Infection dynamics at within-host and between-host scales

Maite Severins

ISBN: 978-90-8891-458-4
Printed by Uitgeverij BOXpress
Cover Design by Maite Severins

Infection dynamics at within-host and between-host scales

Infectie dynamica
op binnen-gastheer en tussen-gastheer niveaus
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op vrijdag 14 september 2012 des middags te 14.30 uur

door

Maite Severins

geboren op 19 juni 1979 te Heerlen

Promotor: Prof. dr. J.A.P. Heesterbeek
Co-promotor: Dr. D. Klinkenberg

This study was financially supported by the Netherlands Organisation for Scientific Research (NWO) under grant number 918.56.620.

Contents

1	**General Introduction**	**3**
2	**The role of Th1/Th2 phenotypes in T-cell vaccination: Insights from a mathematical model**	**21**
3	**The role of Type 1 and Type 2 immune responses in cattle against infection with *Mycobacterium avium* subspecies *paratuberculosis*.**	**45**
4	**Effects of heterogeneity in infection-exposure history and immunity on the dynamics of a protozoan parasite.**	**69**
5	**The success of different variant surface antigen repertoires of malaria parasites competing within and between hosts.**	**89**
6	**How selection forces dictate the variant surface antigens used by malaria parasites.**	**121**
	English Summary	**157**
	Nederlandse Samenvatting	**161**
	Resumen en Español	**165**
	Thesis & Project Facts	**169**
	Acknowledgements	**175**
	Curriculum Vitae	**177**

CHAPTER 1

General Introduction

M. Severins

In order to survive, infectious agents such as viruses, bacteria, protozoa and macroparasites (for convenience we will call them all parasites here) must succeed at two scales: the within-host scale and the between-host scale. Within-host survival involves processes such as reproduction and immune evasion, while between-host survival involves transmission and persistence in the host population. Within the host, the parasite needs to find a balance between exploiting the host's resources to grow, and keeping the host fit enough to transmit to a new host. At the same time, the host's immune system will apply a wide selection of attack mechanisms aimed at killing the parasite as soon as possible. To survive, parasites have evolved intricate techniques allowing them to hide from the immune system or evade getting killed, for as long as possible. At the between-host level the infecting parasites need to be able to transmit to and infect at least one other host to prevent infection from dying-out in the population. To persist in a population for a longer period of time, parasites have evolved mechanisms to find enough susceptible hosts to which they can transmit. For example, some parasites can change their outside appearance so that immune hosts cannot recognise them anymore, or hide inside host cells until transmission conditions are favourable.

When looking for measures to control infection it is important to know which of the many components, interactions and influences that are involved in the above-mentioned infection processes are key in defining the infection. More insight in these allows us to focus our effort on understanding and learning how to influence these factors to our advantage. Using experimental methods alone it can be difficult, if not impossible, to find these key interactions as there are often practical, technical and financial limitations. This is where computational models can help out. For example, turning an interaction on and off, to analyse its function, may be an impossible task in reality, while in a computer model it can be as simple as a click. The difficulty, however, then lies in formulating a model with the limited knowledge available that is both sensible and useful. Models are caricatures of a much more complex reality and to be useful they need to be able to reproduce the essential characteristics of the real processes that are being studied. If that can be achieved, then using models, alongside experiments, opens up a myriad of new possibilities to both generate and test hypotheses.

Models are particularly helpful to study systems where cause and consequence are so entangled that it makes the interpretation of experimental data very hard or experiments hard to perform. This is typically the case in systems where the interactions between the components form one or multiple interacting feedback loops. In the study of infectious disease we find these feedback loops at both the within and between-host level. Two examples are given in Figure 1. The example on the left-hand side shows a typical within-host infection loop for a parasite with intracellular reproduction. On the right-hand side the classical SIR-model, one of the simplest of epidemic models, is shown.

In terms of control measures, the aims for understanding within-host versus

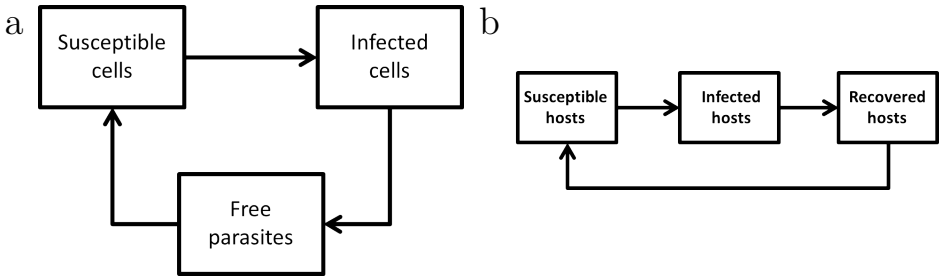


Figure 1: a) Within-host infection feedback loop. b) Between-host infection feedback loop

between-host processes are different. In general, efforts to understand within-host infection and immune processes are aimed at learning how to prevent a host from becoming infected, how to reduce clinical manifestation, or how to cure a host from infection. Efforts to understand between-host spread and infection processes are generally aimed at preventing parasites to cause an epidemic, or to slow down and stop outbreaks. To illustrate, a simple scenario could be that understanding the within-host processes tells us how to develop a vaccine whereas understanding the between-host processes tells us how many and which individuals we need to vaccinate. We therefore need understanding at both scales of infection, and their interaction, to eliminate a parasite from the population.

Within-host and between-host models can be made separately to help us gain such understanding. As such models describe processes at only one scale of the infection they need to include some assumptions as to what happens at the other scale. For example, in a within-host model, infection is usually started by introducing a set dose of infection into the system. In reality, depending on how transmission took place, the initial dose of infection taken-up may vary greatly. As long as the dose of infection does not affect how the infection process continues, this assumption does not affect the results of the within-host model. For between-host models it is usually assumed that all hosts within a compartment, such as a 'Recovered and Immune Hosts' compartment, are equal. In reality, this is not the case, for example during infection individual hosts may use slightly different parts of the infecting parasite to generate an immune response, and therefore show differences in immune memory. Yet, as long as all these individuals become equally resistant against re-infections with that specific parasite, this assumption does not affect the validity of the between-host model.

Parasites can have more complex infection dynamics that entail feedback loops that breach the two scales of infection in such a way that the key processes at one scale are dependent on the processes on the other scale. Such a system of interacting feedback loops across the within-host and between-host scales is illustrated

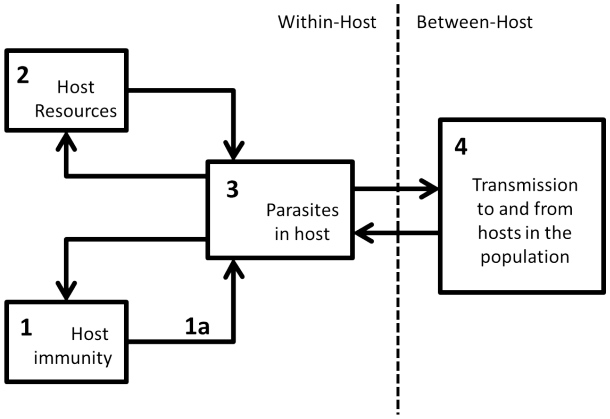


Figure 2: Schematic representation of an infection process that includes feedback loops between the parasites (3), the host resources (2), the host's immunity (1) and transmission (4).

in Figure 2. In infection processes where this is the case, studying infection at only one level may no longer be sensible or will no longer give realistic results [28]. Several studies have shown the importance of models that include feedback loops between infection processes at different scales [12, 20, 4]. In many cases these feedback loops that link the within-host and between-host scale cause heterogeneity in the host population. This can occur, for example, because differences in exposure between individuals cause differences in infection history, leading to heterogeneity in susceptibility and infectivity. Studies that take into account this kind of heterogeneity caused by processes across infection scales are scarce.

The goal of this PhD thesis is to increase our understanding of some of these complex within and between-host infection dynamics through the creation of mathematical and computational models that are able to capture the existing host and/or parasite heterogeneity. This goal is reached through a series of research projects that gradually build up in complexity of both the system modelled and the modelling techniques used. The models created in these projects help to explain a wide range of sometimes contradictory experimental results and are used to predict the effect of control measures. In addition, they generate ideas for the development of new methods of control.

Chapter 2: EAE in mice

The first research project of this thesis focuses on Experimental Autoimmune Encephalomyelitis (EAE), a mouse model of the human auto-immune disease Multiple Sclerosis [31]. In this project the degree of complexity is limited to interacting feedback loops within the immune system itself, without an infectious

parasite. In the context of Figure 2, this project only involves feedback loops within compartment 1.

EAE is caused by auto-immune cells that build a response against a protein of the nervous system. Typically, adaptive immune responses, whether they are built against a foreign pathogen or a protein originating from the body itself, can be of Type 1 (cellular response) or Type 2 (antibody response). These responses can be controlled by regulatory cells that the immune system has at its disposal which is often observed in chronic diseases. A large body of experimental data on EAE shows that the type of the auto-immune cells, and also the type of the cells that can control the autoimmune cells, are key to whether or not mice develop EAE [24]. Therefore, although it is a non-infectious disease, understanding the interactions between the immune cells that are involved in EAE is very relevant to infectious diseases.

The model formulated in this research project was built to better understand these interactions between adaptive immune cells. The model is an ordinary differential equation (ODE) model that describes a set of interacting feedback loops formed by cell populations within the immune system that are involved in EAE. Figure 3 shows a visual representation of these components and highlights the different feedback loops.

There are many more immune cells and interactions that influence EAE than the cells implemented in this model. However, this model can reproduce the results of a range of different vaccination experiments in mice with EAE. The strength of the model lies in just that. Because the results found in the vaccination experiments can be explained by the interactions between only the components used to build the model, it reduces the complexity of the experimental results to only those elements of the immune system implemented in the model.

The fact that regulatory cells could shift the autoimmune response from a Type 1 to a Type 2 following vaccination was initially conceived with surprise [11]. Our modelling results show, however, how such a shift results from the interactions between regulatory cells and Type 1 and Type 2 autoimmune cells, without the need for any direct influence of regulatory cells on the type of autoimmune cells. The shift arises naturally because Type 1 regulatory cells only suppress the Type 1 autoimmune cells, and thus allow the Type 2 autoimmune cells to grow in response to stimulation with self-protein.

Chapter 3: Paratuberculosis in dairy cows

The second research project focusses on a bacterial infection in dairy cows that causes the disease Paratuberculosis. Paratuberculosis, as EAE, is a chronic disease, however in contrast with EAE it is caused by an infectious agent. Once

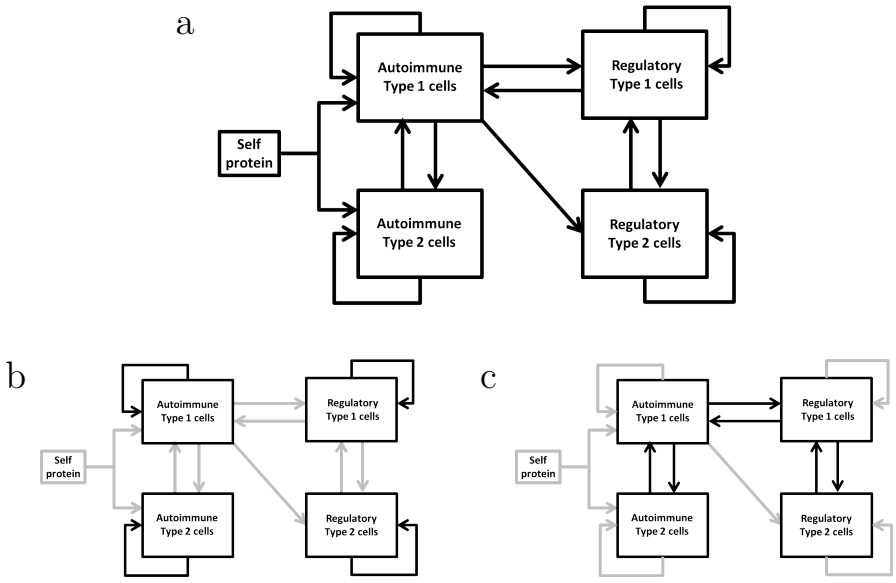


Figure 3: Schematic representation of the EAE model. a) The full model. b) Highlights the positive feedback loops within the immune cell populations. c) Highlights the interacting negative feedback loops between the immune cells.

initiated, infection is viewed as an autonomous process, where subsequent doses of infection later in life have little or no influence on the course of infection within the host. This allows the infection process to be studied at the within-host scale only. In this case we can add a small step to the degree of complexity by having two interacting feedback loops describing within-host infection and immunity. One feedback loop describes parasite reproduction that is influenced differently by Type 1 and Type 2 immune responses, and the other describes the feedback loop between the parasite and the host's resources. The latter in this case also happens to be a part of innate immunity. These interacting feedback loop and the influences on them are illustrate in Figure 4. In the context of Figure 2, this project includes feedback loops between components 2 and 3 under the influence of 1a.

Paratuberculosis is caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP) [3]. When ingested by its host, MAP is able to cross the epithelial barrier lining the gut and infect macrophages inside the gut. Macrophages form part of our innate immune system and under normal conditions engulf and kill bacteria. MAP is able to avoid being killed after it is engulfed [13]. On the contrary, it is able to reproduce inside the macrophages and establish a chronic infection for which no successful treatment is currently available. Cows differ greatly in the duration and the extent to which they can keep this chronic infection under con-

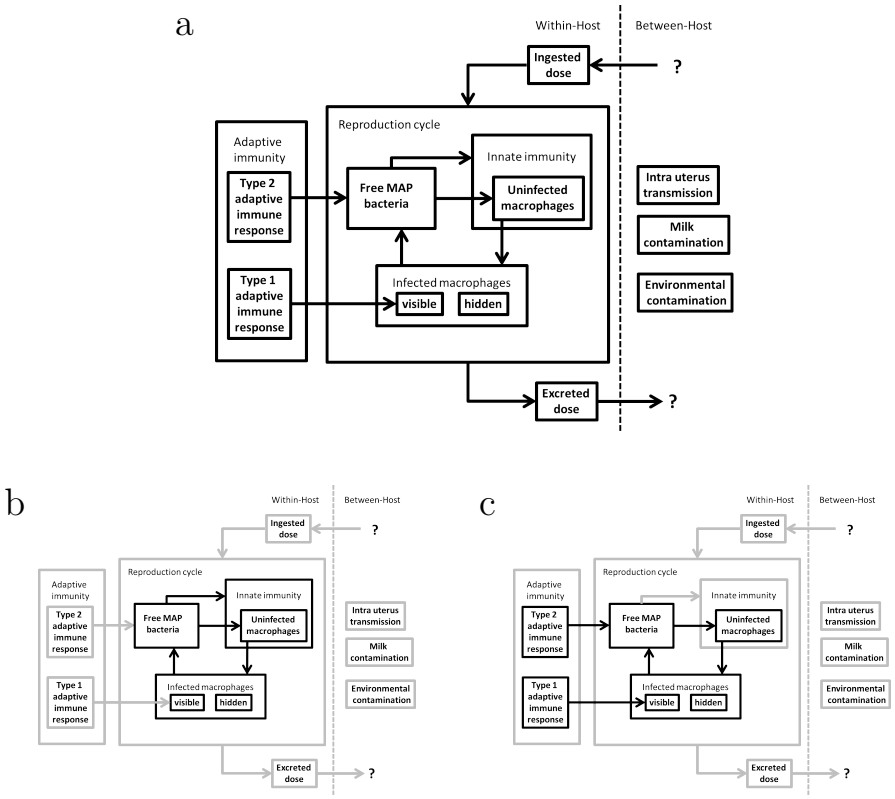


Figure 4: Schematic representation of the Paratuberculosis model. a) The full model. b) Highlight of the within-host parasite reproduction cycle including a feedback loop with the host resources through the attraction of uninfected macrophages to the site of infection by free parasites. c) Highlight of the differential influence of Type 1 and Type 2 adaptive immunity on the reproduction feedback loop.

trol before it eventually progresses into severe disease.

It is unclear what causes the onset of clinical disease. A variety of studies indicate that cows build an early Type 1 immune response that changes to a Type 2 immune response during the course of the disease [34]. It is thought that the Type 1 response is protective while the Type 2 response is permissive, and that the change in response could be the cause of progression into severe disease [35, 21, 33]. However, direct evidence and a thorough understanding why one type of response would be able to control disease while the other fails, is lacking. The model in this chapter was built to clarify the role of the type of the adaptive immune response on the infection progression.

Type 1 adaptive immune responses kill parasite-infected host cells, while Type 2 adaptive immune responses target free parasites in the body. This known functional difference in which the two types of adaptive immune responses combat infection is explicitly included in the model. In the model it is easy to turn each of these responses on and off at any given time which allows for a detailed study of the effects these two responses have on MAP infections in dairy cows.

Through these analyses and by using the results from a longitudinal study in which infected cows were followed for 4 years, the model is able to find that there is no fundamental mechanical impediment to prevent a Type 1 or a Type 2 adaptive response from clearing the infection. The results show that the killing rate of the Type 1 and Type 2 adaptive immune responses together is much stronger than the sum of their individual killing rates. Surprisingly the results also show that the Type 1 response observed experimentally is able to stop the initial growth of infection even if it does not have the critical killing rate needed to control the infection.

Chapter 4: Coccidiosis in broiler chickens

The third research project focusses on coccidiosis in broiler chickens caused by the protozoan parasite *Eimeria acervulina*. This project is the first to include parasite infection processes that breach the within-host and between-host scales. The within-host life cycle of this parasite is such that it allows for a less detailed modelling of the within-host processes than in the previous chapters, and we can therefore keep the degree of complexity more manageable. In the context of Figure 2, this project includes feedback loops between components 1, 2 and 4 without explicitly describing the processes in component 3 through which these feedback loops interact.

Chickens become infected by picking-up *Eimeria* oocysts from their environment which then replicate inside the gut of the animals and are excreted with the chickens' faeces [1]. *Eimeria* reproduction inside the gut is a self-limiting process that

cannot continue to re-infect gut cells without going through an environmental stage. How many parasites on average one *Eimeria* parasite can produce and excrete depends on how many parasites where picked-up at the same time. Excretion also depends on how many were picked-up previously, as these will be occupying some of the gut cells that would otherwise be available for infection [37]. In addition, chickens develop immune responses against *Eimeria* parasites that with every new case of infection protect the chickens better against infection and parasite reproduction, until they are fully immune [10, 26]. Thus, the infection history of the chickens has an effect on the immune status of the chickens and also on the infectiousness of the chickens, i.e. on how many oocysts the chickens will excrete into the environment.

Ingesting small numbers of oocysts over a prolonged period of time will not cause much harm to the chickens. However, in stables of broiler chickens, characterised by a dense population of susceptible hosts, the environmental contamination can quickly grow to levels where chickens ingest high doses of oocysts, causing severe outbreaks of clinical disease. The cleaning intensity between cohorts of broilers is thought to affect the number of clinical cases, yet studies trying to show this find counterintuitive results [17, 16, 20]. The model built in this chapter is used to understand the progression of infection in flocks of commercial broiler chickens and to find if there exist cleaning regimes that minimise the clinical disease cases.

To study the effect of the infection history on the infection processes, the model needs to be able to capture the infection history of the individual chickens, their immune status and the environmental contamination. To allow for this kind of heterogeneity the model is spatial and individual-based, meaning that each chicken can be tracked separately; we use the simulation software Netlogo to implement the model. The two effects that the infection history has on the within-host infection are described by two interconnected feedback loops as illustrated in Figure 5.

The different immune statuses, and ingestion and excretion doses, are limited in the model, which, together with the self-limiting within-host infection process of the *Eimeria* parasite, allows for the modelling of the excreted doses through the use of a limited set of fixed excretion patterns. The model uses specific excretion-templates, supported by experimental data, for every combination of infection and immune status. Furthermore, chickens switch back and forth between templates depending on the oocysts they ingest and their immune status. As a result, the heterogeneity in the chicken population is sufficiently large to capture its effect on the within-host infection and immune processes without having to model the mechanisms behind these processes explicitly.

In the model, different stable cleaning intensities and regimes can be simulated with different initial stable contaminations. The model shows a non-linear 'wave-like' relationship between the intensity of cleaning and the severity of *Eimeria* outbreaks in a flock where intermediate levels of cleaning intensity lead to the

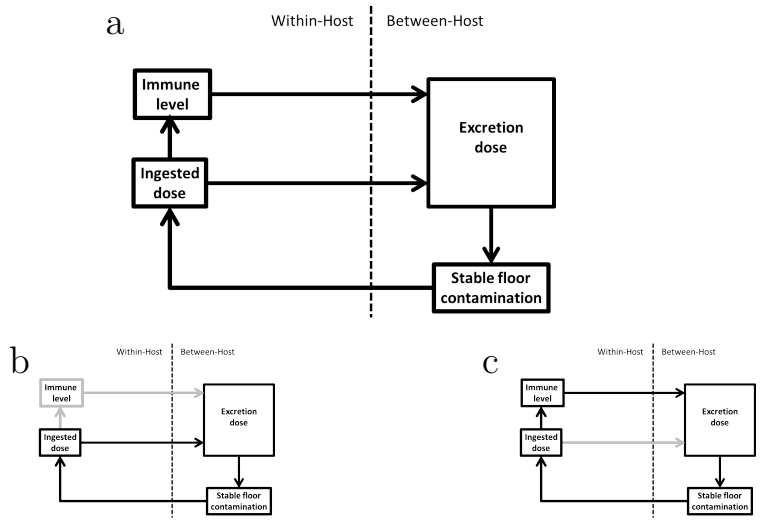


Figure 5: Schematic representation of the coccidiosis model. a) The full model. b) Highlights the within-between host infection feedback loop. c) Highlights the within-between host immunity feedback loop.

least clinical disease. This relationship explains the counterintuitive experimental findings and is caused by an increased heterogeneity in the host population at the peaks of the wave. The heterogeneity originates from stochastic parasite uptake that results in differences in infection between the chickens in the flock. There are cleaning intensities at which the feedback loops across the two infection scales either amplify or dampen the heterogeneity. This relationship could therefore not have been found from a model that does not allow for considerable host heterogeneity or that only models the infection dynamics at one of these scales.

Chapter 5: Malaria in humans I

The fourth research project concentrates on Malaria in humans caused by the eukaryotic protist *Plasmodium falciparum*. Malaria parasites have an intricate infection process with many life stages in two different hosts: mosquitoes and humans. There are many different reactions of the immune system against malaria parasites. In turn, the parasites possess a wide range of immune evasion techniques to avoid the immune systems attacks. In contrast to the previous project, this project describes in detail a within-host feedback loop between infection and adaptive immunity in addition to the interacting set of feedback loops across the within and between-host scales described in the previous project. In the context of Figure 2, this projects includes feedback loops between components 1, 3 and 4 and a large amount of host and parasite heterogeneity to capture the effect of

the interactions between these components.

When a parasite infects a cell, some parts of it usually show on the outside of the surface of the cell. The immune system uses these parts (called antigens) to identify infected cells and then build a response to kill them. This project concentrates on the variant surface antigens (VSA) that malaria parasites use to avoid these immune responses. The particular VSA looked at here, *Plasmodium falciparum* erythrocyte membrane proteins 1 (PfEMP1's), are expressed on the surface of red blood cells (rbc) infected during the parasite's blood-stage. These VSA help parasites to survive in two ways: 1) by frequently changing the type of VSA expressed, the antibodies that the immune system has made against previously expressed VSA become less effective [29, 5], and 2) VSA bind to receptors in organs which prevents the infected rbc from flowing through the spleen where infected rbc are recognised and killed [18, 14]. The strength of this binding affinity of the VSA therefore, at least partially, determines the relative growth rate of the parasites.

Each parasite carries a set of approximately 60 different types of VSA and different parasites carry different sets of VSA [15]. Due to several parasite recombination techniques new types of VSA emerge continuously and therefore the number of different types of VSA that exist is immense [36]. Individuals living in malaria-endemic areas gradually build up a very large repertoire of antibodies against VSA that is associated with protection first from severe disease and eventually from clinical disease, although infections remain common even at old age [2].

Experimental data shows that the different types of VSA can be categorised into genetically similar groups [22, 25]. Data also suggests that every parasite does not carry a random set of VSA but that all VSA groups are proportionally represented in the set [7, 6]. It is unclear why parasites carry representatives of all groups. One line of thought is that parasites need the VSA from the different groups to be able to cope under different circumstances. For example, when infecting young children with a relatively narrow anti-VSA antibody repertoire, the parasite can afford to express a VSA that maximises its growth rate, while when infecting an older host with a broad antibody repertoire, the parasite has to express a very rare VSA for which the host has no antibodies yet. It has been suggested that the difference between a young host and an older host environment would benefit parasites that carry both 'strong' and 'rare' VSA [23, 8]. The model built in this project was developed to understand if under these circumstances parasites with a strong-rare set of VSA are more successful than parasites with other sets of VSA.

To recreate the VSA-related difference in young hosts versus old hosts, the model needs to include the within-host parasite competition and immune responses that determine to which VSA the hosts build up antibodies, and to include the between-host competition for transmission to mosquitoes and susceptible hosts. The model also needs to include a diverse range of VSA-related host, vector and

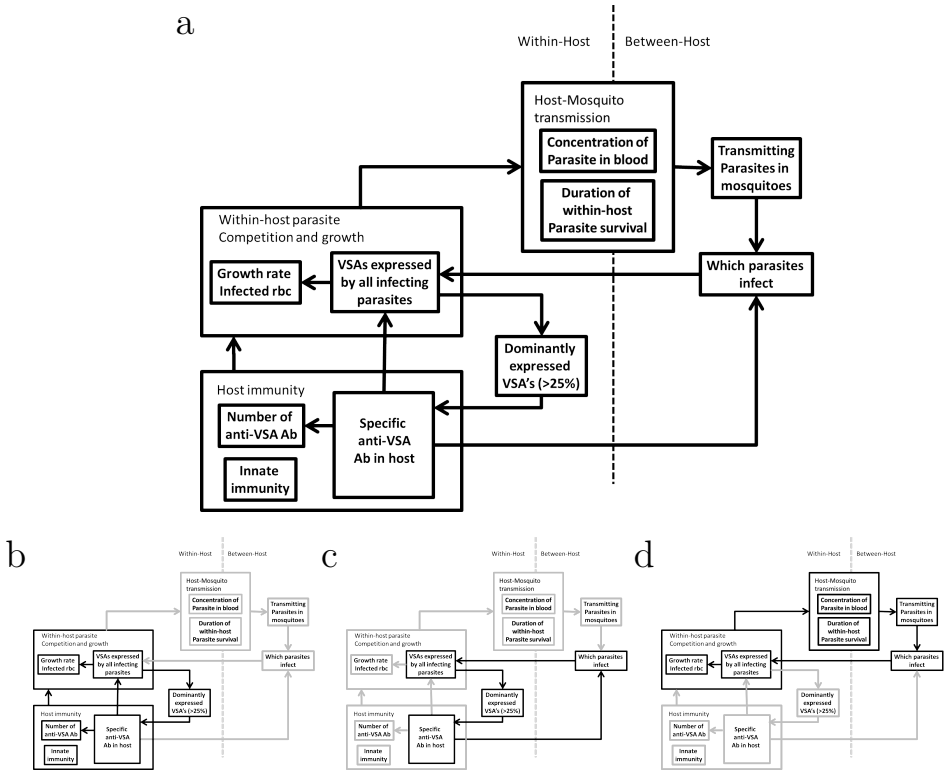


Figure 6: Schematic representation of malaria I model. a) The full model with all interacting feedback loops. b) Highlights the within-host feedback loop between adaptive immunity and parasite reproduction. c) Highlights the feedback loop between host adaptive immunity and transmission. c) Highlights the feedback loop between parasite reproduction and transmission.

parasite heterogeneity. To accommodate all these processes and heterogeneity the object-orientated programming language C++ is used to develop the model and run infection simulations. The model keeps track of, amongst other things, the VSA of each parasite in the parasite population, the VSA expressed by infected rbc and, the anti-VSA antibodies of the hosts. Figure 6 illustrates the multiple interacting feedback loops, across both infection scales.

