of the text.