In the resulting model, the difference in size of the anti-VSA immune repertoire between young hosts and older hosts indeed emerges. Additionally, the model shows that hosts do not acquire antibodies against VSA at random but that there is a general order in which they are acquired in which immunity against strong VSA is acquired earlier. An ordered acquisition of anti-VSA antibodies was also found experimentally [9] and the model now gives a potential explanation for the

origin of this order. However, contrary to what has been suggested, the difference in a young versus older host environment, at least on its own, does not show a selective advantage for parasites with a strong-rare set of VSA over parasites with other, random, sets of VSA.

Chapter 6: Malaria in humans II

The fifth and last research project also concentrates on the VSA used by malaria parasites. However, the focus of this project lies on understanding the key selection forces on the parasites through their VSA and the influence that control measures can have on these. The selection forces work on all interactions shown in Figure 2 and therefore this project includes all components and interacting feedback loops illustrated in Figure 2. As in the previous project, a large amount of host and parasite heterogeneity is allowed for to capture the effect of the feedback loops between these components on infection.

As explained in the previous section, the binding strength of the VSA, at least partially, determines the relative growth rate of the parasites which in turn has an effect on the severity of malaria disease. Disease severity is influenced by the type of VSA expressed in yet another way. The type of VSA expressed determines in which organ(s) the infected rbc can sequester, which has a large impact on the severity of disease, e.g. cerebral malaria is one of the most deadliest forms of malaria [29, 32, 19, 30]. These findings show that the VSA set carried by the parasites impacts on the virulence of the parasite.

VSA are implicated in many of the infection processes of malaria, making it difficult to predict how large scale malaria control programs will influence the virulence of the circulating parasites. This will depend on which parasites are able to survive under all the selection forces that influence the parasites through the VSA they carry. The key selection forces on malaria parasites are: parasite competition (within and between-host), host immunity, mosquito abundance, and host mortality [27]. The model developed in the previous project already included some of these namely, within-host parasite competition, although only the relative competition between parasites, between-host competition (Figure 6c & 6d), host immunity (Figure 6b) and mosquito abundance. For the purposes of the current project the previous model is extended with feedback loops for the within-host parasite competition to infect the host's rbc (Figure 7b) and malaria induced host mortality (Figure 7c), to include all key selection forces that act on the parasites.

Like in the previous project, the model continues to show an ordered acquisition of anti-VSA antibodies. In addition, by including how the parasites compete to infect the available red blood cells, the model gives realistic life-time host parasitemia and provides a framework in which to study the within-host competition through VSA in detail.

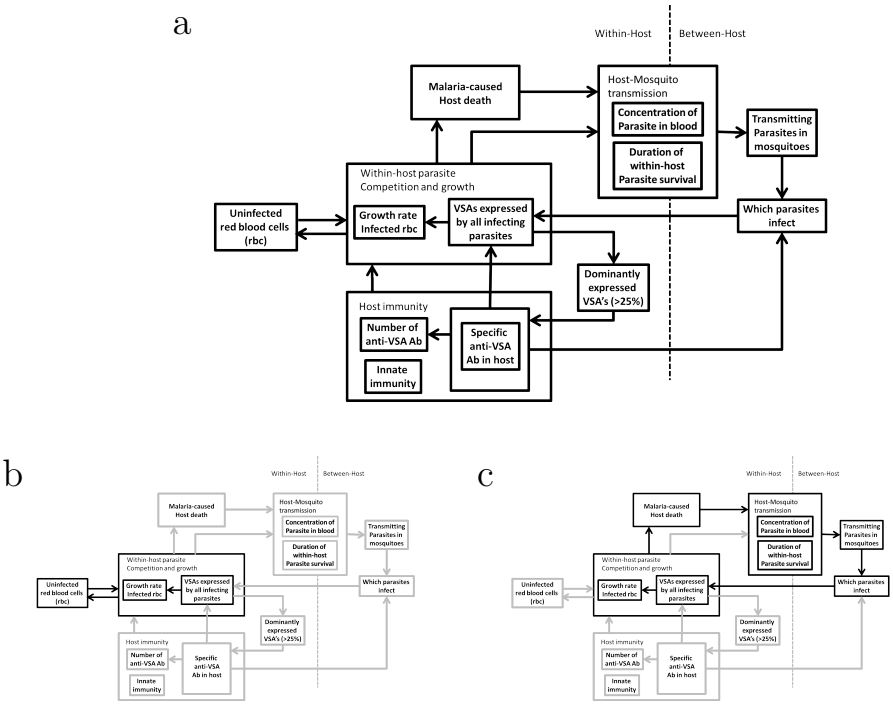


Figure 7: Schematic representation of malaria II model. a) The full model. b) Highlights the feedback loop between parasite reproduction and the host's resources. c) Highlights the feedback loop between malaria-induced host mortality, transmission and parasite reproduction.

The model is used to test the effect that a series of potential and currently implemented control measures can have upon the virulence of parasites circulating in the population. The results show that some seemingly sensible measures, such as reducing the reproduction rate of the parasites, could have the unwanted effect of increasing the frequency of highly virulent parasites, such as those causing cerebral malaria.

Final remarks

As mentioned above, the difficulty in modelling lies in formulating a model that is both sensible and useful. In my experience, the chances for this are greatly improved when modellers and field experts work closely together. To facilitate these collaborations it is important that modellers make models that are appealing to

the expert community. The results and insights from a model that succeeds in being so, are much more likely to reach and be trusted by field experts.

As a biologist, I find that a crucial element for the appeal of a biological model is for its components to have a clear biological meaning or interpretation. In my opinion, the best models are those that do not require an expert in the modelling techniques used to understand it, but those that can be explained in biological terms. Such a direct link between the model and reality makes it easier to use existing data, or gather new data, to parameterise and validate the model. It also makes it straightforward for experimentalists to interpret the model results, give feedback and pinpoint errors.

In this thesis the vast majority of model components have such a direct link with reality. The results have shown some detailed examples of the valuable contribution that models have in understanding infection processes. The most satisfying achievements have come from those models that were able to, in hindsight, make complicated experimental results seem obvious and logical, and where the process of building the model was as insightful as the final results.

References

[1] P. C. Allen and R. H. Fetterer. Recent advances in biology and immunobiology of Eimeria species and in diagnosis and control of infection with these coccidian parasites of poultry. *Clin Microbiol Rev*, 15(1):58–65, Jan 2002.

[2] James G Beeson, Faith H A Osier, and Christian R Engwerda. Recent insights into humoral and cellular immune responses against malaria. *Trends Parasitol*, 24(12):578–584, Dec 2008. doi: 10.1016/j.pt.2008.08.008.

[3] Marcel A. Behr and Desmond M. Collins, editors. *Paratuberculosis: organism, disease, control*. CABI, 2010.

[4] Barbara Boldin and Odo Diekmann. Superinfections can induce evolutionarily stable coexistence of pathogens. *J Math Biol*, 56(5):635–672, May 2008. doi: 10.1007/s00285-007-0135-1.

[5] P. Borst, W. Bitter, R. McCulloch, F. Van Leeuwen, and G. Rudenko. Antigenic variation in malaria. *Cell*, 82(1):1–4, Jul 1995.

[6] Caroline O Buckee, Peter C Bull, and Sunetra Gupta. Inferring malaria parasite population structure from serological networks. *Proc R Soc B*, 276(1656):477–485, Feb 2009. doi: 10.1098/rspb.2008.1122.

[7] Peter C Bull, Matthew Berriman, Sue Kyes, Michael A Quail, Neil Hall, Moses M Kortok, Kevin Marsh, and Chris I Newbold. Plasmodium falciparum variant surface antigen expression patterns during malaria. *PLoS Pathog*, 1 (3):e26, Nov 2005. doi: 10.1371/journal.ppat.0010026.

[8] Peter C Bull, Caroline O Buckee, Sue Kyes, Moses M Kortok, Vandana Thathy, Bernard Guyah, Jos A Stoute, Chris I Newbold, and Kevin Marsh. Plasmodium falciparum antigenic variation. mapping mosaic var gene sequences onto a network of shared, highly polymorphic sequence blocks. *Mol Microbiol*, 68(6):1519–1534, Jun 2008. doi: 10.1111/j.1365-2958.2008.06248.x.

[9] Gerald K K Cham, Louise Turner, John Lusingu, Lasse Vestergaard, Bruno P Mmbando, Jonathan D Kurtis, Anja T R Jensen, Ali Salanti, Thomas Lavstsen, and Thor G Theander. Sequential, ordered acquisition of antibodies to plasmodium falciparum erythrocyte membrane protein 1 domains. *J Immunol*, 183(5):3356–3363, Sep 2009. doi: 10.4049/jimmunol.0901331.

[10] H. D. Chapman. Anticoccidial drugs and their effects upon the development of immunity to Eimeria infections in poultry. *Avian Pathol*, 28(6):521–535, Dec 1999. doi: 10.1080/03079459994317.

[11] I. R. Cohen. T-cell vaccination for autoimmune disease: a panorama. *Vaccine.*, 20:706–710, 2001.

[12] Daniel Coombs, Michael A. Gilchrist, and Colleen L. Ball. Evaluating the importance of within- and between-host selection pressures on the evolution of chronic pathogens. *Theor Popul Biol*, 72(4):576–591, Dec 2007. doi: 10.1016/j.tpb.2007.08.005.

[13] Chantal de Chastellier. The many niches and strategies used by pathogenic mycobacteria for survival within host macrophages. *Immunobiology*, 214(7):526–542, Jul 2009. doi: 10.1016/j.imbio.2008.12.005.

[14] Matthias Frank and Kirk Deitsch. Activation, silencing and mutually exclusive expression within the var gene family of plasmodium falciparum. *Int J Parasitol*, 36(9):975–985, Aug 2006. doi: 10.1016/j.ijpara.2006.05.007.

[15] Malcolm J Gardner, Neil Hall, Eula Fung, Owen White, Matthew Berriman, Richard W Hyman, Jane M Carlton, Arnab Pain, Karen E Nelson, Sharen Bowman, Ian T Paulsen, Keith James, Jonathan A Eisen, Kim Rutherford, Steven L Salzberg, Alister Craig, Sue Kyes, Man-Suen Chan, Vishvanath Nene, Shamira J Shallom, Bernard Suh, Jeremy Peterson, Sam Angiuoli, Mihaela Pertea, Jonathan Allen, Jeremy Selengut, Daniel Haft, Michael W Mather, Akhil B Vaidya, David M A Martin, Alan H Fairlamb, Martin J Fraunholz, David S

Roos, Stuart A Ralph, Geoffrey I McFadden, Leda M Cummings, G. Mani Subramanian, Chris Mungall, J. Craig Venter, Daniel J Carucci, Stephen L Hoffman, Chris Newbold, Ronald W Davis, Claire M Fraser, and Bart Barrell. Genome sequence of the human malaria parasite plasmodium falciparum. *Nature*, 419(6906):498–511, Oct 2002. doi: 10.1038/nature01097.

[16] E. A. Graat, H. W. Ploeger, A. M. Henken, G. De Vries Reilingh, J. P. Noordhuizen, and P. N. Van Beek. Effects of initial litter contamination level with *Eimeria acervulina* on population dynamics and production characteristics in broilers. *Vet Parasitol*, 65(3-4):223–232, Oct 1996.

[17] A. M. Henken, H. W. Ploeger, E. A. Graat, and T. E. Carpenter. Description of a simulation model for the population dynamics of *Eimeria acervulina* infection in broilers. *Parasitology*, 108 (Pt 5):503–512, Jun 1994.

[18] M. Ho and N. J. White. Molecular mechanisms of cytoadherence in malaria. *Am J Physiol*, 276(6 Pt 1):C1231–C1242, Jun 1999.

[19] Mirjam Kaestli, Ian A Cockburn, Alfred Corts, Kay Baea, J. Alexandra Rowe, and Hans-Peter Beck. Virulence of malaria is associated with differential expression of plasmodium falciparum var gene subgroups in a case-control study. *J Infect Dis*, 193(11):1567–1574, Jun 2006. doi: 10.1086/503776.

[20] D. Klinkenberg and J. A P Heesterbeek. A model for the dynamics of a protozoan parasite within and between successive host populations. *Parasitology*, pages 1–10, Feb 2007. doi: 10.1017/S0031182007002429.

[21] A. P. Koets, V. P. Rutten, A. Hoek, D. Bakker, F. van Zijderveld, K. E. Mller, and W. van Eden. Heat-shock protein-specific t-cell responses in various stages of bovine paratuberculosis. *Vet Immunol Immunopathol*, 70(1-2):105–115, Sep 1999.

[22] Susan M Kraemer and Joseph D Smith. Evidence for the importance of genetic structuring to the structural and functional specialization of the plasmodium falciparum var gene family. *Mol Microbiol*, 50(5):1527–1538, Dec 2003.

[23] Susan M Kraemer, Sue A Kyes, Gautam Aggarwal, Amy L Springer, Siri O Nelson, Zoe Christodoulou, Leia M Smith, Wendy Wang, Emily Levin, Christopher I Newbold, Peter J Myler, and Joseph D Smith. Patterns of gene recombination shape var gene repertoires in plasmodium falciparum: comparisons of geographically diverse isolates. *BMC Genomics*, 8: 45, 2007. doi: 10.1186/1471-2164-8-45.

[24] V. Kumar and E. Sercarz. Induction or protection from experimental autoimmune encephalomyelitis depends on the cytokine secretion profile of TCR peptide-specific regulatory CD4 T cells. *J. Immunol.*, 161:6585–6591, 1998.

[25] Thomas Lavstsen, Ali Salanti, Anja T R Jensen, David E Arnot, and Thor G Theander. Sub-grouping of plasmodium falciparum 3d7 var genes based on sequence analysis of coding and non-coding regions. *Malar J*, 2:27, Sep 2003. doi: 10.1186/1475-2875-2-27.

[26] H. S. Lillehoj and E. P. Lillehoj. Avian coccidiosis. A review of acquired intestinal immunity and vaccination strategies. *Avian Dis*, 44(2):408–425, 2000.

[27] M. J. Mackinnon and K. Marsh. The selection landscape of malaria parasites. *Science*, 328(5980):866–871, 2010. doi: 10.1126/science.1185410.

[28] Nicole Mideo, Samuel Alizon, and Troy Day. Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends Ecol Evol*, 23(9):511–517, Sep 2008. doi: 10.1016/j.tree.2008.05.009.

[29] L. H. Miller, M. F. Good, and G. Milon. Malaria pathogenesis. *Science*, 264(5167): 1878–1883, Jun 1994.

[30] Noa D Pasternak and Ron Dzikowski. Pfemp1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite plasmodium falciparum. *Int J Biochem Cell Biol*, 41(7):1463–1466, Jul 2009. doi: 10.1016/j.biocel.2008.12.012.

[31] C. S. Raine. Biology of disease. analysis of autoimmune demyelination: its impact upon multiple sclerosis. *Lab Invest*, 50(6): 608–635, Jun 1984.

[32] Matthias Rottmann, Thomas Lavstsen, Joseph Paschal Mugasa, Mirjam Kaestli, Anja T R Jensen, Dania Mller, Thor Theander, and Hans-Peter Beck. Differential expression of var gene groups is associated with morbidity caused by plasmodium falciparum infection in tanzanian children. *Infect Immun*, 74(7):3904–3911, Jul 2006. doi: 10.1128/IAI.02073-05.

[33] J. R. Stabel. Cytokine secretion by peripheral blood mononuclear cells from cows infected with mycobacterium paratuberculosis. *Am J Vet Res*, 61(7): 754–760, Jul 2000.

[34] J. R. Stabel. Host responses to mycobacterium avium subsp. paratuberculosis: a complex arsenal. *Anim Health Res Rev*, 7(1-2):61–70, 2006. doi: 10.1017/S1466252307001168.

[35] R. W. Sweeney, D. E. Jones, P. Habecker, and P. Scott. Interferon-gamma and interleukin 4 gene expression in cows infected with mycobacterium paratuberculosis. *Am J Vet Res*, 59(7):842–847, Jul 1998.

[36] H. M. Taylor, S. A. Kyes, and C. I. Newbold. Var gene diversity in plasmodium falciparum is generated by frequent recombination events. *Mol Biochem Parasitol*, 110(2):391–397, Oct 2000.

[37] R. B. Williams. Effects of different infection rates on the oocyst production of eimeria acervulina or eimeria tenella in the chicken. *Parasitology*, 67(3):279–288, Dec 1973.

CHAPTER 2

The role of Th1/Th2 phenotypes in T-cell vaccination: Insights from a mathematical model

Maite Severins, José A. M. Borghans and Rob J. de Boer

Published in:
T-Cell Vaccination, Chapter 9, p.p. 139-158.
Editors: Jingwu Zhang and Irun R. Cohen
Publisher: Nova Science Publishers Inc. 2008

Abstract

We give a historical overview of the diverse theoretical models that have been developed to better understand the role of anti-idiotypic regulation in T-cell vaccination (TCV). More recently the importance of the Th1/Th2 phenotype of the T-cells involved in TCV became apparent. To understand the combined role of anti-idiotypic regulation and Th1/Th2 phenotype differentiation in TCV we have developed a new phenomenological model. The model consists of four differential equations, one for each cell type involved: Th1 autoreactive cells, Th2 autoreactive cells, Th1 anti-idiotypic cells and Th2 anti-idiotypic cells. To incorporate the interactions between these cell types we used data from experimental autoimmune encephalomyelitis (EAE), the mouse model of multiple sclerosis. In accordance with experiments, our phenomenological model shows that: (i) challenging the autoreactive cells causes transient EAE followed by protection against re-challenges, (ii) after TCV with the Fr3 peptide of the T-cell receptor (TCR) of the Th1 autoreactive cells, challenging the autoreactive cells does not lead to disease, and (iii) after nasal instillation of the TCR Fr3 peptide, challenging the autoreactive cells leads to exacerbated disease. In addition, the model gives an explanation for the phenotypic shift from a Th1 autoreactive response towards a Th2 autoreactive response seen after TCV, and it shows that a Th2 response to MBP has the potential to regulate EAE. By reproducing the experiments on TCV in EAE, the model provides a transparent framework pinpointing the interactions and mechanisms that result in the complex behaviour observed in TCV experiments.

1 Introduction

T-cell vaccination (TCV) is the process by which the injection of the very cells that cause autoimmune diseases can provide protection against autoimmunity. In order to avoid the induction of autoimmunity, autoreactive cells are either injected at very low concentrations, or are attenuated before injection. It has been shown for several autoimmune diseases, including experimental autoimmune encephalomyelitis (EAE) [2], adjuvant arthritis [17] and insulin-dependent diabetes mellitus [15] that this controlled administration of autoreactive cells can change the state of the immune system from susceptibility to a state in which the animal is protected against induction of the autoimmune disease. The key players in TCV are self antigens, autoreactive T-cells and anti-idiotypic T cells which recognize peptides from the T-cell receptor (TCR) of the autoreactive T-cells.

The coupling of interactions between these key players can give rise to counter-intuitive behaviour, which may be difficult to understand by intuitive reasoning alone. Several mathematical and bioinformatic models have been developed to help interpret the ever growing body of empirical data on TCV. The beauty of these models lies in their simplicity; the largest contribution to our understanding of the role of anti-idiotypic regulation in protection against autoimmunity has come from relatively simple models that have generated transparency in the growing body of experimental results on TCV.

In the late 1990s it became clear that besides the presence of anti-idiotypic T-cells, the Th1/Th2 phenotype of the T-cells involved, plays an important role in the induction of protection against disease [21]. This chapter first gives a historical overview of the diverse theoretical models that have been developed to understand the role of anti-idiotypic cells in TCV. Subsequently, a new phenomenological model is introduced incorporating both anti-idiotypic interactions and the influence of Th1/Th2 switches to study their roles in TCV. To build the model we used data from EAE, the mouse model of multiple sclerosis (MS) that shares several of its pathological and immune dysfunctions. Our phenomenological model, including only four cell types, reproduces all experiments on TCV in EAE. It thereby provides a framework pinpointing the interactions and mechanisms that result in the complex behaviour observed in TCV experiments.

2 Experimental findings

It is still not completely understood how MS develops in human individuals, and which regulatory mechanisms fail to maintain immune tolerance. In animal models of MS, MBP-reactive T-cells of the Th1 phenotype have been shown to be encephalitogenic, while the presence of MBP-reactive T-cells with a Th2 phenotype has been found to protect against EAE [1, 23]. Experiments in BL10.PL mice indicated that the pathological T-cells involved in EAE have a limited clonal di-

versity [26]. During the course of disease there is a dominant T-cell clone reactive to the MBP Ac1-9 peptide, which expands in the lymphoid organs and infiltrates the CNS tissue [26] (see Figure 1a). This T-cell clone, which uses the $V_\beta 8.2 - J_\beta 2.7$ TCR gene segments and has a Th1 phenotype, is lost from these tissues during the spontaneous recovery phase, while many expansions of sub-dominantly Th2 MBP-reactive cells remain [26]. During the recovery phase there is an expansion of anti-idiotypic T-cells which are reactive to the $V_\beta 8.2$ epitope, B5 aa 76-101 (referred to as the TCR Fr3 peptide) of the TCR of the dominant MBP-reactive T-cell clone [21]. After recovery, mice are naturally protected against rechallenge with MBP.

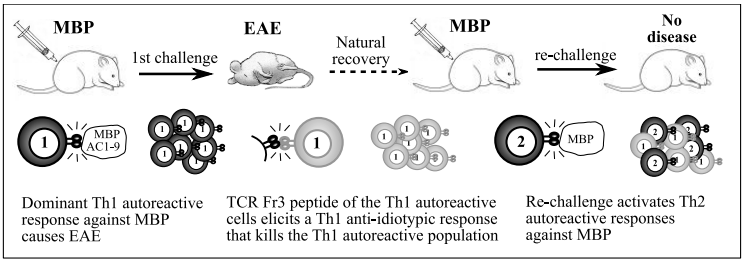
The earliest TCV experiments have shown that autoimmunity can be avoided in naive animals by artificial induction of anti-idiotypic cells, either via their direct injection, or by priming them through TCV with attenuated or low doses of MBP-reactive cells, or by vaccination with the TCR Fr3 peptide of the dominant MBP-reactive clone (Figure 1b). More recently, it has been shown that the Th1/Th2 phenotype of the anti-idiotypic cells is important for the success of the induced anti-idiotypic regulation [21, 29, 22]. It was demonstrated that only anti-idiotypic cells of the Th1 phenotype provide protection to EAE [22]. When anti-idiotypic cells were primed by nasal instillation of the TCR Fr3 peptide, which is known to deviate responses into a Th2 direction [20], no protection from EAE was observed (Figure 1c). Instead this kind of priming led to the contraction of a more chronic and severe form of disease [21]. An important difference between Th1 anti-idiotypic and Th2 anti-idiotypic cells is that the anti-idiotypic cells with a Th1 phenotype can recruit $CD8^+$ anti-idiotypic cells, whereas anti-idiotypic cells with a Th2 phenotype cannot. The $CD8^+$ anti-idiotypic cells have been shown to deplete activated $CD4^+$ T-cells of only the Th1 phenotype that have a $V_\beta 8$ gene segment in their TCR [18]. This explains why private expansions of predominantly Th2 MBP-reactive cells persist after recovery from EAE, despite the presence of anti-idiotypic T-cells [22, 18]. In summary, the Th1/Th2 phenotype of both anti-idiotypic cells and MBP-reactive cells plays a crucial role in the regulation of EAE, but the precise mechanisms by which TCV occurs remain elusive.

3 Historical overview of models for TCV

3.1 Automaton models

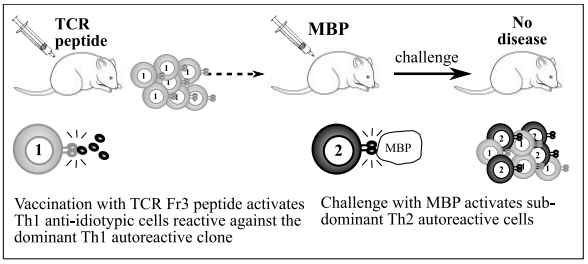
The first theoretical study of T-cell vaccination was the work of Atlan, Weisbuch, and Cohen [11, 35]. This work described several models defining interactions between antigen-specific helper T-cells, antigen-specific effector T-cells, anti-idiotypic suppressor T-cells, anti-idiotypic helper T-cells, and antigen-specific suppressor T-cells. Ignoring the sizes of these different T-cell populations, it was assumed that each T-cell population is either "active" (1) or "in-active" (0). By

a



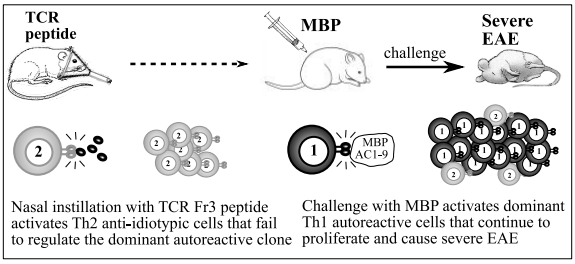
Challenging BL10.PL mice with MBP activates a dominant Th1 clone reactive against MBP peptide Ac1-9, and causes transient disease. Concomitant to recovery there is an expansion of Th1 anti-idiotypic cells reactive to the $V_\beta 8.2$ TCR peptide B5 aa 76-101 (also referred to as the TCR Fr3 peptide) of the dominant autoreactive clone [21]. Re-challenge with MBP is not associated with disease and causes expansion of sub-dominant Th2 autoreactive cells.

b



TCV with the TCR Fr3 peptide activates a Th1 anti-idiotypic clone. MBP challenges after TCV lead to the expansion of sub-dominant Th2 autoreactive clones and are not associated with disease [21].

c



Nasal instillation of the TCR peptide Fr3 activates a Th2 anti-idiotypic clone. Challenge with MBP activates the dominant Th1 autoreactive clone and leads to severe EAE [21].

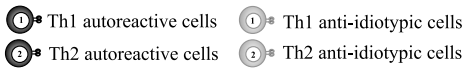


Figure 1: TCV experiments in mice.

keeping a fixed order of the different T-cell populations, the state of the immune system was subsequently described by a vector, in which e.g. "11000" corresponds to a state of autoimmunity in which antigen-specific helper cells and effector cells are active. The model consisted of a series of transition rules between the different states of the immune system.

TCV experiments were modelled by introducing effectors in the normal state "00000" and by applying the transition rules of the model until a stable state of the system was attained. A model like this would account for TCV if the introduction of antigen-specific effector T-cells can push the system to a state in which the antigen-specific effector cells are suppressed. It was shown that autoimmunity could never be a stable steady state in this model. Disease was therefore assumed to be due to an inadequate network connection. Cohen & Atlan [11] proposed that TCV might prevent autoimmunity by strengthening the connections in the network.

Our first criticism on this work is that the model is rather loosely connected to the data. It is based upon quite a large number of cell types and it is not obvious which of the five cell types corresponds to which cell type in TCV experiments. Secondly, autoimmune disease can also be interpreted as a "slow transient" leading to a state of vaccination, and need not be due to an inadequate network connection. A slow transient, however, is ill-defined in an automaton model.

3.2 Reverse engineering

Segel et al. [30, 31] used a very different approach by modelling the phenomenology of TCV, rather than the immunological processes underlying TCV, a method called "reverse engineering." The idea behind this approach is to investigate which set of mathematical models can describe a set of immunological phenomena. These models can subsequently serve as a first step in developing more realistic, data-based models.

The first step was to identify a set of stable steady states that suffices to describe all phenomena related to TCV. Segel et al. [30, 31] defined a normal state in which the number of autoreactive T-cells is small, a disease state in which the number of autoreactive T-cells is large, and a vaccinated state in which the number of anti-idiotypic T-cells is large. They subsequently constructed different mathematical models that could account for these three stable steady states. Secondly the set of initial conditions that led to each of the three stable steady states was defined. Experiments have demonstrated that the effect of the injection of autoreactive cells into a healthy individual is dependent on the dose of the injected cells: the injection of autoreactive cells should take the immune system across a first "critical boundary" beyond which it approaches the vaccinated state, while the injection of even more autoreactive cells would take the system across a second boundary beyond which it approaches the disease state. Based on such

constraints, Segel et al. [30, 31] developed a set of models that can account for TCV. Although this reverse engineering approach is quite unusual, it has helped to identify the different types of models that possibly provide insights into TCV, and has suggested novel experiments. The model suggested that inoculating too large a dose of autoreactive cells in an animal suffering from autoimmunity could cure the animal without leading to vaccination [30], and has thereby allowed us to "think the unthinkable" [31].

3.3 Idiotypic network models

Interactions in idiotypic networks have extensively been studied for B cells [36, 12, 13, 32], and were traditionally based on the assumption that idiotypic interactions are symmetric, i.e. that a cell stimulates or inhibits its anti-idiotypic partner as strongly as the anti-idiotypic partner stimulates or inhibits the specific cell. TCV has briefly been investigated in an anti-idiotypic network model [28]. However, the assumption of symmetric interactions does not hold for T-cells, because a T-cell that recognizes the TCR of another T-cell via its presentation on MHC molecules, is unlikely to be recognized by the latter T-cell.

Borghans et al. [4] developed a mathematical model of EAE that did not assume symmetric interactions between autoreactive and anti-idiotypic T-cells. The model showed how complex features of the disease and induction of immunity by TCV could be explained by a few relatively simple interactions between autoreactive and anti-idiotypic T-cells. The model was defined in terms of differential equations describing the rate of change of these two T-cell populations. The core of the model consisted of two feedback loops: a positive feedback loop between autoreactive T-cells and the expression of self antigen, and a negative feedback loop between autoreactive T-cells and anti-idiotypic T-cells. The negative feedback loop was based on the experimental observation that anti-idiotypic T-cells proliferate in response to autoreactive T-cells, while autoreactive T-cells are inhibited by anti-idiotypic T-cells. The positive feedback loop was based on the assumption that activated autoreactive T-cells produce cytokines that enhance the presentation of self antigen [5]. Hence autoreactive cells stimulate their own proliferation, leading to clonal expansion and autoimmunity.

The model has two stable steady states, a passive state and an active state of self-tolerance, denoted as the naive state and the vaccinated state, respectively. In the naive state both the autoreactive T-cell clone and the anti-idiotypic T-cell clone are non-activated, while in the vaccinated state autoreactive T-cells are actively suppressed by anti-idiotypic T cells. Disease was defined as a transient of very high autoreactive T-cell numbers which could be reached by increasing the number of autoreactive T-cells in the naive state beyond a certain threshold. If the increase was sufficiently large (e.g. due to the presence of a self-mimicking antigen), the autoreactive T-cells would continue to proliferate and clonally expand until the anti-idiotypic T-cell clone would expand as well and suppress the

autoreactive T-cell clone.

Vaccination in the model corresponded to a switch from the naive state to the vaccinated state without going through this transient of disease. The model predicted qualitatively different results for vaccination with a low-dose of autoreactive T cells, leading to long lasting protection, and vaccination with attenuated autoreactive T-cells or anti-idiotypic T-cells, both leading to a temporary increase in anti-idiotypic T-cell numbers resulting in dose-dependent transient protection from disease.

4 New TCV model including Th1/Th2 differentiation

None of the above models took into account the more recent finding that the Th1/Th2 phenotype of both autoreactive and anti-idiotypic T-cells plays an important role in the outcome of TCV. We therefore developed a new model to combine the experimental findings on Th1/Th2 regulation in EAE with the interactions between autoreactive and anti-idiotypic T-cells that were described in the previous model [4]. In the previous model [4], T-cell clones required sustained antigen presentation to remain at an elevated population size. Instead, we modeled a state of immunological memory giving distinct primary and secondary immune responses without the need for sustained exposure to antigen to maintain the memory cells. Below we first describe the phenomenological basis of the new model, then show how we extended the model with differentiation of naive T-cells into a Th1 or Th2 phenotype, and finally illustrate how regulation of autoreactive T-cells by anti-idiotypic T-cells was incorporated to study TCV.

4.1 General T-cell model

Our general T-cell model combines naive, memory, and effector cells of one particular T-cell clone into a single variable T. Changes in the number of cells of this clone are described by:

$$\frac{\mathrm{d}T}{\mathrm{d}t} = \text{influx} + \text{antigen-dependent proliferation} + \text{memory renewal} - \text{cell death} \tag{1a}$$

$$\frac{\mathrm{d}T}{\mathrm{d}t} = i + pTS + m\frac{T^2}{h^2 + T^2} - dT \tag{1b}$$

Here i is a small daily production of naive T-cells by the thymus, and possibly naive T-cell renewal. Cells die naturally at a rate d per day. Foreign antigenic stimulation, S, which may vary between zero and one, decays exponentially at rate r per day, i.e. $S(t) = e^{-rt}$, reflecting the rate at which the foreign antigen is cleared. Antigen-dependent T-cell proliferation occurs at maximal rate p per

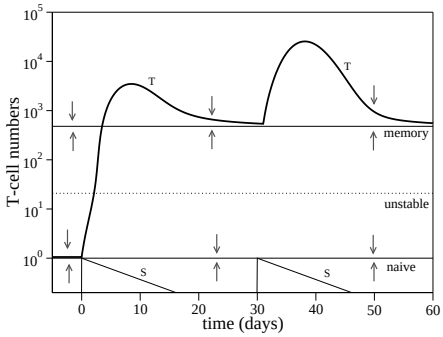


Figure 2: Primary and secondary responses in the general T-cell model. Steady states are depicted by horizontal lines; solid lines denote the stable naive and memory states and the dotted line denotes the unstable state. The unstable steady state forms the boundary (separatrix) between the basins of attraction of the stable steady states. Immune responses after primary stimulation at day 0 (S=1), and after re-stimulation at day 30 (S=1) are shown by the thick curves. Arrows indicate the attraction towards the stable steady states and antigen stimulation is denoted by the thin curves. Parameters: $i = 1$ cell/day, $p = 2$ per day, $m = 500$ cells/day, $h = 100$ cells, $d = 1$ per day, $f = 0.9$, $r = 0.1$, $S(t) = e^{-rt}$.

day when antigen stimulation is maximal, i.e. $S = 1$. If the T-cell population size becomes large enough, the threshold for memory T-cell induction is breached. At a clonal population size of $T = h$, memory renewal division proceeds at half its maximal rate of m cells per day.

When parameters are set properly, this equation allows for maximally three steady states. The first steady state is the state in which the T-cell clone is naive and there is no T-cell memory because it has not yet encountered (enough) antigen ($T \ll h$). Because the contribution of the proliferation term and the memory renewal term can be neglected in this state, the number of T-cells in the naive steady state is approximately $T = i/d$ cells. When a naive T-cell clone is temporarily exposed to antigen with the appropriate co-stimulatory signals, antigen-dependent proliferation sets in and initiates clonal expansion. As a result, the population size breaches the threshold for memory induction. When memory sets in, another steady state can be reached. Once the antigen has been cleared, and the T-cell clone is sufficiently large, the model can be simplified into $\mathrm{d}T/\mathrm{d}t = m\frac{T^2}{h^2+T^2} - dT$ because the small production i by the thymus has a negligible contribution compared to memory T-cell renewal divisions. Calculating the steady states of this equation shows that there is a large stable steady state of $T \simeq m/d - \frac{d \cdot h^2}{m}$ cells and a smaller unstable steady state of $T \simeq \frac{d \cdot h^2}{m}$ cells. The above described total of three steady states are obtained if the threshold for memory T-cell induction is substantially higher than the naive clonal population size, i.e. if $h \gg i/d$.

To illustrate that the above model correctly describes naive and memory responses to antigen, we simulated the response of a T-cell clone to repeated antigenic stimulation (Figure 2). When a naive T-cell clone was challenged with antigen at day 0, the T-cell clone expanded and breached the threshold for the induction of memory T-cells. A primary immune response was initiated reaching a peak of ± 4000 cells at approximately seven days post-stimulation. When the stimulus decreased, memory renewal divisions took over and maintained the T-cell clone at approximately 500 cells. When this T-cell clone was rechallenged with the same antigen at day 35, the secondary immune response reached a peak of ± 20000 T-cells and eventually returned to the same memory state.

4.2 A Phenomenological Th1/Th2 model

Model description

The balance between Th1 and Th2 responses has been modeled before, among others by Fishman et al. [14], Louzoun et al. [25] and Yates et al. [37]. These earlier models were quite detailed, in that they were based on the specific influences of various cytokines. In order to understand the effect of the Th1/Th2 switch on TCV, we prefered a simpler caricature model for the Th1/Th2 balance; this allowed us to focus on the effect of T-cell phenotypes in a complex multi-cellular regulatory system, rather than on the precise Th1/Th2 molecular machinery.

Differentiation of a naive T-cell towards a Th1 or Th2 phenotype is dependent on its cytokine environment. Since T-helper cells secrete cytokines that enhance differentiation into their own phenotype and inhibit differentiation into the opposite phenotype, we made T-cell differentiation dependent on the fraction of Th1 and Th2 cells present. To keep track of the number of Th1 and Th2 cells, the general T-cell model was applied to Th1 and Th2 cells separately:

$$\frac{\mathrm{d}T_1}{\mathrm{d}t} = i + pT_1Sv \qquad + m\frac{T_1^2}{h^2 + T_1^2} - dT_1 \tag{2a}$$

$$\frac{\mathrm{d}T_2}{\mathrm{d}t} = i + pT_2S(1-v) + m\frac{T_2^2}{h^2 + T_2^2} - dT_2 \tag{2b}$$

Here, T_1 and T_2 represent the number of T-cells with a Th1 or Th2 phenotype, respectively, and v $(0 < v < 1)$ is a function representing the influence of the Th1 and Th2 cells on the differentiation of naive cells into Th1 cells:

$$v = \frac{T_1}{T_1 + fT_2} \tag{3}$$

Because the influence of cytokines is most important during the onset of the initial immune response, when the activated T-cell clone is maturing into one particular phenotype, the function v only influences antigen-dependent proliferation. We assumed that daughter cells arising by memory renewal take over the same phenotype as their progenitors. Indeed, it has been shown that memory T-cells of either

the Th1 or Th2 phenotype, are much less likely to switch to the opposite phenotype than naive T-cells, when exposed to changes in the cytokine environment [9, 19]. The parameter f was used to favour differentiation into either of the two phenotypes. Because the dominant T-cell response in mice with MBP-induced EAE is typically of the Th1 phenotype [1] we set $f = 0.9$, thereby introducing a small bias in favour of Th1 differentiation (i.e. at equal Th1 and Th2 densities, $v = 0.53$). Every clone alone has its own independent memory-renewal term that will maintain a population of memory cells once the threshold for memory induction is breached.

Steady state description

Since the v-function is the only difference between the the general T-cell model and the equations of the phenomenological Th1/Th2 model, each phenotype equation has the same two stable steady states as the general T-cell model: the naive state and the memory state. We will represent the different steady states by assigning '0' to naive clones and '1' to memory clones, in the order 'T_1T_2'. Thus, the naive state is denoted by '00', the Th1-dominated memory state by '10', the Th2-dominated memory state by '01', and the state with both memory phenotypes by '11'. In our model the latter state cannot be attained through antigenic stimulation, and is only reached if memory Th1 and Th2 cells are artificially brought together.

Like before, we simulated the response to repeated stimulation with antigen, now following both the Th1 and the Th2 cells. After antigen stimulation at day 0 ($S = 1$), at first both clones expanded, but since the model favours differentiation into the Th1 phenotype, the Th2 clone was rapidly inhibited by the more dominant Th1 response. Stimulation was strong enough to drive the Th1 population over the threshold for memory induction, so that a stable Th1-dominated response (10) was maintained in the absence of stimulation (Figure 3a). The same behaviour can be depicted by plotting the number of Th1 autoreactive cells against the number of Th2 autoreactive cells (Figure 3b). Antigenic stimulation pulled the system out of the basin of attraction of the naive steady state, into the basin of attraction of the Th1 dominated memory state 10. In this 10 memory state the function v approximated one, making it impossible to obtain a Th2 response by re-stimulation with antigen as shown in Figure 3a where the system was re-stimulated at day 30.

4.3 TCV model with anti-idiotypic regulation and Th1/Th2 differentiation

We used the phenomenological Th1/Th2 model described above to extend the model for anti-idiotypic regulation in EAE [4]. The model consists of four $CD4^+$ T-cell clones, a Th1 autoreactive clone (A_1), a Th2 autoreactive clone (A_2), a Th1 anti-idiotypic clone (R_1) and a Th2 anti-idiotypic clone (R_2). The Th1 au-

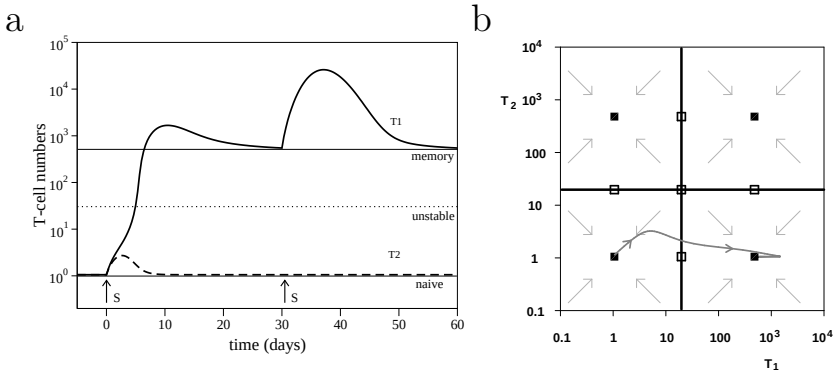


Figure 3: a: Primary and secondary responses in the phenomenological Th1/Th2 model. Steady states are depicted by horizontal lines; solid lines denote the stable naive and memory states and the dotted line denotes the unstable state. The unstable steady state forms the boundary (separatrix) between the basins of attraction of the stable steady states. Immune responses after primary stimulation at day 0 (S=1), and after re-stimulation at day 30 (S=1) are shown by the thick curves. Antigen stimulation is denoted by the arrows and the thick solid curve denotes the Th1 cells and the dashed thick curve denotes the Th2 cells. b: The number of Th1 cells (T1) plotted against the number of Th2 cells (T2) for the Th1/Th2 phenomenological model. Stable and unstable steady states are shown as solid and open squares, respectively. Arrows indicate the attraction towards the stable steady states and bold lines represent the separatrices between the basins of attraction of the steady states. The grey curve depicts the trajectory the model makes after primary stimulation ($S = 1$). Parameters: $i = 1$ cell/day, $p = 2$ per day, $m = 500$ cells/day, $h = 100$ cells, $d = 1$ per day, $f = 0.9$, $r = 0.1$, $S(t) = e^{-rt}$.

toreactive clone, A_1, is the dominant $V_\beta 8.2 - J_\beta 2.7$ T-cell clone, which induces a response against the immuno-dominant MBP peptide Ac1-9 in BL10.PL mice. The Th2 autoreactive cells are stimulated by other, sub-dominant, MBP epitopes, and have a different T-cell receptor [26]. In contrast, both Th1 anti-idiotypic cells and Th2 anti-idiotypic cells are of the same clonal origin and reactive to the same peptide, the dominant TCR peptide Fr3 of the Th1 autoreactive T-cells. As a consequence, only the Th1 autoreactive cells are down-regulated by anti-idiotypic cells. Since only Th1 anti-idiotypic cells can activate $CD8^+$ cytotoxic T-cells, they are the ones that have an inhibitory effect on the autoreactive Th1 cells while the Th2 anti-idiotypic cells do not have any direct influence on the autoreactive Th1 clone [26].

Because both Th1 autoreactive cells and Th2 autoreactive cells are activated by MBP on probably the same APCs, the cytokine environment in which both clones are activated is expected to be very similar. We therefore assumed that upon activation, Th1 autoreactive cells and Th2 autoreactive cells exert an inhibitory effect on each other as described in the phenomenological Th1/Th2 model. The same was done for the inhibitory effect between the Th1 anti-idiotypic cells and Th2 anti-idiotypic cells which are both activated by the same peptide. The above mentioned interactions between the autoreactive Th1 and Th2 clones and the anti-idiotypic Th1 and Th2 clones are represented schematically in Figure 4, and are mathematically described by the following differential equations:

$$\frac{\mathrm{d}A_1}{\mathrm{d}t} = i + pA_1Sv_A + m\frac{A_1^2}{h^2 + A_1^2} - dA_1 - kA_1R_1 \tag{4a}$$

$$\frac{\mathrm{d}A_2}{\mathrm{d}t} = i + pA_2S(1 - v_A) + m\frac{A_2^2}{h^2 + A_2^2} - dA_2 \tag{4b}$$

$$\frac{\mathrm{d}R_1}{\mathrm{d}t} = i + pR_1S_Rv_R + m\frac{R_1^2}{h^2 + R_1^2} - dR_1 \tag{4c}$$

$$\frac{\mathrm{d}R_2}{\mathrm{d}t} = i + pR_2S_R(1 - v_R) + m\frac{R_2^2}{h^2 + R_2^2} - dR_2 \tag{4d}$$

where

$$v_A = \frac{A_1}{A_1 + fA_2} \tag{5}$$

and

$$v_R = \frac{R_1}{R_1 + fR_2} \tag{6}$$

Note that anti-idiotypic $CD8^+$ cells are not modelled explicitly; instead, their recruitment is assumed to be proportional to the number of Th1 anti-idiotypic cells (A_1). Therefore the anti-idiotypic suppression term is $-kA_1R_1$, with a suppression rate k per R_1 cell per day.

To model the antigenic stimulation of the anti-idiotypic cells we adopted the stimulation function proposed by Borghans et al. [4], which saturates as a function of

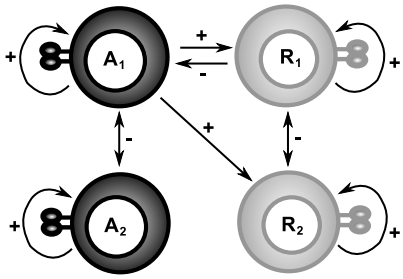


Figure 4: Schematic representation of the interactions between the T-cells of the new TCV model including Th1/Th2 differentiation. A_1= MBP-reactive Th1 cells, A_2= MBP-reactive Th2 cells, R_1= anti-idiotypic Th1 cells and R_2= anti-idiotypic Th2 cells.

the number of Th1 autoreactive cells, and which involves competition for antigen binding between anti-idiotypic cells:

$$S_R = \frac{A_1}{A_1 + c(R_1 + R_2) + h_R} \tag{7}$$

Here, c determines the strength of the competition, and prevents unrealistically high levels of anti-idiotypic cells during sustained antigen stimulation. In the absence of competition ($c(R_1 + R_2) = 0$), half-maximal stimulation is reached at $A_1 = h_R$ Th1 autoreactive cells. The parameter h_R thus determines the sensitivity of the anti-idiotypic cells for activation by the Th1 autoreactive cells. There is no direct interaction between the Th2 autoreactive cells and any of the anti-idiotypic cells (see Figure 4). In fact, the only influence the Th2 autoreactive population exerts on the system is its inhibitory effect on the differentiation of autoreactive cells into Th1 cells.

Steady state description

Just like in the general T-cell model, every clone can only be in either of two stable steady states, the stable naive or the stable memory state. Without interactions between the clones this yields a maximum of $2^4 = 16$ stable steady states. Again, steady states are represented by assigning '0' to naive clones and '1' to memory clones; 'x' is used to denote that a clone can be either naive or memory, and the states of the clones are given in the order '$A_1A_2R_1R_2$'.

In the presence of interactions between the clones there are fewer stable steady states. We chose h_R high enough to ensure that a naive Th1 autoreactive cell population (A_1) would not cause expansion of the anti-idiotypic cells ($h_R \gg i/d$) and small enough to allow the anti-idiotypic cell population to expand beyond the memory threshold when the Th1 autoreactive cells approach their memory state ($h_R \ll m/d - \frac{d \cdot h^2}{m}$). By doing so, the 1x00 states no longer existed, because the

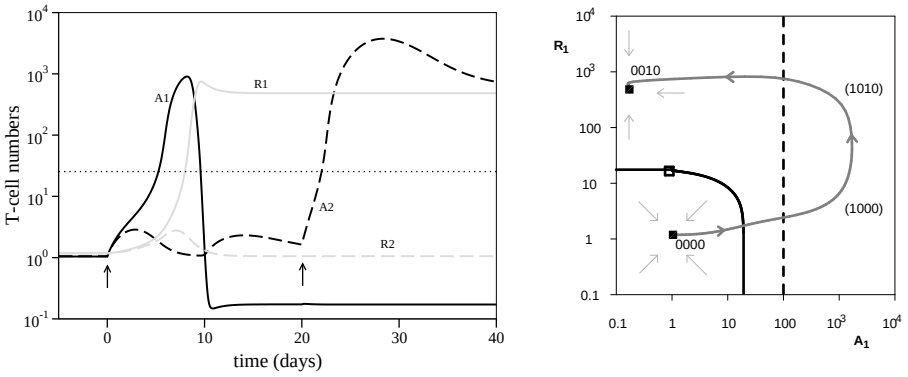


Figure 5: Primary stimulation and secondary challenge with MBP in the new model with anti-idiotypic regulation and Th1/Th2 differentiation. The left-hand panel depicts a time plot, the right-hand panel depicts a phase diagrams. Right-hand side: Only primary stimulation is shown. In the time plot a small arrows depicts A1 cells stimulation and the dotted line depicts the threshold for memory. In the phase diagram solid and open squares depict stable and unstable steady states respectively. Thick solid lines depict the separatrices between the stable steady states and the arrows indicate the direction of the attraction. The region in which the Th1 autoreactive population is large is interpreted as "disease" and reached when the thick dashed line is crossed. Trajectories after stimulation ($S = 1$) are depicted by the grey curves. Parameters: $i = 1$ cell/day, $p = 2$ per day, $f = 0.9$, $m = 500$ cells/day, $h = 100$ cells, $d = 1$ per day, $r = 0.1$ per day, $c = 0.1$, $h_R = 10$ cells, $k = 0.01$ per day.

memory Th1 autoreactive cells would stimulate the anti-idiotypic cells to expand.

The suppression parameter was set such that once the Th1 anti-idiotypic cells (R_1) have reached their memory state, their influence would be strong enough to keep the Th1 autoreactive cells (A_1) at a very low level. Consequently, the 1x1x states were also no longer steady states of the model, because the anti-idiotypic suppression ensures that memory Th1 autoreactive cells and memory Th1 anti-idiotypic cells cannot exist together.

The above parameter setting removed six steady states, i.e., 2× (1x00) + 4× (1x1x), from the model leaving 10 stable steady states (see Table 1).

4.4 Results

Antigen-induced EAE and re-challenge

We simulated antigen-induced EAE (experiment in Figure 1a) by starting the system in the naive steady state (0000) and setting $S = 1$ at day 0, which stimulates both autoreactive clones (Figure 5). Immediately after the antigenic stimulation,

clonal expansion of both autoreactive clones started at maximal proliferation rate p. Due to the imposed Th1 bias ($f = 0.9$), the autoreactive response differentiated into a Th1 phenotype, and within a few days the Th2 autoreactive cells fell back to their naive level (Figure 5). Stimulated by the expansion of the Th1 autoreactive cells both anti-idiotypic populations started to expand. Just like the autoreactive response, the anti-idiotypic response differentiated into a Th1 phenotype within a few days, and the Th2 anti-idiotypic (R_2) cells started to fall back to their naive level. Subsequently, the expanding Th1 anti-idiotypic (R_1) cells started to suppress the Th1 autoreactive (A_1) cells. By the time the Th1 anti-idiotypic population approached its memory level, the Th1 autoreactive population had decreased dramatically from $\pm$ 1000 cells until below the naive level. Deprived of antigenic stimulation due to this drop in the number of Th1 autoreactive cells, the Th1 anti-idiotypic population started to fall back to its memory level. This stable steady state 0010 was subsequently maintained by memory renewal divisions of the Th1 anti-idiotypic cells. The same behaviour was observed when the number of Th1 autoreactive cells (A_1) was plotted against the number of Th1 anti-idiotypic cells (R_1) (Figure 5, right-hand side). Antigenic stimulation pulled the system out of the basin of attraction of the naive steady state, through the region of disease and then towards steady state 0010. Stimulating the autoreactive Th1 cells with antigen thus led to a temporal increase in the number of autoreactive cells, which got cured by the induction of anti-idiotypic Th1 cells, in line with the experimental observations in mice (Figure 1a).

Since most animals that have contracted EAE are subsequently resistant to the disease, we tested what happens if the autoreactive cells in the 0010 steady state are re-stimulated by antigen at day 20 (Figure 5 left-hand side). The Th2 autoreactive cells could expand in response to the stimulation, because they were not inhibited by Th1 autoreactive cells, which were still suppressed by the Th1 anti-idiotypic clone. The stimulation brought the Th2 autoreactive cells, which are not suppressed by anti-idiotypic cells, over the threshold for memory induction. However, the presence of Th2 autoreactive cells failed to influence the other T-cell populations, and once the antigenic stimulation had faded, the system approached the steady state with Th2 autoreactive and Th1 anti-idiotypic memory cells (0110, Figure 5 left-hand side, day 40). Thus re-challenge did not lead to an expansion of disease-causing Th1 autoreactive cells (Figure 5), in accordance with the experimental results (Figure 1a). Follow-up stimulations of the autoreactive cells caused the Th2 autoreactive memory cells to expand temporarily, as long as the stimulation lasted, but the system always returned to the same steady state (0110, not shown).

T-cell vaccination with TCR Fr3 peptide

TCV with the TCR Fr3 peptide of the autoreactive cells, activates Th1 anti-idiotypic T-cells (R_1). We therefore simulated TCV in a Th1 context by initializing the Th1 anti-idiotypic cells at their memory level, i.e. by initiating the system

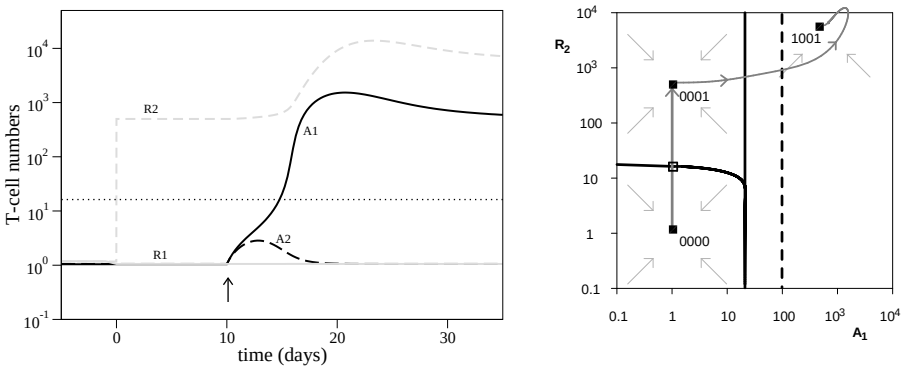


Figure 6: Nasal instillation with the TCR Fr3 peptide and subsequent challenge with MBP on day 10 in the new model with anti-idiotypic regulation and Th1/Th2 differentiation. The left-hand panel depicts a time plot, the right-hand panel depicts a phase diagrams. Nasal instillation activates Th2 anti-idiotypic cells and was simulated by setting the Th2 anti-idiotypic cells at their memory state at t=0. (for legend and parameter values see caption Figure 5)

at steady state 0010. When the autoreactive cells in the system were subsequently stimulated ($S = 1$), only the Th2 autoreactive cells responded to the stimulation and after the antigen was cleared the system approached state 0110 (similar to Figure 5, left-hand side, day 20). Thus, in accordance with the experimental findings (Figure 1b), TCV in a Th1 context did not cause autoimmune disease.

Nasal instillation with TCR Fr3 peptide

We also simulated nasal instillation with TCR Fr3 peptide, by setting the Th2 anti-idiotypic cells at their memory level, i.e. by starting the system in state 0001 (Figure 6). When the autoreactive cells were subsequently stimulated with antigen on day 10 ($S = 1$), the autoreactive response rapidly differentiated into a Th1 phenotype. Stimulated by the expansion of the Th1 autoreactive cells, the Th2 anti-idiotypic memory population started to expand suppressing the activation of the Th1 anti-idiotypic cells. When the antigen was cleared, the Th1 autoreactive population fell back to its memory level. The Th1 autoreactive memory population continued to stimulate the Th2 anti-idiotypic population, leading the system to state 1001. The same behaviour can be depicted by plotting the number of Th1 autoreactive cells (A_1) against the number of Th2 anti-idiotypic cells (R_2) (Figure 6 right-hand side). Antigenic stimulation in state 0001, pulled the system out of the basin of attraction of the 0001 state, into the disease region, towards steady state 1001.

Antigen challenge after TCV by nasal instillation thus led the system to a state of permanent autoimmunity in the presence of Th2 anti-idiotypic cells (Figure 6).

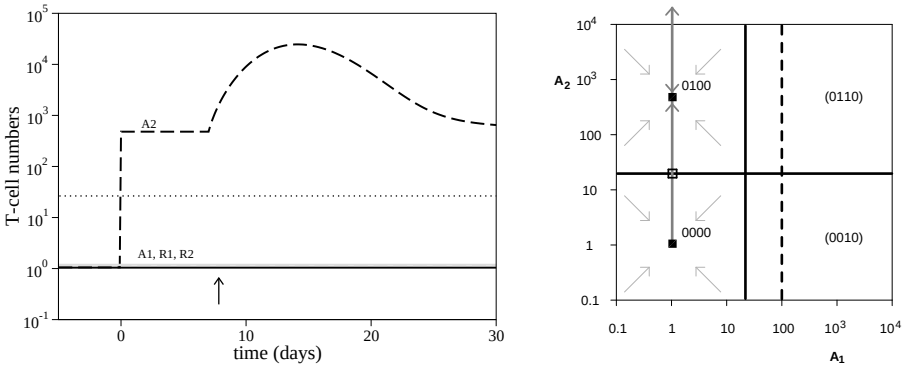


Figure 7: Stimulation at day 7, in the presence of Th2 autoreactive memory cells in the new model with anti-idiotypic regulation and Th1/Th2 differentiation. The left-hand panel depicts a time plot, the right-hand panel depicts a phase diagrams. (for legend and parameter values see caption Figure 5)

This is in line with the experimental observation that mice that are challenged with antigen after TCV through nasal instillation get severe EAE and do not recover (Figure 1c). Re-stimulation of the autoreactive cells in this state led to an even further, temporary, expansion of the Th1 autoreactive population followed by a temporary re-expansion of the Th2 anti-idiotypic population. After the antigen had been cleared the system returned to the 1001 state (not shown).

Protection by Th2 autoreactive cells

From experiments on EAE in mice it became clear that Th2 responses to MBP do not cause disease [26, 38], and can even be protective [26, 29]. To study the influence of Th2 autoreactive cells in our model, the system was initiated in the state 0100, in which only autoreactive Th2 memory cells are present. When both autoreactive T-cell clones were subsequently stimulated by antigen, only the Th2 autoreactive T-cell clone expanded, because of the memory Th2 cells already present (Figure 7 left-hand side). The autoreactive Th1 cells failed to respond because they were suppressed by the autoreactive Th2 memory cells. Both anti-idiotypic clones R_1 and R_2 remained in their naive state, because the Th2 autoreactive cells failed to stimulate them. The same behaviour was depicted by plotting the number of Th1 autoreactive cells (A_1) against the number of Th2 autoreactive cells (A_2) (Figure 7 right-hand side). The system was pushed out of state 0100 by the antigenic stimulation but could not leave the basin of attraction of state 0100, and was therefore pulled back to state 0100 after the stimulation faded away. These simulations suggest that 0100 is an alternative "vaccinated" state in the absence of anti-idiotypic regulation, in which Th2 autoreactive cells avoid the outgrowth of Th1 autoreactive cells.

A All steady states	**B** Naive/Vulnerable
0000	$0000 \Rightarrow 1000 \rightarrow 1010 \rightarrow 0010$
0001	$0001 \Rightarrow 1001$
	Vaccinated
$1000 \rightarrow 1010 \rightarrow$ **0010**	$0010 \Rightarrow 0110$
$1011 \rightarrow$ **0011**	$0011 \Rightarrow 0111$
0100	$0100 \Rightarrow 0100$
0101	$0101 \Rightarrow 0101$
$1100 \rightarrow 1110 \rightarrow$ **0110**	$0110 \Rightarrow 0110$
$1111 \rightarrow$ **0111**	$0111 \Rightarrow 0111$
	Disease
1001	$1001 \Rightarrow 1001$
1101	$1101 \Rightarrow 1101$

Table 1: Behavior of the four-dimensional model in a logical framework. A: Stable steady states approached when the system is initiated in either of the 16 steady states of the model. The stable steady states are shown in boldface. The arrows indicate the states passed when converging from each of the possible steady states to the stable steady states. B: The response of the model to antigenic stimulation of the MBP-reactive cells, $S(t) = e^{-rt}$, starting from the ten stable steady states. Stimulation is indicated by $\Rightarrow$.

Summary of the results

The above in silico experiments show that the effect of stimulating autoreactive cells depends crucially on the initial state of the system. To summarize all possible effects, we have stimulated the autoreactive T-cells in all stable steady states, and categorized them according to the result of the stimulation (Table 1). The stable steady states of this four-dimensional model can be categorized as follows: 'Naive/Vulnerable' (0000, 0001), 'Vaccinated' (0100, 0010, 0110, 0101, 0011, 0111) and 'Disease' (1001, 1101). In the vaccinated states the system is protected from disease since stimulation does not lead the system through the disease region. The vulnerable states do not lie within the disease region, but stimulation starting from such a state will bring the system to a state in the disease region. Within the logic framework of Table 1 we depicted the activation or depletion of the four different T-cell populations during these stimulations without drawing a full time plot (like was done in the early automaton models [11, 35]).

5 Discussion

Over the last years the use of models that attempt to provide a full description of the immune system has gained considerable popularity (eg. IMMSIM [8, 3] and Simmune [27]). Such general models attempt to include all of the detailed immunological knowledge that is available. There is no chance, however, that any model will ever be complete, no matter how much detail is incorporated,

and a great disadvantage of this complex modelling approach is that it hampers the interpretation of the results, and thereby often fails to deliver insights into the experimental findings. We are convinced that the beauty of more conceptual models lies in their simplicity and transparency, which truly helps to improve our insights into complex systems like the immune system [24]. In that respect we fully agree with the previously made statement: "All models are wrong, but some are useful" [6], and although many models "certainly do not capture the reality in full, and no model does ... they capture certain aspects and give a general direction on how to understand the issue better [7]."

Despite the relative simplicity of our mathematical model for the role of Th1/Th2 phenotypes in TCV, all of the in silico experiments are in good agreement with the experimental observations on TCV. The model accounts for i. transient disease after a first encounter with an MBP-mimicking antigen, ii. protection from EAE after TCV in a Th1-mode, iii. exacerbation of disease after TCV in a Th2 context and, iv. the shift from a Th1-dominated MBP-reactive response to a Th2-dominated response after TCV. All of these phenomena are apparently a natural consequence of the interactions between the different T-cell populations involved.

The fact that anti-idiotypic cells could shift the phenotype of autoreactive cells from a Th1 to a Th2 phenotype following TCV was initially conceived with surprise [10]. Our modelling results show, however, how such a Th1/Th2 shift of autoreactive cells results from the interactions between anti-idiotypic cells and Th1 and Th2 autoreactive cells, without the need for any direct influence of anti-idiotypic cells on the phenotype of autoreactive cells. The shift arises naturally because Th1 anti-idiotypic cells only suppress the Th1 autoreactive cells, and thus allow the Th2 autoreactive cells to proliferate in response to MBP stimulation.

Our previous model [4] already showed the general features of TCV in EAE by modeling anti-idiotypic regulation alone. There are, however, several important differences between the current and the previous model. In the new model autoreactive T-cells can be maintained by memory renewal, while in the previous model they could only be maintained by the positive feedback between autoreactive cells and the presentation of their self peptide. As a result, the previous model showed a qualitative difference between vaccination with attenuated autoreactive cells, and vaccination with a low dose of live autoreactive T-cells [4]. Vaccination with attenuated autoreactive cells could only give rise to transient protection from disease, because it failed to trigger the positive feedback loop between autoreactive cells and self antigen, while low-dose vaccination with live autoreactive cells could give rise to long-lasting protection from disease. In the new model both methods of vaccination give rise to long-lasting protection, because Th1 anti-idiotypic cells can be maintained by memory renewal.

Another characteristic of the previous model was that it could explain relapsing-remitting disease by changing only one parameter of the model [4]. Relapsing-

remitting disease characterized by sporadic attacks (relapses) followed by a period of partial or total recovery (remission) is commonly seen in MS patients. In the current model we have not yet found a parameter setting that could give rise to such relapses, which does not exclude the possibility that a more extensive parameter search would have revealed this kind of dynamic behaviour. Alternatively, relapses may be due to other mechanisms that are not included in the models, such as the emergence of new autoreactive clones. Indeed, in clinical trials of TCV in MS patients, new T-cell responses against MBP with similar functional properties but different clonal origins compared to the cells that were used for TCV have repeatedly been observed a few years after TCV [16]. Such novel autoreactive responses against MBP have, however, never been observed after TCV in mice. A possible explanation for this discrepancy is that most experiments on TCV in mice have been performed on H-2^u strains of BL10.PL mice. These mice have a T-cell repertoire of which 60% expresses the $V_\beta 8$ chain (personal communication V. Kumar). When Th1 anti-idiotypic cells become activated they can down-regulate all activated Th1 cells expressing this TCR chain, which should leave the mice with only 40% of their repertoire.

Summarizing, simple mathematical models have contributed significantly to our understanding of the role of anti-idiotypic regulation in TCV. By also including Th1/Th2 phenotype switches, the current TCV model helps to structure TCV experiments by showing how the large variety of experimental observations on TCV naturally results from a few relatively simple interactions between Th1 and Th2 autoreactive and anti-idiotypic cells.

Very recently, a novel population of nonintestinal CD8alphaalpha+TCRalphabeta+ regulatory T cells (Treg) has been shown to be capable of controlling the activated Vbeta8.2+ CD4+ T cells that mediate EAE. Adoptive transfer or in vivo activation of these CD8+ Treg cells was shown to prevent the induction of EAE [33, 34]. Extension of the here proposed model with this class of CD8+ Treg cells may shed light on the relative contribution of anti-idiotypic CD4+ T cells and these CD8+ Treg cells in TCV experiments.

acknowledgements: *The authors would like to thank Vipin Kumar and Eli Sercarz for extensive discussions on the topic of TCV, and are grateful to Vipin Kumar for critically reading the manuscript.*

References

[1] D.G. Ando, J. Clayton, D. Kono, J.L. Urban, and E.E. Sercarz. Encephalitogenic T cells in the B10.PL model of experimental allergic encephalomyelitis (EAE) are of the Th-1 lymphokine subtype. *Cell Immunol*, 124(1):132–43, 1989.

[2] A. Ben-Nun, H. Wekerle, and I. R. Cohen. Vaccination against autoimmune encephalomyelitis with t-lymphocyte line cells reactive against myelin basic protein. *Nature*, 292(5818):60–61, 1981.

[3] M. Bernaschi and F. Castiglione. Design and implementation of an immune system simulator. *Comput Biol Med*, 31(5):303–331, 2001.

[4] J. A. Borghans, R. J. De Boer, E. Sercarz, and V. Kumar. T cell vaccination in experimental autoimmune encephalomyelitis: a mathematical model. *J. Immunol.*, 161:1087–1093, 1998.

[5] G. F. Bottazzo, R. Pujol-Borrell, T. Hanafusa, and M. Feldmann. Role of aberrant hla-dr expression and antigen presentation in induction of endocrine autoimmunity. *Lancet*, 2(8359):1115–1119, 1983.

[6] G.E.P. Box. *Robustness in the strategy of scientific model building.* Academic Press: New York., 1979.

[7] D. Butler and J. Hogan. Modellers seek reason for low retraction rates. *Nature*, 447:236–7, 2007. doi: doi:10.1038/447236b.

[8] F. Celada and P. E. Seiden. A computer model of cellular interactions in the immune system. *Immunol Today*, 13(2):56–62, 1992.

[9] R. L. Coffman. Origins of the t(h)1-t(h)2 model: a personal perspective. *Nat Immunol*, 7(6):539–541, 2006. doi: 10.1038/ni0606-539.

[10] I. R. Cohen. T-cell vaccination for autoimmune disease: a panorama. *Vaccine.*, 20:706–710, 2001.

[11] I. R. Cohen and H. Atlan. Network regulation of autoimmunity: an automation model. *J Autoimmun*, 2(5):613–625, 1989.

[12] R. J. de Boer and A. S. Perelson. Size and connectivity as emergent properties of a developing immune network. *J Theor Biol*, 149(3):381–424, 1991.

[13] R. J. de Boer, M. Boerlijst, B. Sulzer, and A. S. Perelson. A new bell-shaped function for idiotypic interactions based on cross-linking. *Bulletin of Mathematical Biology*, 58(2):285–312, 1996.

[14] M. A. Fishman and A. S. Perelson. Th1/th2 cross regulation. *J Theor Biol*, 170(1):25–56, 1994. doi: 10.1006/jtbi.1994.1166.

[15] B. Formby and T. Shao. T cell vaccination against autoimmune diabetes in nonobese diabetic mice. *Ann Clin Lab Sci*, 23(2): 137–147, 1993.

[16] G. Hermans, R. Medaer, J. Raus, and P. Stinissen. Myelin reactive T cells after T cell vaccination in multiple sclerosis: cytokine profile and depletion by additional immunizations. *J. Neuroimmunol.*, 102:79–84, 2000.

[17] J. Holoshitz, Y. Naparstek, A. Ben-Nun, and I. R. Cohen. Lines of t lymphocytes induce or vaccinate against autoimmune arthritis. *Science*, 219(4580):56–58, 1983.

[18] H. Jiang, N. S. Braunstein, B. Yu, R. Winchester, and L. Chess. CD8+ T cells control the th phenotype of mbp-reactive CD4+ T cells in eae mice. *Proc Natl Acad Sci U S A*, 98(11):6301–6306, 2001. doi: 10.1073/pnas.101123098.

[19] X. Jiao, R. Lo-Man, N. Winter, E. Dériaud, B. Gicquel, and C. Leclerc. The shift of th1 to th2 immunodominance associated with the chronicity of mycobacterium bovis bacille calmette-guérin infection does not affect the memory response. *J Immunol*, 170(3):1392–1398, 2003.

[20] H. P. Jones, L. M. Hodge, K. Fujihashi, H. Kiyono, J. R. McGhee, and J. W. Simecka. The pulmonary environment promotes Th2 cell responses after

nasal-pulmonary immunization with antigen alone, but Th1 responses are induced during instances of intense immune stimulation. *J. Immunol.*, 167:4518–4526, 2001.

[21] V. Kumar and E. Sercarz. Induction or protection from experimental autoimmune encephalomyelitis depends on the cytokine secretion profile of TCR peptide-specific regulatory CD4 T cells. *J. Immunol.*, 161:6585–6591, 1998.

[22] V. Kumar, J. Maglione, J. Thatte, B. Pederson, E. Sercarz, and E. S. Ward. Induction of a type 1 regulatory CD4 T cell response following V *beta* 8.2 DNA vaccination results in immune deviation and protection from experimental autoimmune encephalomyelitis. *Int. Immunol.*, 13:835–841, 2001.

[23] R. S. Liblau, S. M. Singer, and H. O. McDevitt. Th1 and Th2 CD4+ T cells in the pathogenesis of organ-specific autoimmune diseases. *Immunol. Today*, 16: 34–38, 1995.

[24] Y. Louzoun. The evolution of mathematical immunology. *Immunological Reviews*, 216(1):9–20, 2007. doi: 10.1111/j.1600-065X.2006.00495.x.

[25] Y. Louzoun, H. Atlan, and I. R. Cohen. Modeling the influence of TH1- and TH2-type cells in autoimmune diseases. *J. Autoimmun.*, 17:311–321, 2001.

[26] L. T. Madakamutil, I. Maricic, E. Sercarz, and V. Kumar. Regulatory T cells control autoimmunity in vivo by inducing apoptotic depletion of activated pathogenic lymphocytes. *J. Immunol.*, 170:2985–2992, 2003.

[27] M. Meier-Schellersheim and G. Mack. Simmune, a tool for simulating and analyzing immune system behavior. *ArXiv Computer Science e-prints*, 1999.

[28] A.S. Perelson and R. J. de Boer. Autoimmunity and t-cell vaccination in a model idiotypic network. *J Cell Biochem*, 20th annual meetings:Supplement 15A, 1991.

[29] F. Ramirez and D. Mason. Induction of resistance to active experimental allergic encephalomyelitis by myelin basic protein-specific Th2 cell lines generated in the presence of glucocorticoids and IL-4. *Eur. J. Immunol.*, 30:747–758, 2000.

[30] L. A. Segel and E. Jäger. Reverse engineering: a model for t-cell vaccination. *Bull Math Biol*, 56(4):687–721, 1994.

[31] L. A. Segel, E. Jäger, D. Elias, and I. R. Cohen. A quantitative model of autoimmune disease and t-cell vaccination: does more mean less? *Immunol Today*, 16(2): 80–84, 1995.

[32] B. Sulzer, R. J. De Boer, and A. S. Perelson. Cross-linking reconsidered: binding and cross-linking fields and the cellular response. *Biophys J*, 70(3):1154–1168, 1996.

[33] X. Tang, I. Maricic, N. Purohit, B. Bakamjian, L. M. Reed-Loisel, T. Beeston, P. Jensen, and V. Kumar. Regulation of immunity by a novel population of qa-1-restricted cd8alphaalpha+tcralphabeta+ t cells. *J Immunol*, 177(11):7645–7655, 2006.

[34] X. Tang, I. Maricic, and V. Kumar. Anti-tcr antibody treatment activates a novel population of nonintestinal cd8alphaalpha+tcralphabeta+ regulatory t cells and prevents experimental autoimmune encephalomyelitis. *J Immunol*, 178(10):6043–6050, 2007.

[35] G Weisbuch and H Atlan. Control of the immune response. *Journal of Physics A: Mathematical and General*, 21(3):L189–L192, 1988.

[36] G. Weisbuch, R. J. De Boer, and A. S. Perelson. Localized memories in idiotypic networks. *J Theor Biol*, 146(4):483–499, 1990.

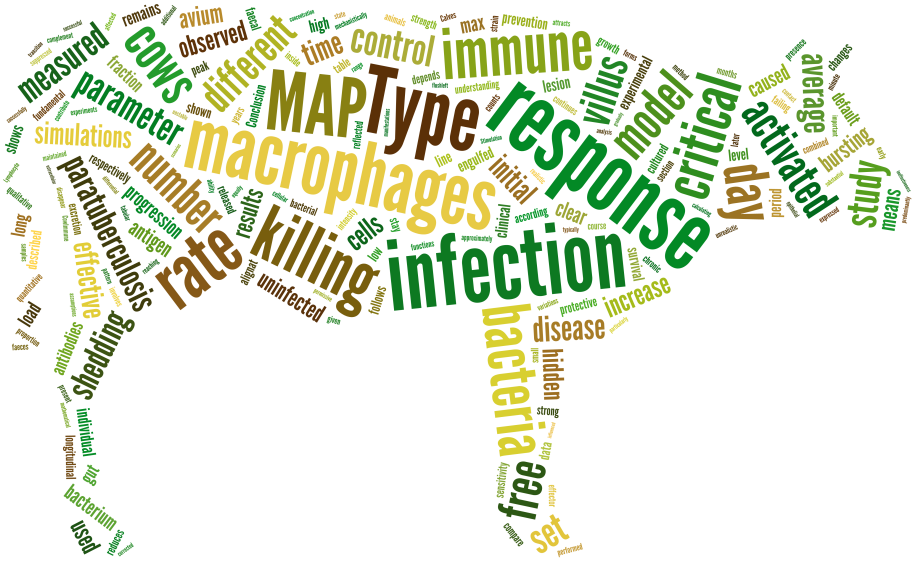
[37] A. Yates, C. Bergmann, J. L. Van Hemmen, J. Stark, and R. Callard. Cytokine-modulated regulation of helper T cell populations. *J. Theor. Biol.*, 206:539–560, 2000.

[38] Y. C. Zang, J. Hong, M. V. Tejada-Simon, S. Li, V. M. Rivera, J. M. Killian, and J. Z. Zhang. Th2 immune regulation induced by T cell vaccination in patients with multiple sclerosis. *Eur. J. Immunol.*, 30:908–913, 2000.

CHAPTER 3

The role of Type 1 and Type 2 immune responses in cattle against infection with *Mycobacterium avium* subspecies *paratuberculosis*.

M. Severins, A. Koets, Merel Langelaar, J.A.P Heesterbeek and D. Klinkenberg



Manuscript in preparation

Abstract

Paratuberculosis in cows is a chronic disease caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP) with substantial economic and welfare effects. The course of infection and progression into disease differs greatly between cows although typically the cows initiate predominantly Type 1 responses against infection which later change into predominantly Type 2 responses. Longitudinal studies suggest that Type 1 immune responses against MAP are protective while Type 2 response are permissive. However, direct evidence and a thorough understanding of why one type of response would be able to control disease while the other fails, is lacking. In this study we aim to find if there is a fundamental mechanistic reason why a Type 1 response would be protective and a Type 2 response would not. For this aim we construct a basic mathematical model that describes known interactions between MAP bacteria, macrophages and the two mechanistically different Type 1 and Type 2 adaptive immune responses. The model allows us to study the differential effect (both qualitative and quantitative) of the two responses on disease progression. In addition, we input the immune response measurements from a 4-year longitudinal study into the model and compare the results of the model with faecal shedding measurements from the same longitudinal study. From the model we find no mechanical or biological impediment preventing a Type 1 or a Type 2 adaptive response from controlling and clearing the infection. By comparing model results to the faecal excretion data we find that the Type 1 response in the cows can stop the initial growth of infection even if it does not have the critical killing rate needed to control the infection. Our results suggest that the killing rate of the Type 2 response in these cows is not strong enough to control infection. We find that in order to control infection, both a Type 1 or a Type 2 response needs to keep up a critical bacterial killing rate for a very long time. For both responses individually it seems unrealistic that they would be able to clear infection within the lifetime of the cow. The model shows that the two responses complement each other very well and when both are active, only a fraction of their individual critical killing rate is needed to keep infection under control. These findings contribute to our understanding of the immune responses during paratuberculosis which is essential for the successful development of treatment and prevention measures, such as vaccines.

1 Introduction

Paratuberculosis (Johne's disease) is a chronic infection of the small intestine of ruminants with substantial economic and welfare effects on livestock worldwide, particularly in cattle. Infection is caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP). MAP typically initiates a subclinical period of 2 to 5 years after which a proportion of the infected animals develops an effectively incurable progressive form of protein loosing enteropathy, clinically recognized by chronic diarrhea and emaciation, ultimately leading to death [25].

Calves acquire the infection in the first months of life through uptake of MAP contaminated milk or faeces from their environment. Once inside the gut, to cross the epithelial barrier, MAP exploits M-cells lining Peyer's patches (lymphoid tissue of the gut) that transport organisms and particles from the gut lumen to immune cells in the lamina propria that lies behind the epithelial barrier [16]. Susceptibility to infection is reduced with age in concordance with the regression of the ileal Peyer's patches [4].

The first line of defence against MAP infection is formed by the macrophages in the Peyer's patches. Macrophages phagocytise MAP for the purpose of destroying the foreign invader. Similar to tuberculosis in humans, for reasons that are not fully understood, in some macrophages MAP interferes with phagosome maturation, hereby preventing bacterial destruction. The interference with macrophage functions also hampers proper antigen presentation and the development of effective immune responses [3]. These immune evasive mechanisms help MAP to grow while hiding from the host's immune response within the macrophages it infects.

MAP continues to multiply slowly until it eventually kills the macrophage, causing it to burst so that the released bacteria spread and infect other nearby macrophages. The release of free bacteria after bursting attracts more macrophages to the site and leads to the formation of diffuse and extensive lesions with predominantly macrophage infiltrate with high intracellular bacterial burden (lepromatous lesions) [14]. These lesions usually start forming within one villus and grow extensively due to the accumulation of macrophages and immune cells causing a substantial thickening of the gut.

MAP is shed from the lamina propria into the lumen of the gut and excreted with faeces, an important route of transmission. The mechanism for this shedding is unknown since epithelial erosion is not a feature of paratuberculosis. Shedding may occur quite early in the course of the disease followed by a 'silent' phase. After the silent phase, the length of which varies greatly between cows, cows progress into a phase of real or apparent intermittent faecal shedding but as the disease develops, shedding in faeces becomes continuous [23].

It is unclear what causes the onset of clinical disease and the shedding pattern

of MAP in the cow's faeces. One potential cause involves the next line of defence against MAP; the cow's adaptive immune response. The two classic types of adaptive immune responses are Type 1 and Type 2 responses. Type 1 and Type 2 responses are fundamentally different in that Type 1 responses are aimed at killing infected cells in this case to kill MAP-infected macrophages and Type 2 responses aim to kill extra-cellular pathogens in this case free MAP bacteria. A variety of studies measuring adaptive immune responses in MAP infected cows indicate an early Type 1 that switches to a non-protective Type 2 response during the course of the disease. Collectively, the data from these studies have been used to argue that chronic progressive forms of paratuberculosis involve a switch in the host immune response from a protective Type 1 immune response to a permissive Type 2 immune response [19, 12, 18], according to the murine Th1-Th2 paradigm [17]. Most of the data leading to these conclusions have been generated using cross-sectional studies. Whether the observed changes in type of immune response are linked to disease exacerbation is largely unknown, especially for later stages of infection [22]. More recent studies indicate that both Type 1 and Type 2 responses may be impaired during later stages of disease [13, 14]. Nonetheless, as MAP infection is intra-cellular, it is thought that, in analogy with human tuberculosis, Type 1 immune responses are protective and inhibiting MAP infection while Type 2 immune responses are ineffective and permissive.

A better understanding of the effectiveness of the immune responses is crucial for the successful development of treatment and prevention measures, such as vaccines. Unfortunately, direct evidence and a thorough understanding of why one type of response would be able to control disease while the other fails, is lacking.

In this study we aim to understand if there is a fundamental mechanistic reason why a Type 1 response would be protective and a Type 2 response would not, and to determine if transition into clinical disease can be caused by a change in type of response.

For this aim we construct a basic mathematical model that describes known interactions between MAP bacteria and macrophages in a villus of the gut. To this model we add the two mechanistically different Type 1 and Type 2 adaptive immune responses which allows us to study in detail the differential effect (both qualitative and quantitative) of these two mechanisms on disease progression. Because the model includes several of the immune evasion mechanisms of MAP it also allows us to study the sensitivity of both types of immune responses to changes in MAP virulence.

To study the role of the two responses in the transition to clinical disease we use data from a 4 year longitudinal study in which young calves were infected and markers for Type 1 and Type 2 responses were measured along with measurements of MAP shedding in their faeces. The adaptive immune responses measured in these animals are used as input to our model, after which we compare the disease

progression and infection load from model simulations with the faecal shedding measurements.

2 Methods: Mathematical Model & Clinical Experiments

2.1 Clinical experiments

Animals: Twenty Holstein-Friesian calves were purchased at birth from different commercial farms and housed at the SPF facilities of the Central veterinary Institute (CVI) in Lelystad, The Netherlands, throughout the experimental period. Animals were kept on a regular feed regimen according to their age and lactation status, but never received fresh grass. Experimental procedures were approved by the Ethical Committee of the CVI. Calves where followed over an experimental period of 55 months (31-01-1999 until 27-08-2003). During the course of the investigation period 7 out of the initial 20 cattle survived to the end. Post-mortem examinations of animals that died during the study showed a diverse number of causes of death, none of which were due to the experimental infection with MAP.

Experimental infection: Calves were infected by way of 20 grams of MAP contaminated feces given orally, three times a week for a period of four weeks during the first month of life. The inoculum was obtained from a cow with clinical signs of MAP infection consistently shedding IS900 positive MAP.

Fecal shedding measurements
Rectal samples for fecal culture were taken every two weeks. Bacteria were cultured according to a modified method of Joergenson [10]. Growth of MAP was mycobactin dependent and checked every 4 weeks. If no growth was observed after 6 months of culture, the sample was considered negative. The presence of MAP in positive cultures was confirmed by amplification of the MAP specific IS900 by PCR [21]. Data was expressed semi-quantitatively. 0=negative, 1=1-10 cfu/slant, 2=10-100 cfu/slant, 3=>100 cfu/slant.

Blood sampling: Blood was collected every two weeks from the jugular vein into heparinised tubes and into serum tubes (BD Vacutainer, Becton, Dickinson, Europe). Heparinised blood was used for the isolation of peripheral blood mononuclear cells (PBMC). Serum was stored at -20C and processed at a later time-point.

Antigen: PPD-P (Johnine) antigen was used in the Lymphocyte Stimulation Test and the ELISA. The PPD-P was produced at CVI, Lelystad, as previously described, from MAP strain 3+5/C (PPD-P) [5].

Type 1 response measurements: Peripheral blood mononuclear cells (PBMC) were isolated and cultured according to methods described in detail elsewhere [12]. Lymphocyte Stimulation Tests (LST) were performed according to methods described in detail previously [12]. In short, cells were cultured in 96-well microtitre plates using 100 μl of the PBMC suspension and 100 μl of antigen per well in triplicates. The PPD-P antigen was used in predetermined optimal concentrations of 10 g/ml. Concanavalin A (ConA) was used as a positive control (2.5 μg/ml) and medium alone as a negative control. Cells were cultured at 37 °C and 5 % CO2 in a humidified incubator for 3 days. Then 0.4 μCi 3H thymidine (Amersham International) was added to each well and cells were cultured for an additional 18 hrs. Subsequently, cells were harvested onto glass fibre filters. Incorporation of 3H thymidine was measured by liquid scintillation counting, and expressed as counts per minute (cpm). Cpm is used as a measure for the intensity of the Type 1 response.

Type 2 response measurements Antibodies (total IgG) specific for PPD-P were detected by ELISA, using the above described antigen, according to the method described earlier [13]. All sera were 10 x diluted in blocking buffer. Results are expressed as background corrected mean optical densities, measured at 405 nm wavelength (OD405nm). OD405nm is used as a measure for the intensity of the Type 2 immune response.

2.2 The Model

We devised a model to describe the dynamics of uninfected macrophages M_u, infected macrophages M_i, and MAP bacteria P per initially infected villus of the gut:

$$\frac{dM_u}{dt} = i - dM_u + aP - cPM_u \tag{1a}$$

$$\frac{dM_i}{dt} = fcPM_u - bM_i \tag{1b}$$

$$\frac{dP}{dt} = bAM_i - cmPM_u - eP \tag{1c}$$

When there is no infection, i uninfected macrophages (M_u) enter the villus per day and leave again at a rate of d per macrophage per day. In this disease-free state the number of uninfected macrophages M_u is at an equilibrium of i/d, where $1/d$ is the average time each uninfected macrophage spends in the villus.

When there are free MAP bacteria (P) present in the villus, each bacterium attracts an additional number of a uninfected macrophages to the site per day. Uninfected macrophages phagocytise free MAP bacteria at a contact rate c per free MAP bacterium per day. Macrophages engulf on average m free bacteria per day.

MAP inhibition of the bacterial killing mechanisms of macrophages [20] is successful in a fraction f of the macrophages that have engulfed bacteria and this fraction become infected [7, 6]. These infected macrophages (M_i) are effectively paralysed by MAP to stay in the gut and will accumulate in the villus, contributing to the development of the lesion [2]. The remaining fraction $(1 - f)$ of macrophages that have engulfed MAP disappear from the villus either because they successfully digest MAP and go to the lymph node to present MAP antigens, or because they go into apoptosis and are killed together with the MAP bacteria inside them.

Infected macrophages disappear from the lesion when they burst, which happens at a rate of b per day, where $1/b$ is the average time an infected macrophage survives. The bursting of infected macrophages contributes to the number of free MAP bacteria in the villus with on average A MAP bacteria being released from each bursting macrophage. Free MAP bacteria disappear from the villus at a rate of e per day which is a combination of complement activity, passive efflux to the lymph with the natural flow of the interstitial fluid and efflux into the lumen of the gut.

To this model we then added the two fundamentally different types of adaptive immune responses:

$$\frac{dM_u}{dt} = i - dM_u + aP - cPM_u \tag{2a}$$

$$\frac{dM_i}{dt} = (1-h)fcPM_u - bM_i - T_1M_i \tag{2b}$$

$$\frac{dM_h}{dt} = hfcPM_u - bM_h \tag{2c}$$

$$\frac{dP}{dt} = bA(M_i + M_h) - cmPM_u - eP - T_2P \tag{2d}$$

The Type 1 response reduces the number of infected macrophages at a rate T_1 per day while the Type 2 response reduces the number of free MAP bacteria at rate T_2 per day. This means that $1/T$ is the average time the antigen (either M_i or P) survives in the lesion before it is removed by one of the two responses.
When MAP bacteria infect a macrophage they try to suppress the extracellular signals the cell uses to flag that it is infected and prompt the Type 1 response to kill the infected macrophage. To model this we introduce a subset of macrophages in which MAP has successfully suppressed these flagging signals and remains hidden to the Type 1 response (M_h). A fraction, h, of the macrophages that have engulfed MAP bacteria will be hidden and not affected by the Type 1 response, while the rest of the infected macrophages $(1 - h)$ do not escape the Type 1 response.

i	$\frac{M_u}{day}$	750	Influx of uninfected macrophages per day.
d	$\frac{1}{day}$	5	Migration rate per day.
a	$\frac{M_u}{day*P}$	0.5	Number of uninfected macrophages attracted to the lesion per day per free bacterium.
c	$\frac{1}{day*P}$	0.01	Contact of bacteria with macrophages, rate per day per free bacterium.
f	n.a.	0.3	Fraction of engulfing macrophages that becomes infected.
h	n.a.	0.1	Fraction of infected macrophages that remains hidden from immune response.
b	$\frac{1}{day}$	0.01	Bursting rate of infected macrophages per day.
A	$\frac{P}{M_i}$	1000	Average number of bacteria released from a bursting macrophage.
m	$\frac{P}{M_u}$	20	Number of bacteria on average engulfed per uninfected macrophage.
e	$\frac{1}{day}$	30	Efflux rate per day of free bacteria from lesion.
T_1	$\frac{1}{day}$	0-0.286	Killing rate per day of Type 1 response (variations from 0-110% of critical rate).
T_2	$\frac{1}{day}$	0-429	Killing rate per day of Type 2 response (variations from 0-110% of critical rate).
r_1	n.a.	0.005	Decay rate per day of the simulated Type 1 response.
r_2	n.a.	0.00083	Incremental rate per day of the simulated Type 2 response.

Table 1: Parameters, their units, their default values and description.

2.3 Default model parameter values

To set realistic parameter values for model simulations, where possible, we use counts and observations from in-house pathological microscopic slides and data from the literature. A list of the parameters, their units, their default values and a description can be found in Table 1. Parameter d, the daily migration rate of uninfected macrophages is set to 5 such that the average time a macrophage spends in a villus is around 5 hours ($1/d$ day). If we take the average length of a villus to be approximately 400 micron and that a macrophages travels 800 micron to get in and out of a villus, then the average speed of a macrophages needs to be 2.8 micron per minute for the macrophages to pass though the villus in 5 hours. This is well within the range of straight line leucocyte velocity as measured in [11]. From macrophage counts in coupes of uninfected cows villi we estimated that the average number of uninfected macrophages per villus is approximately 150. As we set the migration rate of macrophages to 5 per day, the daily influx of macrophages into a villus, i, is set to 750 which together maintain a constant 150 macrophages per villus in uninfected cows. When cows are infected, inflammation and cytokines cause more macrophages to be attracted to the site of infection. This is reflected in parameter, a, which we set at 0.5 extra macrophages per day per bacterium. This value ensures that within realistic infection loads the added volume to the villus and lamina propria due to increased cellular influx does not exceed the approximately 20-fold maximum increase in gut thickness observed.

We have set the contact rate of macrophages and bacteria to 0.01 per day per bacterium. The number of bacteria that on average is taken up by one uninfected macrophage in one day, m, is set to 20 as observed in [6]. This means that when there is a concentration of 100 free MAP bacteria per villus it takes one uninfected macrophage, one day to engulf 20 bacteria. Several studies have been performed that quantify the uptake and infection of MAP bacteria by macrophages [24, 6]. Based on these studies, we have set parameter, f, the fraction of macrophages that becomes infected upon engulfing MAP within the observed range to 0.3. The fraction of infected macrophages that remains hidden for the Type 1 immune system, h, is set to 0.1.
The bursting rate of infected macrophages, b, is set to 0.01 as the average lifetime of a macrophage ($1/b$) is approximately 100 days. The average number of free MAP bacteria released from an infected macrophage, A, is set to 1000. This value comes from intra-macrophage bacterial counts in clinical histopathological samples from cows with lepromatous type pluribacillary type lesions (unpublished observations). The rate at which free MAP bacteria disappear, mostly due to innate immune activity, but also leakage to the gut lumen and lymph is set to 30 per day. This means that on average a free MAP bacterium does not 'survive' more than 48 minutes in the interstitial fluid of the villus.

2.4 Model analyses

Quantitative differences between responses

The system of equations in our model only has one equilibrium, namely the infection-free state. Without the presence of an adaptive immune response this equilibrium is unstable, which means that any initial number of free MAP bacteria or infected macrophage will lead to expansion of infection. The unstable infection-free equilibrium can become a stable equilibrium in the presence of adaptive immunity when the killing rate of the response is large enough. We have calculated the minimum killing rate needed for each response on its own to make the infection-free equilibrium a stable one, by carrying out a linear stability analysis of the infection-free equilibrium (see, for example [9])

Differential effect of bacterial survival mechanisms on Type 1 and Type 2 response

MAP bacteria have evolved several mechanisms to survive and evade their hosts' immune responses. In our model we have several parameters that represent the capacity of MAP bacteria to use those mechanisms. Some strains of MAP are better at exploiting their hosts and evading host immunity than others and changes in the values of these parameters will reflect that. We study if there is a difference in how well the two adaptive immune responses cope with changes in these survival mechanisms of MAP by performing a sensitivity analysis for those parameters. The analysis consists of increasing a parameter by 10% and calculating the new T1 and T2 critical killing rates. The increase in critical killing rates needed for each response are then compared.

In total we analyse three parameters that represent key MAP survival mechanisms: i) How well an infecting MAP bacterium can grow and reproduce inside a macrophage is reflected by parameter A that sets the average number of free MAP bacteria that are released from a bursting macrophage. ii) The ability of MAP bacteria to infect macrophages rather than being killed by them is reflected in the value of parameter f, which is the proportion of macrophages that become infected upon engulfing bacteria as opposed to the macrophage killing the engulfed bacterium. iii) MAP bacteria have the ability to suppress the extracellular signals macrophages use to flag to the immune system that they are infected. This ability allows MAP bacteria to hide inside macrophages undetected. The fraction of infected macrophages in which MAP bacteria succeed in to hide from the host's immunity is given by parameter h.

Combined Type 1 and Type 2 response

In reality, although one of the two responses tends to be dominant over the other, we find that both responses become active during infection [1], particularly when it concerns a long-term or chronic infection. To study how the two responses work

in each other's presence we calculate the critical killing rate of each response while the other response is present. We vary the rate at which the other response is present between 0 and its individual critical killing rate. The resulting critical killing rates for T1 and T2 together are denoted as the T1T2 joint critical killing rates.

2.5 Model Simulations

Model simulations are done using the default parameter settings unless denoted otherwise.

Qualitative differences between responses

To assess the qualitative differences that Type 1 and Type 2 responses have on infection progression, simulations were performed that start with one infected macrophage at day zero. After running the model for a 100 days without any adaptive immunity (T1=0 and T2=0) one of the two types of adaptive immunity is set to 110% of its critical killing rate for the rest of the simulation.

Cow immune response simulations

To assess the effect of more realistic Type 1 and Type 2 responses on infection progression we have constructed two functions (see Figure 1) that, respectively, mimic the shape of the average Type 1 and Type 2 immune responses, as observed in our experiments (see section 3.1, Figure 3a & 3b):

$$T_1 = e^{(-r_1 t)}\, T_1^{max} \tag{3a}$$

$$T_2 = r_2 t\, T_2^{max} \tag{3b}$$

The parameters T_1^{max} and T_2^{max} set the maximum killing rate at the highest point of the Type 1, and Type 2 response, respectively. Four variations of combined responses are tested, one with both T_1^{max} and T_2^{max} at 110% of their critical killing rates, one with both at 50%, and two with either one at 110% and the other at 50%. As it is unclear how the infection load is related to MAP excretion, we look at all manifestations of infection (M_i, M_h and P) to compare with the shedding data from the cows in our longitudinal study.

3 Results & Discussion

3.1 Clinical Data: Shedding and immune results

The average MAP in faecal shedding (Figure 2) shows an initial peak in infection that starts in the second month post infection and lasts until approximately 6

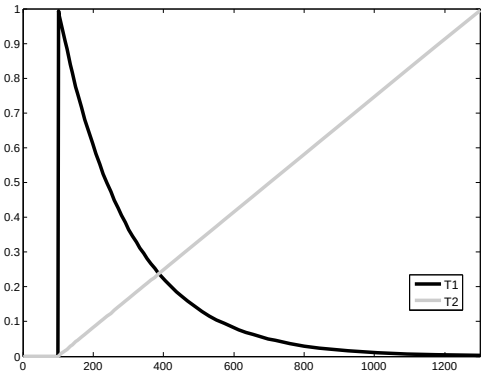


Figure 1: Killing rate of T1 and T2, normalised between zero and one, during the course of the model simulations.

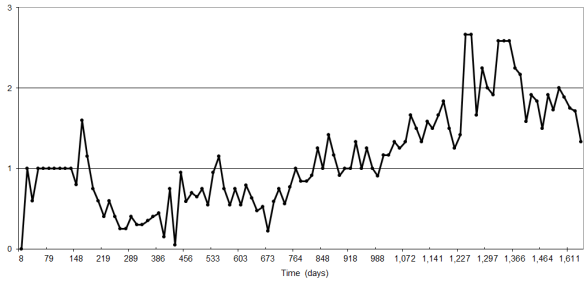


Figure 2: Average ordinal score of MAP bacteria in faecal shedding of a group of 20 cows over a 4 year period.

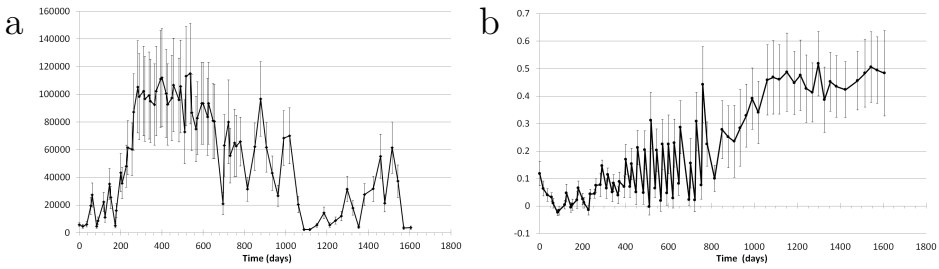


Figure 3: a) Average (+/- SEM) intensity of Type 1 response against MAP antigens in a group of 20 cows over a period of 4 years, measured as background corrected 3H incorporation expressed as counts per minute (CPM). b) Average (+/- SEM) intensity of Type 2 response against MAP antigens in a group of 20 cows over a period of 4 years, measured as backround corrected OD405nm.

months post infection. After the initial peak, cows enter a period of low shedding. With time, shedding gradually increases again reaching high levels which ultimately lead to the onset of clinical disease. During our simulated infections we will refer to these phases as: 'the initial peak', 'the controlled shedding phase' and 'the high shedding phase'.

The average intensity of the Type 1 and Type 2 responses that were measured in the MAP infected cows are shown in Figure 3a and 3b, respectively. Both responses have large standard errors at almost all measure points reflecting the large variation in response intensity between individual cows. In general we can see that both responses do not rise above normal background activity until 200 days post infection. At this point a Type 1 response grows fast and is maintained at a high level for a period of time until it gradually fades down. The Type 2 response on the other hand, from 200 days post infection onwards, very slowly increases and continues to do so for the duration of the experiment.

3.2 Model: Type 1 Response versus Type 2 Response, the fundamental differences.

Quantitative differences between responses

The system of equations in our model only has one equilibrium, namely the disease-free state. Without the presence of an adaptive immune response this equilibrium is unstable and any number of free MAP bacteria or infected macrophage will lead to expansion of infection. The unstable disease-free equilibrium becomes a stable equilibrium in the presence of adaptive immunity when the killing rate of the response is large enough. Hence the rate at which the adaptive immune effectors (cells, antibodies or cytokines) can clear antigen (either infected macrophages or free bacteria) will determine whether once infection has kicked off the response

will be able to bring the system back to the disease-free state. The minimal killing rate or critical value for both immune responses to be effective is given by:

$$T_1^{critical} = \frac{fcibA - b(cim + ed)}{cim + ed - hfciA} \tag{4a}$$

$$T_2^{critical} = \frac{fcibA}{d} - \frac{cim}{d} + e \tag{4b}$$

With our default parameter settings the critical value for T_1 is 0.26. This means that the response should be strong and effective enough to, on average, kill each non-hidden infected macrophages within 3.8 days ($\frac{1}{T_1^{critical}}$) of its infection. The critical value for T_2 is 390. As the Type 2 response involves antibodies which do not kill antigen themselves the interpretation of this value is as follows. Within, on average, 9 minutes ($\frac{1}{T_2^{critical}}$) after an infected macrophage bursts, enough antibodies should bind to each free MAP bacterium to ensure that the bacterium can no longer successfully infect a macrophage.

The critical killing rates for T1 and T2 show that the rate of antigen binding for a Type 2 response should be at least 1500 times as fast, or as much, as the rate for a Type 1 response to be equally effective. If we assume that the effector units of both responses are equally effective and equally capable of reaching their antigen, this would mean that to be equally strong a Type 2 response needs to comprise of 1500 times more effector units than a Type 1 response. Given that a Type 2 response involves the production of large numbers of antibodies a 1500-fold difference cannot be discarded as unrealistic, however, the more antibodies that need to bind to a bacterium to neutralise it, the larger this difference becomes. On the other hand, antibodies are much smaller than the cells active in a Type 1 cellular response and will find it easier to reach their antigen than Type 1 effector cells.

Conclusion I: Under the condition that the responses work properly, we find no mechanical or biological impediment, which would prevent a Type 1 or a Type 2 adaptive response from controlling and clearing the infection.

Qualitative differences between responses

Figure 4 shows that MAP infection grows exponentially until one of the two responses is activated. When T1 is activated (black line, Figure 4) above its critical killing rate the growth of all manifestations of infection is prevented (M_i, M_h and P). The rate of infection clearance by T1 is very different for the different manifestations of infection. After activation, the regular infected macrophages (M_i) are cleared very quickly and reduced to very low numbers. The faster turn-over in P means that the number of free bacteria depends on the number of bursting infected macrophages. Therefore, although not directly cleared by T1, P follows this drop in M_i. Hidden infected macrophages are not cleared by T1. The rate

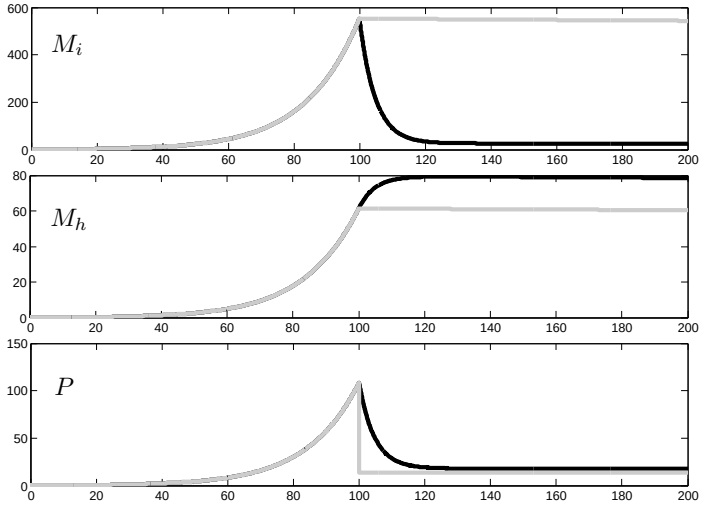


Figure 4: Default parameter settings, T1 (black line) and T2 (grey line) on at 110% of their critical killing rate from t=100 days onwards. x-axes show time (t) in days.

at which they disappear is mostly set by their bursting rate, which is even slowed down by the low level of bacteria maintained by this bursting that still manage to infect some macrophages.

Our model shows that in a short time the Type 1 response causes a large drop in the overall infection load. After this drop a low level of 'hidden' infection remains and therefore the Type 1 response needs to stay activated above its critical killing rate for a very long time to continue to prevent an exponential increase of the infection. How long the Type 1 response needs to stay activated depends on the number of M_h at the time of the Type 1 response activation. M_h at the time of activation will depend on parameter, h, which is influenced by both MAP strain and cow characteristics.

When a Type 2 response is activated (grey line, Figure 2) we also see that it prevents the growth of all manifestations of infection (M_i, M_h and P). The Type 2 response directly kills free MAP bacteria and it is therefore unsurprising that immediately after activation P plummets to very low numbers and stays low as long as the Type 2 response remains active. It is this very low number of free MAP bacteria that stops the infected macrophages from increasing because the Type 2 response does not kill any infected cells itself. After Type 2 response activation the number of infected macrophages, both regular and hidden, only slowly declines at the rate at which the cells burst. Compared to the Type 1 response, the Type 2 response is not as effective at reducing the overall infection load.

	T1 critical killing rate	T2 critical killing rate
default	0.26 (100%)	390 (100%)
A (1100)	0.414 (159%)	435 (112%)
f (0.33)	0.414 (159%)	435 (112%)
h (0.11)	0.375 (143%)	390 (100%)

Table 2: Critical killing rates of T1 and T2 under default parameter values and when one of the parameter values is increased by 10% if its default value.

Conclusion II: A Type 1 response quickly reduces the overall infection load by clearing all but the 'hidden' infected macrophages and some free MAP bacteria associated with the bursting of these. A Type 2 response quickly reduces free MAP bacteria to very low levels; the reduction of infected macrophages is slow. Both responses take a very long time to completely clear infection.

Difference in sensitivity between responses

Next, we have performed a sensitivity analysis to study if there is a difference in how well the two adaptive immune responses cope with changes in survival mechanisms of MAP. The results of the analyses are shown in Table 2. Table 2 shows that when either parameter A or f is increased by 10%, the critical killing rates of the Type 1 and Type 2 immune responses respectively become 0.414 (159% of default) and 435 (112% of default). This implies that if a MAP strain either improves its growth rate or improves its ability to infect macrophages, the Type 1 response needs to increase the number of its effector cells (or the affectivity of their killing) by more than 50%, while the Type 2 response only needs to increase its efforts by 12%.

Increasing the ability of MAP bacteria to hide inside macrophages by 10% has no impact at all on the critical killing rate of the Type 2 response because both hidden and non-hidden infected macrophages do not differ in the number of free MAP bacteria they release upon bursting. In contrast the Type 1 response needs to increase the rate at which it kills non-hidden infected macrophages by 43% to compensate for the increase in hidden MAP bacteria.

Conclusion III: In general a Type 1 response is more vulnerable and more likely to fail in cows that become infected with more virulent MAP strains.

Combined Type 1 response and Type 2 response:

To study how the Type 1 and Type 2 responses complement each other we plot the range of critical killing rates of T1 and T2 at which their combined effect are able to bring the system back to the infection-free state (Figure 5).

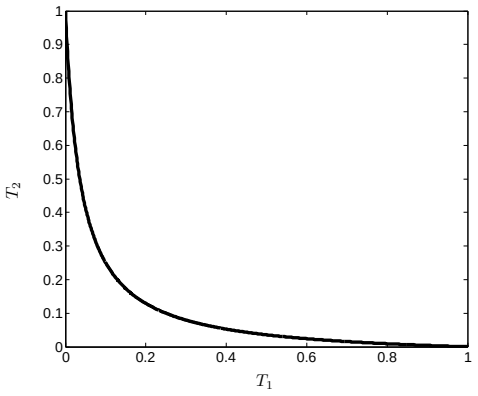


Figure 5: The range of T1T2-joint critical killing rates given with respect of their individual critical killing rates.

The range of T1T2-joint critical killing rate with respect to their individual critical killing rates is show in Figure 5. It shows that the two types of responses enhance each other's effectivity, i.e., when both responses are at less than half of their individual critical killing rate, they can still clear infection when active together.

Conclusion IV: The killing rate of the Type 1 and Type 2 adaptive immune responses combined is stronger than the sum of their individual killing rates.

3.3 Model: Simulation infection progression

Next, we used the model to see if the experimentally observed immune responses can explain the observed MAP excretion pattern and to see if the results can indicate if one or both of the responses fail to control infection. Figures 6 and 7 show the resulting infection progression for four variations of the maximal killing rates by the two immune responses.

When adaptive immunity is activated, all simulations, also those where the maximum killing rate of the Type 1 response is only at 50% of its critical value, show a quick drop in the infection load. This means that the initial decline in shedding that is observed in all our experimental cows does not necessarily mean that the initial Type 1 response is above the threshold. In other words, when we observe that during a Type 1 response shedding declines, it does not mean that this response is capable of controlling and eliminating the infection, even if continued at the same strength for the duration of infection. As a result, from these simulations we cannot tell if the cows in our longitudinal study are capable of building a Type 1 immune response that is strong enough to control and eliminate infection.

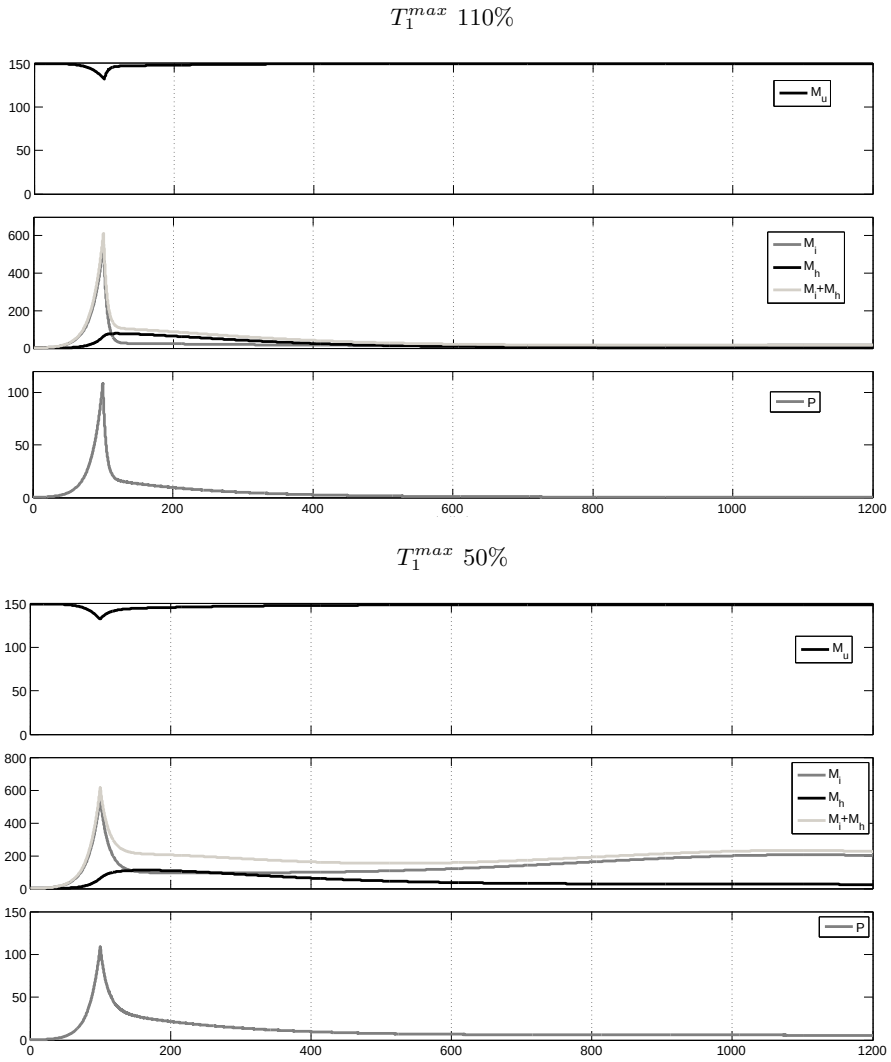


Figure 6: Cow immune response simulations for T_2^{max} set to 110% of its critical killing rate. Infection is initiated at day 0 with one infected macrophage. Infection is followed for 1200 days by means of tracking M_u, M_i, M_h and P. Adaptive immunity is activated at day 100 and follows functions 3a and 3b. Two different scenarios for the absolute strength of the Type 1 response are shown. Simulations of cows where the T_1^{max} response is set to 110% and 50% of its critical killing rate are shown in the top sub-figure and bottom sub-figure, respectively. The x-axes plots time in days, the y-axes shows numbers per villus.

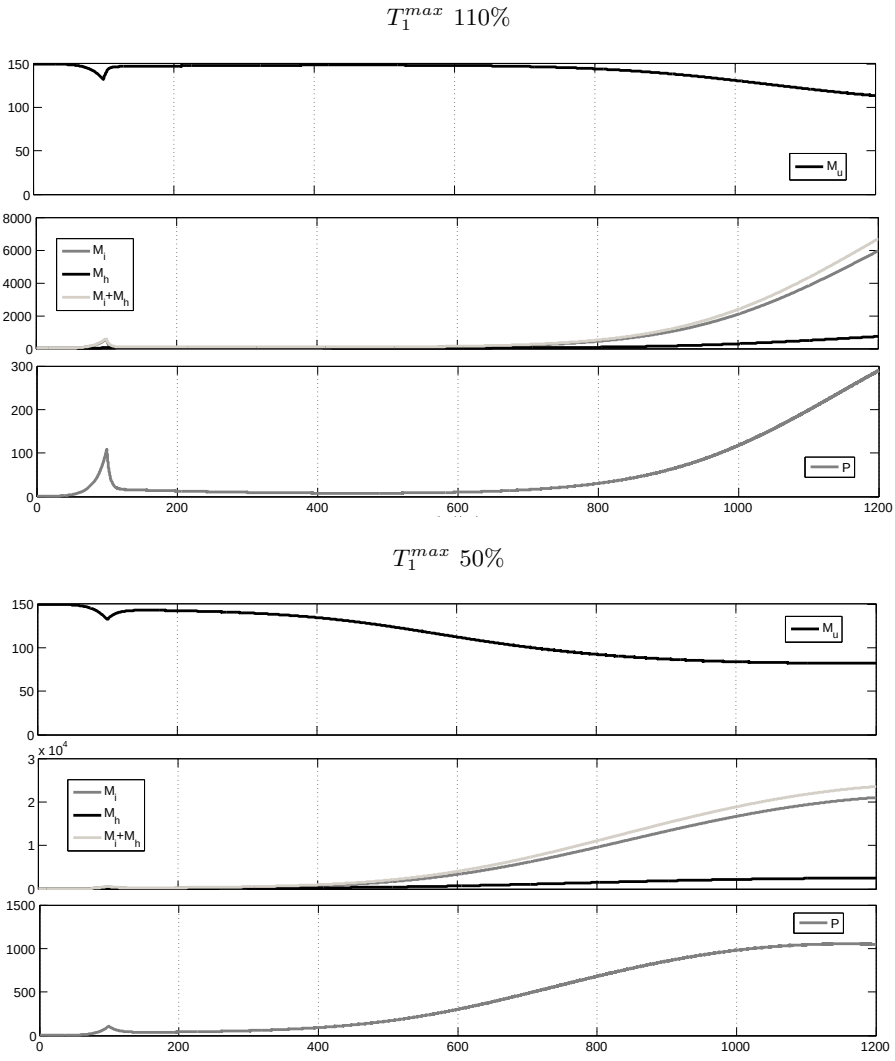


Figure 7: Cow immune response simulations for T_2^{max} set to 50% of its critical killing rate. Infection is initiated at day 0 with one infected macrophage. Infection is followed for 1200 days by means of tracking M_u, M_i, M_h and P. Adaptive immunity is activated at day 100 and follows functions 3a and 3b. Two different scenarios for the absolute strength of the Type 1 response are shown. Simulations of cows where the T_1^{max} response is set to 110% and 50% of its critical killing rate are shown in the top sub-figure and bottom sub-figure, respectively. The x-axes plots time in days, the y-axes shows numbers per villus. NB. The scales on the y-axes differ between the top and bottom panels.

Figure 7 shows that only when the maximum killing rate of the Type 2 response is below critical value, infection progresses from the controlled infection phase into a high shedding phase as we observe in our cows. This means that the Type 2 response in the cows of our longitudinal study do not build a Type 2 response that is strong enough to keep infection under control.

Conclusion V: Our results show that the Type 1 response in our cows is able to stop the initial growth of infection even if it does not have the critical killing rate needed to control the infection. Our results suggest that in these cows the Type 2 response is below the critical killing rate.

Factors involved in the controlled infection phase

Right after the initial drop in infection a low amount of infection remains, the level of which mostly depends on the strength of the Type 1 response. If the Type 1 killing rate is above its critical rate, only those infected macrophages that are 'hidden' at the time the response was activated remain (Figure 6, top sub-figure, middle panel, black line). If the Type 1 is below its critical killing rate, in addition to the hidden macrophages, also a proportion of the non-hidden macrophages will survive (Figure 6, bottom sub-figure, middle panel). As long as the Type 1 and Type 2 killing rates stay above their joint critical rate, the number of infected macrophages remaining will slowly decrease, as they gradually burst without being able to cause enough new infections.

The cow immune simulations have shown us that the number of infected macrophages after the initial peak depends on the strength of the Type 1 response and the number of hidden infected macrophages at the time the Type 1 response becomes active. This number in turn depends on i), the proportion of infected macrophages in which bacteria succeed to hide ii), how long infection has been able to expand before the Type 1 response is activated and, as our model reflects the infection burden caused per initially infected villus, iii) the number of initially infected villi.

Conclusion VI: The factors that, according to our study, contribute to the infection load during the controlled infection phase are: The number of initially infected villi (related to the dose of infection), the MAP strain the cow is infected with, how long it takes before the adaptive immune response becomes activated, and the strength of the early Type 1 response.

Dose of infection

In our model we have tried to explain the MAP excretion pattern by immune responses, but [15] have shown that infection dose can also be of influence, though only for the duration of the initial peak. Our model describes MAP dynamics

per villus, therefore the dose would just be reflected in the number of initially infected villi. As a consequence, only the level of excretion would be affected, not the pattern. On the cow level, this is likely to be observed as an effect on the shedding level of the first peak, but not so much of the second peak, because that increases indefinitely irrespective of the dose. A different dose may be observed as a difference in start of the second peak, but that is not likely as this is influenced by many cow-related differences.

The failing Type 2 response

It is important to note a few model assumptions that affect the simulation results at high infection loads during the later stages of disease:

1. Our model simulates the progress in infection load caused per initially infected villus. We do not explicitly take into account that infection spreads from one villus to another one. As long as uninfected macrophages are abundant this does not affect our results. However when the infectious load increases and the number of uninfected macrophages are depleted, the role of infection that has moved to a nearby villus where uninfected macrophages are still abundant becomes important.

2. In our model the immune response is proportional to infection without a limit. It means that when we set the killing rate of a response, this rate is maintained no matter how high the infectious load is. At very high infection levels, it is unrealistic that there are enough immune effector cells to keep up the killing rate.

3. Finally, in our model we have not taken into account the spatial distribution of the infected macrophages. MAP typically forms a spherical infection lesion of infected macrophages that can be penetrated by antibodies of a Type 2 response but not by the lymphocytes of a Type 1 response (observations by A. Koets). This means that with growing numbers of infected macrophages there will be more infected cells that cannot be reached i.e., that are like 'hidden' cells from the T1 response than we take into account in our model.

All three of the above model assumptions make that the rate of infection increase after the transition into the high shedding phase is likely to be faster in reality than shown by our model.

Our results show that when the Type 2 response increases in strength up to its critical killing rate it should be able to clear infection (Figure **??**, left panels). Because this is not what we observed in our experimental study, we conclude that the Type 2 response in the cows must have stayed below the critical killing rate. However, as argued above, the ineffectiveness may be due to the infection already having grown out of control when the T2 response is strongest. The same

response may be effective in an earlier phase of the infection.

Our results also show that there is no fundamental impediment why an (early) Type 2 response should not be able to reach the critical killing rate needed to clear infection. However in calculating this rate we have assumed that the Type 2 response is functioning properly. Several scenarios have been suggested in which this is not the case. For example, if the antibodies of the Type 2 response are built against non-exposed parts of the bacterium, or if the Type 2 response is not stimulated enough to keep the concentration of antibodies high enough because the bacterium is mostly hidden inside macrophages, or, if the increased uptake of antibody-covered bacteria by macrophages (as observed by [7]) aggravates infection, by increasing the contact rate, to such extent that it outweighs the killing rate of the Type 2 response.

3.4 Implications for MAP prevention and control

Ideally, to control MAP infection both the Type 1 and Type 2 response should be strong and remain activated until all infection is cleared. This is however not realistic as particularly a Type 1 response is a very aggressive response the prolonged strong activation of which is likely to cause harm in itself to the animal. Also, it is know that during gestation Type 1 responses are suppressed to protect the unborn calf [8]. Therefore, it would be best if an effective Type 2 response is also activated early during the infection, or at least immediately after the Type 1 response begins to fall, and stays active until infection is fully cleared. It is this Type 2 response that we don't see in our cows, or at least not timely enough. An important next step in the development of effective control measures is to understand why exactly the T2 response in these cows is failing. If a vaccine could be developed resulting in an earlier and stronger Type 2 response the observed progression to disease could be avoided.

In conclusion, for both the Type 1 and Type 2 responses it holds that the response needs to continue above this critical rate for a very long time if infection is to remain under control. For both responses individually it seems unrealistic that they would be able to clear infection within the lifetime of the cow. The two responses complement each other very well and when both are active, even at a fraction of their individual critical killing rate, it is likely that infection can be kept under control more effectively.

References

[1] John P Bannantine, Darrell O Bayles, W. Ray Waters, Mitchell V Palmer, Judith R Stabel, and Michael L Paustian. Early antibody response against mycobacterium avium subspecies paratuberculosis antigens in subclinical cattle. *Proteome Sci*, 6:5, 2008. doi: 10.1186/1477-5956-6-5.

[2] C. Brady, D. O'Grady, F. O'Meara, J. Egan, and H. Bassett. Relationships between clinical signs, pathological changes and tissue distribution of mycobacterium avium subspecies paratuberculosis in 21 cows from herds affected by johne's disease. *Vet Rec*, 162(5):147–152, 2008.

[3] Paul M. Coussens, Sue Sipkovsky, Brooke Murphy, Jon Roussey, and Christopher J. Colvin. Regulatory t cells in cattle and their potential role in bovine paratuberculosis. *Comp Immunol Microbiol Infect Dis*, 2012. doi: 10.1016/j.cimid.2012.01.004.

[4] M. M. Dennis, M. C. Antognoli, F. B. Garry, H. L. Hirst, J. E. Lombard, D. H. Gould, and M. D. Salman. Association of severity of enteric granulomatous inflammation with disseminated mycobacterium avium subspecies paratuberculosis infection and antemortem test results for paratuberculosis in dairy cows. *Vet Microbiol*, 131(1-2):154–163, 2008. doi: 10.1016/j.vetmic.2008.02.017.

[5] N J L Gilmour and G W Wood. *OIE Manual of Standards for Diagnostic Tests and Vaccines*. Office International des Epizooties, Paris, third edition edition, 1996.

[6] Nicole S. Gollnick, Rebecca M. Mitchell, Martin Baumgart, Harish K. Janagama, Srinand Sreevatsan, and Ynte H. Schukken. Survival of mycobacterium avium subsp. paratuberculosis in bovine monocyte-derived macrophages is not affected by host infection status but depends on the infecting bacterial genotype. *Vet Immunol Immunopathol*, 120(3-4):93–105, 2007. doi: 10.1016/j.vetimm.2007.07.017. p MAP multiplication and

[7] J. Hostetter, R. Kagan, and E. Steadham. Opsonization effects on mycobacterium avium subsp. paratuberculosis–macrophage interactions. *Clin Diagn Lab Immunol*, 12(6):793–796, 2005. doi: 10.1128/CDLI.12.6.793-796.2005.

[8] Elisabeth A Innes. The host-parasite relationship in pregnant cattle infected with neospora caninum. *Parasitology*, 134(Pt 13):1903–1910, 2007. doi: 10.1017/S0031182007000194.

[9] D. W. Jordan and P. Smith. *Mathematical Techniques*. Oxford University Press Inc., New York, 4th edition, 2008.

[10] J. B. Jørgensen. An improved medium for culture of mycobacterium paratuberculosis from bovine faeces. *Acta Vet Scand*, 23(3):325–335, 1982.

[11] Alexander Georg Khandoga, Andrej Khandoga, Christoph Andreas Reichel, Peter Bihari, Markus Rehberg, and Fritz Krombach. In vivo imaging and quantitative analysis of leukocyte directional migration and polarization in inflamed tissue. *PLoS One*, 4(3):e4693, 2009. doi: 10.1371/journal.pone.0004693.

[12] A. P. Koets, V. P. Rutten, A. Hoek, D. Bakker, F. van Zijderveld, K. E. Mller, and W. van Eden. Heat-shock protein-specific t-cell responses in various stages of bovine paratuberculosis. *Vet Immunol Immunopathol*, 70(1-2):105–115, 1999.

[13] A. P. Koets, V. P. Rutten, M. de Boer, D. Bakker, P. Valentin-Weigand, and W. van Eden. Differential changes in heat shock protein-, lipoarabinomannan-, and purified protein derivative-specific immunoglobulin g1 and g2 isotype responses during bovine mycobacterium avium subsp. paratuberculosis infection. *Infect Immun*, 69(3):1492–1498, 2001. doi: 10.1128/IAI.69.3.1492-1498.2001.

[14] Ad Koets, Victor Rutten, Aad Hoek, Frans van Mil, Kerstin Mller, Douwe Bakker, Erik Gruys, and Willem van Eden. Progressive bovine paratuberculosis is associated with local loss of cd4(+) t cells, increased frequency of gamma delta t cells, and related changes in t-cell function. *Infect Immun*, 70(7):3856–3864, 2002.

[15] R. M. Mitchell, G. F. Medley, M. T. Collins, and Y. H. Schukken. A meta-analysis of the effect of dose and age at exposure on shedding of mycobacterium avium subspecies paratuberculosis (map) in experimentally infected calves and cows. *Epidemiol Infect*, 140(2):231–246, 2012. doi: 10.1017/S0950268811000689.

[16] E. Momotani, D. L. Whipple, A. B. Thiermann, and N. F. Cheville. Role of m cells and macrophages in the entrance of mycobacterium paratuberculosis into domes of ileal peyer's patches in calves. *Vet Pathol*, 25(2):131–137, 1988.

[17] T. R. Mosmann and S. Sad. The expanding universe of t-cell subsets: Th1, th2 and more. *Immunol Today*, 17(3):138–146, 1996.

[18] J. R. Stabel. Cytokine secretion by peripheral blood mononuclear cells from cows infected with mycobacterium paratuberculosis. *Am J Vet Res*, 61(7): 754–760, 2000.

[19] R. W. Sweeney, D. E. Jones, P. Habecker, and P. Scott. Interferon-gamma and interleukin 4 gene expression in cows infected with mycobacterium paratuberculosis. *Am J Vet Res*, 59(7):842–847, 1998.

[20] M. Z. Tessema, A. P. Koets, V. P. Rutten, and E. Gruys. How does mycobacterium avium subsp. paratuberculosis resist intracellular degradation? *Vet Q*, 23 (4):153–162, 2001.

[21] P. H. Vary, P. R. Andersen, E. Green, J. Hermon-Taylor, and J. J. McFadden. Use of highly specific dna probes and the polymerase chain reaction to detect mycobacterium paratuberculosis in johne's disease. *J Clin Microbiol*, 28(5):933–937, 1990.

[22] W. R. Waters, J. M. Miller, M. V. Palmer, J. R. Stabel, D. E. Jones, K. A. Koistinen, E. M. Steadham, M. J. Hamilton, W. C. Davis, and J. P. Bannantine. Early induction of humoral and cellular immune responses during experimental mycobacterium avium subsp. paratuberculosis infection of calves. *Infect Immun*, 71(9): 5130–5138, 2003.

[23] R. J. Whittington and E. S. Sergeant. Progress towards understanding the spread, detection and control of mycobacterium avium subsp paratuberculosis in animal populations. *Aust Vet J*, 79(4): 267–278, 2001.

[24] Seng Ryong Woo and Charles J Czuprynski. Tactics of mycobacterium avium subsp. paratuberculosis for intracellular survival in mononuclear phagocytes. *J Vet Sci*, 9(1):1–8, 2008.

[25] Chia-wei Wu, Michael Livesey, Shelly K Schmoller, Elizabeth J B Manning, Howard Steinberg, William C Davis, Mary Jo Hamilton, and Adel M Talaat. Invasion and persistence of mycobacterium avium subsp. paratuberculosis during early stages of johne's disease in calves. *Infect Immun*, 75(5):2110–2119, 2007. doi: 10.1128/IAI.01739-06.

CHAPTER 4

Effects of heterogeneity in infection-exposure history and immunity on the dynamics of a protozoan parasite.

M. Severins, D. Klinkenberg, and J.A.P Heesterbeek

Published in:
Journal of the Royal Society Interface (2007) **4**(16): 841-849

Abstract

Infection systems where traits of the host, such as acquired immunity, interact with the infection process can show complex dynamic behaviour with counter-intuitive results. In this study we consider the traits 'immune status' and 'exposure history' and our aim is to assess the influence of acquired individual heterogeneity in these traits. We have built an individual-based model of *Eimeria acervulina* infections, a protozoan parasite with an environmental stage, that causes coccidiosis in chickens. With the model we simulate outbreaks of the disease under varying initial contaminations. Heterogeneity in the traits arises stochastically through differences in the dose and frequency of parasites that individuals pick up from the environment. We find that the relation between the initial contamination and the severity of an outbreak has a non-monotonous 'wave-like' pattern. This pattern can be explained by an increased heterogeneity in the host population caused by the infection process at the most severe outbreaks. We conclude that when dealing with these types of infection systems, models that are used to develop or evaluate control measures can not neglect acquired heterogeneity in the host population traits that interact with the infection process.

1 Introduction

Understanding the dynamics of infectious diseases in populations is important for the development and improvement of intervention methods. Much understanding has been obtained by use of mathematical and simulation models. In the most recent decades such models have increasingly been able to incorporate population structure, interpreted as differences between individuals in traits that influence transmission [6]. Where the infection process is, in some way, influenced by all relevant traits, many of these traits will not in turn be influenced by the infection dynamics (think of age, sex), or at least this is usually implicitly assumed (for example: spatial location might change, but decision to move or movement itself is usually taken independent of infection dynamics).

There are important situations where the interaction between trait and infection dynamics does work both ways. A class of systems where this happens are those systems where the heterogeneity between individuals lies in the somewhat implicit trait 'infection history' and the more explicit trait, 'immune status'. Examples of such systems are most host-parasite systems with protozoan parasites or with worm parasites, where the heterogeneity in infection history and immune status (however defined) exists due to stochastic differences in exposure. The circle of interaction consists of the exposure of an individual host shaping its infection history and immune status (acquired immunity), which in turn will influence the infectiousness of that individual to others. It is this type of *acquired heterogeneity* that is not well understood in the context of pathogen transmission, and notably the effects that the interaction has on the dynamics and on the effectiveness of control measures. In a very caricatural model of an unspecific parasite-host system Roberts and Heesterbeek [21] showed that even including the two-way interaction between immunity and transmission (acquired immunity) alone, can have substantial and counterintuitive effects. In that model all individuals responded the same; there was acquired immunity but not yet acquired heterogeneity.

In this paper we present an initial study of the effects of including acquired heterogeneity in exposure into an epidemiological model. For this we model the host dynamics of a protozoan parasite, including acquired immunity and repeated infection through a spatially structured environment. We have chosen to develop our ideas using the specific parasite-host system of *Eimeria acervulina* in chickens for which a large body of experimental results is available and for which we have gained previous insight combining data and (within and between-) host models [15, 16]. We use an individual-based model to allow for stochastically emerging heterogeneity. The dynamics are modelled with only a limited number of infection and immunity levels, which allows us to easily track the immune status, stochastic exposure history, and pathogen excretion of each individual host through time.

We describe the biological system, give the model description and assumptions in section 2. The simulations that we have performed with the system to gauge

the influence of various sources of heterogeneity are shown in section 3. Notably we study variation in the immune response, in transmission, and in movement of individuals. We discuss our results in section 4.

The system is studied by relating initial levels of oocysts in the environment to the average cumulative excretion of the chickens in the weeks thereafter as a measure of unprotected infection. Experiments have shown a non-monotonous relation between the initial oocyst levels, disease and production loss, with intermediate oocyst levels giving optimal results. We also observe this in our model analysis and we can explain this pattern from the variation in exposure history and immunity between individuals.

2 Biological system model description

2.1 *E. acervulina* life cycle

We model the dynamics of *Eimeria acervulina* infections, one of the milder and most common species of the coccidium *Eimeria*, in a commercial broiler chicken setting. The disease that these *Eimeria* parasites cause is called coccidiosis, and is economically very important because of reduced growth of the animals. We first briefly describe the life cycle. Chickens continuously peck on the stable floor and can thereby pick up oocysts, the infectious stage of the parasite. Within the cells of the gut epithelium, the parasite undergoes a series of replications resulting in the formation of new oocysts that are excreted in the faeces [2]; after excretion, oocysts have to sporulate to become infectious. Excretion of oocysts starts four days after infection and may last for ten days, whereas sporulation takes another two days to complete the cycle [8, 1]. Under favourable circumstances oocysts may survive a long time [20], however in chicken manure they have a half-life of approximately two days [20, 28].

2.2 The model

The model environment

We model the complex chicken-environment interaction with an individual-based simulation model within the modelling environment of Netlogo [26] because it allows for stochastically emerging heterogeneity and studying the effect of space. The model algorithm consists of an initial step followed by a set of rules that together determine one day of the chickens, and change the environment; these rules are repeated until the simulation stops. The set consists of model rules for moving, taking up oocysts, excretion of oocysts, and the development of immunity:

1. **Setup** Initiate all model variables
   (a) Create a regular grid of square patches
   (b) Initiate a percentage of patches with a certain oocyst contamination level, and specify the initial spatial distribution of contaminated patches.

   (c) Create a number of chickens with no uptake history or immunity.
   (d) Randomly place the chickens onto the patches with no more than one chicken per patch allowed.
   (e) Set $day = 1$
2. **Chicken Movement** Let each chicken displace to a random patch (again only one chicken per patch), and let the chickens pick on the patch. Repeat 24 times, this makes up one day of movement and exposure.
3. **Oocyst uptake** Let every chicken add the highest encountered level of contamination during one day to its uptake history.
4. **Oocyst excretion** Adjust the contamination level of the patch on which a chicken ended in step 2, according to the excretion corresponding to the uptake history and immune level of the chicken.
5. **Immunity** Adjust the immune level of each chicken according to its uptake history.
6. If $day = 49$, then stop, otherwise $day = day + 1$ and go to step 2.

Below we will explain all steps in detail and discuss the biological fundaments of these rules and which assumptions and simplifications where made.

Chicken movement

We simulate a 4 x 4 meters section of a broiler shed represented by a regular grid of square patches which we define to be 10 x 10 cm. During the set-up rule chickens are randomly placed onto the patches, with only one chicken per patch allowed. The movement rule lets chickens move 24 times, i.e. once per hour. Chickens move through the shed section by randomly changing position to another patch. Random mixing between the chickens and the patches is supported by observations in broiler sheds [19, 17] where chickens show no sign of territorial behaviour and have an average displacement of 8 m/h which is large compared to our model settings.

Floor contamination

Each patch is characterised by its oocyst contamination level, either empty, low, medium, or high, reflecting oocyst numbers of the order 0, 10^3, 10^5, and 10^7. During the set-up the initial contamination level of choice is implemented as a percentage of contaminated patches. The initially contaminated patches can be randomly distributed over the field or concentrated in a single position in the field. Patch contamination levels increase when a chicken on a patch excretes a higher level of oocysts on that patch than already present. Patch contamination levels decrease if the last excretion on that patch was 14 days ago resulting from the two days half-life of oocysts [20, 28]. We neglect the effect of decreasing oocyst levels due to uptake because it is very small relative to the natural environmental decay of oocysts.

Oocyst uptake

The movement rule lets chicken 'remember' the highest level of contamination on a patch they have been exposed to while moving through the shed during the day. The oocyst uptake rule determines that the dose of oocysts that chickens pick up from the patches is 1% of this highest contamination level that they have been exposed to yielding low, medium, and high ingestion doses, equivalent to 10, 10^3, and 10^5 ingested oocysts per day. This implies that, within one day, lower levels of contamination and repeated encounters with the same contamination are neglected, justified because of the 100-fold difference between categories. Thus after the movement rule, 1% of the highest level of contamination encountered is saved by the uptake rule in the chickens individual uptake record. This record is then later used by the oocyst excretion rule and the immunity rule to respectively determine the oocyst excretion of the chicken and the chicken's immunity level.

Immunity

Various experiments have shown that past infections with *E acervulina* induce protection against disease and oocyst production resulting from later infections [4, 18], but no single immune variable has yet been identified as a measure of protection. As we use immunity in our model to affect the dynamics of the parasite, we only regard immunity in relation to a reduction in oocyst excretion. Three immune levels are distinguished: no, partial, and full immunity. A new level of immunity is reached when a certain dose and/or frequency of oocysts are ingested. In addition to the dose and frequency conditions, it takes a fixed time before a new level is reached reflecting that it takes time to build an immune response. Every day the immune rule verifies if chickens have satisfied the conditions required to reach a new level of immunity and assigns a new immunity level to those chickens that meet the conditions.

Oocyst excretion

Essential for the dynamics is the relation between uptake and excretion of oocysts. This relation is mainly determined by the dose and the frequency with which the oocysts are ingested, the immune status, age and to a lesser extent the breed of the chicken. In our model we take into account dose, frequency and immune status, neglecting breed and age due to limitation of data; most experiments we used to fit the excretion patterns were done with chickens within the age range of the model chickens. We do not explicitly model the within-host dynamics, but instead define excretion templates as a function of the uptake history, derived from experiments in the literature (table 1). We recognise nine templates, one for each combination of the three immune levels and the three doses that can be ingested. The dose of oocysts excreted always relates to the present immunity level and the uptake dose of four days earlier, i.e. the first template is started four days after the first dose. Chickens follow their excretion template in the days thereafter, until it finishes or is interrupted by the uptake of an equal or higher dose, or when a higher immune

Table 1: Templates showing daily excretion dose of naive, partial and full immune birds after ingestion of a low, middle or high dose of oocysts starting from the fourth day after ingestion. In the templates, 2, 3 and 4 respectively represent low, middle and high doses of oocyst excretion.

immune state	ingested dose	template
naive	low	[2 2 2]
	middle	[3 3 3 3 3 3 3 2 2]
	high	[4 4 4 4 4 4 4 3 3 2]
partial	low	[ ]
	middle	[2 3 2 2 2]
	high	[3 3 3 2 2 2]
full	low	[ ]
	middle	[ ]
	high	[0 2]

level is reached. In these cases, the template is re-started or a new template starts.

Single dose experiments were used to translate the ingested and excreted numbers of oocysts to the categories in our model, i.e. low, medium, and high. The excretion patterns thus obtained were roughly averaged (table 2) to become the excretion templates as shown in table 1. Single dose experiments performed at our group by Velkers et al. [23] were also used (data not shown), in particular for the low-dose template.

We explicitly distinguish between excretion after uptake of an isolated single dose and so-called trickle infections where chickens repeatedly take up doses on consecutive days causing accumulation of excreted oocysts. It has been shown experimentally and in mechanistic models that immune reaction and excretion patterns are different for a same total dose of oocysts, depending on uptake as either a single or as a trickle dose [27, 15, 22]. The same averaging of excretion patterns was done for trickle infection experiments (table 3), from which it was concluded that the same templates could be used, rather than defining separate excretion templates for trickle infections. We let excretion induced by a low dose trickle infection follow the medium dose template, and in chickens with no immunity we let a medium dose trickle infection result in excretion according to the high dose template. Because high dose trickle infections have not been described and are not likely to cause a large accumulation of excreted oocysts due to the so-called crowding effect [29], in the model, chickens that undergo such a high dose trickle infection are assigned with the corresponding single dose template. The separate set of excretion templates for partially and fully immune chickens were constructed according to immunization experiments shown in table 4.

Table 2: Experiments from the literature and model simulations where birds are inoculated with one dose of oocysts and the number of daily shed oocysts are recorded. The first column shows the inoculation dose, the experimental inoculation dose is grouped into the model categories for low (until 100 oocysts), middle (more than 100 until 10^4 oocysts) and high(more than 10^4 until 10^6 oocysts) uptake. The second column shows the excreted oocysts in the days subsequent to the inoculation. The experimental excretion data is translated into the model categories for low (until 10^4 oocysts, depicted as 2), middle (more than 10^4 oocysts until 10^6 oocysts, depicted by 3) and high (more than 10^6 oocysts until 10^8 oocysts, depicted by 4) dose of excretion. * denotes less than $500,000$ oocysts.

inoculation dose (oocysts)	excreted	origin
low	**0 0 0 0 2 2 2**	model
middle	**0 0 0 0 3 3 3 3 3 3 3 2 2**	model
150	0 0 0 0 2 3 3 3 3 3 3 3	[14]
1,250	0 0 0 0 3 4 4 4 3 3 3	[10]
5,000	0 0 0 0 3 4 4 4 4 4 3	[10]
high	**0 0 0 0 4 4 4 4 4 4 4 4 3 3 2**	model
20,000	0 0 0 0 4 4 4 3 3 3 3 3 2 2 2	[7]
1,000,000	0 0 0 0 4 4 4 4 4 4 4 3 3 2 2	[7]
20,000	0 0 0 0 4 4 4 4 4 4 4 4 3	[12]
20,000	0 0 0 0 3 4 4 4 4 4 4 *	[10]
80,000	0 0 0 0 3 4 4 4 4 4 4 *	[10]
320,00	0 0 0 0 3 4 4 4 4 4 4 4 * *	[10]
80,000	0 0 0 0 4 4 4 4 4 4 4 * * * *	[11]
320,000	0 0 0 0 4 4 4 4 4 4 4 * * *	[11]
50,000	0 0 0 0 3 4 4 3 3 2	[25]

Table 3: Trickle infection experiments from the literature and model simulations. The first column shows the inoculation dose the birds receive per day and for how many days. The second column shows the excreted oocysts in the days subsequent to the inoculation. The experimental excretion data is translated into the model categories for low (until 10^4 oocysts, depicted as 2), middle (more than 10^4 oocysts until 10^6 oocysts, depicted by 3) and high (more than 10^6 oocysts until 10^8 oocysts, depicted by 4) dose of excretion.

trickle infection (days x oocysts)	excreted	origin
25 x low	**0 0 0 0 3 3 3 3 3 3 3 3 2 2 2 2 2 2**	model
25 x 25	0 0 0 0 2 3 3 3 3 3 3 3 3 3 3 3 3 3	[14]
20 x middle	**0 0 0 0 4 4 4 4 4 4 4 4 2 2 2 2 2**	model
20 x 1,000	0 0 0 0 4 4 4 4 4 4 4 3 3 3 2 2	[7]

Table 4: Immunization experiments from the literature and model simulations. The first column shows the immunizing dose(s) the birds receive per day and for how many days. The challenge dose is shown in the second column. The third column shows the excreted oocysts in the days subsequent to the challenge. The experimental excretion data is translated into the model categories for low (until 10^4 oocysts, depicted as 2), middle (more than 10^4 oocysts until 10^6 oocysts, depicted by 3) and high (more than 10^6 oocysts until 10^8 oocysts, depicted by 4) dose of excretion.

immunization dose (days x oocysts)	challenge (oocysts)	excreted	origin
20 x middle	high (day 21)	**0 0 0 0 0 2**	model
20 x 1,000	1,000,000 (day 21)	0 0 0 0 3 2	[7]
1 x high	high (day 15)	**0 0 0 0 3 3 3 2 2 2**	model
1 x 80,000	160.000	0 0 0 0 3 3 3	[11]
high (day 1 and 15)	high (day 28)	**0 0 0 0 3 3 3 2 2 2**	model
80,000 160,000	80,000 (day 28)	0 0 0 0 3 3 3 3 3 3 3 3 3 3	[11]
80,000 160,000	5,120,000 (day 28)	0 0 0 0 0 3 3 3 3 3 3 3	[11]

2.3 Model output

Simulations start with an initial situation of only naive chickens and a user-given percentage of low-level contaminated patches in the shed. We study the model by simulating the dynamics for 49 days, a common rearing time for broiler chickens, and which proved sufficiently long to capture the dynamics of the epidemic under all initial conditions studied. During the simulations we monitor the numbers of chickens ingesting low, medium, or high oocyst levels, excreting low, medium, or high oocyst levels, and becoming partial or fully immune. In our analysis we will focus on the average number of high level excretions during one simulation, i.e. the mean high excretion per chicken during its time in the shed. This output variable gives an indication of the severity of an outbreak since the excretion of high amounts of oocysts is more likely to be related to problems caused by clinical coccidiosis, such as intestinal damage and individual-level growth reduction [5].

2.4 Default settings for simulation

The default setting of the model for simulation is a grid of 1681 patches of 10 by 10 cm on which we let 353 chickens move yielding a chicken density of 21 chickens per square meter. During set-up, initial patch contamination is set to low levels of oocysts and concentrated in a single position in the field. An ingested dose of oocysts is defined to be part of a trickle infection when, within four days after uptake, a chicken picks up at least two more doses of equal or higher level. Chickens reach partial immunity nine days after the second of two low or medium doses, or nine days after one single high dose. Full immunity is reached after 13 doses with a cumulative oocyst exposure of 33 (low doses counting for two,

medium for three, and high for four), but no earlier than six days after becoming partially immune. To asses the stochastic variation for one set of initial conditions we vary the random seed of the model.

3 Results

3.1 Default result

The average number of high-level excretions per chicken is shown for a range of initial contaminations in Figure 1a. At each initial contamination depicted, simulations were executed in three-fold with varying random seeds, which proved sufficient to capture a large part of the stochastic variation. The graph shows a non-monotonous 'wave-like' relation with alternating peaks and troughs between the percentage of patches initially contaminated and the average number of high excretions per individual. The effect of field size was studied by increasing the field to 14 x 14 meters, which also allowed us to simulate with even lower initial levels, and observe that the wave-like pattern continued. The time course of the infection with respect to exposure and immunity at the peaks in excretion (initial contamination of 0.01% and 0.8%) and the trough (initial contamination of 0.1%) are shown in Figures 1b,c and d.

The fraction of chickens that pick up a low-dose trickle infection during a simulation is shown in Figure 1b for one random seed. Middle-dose trickle ingestion and high dose ingestion are the ones that lead to high-dose excretions in non-immune chickens. We show the fraction of chickens that pick up either one of these in Figure 1c for the same simulation as in Figure 1b. Both partial and full immunity preserve chickens from high-dose excretions. The fraction of chickens that are partially or fully immune is shown in Figure 1d.

3.2 Robustness and sensitivity

We carried out a sensitivity analysis with respect to the model settings, to assess the robustness of the wave-like pattern and investigate the underlying mechanism. These analysis were performed for initial contaminations of 0.05 % up till 40% which we tested to be sufficient to capture the sensitivity of the parameters.

Density

The maximal density of broilers for conventional broiler farms is 23 birds per square meter [3]. To asses the effect of broiler density on the results we have studied densities of 6, 10, 13 and 23 birds per square meter additional to the default settings. No or negligible difference with the default result was found.

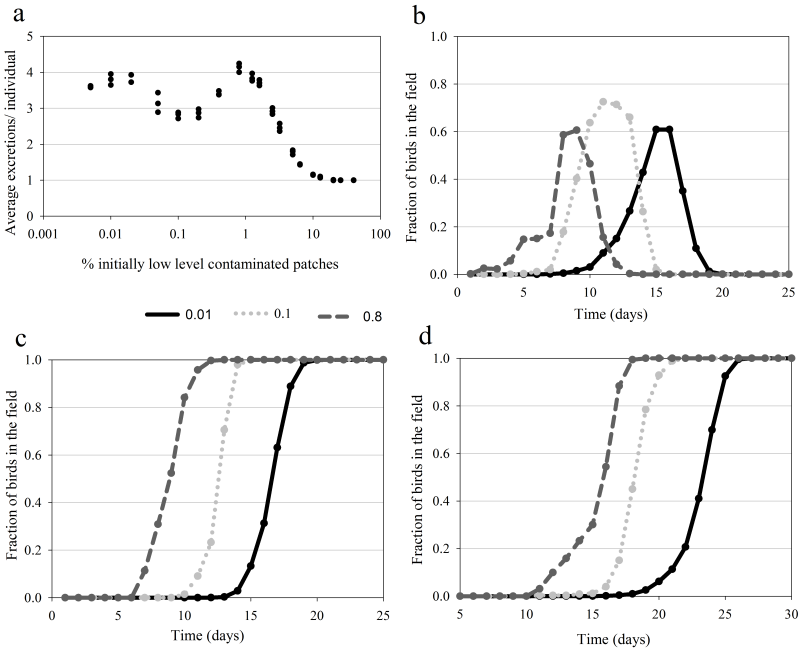


Figure 1: (a) The average number of high-dose excretions per individual for the default model settings plotted for a range of initial percentage of contaminated patches with a low-level of oocysts. For each initial contamination three simulations with different random seeds are plotted. The time course of the fraction of individuals that have ingested a low-dose trickle infection (b), ingested either middle-dose trickle or high-dose infection (c) and have reached either partial or full immunity (d) is plotted for the two peaks in high-dose excretion (of panel a) at 0.01% and 0.8% initial contamination and the trough in excretion at 0.1% initial contamination. NB. Note the difference in the time range in panels b, c and d.

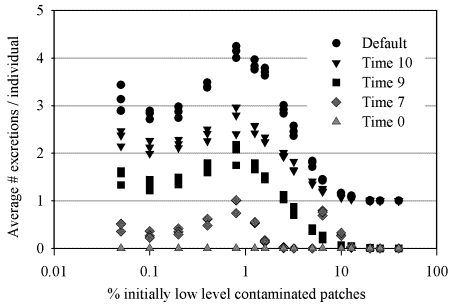


Figure 2: The average number of high-dose excretions per individual is plotted for a range of initial contaminations using different immune rules. The upper graph (full black circles), depicts simulations with the default model settings for immunity, i.e. individuals reach immunity nine days after the second ingestion of a low or middle dose of oocysts or nine days after the first ingestion of a high dose of oocysts. The four lower graphs denoted with Time 10 to Time 0, depict simulations were individuals reach immunity respectively 10 days, 9 days, 7 days, or immediately after the first ingestion of any dose of oocyst. For each graph at each initial contamination three simulations with different random seeds are plotted.

Immune rules

Because it is the second dose of oocysts that triggers the build-up of immunity (starting 9 days after the second dose), heterogeneity in immune status arises due to the different waiting times for chickens to encounter their second dose. If the oocyst level were constant, these waiting times would follow a hypergeometric distribution with a variance of 2r(1-r), r being the daily ingestion probability. To decrease the variance (to r(1-r) in a constant environment), we considered the case with immunity starting 0, 7, 9, or 10 days after the first oocyst dose (Figure 2). Not surprising, excretion levels decreased over the whole range due to earlier immunity, but besides that, the wave amplitude also decreases significantly. This suggests that heterogeneity in immunity is an important factor causing the wave-like pattern.

Trickle rules

A trickle infection occurs when ingestion of an oocyst dose is followed by two more equal or higher-level doses in the subsequent four days, and it results in elevated excretion levels. If the definition for a trickle infection is changed to only one more dose in the subsequent four days, this obviously results in a faster increase in environmental oocyst levels, and thus in more high excretions (Figure 3). Conversely, if three or four more doses are required, excretion levels considerably decrease, but also the wave amplitude decreases. Because these stronger requirements for

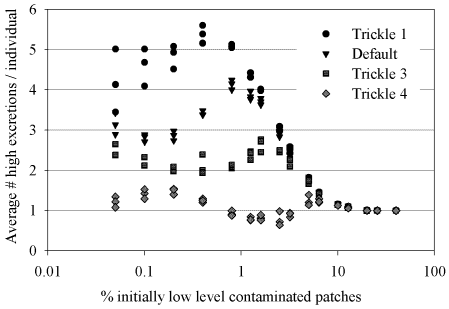


Figure 3: The average number of high-dose excretions per individual is plotted for a range of initial contaminations using different trickle definitions. The second graph (triangles), depicts simulations with the default trickle definition, i.e. within four days after the uptake under consideration, an individual must pick up at least two more doses of equal or higher level to acknowledge a trickle infection. The three other graphs denoted with Trickle 1, Trickle 3 and Trickle 4, depict simulations were respectively one, three, or four additional doses of equal or higher level must be ingested within four days after the uptake under consideration to acknowledge a trickle infection. For each graph at each initial contamination three simulations with different random seeds are plotted.

trickle infections make it less likely that there will be chickens experiencing trickle infections and chickens not ingesting any oocyst at the same time, this suggests that heterogeneity in exposure is crucial to the wave-like pattern.

Multiple dose initial contamination

One can expect that at higher initial contamination levels, patches are not only low-level contaminated but that also some patches will be contaminated with a higher level of oocysts. These more natural conditions were simulated by letting the initially contaminated patches have a 90% probability of becoming low-level contaminated, 9% of becoming middle-level contaminated and, 1% of becoming high-level contaminated (Figure 4). This results in an increase in the average number of high excretions at high initial contamination but does not change the wave-like pattern.

Spatial effects

The initially contaminated patches can be randomly distributed over the field or concentrated in a single position in the field. In both cases the results remain the same because of the random mixing of the chickens. Although biologically unrealistic, we turned off the random mixing and limited the movements of the chickens to improve our understanding about spatial effects that could occur in

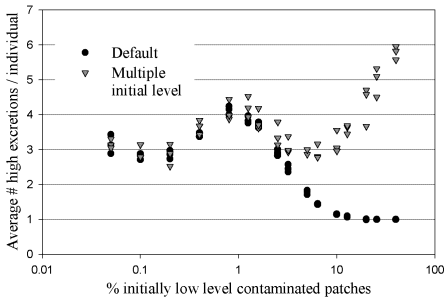


Figure 4: The average number of high-dose excretions per individual is plotted for a range of initial contaminations using different initial conditions. The graph denoted with the full black circles, depicts simulations were all initially contaminated patches are low-level contaminated (the default model settings) The graph denoted with the triangles, depicts simulations were the initially contaminated patches have a 90% probability of being low-level contaminated, 9% of being middle-level contaminated and, 1% of being high-level contaminated. For each graph at each initial contamination three simulations with different random seeds are plotted.

this disease system under other circumstances. Restricted movement is implemented by letting chickens displace to one of their empty neighbouring patches only once every time step (i.e. movement is restricted to 10 cm per hour). If no empty patch is available they did not move. Restricted movement is studied with local initial contamination (Figure 5), which resulted in much higher excretion and disappearance of the wave-like pattern. Because spatial effects are likely to induce much heterogeneity that was not there in the default model - by enabling trickle infections even with few oocysts as well as by keeping some birds away from oocysts even if there is quite many contamination - Figure 5 suggests that the spatial heterogeneity mechanism is much stronger than the one causing the waves (which we will discuss shortly).

Vaccination

Vaccines against coccidiosis that can sprayed over the feathers of newly hatched chicks are available of attenuated and/or wildtype *Eimeria* oocysts causing them to ingest a low dose of oocysts while grooming. In the model vaccination is simulated by letting all chickens ingest a low dose of oocysts on the first day (Figure 6). At low initial contamination levels broilers excrete much less high levels of oocysts than without vaccination but at higher initial contamination levels the difference becomes much smaller.

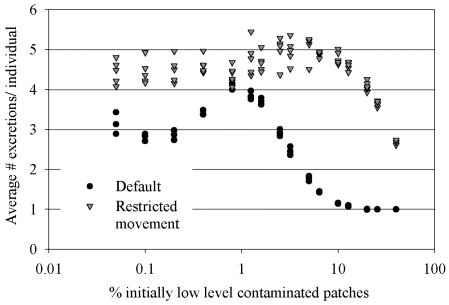


Figure 5: The average number of high-dose excretions per individual is plotted for a range of initial contaminations using different displacement rules for the individuals. The graph denoted with the full black circles, depicts simulations were individuals move by randomly changing position to another empty patch in the field once every hour (the default model settings). The graph denoted with the triangles, depicts simulations were the movement is restricted to only displace to one of the empty neighboring patches every hour. If no empty patch is available an individual will not move. At each initial contamination three simulations with different random seeds are plotted for the default graph and five simulations with different random seeds are plotted for the restricted movement graphs.

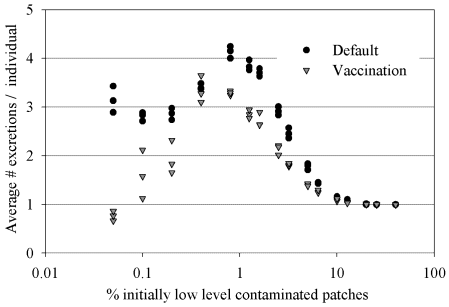


Figure 6: The average number of high-dose excretions per individual is plotted for a range of initial contaminations with and without vaccination. The graph denoted with the full black circles, denotes the the default model settings (no vaccination). The graph denoted with the triangles, depicts simulations were all individuals were vaccinated by giving them a low-dose of oocysts on the first day. For each graph at each initial contamination three simulations with different random seeds are plotted.

4 Discussion

Previously, a simple caricatural model for the dynamics of parasites of farmed animals including a two-way interaction between immunity and transmission (acquired immunity) showed that such a disease system can exhibit complicated dynamics [21]. The result of the complexity is that small differences in e.g. initial contamination can lead to large differences in the severity of outbreaks. The more biologically realistic mechanistic model of Klinkenberg and Heesterbeek [16] incorporated the interaction between immune status and infection history for a specific system, *Eimeria* infections in chicken. They show similar bifurcation behaviour as in the caricatural model. Moreover this model showed dynamics resulting in a non-monotonous wave-like relation between initial contamination and the severity of outbreaks. The direct consequence of this pattern in terms of control measures is counter intuitive, better cleaning to remove more of the parasites can lead to more severe outbreaks. Both these analysis regarded all individuals to be identical and did not yet include acquired heterogeneity.

In this paper we modelled the spread of *E. acervulina* in chickens as a parasite-host system where traits of the host influence the infections dynamics and vice-versa making explicit differences between individuals. Our model description was based on empirical research and can also be validated empirically. We have simulated transmission experiments performed at our group by Velkers et al. [24]. In these experiments one susceptible naive 1-day-old chick and one inoculated with a dose of oocysts, are reared together in a confined area. The excretions of the two chickens are recorded to study the transmission dynamics of the infection. The result of their experiments look promisingly consistent with our results.

Previous studies both empirical and theoretical, have shown a non-monotonous wave-like relation between initial contamination and the oocysts produced during an outbreak of coccidioses [13, 9, 16]. With our model we indeed also observe this wave-like relation between initial oocyst contamination and the numbers of oocysts excreted in the weeks thereafter. In addition we have investigated the robustness of the wave-like relation to various model settings and we find the pattern to be robust. In contrast to previous studies our analysis allows us to give a possible mechanism behind the pattern. The wave-like relation can be explained by an increased heterogeneity in the host population at the peaks, caused by the infection process. Subsequent waves arise because at lower initial contaminations, similar heterogeneity can be reached through additional oocyst generations.

Generally, the amount of high oocyst excretions is determined by a race between the chickens trying to be immune as fast as possible, and the parasite trying to induce high excretions as early as possible. When more chickens are immune earlier while oocyst accumulation remains unaltered, excretion decreases (Figure 2), and when oocysts accumulate faster or slower, excretion increases or decreases, respectively (Figure 3). However, the initial contamination level can also influ-

ence this race, maximising excretion by letting a minority of chickens accumulate oocyst levels whereas most are not yet building up immunity. This can clearly be seen in Figure 1b through d, where the time course for three initial conditions is shown of the fraction of chickens taking up low trickle infections (that will give rise to medium excretion), taking up medium trickle and high infections (that may give rise to high excretion if not yet immune), and being partially or fully immune.

The interesting difference occurs after an initial contamination-dependent start-up phase needed to accumulate low doses to enable trickle infections. Then, in the simulations with 0.01% and 0.8% contamination, the increase in the numbers of chickens taking up low trickle infections occurs over a much longer time span than with 0.1%, mainly due to a slow start-up phase (Figure 1b). However, the reduced number of initially infected chickens do excrete sufficient medium doses for all birds, because the uptake of medium trickle (and subsequent high) infections is essentially the same for all three simulations (Figure 1c). As a result, all chickens that were not infected in the slow start-up, and thus could not start to build-up immunity (Figure 1d), loose a couple of days in the 'race against the parasite' and excrete more high doses.

The importance of heterogeneity in causing problems, or conversely, of homogeneity reducing problems is seen if a more homogeneous population is created by means of vaccination (Figure 6) or by starting with a very high percentage of low oocyst level coverage (Figure 1a). On the other hand, spatial effects are likely to enhance heterogeneity, because some birds will remain near infected patches, thus accumulating oocyst levels, whereas many other chickens will not yet be able to mount an immune response (Figure 5).

In our relatively simple model with only a minimum number of oocyst and immunity levels we show the large effect that acquired heterogeneity can have. This effect is relevant in any disease system where heterogeneity can be acquired in traits of the host that interact with the infection process. For example diseases where more resistant or more virulent strains arise for infectious agents where there is some cross-immunity between strains or serotypes, and where the internal dynamics are then influenced by the memory of previous exposures. In this case as well, it is the exposure history which shapes the immune status, however defined, and it is this immune status that in turn influences the infectiousness of that individual to others. When using models to develop or evaluate control measures for these type of disease systems it is always crucial to incorporate the effect of acquired heterogeneity.

References

[1] *The Merck Veterinary Manual*. Merck & Co., Inc. and Merial Limited, Whitehouse Station, NJ, USA, 9th edition edition, 2006.

[2] P. C. Allen and R. H. Fetterer. Recent advances in biology and immunobiology of Eimeria species and in diagnosis and control of infection with these coccidian parasites of poultry. *Clin Microbiol Rev*, 15(1):58–65, 2002.

[3] Anonymous. *The Welfare of Chickens Kept for Meat Production*. Scientific Committee on Animal Health and Animal Welfare, Health& Consumer Protection Directotate-General, European Commission, 2000.

[4] H. D. Chapman. Anticoccidial drugs and their effects upon the development of immunity to Eimeria infections in poultry. *Avian Pathol*, 28(6):521–535, 1999. doi: 10.1080/03079459994317.

[5] H. D. Chapman, P. L. Matsler, V. K. Muthavarapu, and M. E. Chapman. Acquisition of immunity to *Eimeria maxima* in newly hatched chickens given 100 oocysts. *Avian Dis*, 49(3):426–429, 2005.

[6] O. Diekmann and J.A.P. Heesterbeek. *Mathematical Epidemiology of Infectious Diseases*. John Wiley & Sons Ltd, 2000.

[7] M. M. Galmes, C. C. Norton, and J. Catchpole. Comparison of resistance level and circulating IgG response in chickens experimentally inoculated with a multiple or single immunizing doses of *Eimeria acervulina*. *Ann Parasitol Hum Comp*, 66(4):144–148, 1991.

[8] E. A. Graat, A. M. Henken, H. W. Ploeger, J. P. Noordhuizen, and M. H. Vertommen. Rate and course of sporulation of oocysts of Eimeria acervulina under different environmental conditions. *Parasitology*, 108 (Pt 5):497–502, 1994.

[9] E. A. Graat, H. W. Ploeger, A. M. Henken, G. De Vries Reilingh, J. P. Noordhuizen, and P. N. Van Beek. Effects of initial litter contamination level with *Eimeria acervulina* on population dynamics and production characteristics in broilers. *Vet Parasitol*, 65(3-4):223–232, 1996.

[10] H. Hein. The pathogenic effects of Eimeria acervulina in young chicks. *Exp Parasitol*, 22(1):1–11, 1968.

[11] H. Hein. Resistance in young chicks to reinfection by immunization with two doses of oocysts of Eimeria acervulina. *Exp Parasitol*, 22(1):12–18, 1968.

[12] H. E. Hein. *Eimeria acervulina*, *E. brunetti*, and *E. maxima*: pathogenic effects of single or mixed infections with low doses of oocysts in chickens. *Exp Parasitol*, 39(3):415–421, 1976.

[13] A. M. Henken, H. W. Ploeger, E. A. Graat, and T. E. Carpenter. Description of a simulation model for the population dynamics of *Eimeria acervulina* infection in broilers. *Parasitology*, 108 (Pt 5):503–512, 1994.

[14] L. P. Joyner and C. C. Norton. The immunity arising from continuous low-level infection with *Eimeria maxima* and *Eimeria acervulina*. *Parasitology*, 72(1):115–125, 1976.

[15] D. Klinkenberg and J. A P Heesterbeek. A simple model for the within-host dynamics of a protozoan parasite. *Proc. R. Soc. B*, 272(1563):593–600, 2005. doi: 10.1098/rspb.2004.2987.

[16] D. Klinkenberg and J. A P Heesterbeek. A model for the dynamics of a protozoan parasite within and between successive host populations. *Parasitology*, pages 1–10, 2007. doi: 10.1017/S0031182007002429.

[17] N. J. Lewis and J. F. Hurnik. Locomotion of broiler chickens in floor pens. *Poult Sci*, 69(7):1087–1093, 1990.

[18] H. S. Lillehoj and E. P. Lillehoj. Avian coccidiosis. A review of acquired intestinal immunity and vaccination strategies. *Avian Dis*, 44(2):408–425, 2000.

[19] A. P. Preston and L. B. Murphy. Movement of broiler chickens reared in commercial conditions. *Br Poult Sci*, 30(3): 519–532, 1989.

[20] P. S. Reyna, L. R. McDougald, and G. F. Mathis. Survival of coccidia in poultry litter and reservoirs of infection. *Avian Dis*, 27(2):464–473, 1983.

[21] M. G. Roberts and J. A. Heesterbeek. A simple parasite model with complicated dynamics. *J Math Biol*, 37(3):272–290, 1998. doi: 10.1007/s002850050129.

[22] W. J C Swinkels, J. Post, J. B. Cornelissen, B. Engel, W. J A Boersma, and J. M J Rebel. Immune responses in *Eimeria acervulina* infected one-day-old broilers compared to amount of *Eimeria* in the duodenum, measured by real-time pcr. *Vet Parasitol*, 138(3-4):223–233, 2006. doi: 10.1016/j.vetpar.2006.02.011.

[23] F. C. Velkers, D. P. Blake, E A M. Graat, J C M. Vernooij, A. Bouma, M C M. de Jong, and J. A. Stegeman. Quantification of eimeria acervulina in faeces of broilers: comparison of mcmaster oocyst counts from 24h faecal collections and single droppings to real-time pcr from cloacal swabs. *Vet Parasitol*, 169(1-2):1–7, 2010. doi: 10.1016/j.vetpar.2010.01.001.

[24] F. C. Velkers, A. Bouma, E A M. Graat, D. Klinkenberg, J. A. Stegeman, and M C M. de Jong. Eimeria acervulina: the influence of inoculation dose on transmission between broiler chickens. *Exp Parasitol*, 125(3):286–296, 2010. doi: 10.1016/j.exppara.2010.02.005.

[25] B. Vermeulen, H. W. Peek, J. P. Remon, and W. J M Landman. Effect of ibuprofen on coccidiosis in broiler chickens. *Avian Dis*, 48(1):68–76, 2004.

[26] U Wilensky. Netlogo. *Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.*, http://ccl.northwestern.edu/netlogo/, 1999.

[27] R. B. Williams. Effects of different infection rates on the oocyst production of *Eimeria acervulina* or *Eimeria tenella* in the chicken. *Parasitology*, 67(3):279–288, 1973.

[28] R. B. Williams. Epidemiological studies of coccidiosis in the domesticated fowl (*Gallus gallus*): II. Physical condition and survival of *Eimeria acervulina* oocysts in poultry-house litter. *Appl Parasitol*, 36 (2):90–96, 1995.

[29] R. B. Williams. Quantification of the crowding effect during infections with the seven *Eimeria* species of the domesticated fowl: its importance for experimental designs and the production of oocyst stocks. *Int J Parasitol*, 31(10):1056–1069, 2001.

CHAPTER 5

The success of different variant surface antigen repertoires of malaria parasites competing within and between hosts.

M. Severins, D. Klinkenberg, and J.A.P Heesterbeek

Abstract

Red blood cells infected by the malaria parasite *Plasmodium falciparum*, express variant surface antigens (VSA) to evade host immunity. The best-studied malaria VSA are coded by the large and extremely diverse *var* gene family. The population-level *var* gene pool is organized into groups that preferentially recombine within-group and each parasite carries a proportional representation of *var* genes from these groups in its repertoire. One proposed mechanism for this structuring (both on population and individual level) to have evolved is the opposing selection pressures for optimal functionality and for maximal diversity. In young hosts with little immunity, what we here denote by, 'strong' VSA can confer optimal functionality while in older hosts with more immunity, 'rare' VSA generate maximal diversity. The aim of this study is to examine whether these two selection pressures are sufficient to explain the existence of 'strong-rare' VSA repertoires and to learn more about the infection dynamics of the parasites. We use an individual-based model, including within and between-host parasite competition which we partially validate with published field data. We find that partitioning is not the optimal structure that emerges for the parasites.

1 Introduction

The malaria parasite *Plasmodium falciparum*, expresses variant surface antigens (VSA) on the surface of infected red blood cells (rbc) to avoid recognition by host antibodies [8, 39]. *P falciparum* has several VSA, the best-studied being *P. falciparum* erythrocyte membrane protein 1 (PfEMP1), coded by the *var* gene family [4, 53]. These *var* genes make *P. falciparum* unique amongst *Plasmodium* species because they enable the parasites to adhere to postcapillary endothelial cells (cytoadherence) and to non-infected red blood cells (rosetting) [32, 24, 47], leading to sequestration of the cells in different tissues. This avoids their passage through the spleen where they would be recognised and destroyed [32, 25] and is responsible for many of the pathological features of malaria [46].

Var genes are extremely diverse on a population level, with each parasite carrying approximately 60 different copies of the gene [51, 8, 27]. Due to frequent recombination new parasites with different sets of *var* genes emerge in the parasite population continuously [52, 37, 24, 26]. Parasites express their *var* genes sequentially in a mutually exclusive manner [15, 24], resulting in a persistent *P. falciparum* infection with repeated waves of parasitemia [45, 23]. In this process of antigenic variation each wave represents a distinct sub-population of parasites that dominantly expresses a different VSA on the surface of infected rbc [25, 47]. In areas with high transmission, individuals build up a broad repertoire of antibodies specific for these VSA [6] which are thought pivotal for naturally acquired immunity [10, 28, 20, 26, 54]. Although hosts continue to become infected throughout life, after relatively few infections children are protected from severe malaria disease and eventually against all clinical malaria [12, 6].

Studies by Bull et al. [11, 12] show that parasite isolates taken from young patients with severe disease are more frequently recognised by antibodies from other individuals than parasites isolates from older children, suggesting an order in the VSA expression of infecting parasites. This initiated the idea that there might be a limited number of *var* genes with corresponding antigens, expressed earlier in life and responsible for disease [31, 13]. Using genomic data, *var* genes have been classified into approximately five groups [36, 41, 39] and the proportional representation of *var* genes from each group in individual parasite genomes seems to be preserved [13, 9]. The expression of *var* genes from specific groups has been associated with distinct features of malaria infection and pathology [33, 37]. Recombination of the *var* genes, in general, appears to be more frequent within the groups than between the groups [38, 14], in particular for the group of *var* genes that is most associated with severe disease [14].

Directly testing the effect of a partitioned VSA repertoire on host immunity and *vice versa*, is difficult due to the missing link between *var* genotype and phenotype (e.g. antigen), the extreme *var* gene diversity and other technical constraints [16, 9]. Mathematical and computational studies have helped to increase our un-

derstanding of the currently available data. For example work of Recker et al. [48] showed how a partitioned VSA repertoire can protect from severe malaria after relatively few infections and reduce strain diversity. Work by Gupta et. al [30, 29] showed that under the influence of immune selection, the parasite population can become organized into groups of strains with non-overlapping antigenic repertoires that are static or have cyclical dynamics. Buckee et al. [9] showed that parasite strains that have non-overlapping repertoires for VSA from one small group and a random selection of VSA taken from a large group, better fit serological data than strains with a completely random set of VSA, or strains that are completely non-overlapping in their VSA repertoire.

At present, the role of host immunity in the emergence and maintenance of a grouped structure of the *var* genes remains unknown. For the development of vaccines or other control strategies and for the prediction of their (long-term) effectiveness, it is important to understand how the parasite population structure is shaped by host immune selection [16, 40]. Due to the gradual build up of immunity parasites are thought to experience different selection forces with host age. In young children, with no or few antibodies, selection favours parasites expressing *var* genes with optimal functionality, while in older (more immune) children, selection favours parasites expressing rare *var* genes with VSA that are not recognised by host immunity [11, 9]. It has been suggested that a grouping in the *var* genes is maintained because each group contains genes that are functionally specialised [49, 34, 37] and structuring helps maintain a balance between function optimization and antigenic diversity [38, 14].

In this study we use an individual-based model where VSA have only one trait, for which each VSA is competent to a greater or lesser extent. Because of the single trait we expect a two-group structure, where parasites have a small set of VSA with optimal function to express in young hosts, referred to as the 'strong' VSA, and a larger set of very diverse VSA with less optimal characteristics to express in older hosts, referred to as the 'rare' VSA, to have an evolutionary advantage over other structures. Parasites with such a structure we will refer to as parasites with a *strong-rare* structure. We do not a priori impose any partitioning onto the parasites, or grouping onto the VSA but rather allow for parasite structures to emerge freely under the influence of infection and immunity. The hosts antibody responses that are built up as a result of infections by the emerging parasite strains are partially validated using published field data from Kenya [11, 12, 13]. We examine whether under the proposed selection pressures parasite strains with a *strong-rare* structure as described above are indeed more successful, thus providing evidence for the proposed mechanism for structure existence. We find that under a variety of conditions examined in the model, the selection mechanism proposed is not sufficient to explain the existence or emergence of a *strong-rare* structure in the parasites' VSA repertoire.

2 The model and methods

The model is written in the C++ programming language, a detailed description of the model can be found in the Supplementary Material. The main actors in our model are hosts, parasite strains and VSA. Populations of these are kept constant (2000 hosts, 50 strains, 750 VSA) during simulation. A parasite strain is defined as a set of 10 different VSA. At initiation hosts are created, each with age zero, and without antibodies. Model simulations run with time steps that are equivalent to days in real life. Once every year (365 time steps) either the host's age is incremented by one or the host dies. When a host dies it is replaced by an uninfected newborn host, with age zero and no antibodies. To obtain a realistic host demography we have imposed age-dependent mortality probabilities from a demographic study in a high-transmission malaria area characterised by high infant mortality, for a large part caused by malaria [1, 2, 5]. Simulations last the equivalent of 300 years. Since at initiation all hosts are newborns, the model is run without generating output for a transient period of 150 years to allow for the establishment of a realistic host demography and antibody repertoire.

We assume that each *var* gene corresponds to only one VSA and that this VSA is located at the binding site of the surface protein. Vsa of other non-*var* gene surface proteins are not included in the model. Each VSA is defined by the value it has for its single trait, its 'binding value'. The binding value is an integer number which we take to be inversely correlated with the ability of expressed VSA to cytoadhere. Thus the lower the binding value, the 'stronger' the VSA. We assume that stronger VSA are better at avoiding splenic clearance and therefore the binding value of the expressed VSA determines the net within-host growth rate of an infecting parasite (as suggested in [11]).

2.1 Within-host dynamics

We only explicitly model the dynamics of the parasite that takes place in the general bloodstream. Liver-stage parasite proliferation is implicitly taken into account by assuming that a large number of new parasites enters the general bloodstream and infects red blood cells. Infected rbc rupture and release a new generation of parasites into the blood stream approximately every 48 hours [46, 24]. Each rbc-infecting parasite expresses only one of its 10 VSA and we assume that epigenetic mechanisms within the parasites ensure that new generations of infected rbc remember the *var* gene that was expressed by their parental rbc [17]. Supported by findings of Lavstsen et al. [42] we assume that initially every VSA is expressed with equal probability on the surface of the infected rbc, i.e. in 10% of the first generation.

Since the only difference between parasites of an infecting strain is the VSA they are expressing, we can model the dynamics by keeping track of the proportion with which each VSA is expressed by the infected rbc. Because we do not model

the absolute number of infected rbc or the production of gametocytes (sexual stage of the parasite which is taken up by the mosquitos [21]), the probability of transmission remains equal during infection.

Infection is updated once every two days by calculating the new proportions with which infected rbc express each VSA. This calculation is based on the relative differences between the binding values of the expressed VSA (for the equations used in the calculations see Supplementary Material). Due to the differences in net growth rate inferred by the binding values, very quickly, the majority of infected rbc will express the strongest VSA for which no immunity exists.

2.2 Co-infections

Co-infections with different strains are common [50, 22]. In the model, the proportion with which each VSA of a newly infecting strain is expressed on the surface of first generation infected rbc is set to three percent of all infected rbc. The proportional VSA expression in infected rbc in the next generation then depends on the proportion of expression in the current generation and differences in binding values between the VSA of all resident strains. In the default model, all infecting parasites strains are equally likely to become transmitted.

We model within-host strain competition by saying that a parasite strain must infect a minimal proportion of rbc to survive into the next generation. If all of the VSA of a parasite strain are expressed on less than five percent of all infected rbc, the parasite strain does not survive and is cleared from the host (see Figure 5b Supplementary Material).

2.3 Immunity

When a VSA is expressed on more than 30% of the infected rbc, the immune system of the infected host starts to build a specific antibody response against the VSA. Antibodies become effective 22 days after initiation of the specific response (as suggested in [18, 19]). As a result, the proportion of infected rbc expressing the VSA is set to zero. When a host has built up immunity against all VSA that a particular parasite can express, the parasite can no longer survive in the host and is cleared (see Figure 5a Supplementary Material).

2.4 Transmission

Hosts have a daily probability of becoming infected of 0.679, in agreement with infective biting rates measured in high-transmission areas [7]. The infecting parasite strain is selected from all strains with a weight of $1 + strain\,prevalence$ in the human host population. We thereby implicitly assume that the prevalence of the parasite strains in the host population is proportional to the prevalence of

the parasite strains in the mosquito population and that the mosquito population is sufficiently large. By adding one to the strain prevalence, parasites with a prevalence of zero, still have a small probability of infecting a host. This keeps the effective parasite strain population constant and it also allows newly formed parasites (see below) to infect hosts.

2.5 New strain formation

At initialisation strains are created by randomly selecting a set of VSA from the VSA pool. Malaria parasites are known to undergo frequent recombination of their *var*-gene sequences [52, 3]. Recombination takes place during sexual reproduction of the parasite inside the mosquitoes. In the model we allow new strains to form at a daily probability of 0.12. The VSA of the new parasite strain are selected from the pool of all 750 VSA with a (within-host) prevalence-weighted probability of: $1+VSA prevalence$. This reflects that VSA that are more common in the host population are also more likely to meet each other in mosquitoes and recombine into a new strain. There is a small probability for new VSA, with zero prevalence, to be included into the new parasite. The new parasite strain replaces the resident strain infecting the least hosts (typically none).

2.6 Patient sampling and agglutination assays

In agglutination assays, antibodies and infected rbc from the same or different patients are brought together. If the cells agglutinate it indicates that the antibodies recognize (bind to) the VSA used by the parasite isolate. In agglutination assays from hospital studies by Bull et al. [11, 12, 13], antibodies and parasite isolates originate from young patients suffering from malaria disease. The agglutination assays are used to determine the frequency by which parasite isolates of a group of young patients are recognised by antibodies taken three years previously from another group of patients. The frequencies of recognition are then plotted against the age of the patients where the parasites isolates originate from (see figure 2 in [12]). In the model we can distinguish between infected and non-infected individuals but we have no measure for clinical malaria. To have a similar study group as in the hospital studies we have randomly included infected hosts under the age of 21 at seven points in time, a total of 3706 unique hosts (see Table 4 Supplementary Material). We assumed that each host sample contained parasites expressing a single VSA –the most dominant at the time of sampling– and all antibodies of the host. An agglutination event was observed if the expressed VSA of one host was recognised by any of the antibodies of another host. The model is validated by comparing the trends in frequencies of recognition observed in the patients studies with the ones from our study groups.

2.7 Definitions of prevalence measures and parasite success

The within-host success of a parasite strain is defined as the average duration of infection of all the infections caused by the strain. The between-host success, which we use as a measure for the evolutionary success of a parasite strain, is defined as the total number of infections caused by a parasite strain during its existence. Note that this is not the same as the number of different hosts it has infected, as a strain can infect a host more than once. The relative within-host prevalence per strain is defined as the proportion of rbc infected by the same strain (which is one, if the host is only infected by one strain). The prevalence of parasite strains in the host population is defined as the number of hosts infected by a specific strain at a given time. We define four measures of VSA prevalence (the maximum values are for the default model):

- **Expression** The number of hosts, at a given time, in whom the VSA is dominantly expressed (maximum: 2000 hosts, if all hosts dominantly express that VSA).
- **Host** The number of infecting parasite strains carrying the VSA (maximum: 2000 hosts x 50 strains, if all hosts are infected by all strains and all strains carry the VSA).
- **Parasite** The number of parasite strains that carry the VSA (maximum: 50, if all strains carry the VSA).
- **Immune** The number of hosts that have immunity against the VSA (maximum: 2000, if all hosts are immune).

During simulations all four VSA prevalence measures are recorded once every two days. When we look at average VSA prevalences we distinguish between a long-term average and a strain-relative average. The long-term prevalence of a VSA is defined as the average of all the prevalences recorded during a simulation (not including the 150 years transient period). The strain-relative VSA prevalence is meant to capture the average VSA prevalence experienced by a VSA in a specific parasite strain. It is defined as the average of all the VSA prevalences recorded for this VSA during the time of the strain's existence.

2.8 Emerging parasite strain VSA structure

To examine whether structured strains have an evolutionary advantage, we compare the success of several structured strains with each other and with the average parasite strain. First, we look whether strains that have a specific VSA structure are more successful between-host than the average parasite strain ('*structure* → *success*') and second, we look whether strains that are very successful between-host, more often have a specific VSA structure than less successful strains ('*success* → *structure*').

Analyses are performed using the following protocol. At the end of a simulation, for all parasite strains, the strain-relative prevalence of each of their VSA is calculated for all four prevalence measures (a total of 4x10 measures per strain). Then, for every prevalence measure separately, the median and third quartile value of the strain-relative prevalences of all VSA (10 x number of strains) is determined. Also the first quartile of all the lowest binding values (i.e. strongest VSA) of each strain is determined. The strains are then classified as having a *common-rare*, *strong-rare* or a *rare* VSA structure, with respect to each of the four prevalence measures. For this we made the following choices. *Common-rare* parasite strains are defined to have nine out of their ten VSA with a strain-relative prevalence lower than the median value and one VSA with a value higher than the third quartile value. *Strong-rare* parasite strains are defined to have at least nine out of their ten VSA with a strain-relative prevalence lower than the median value and also carry at least one VSA that has a lower binding value (i.e. is stronger) than the first quartile value. *Rare* parasite strains carry VSA that all have a strain-relative prevalence lower than the median value.

For the *success* $\rightarrow$ *structure* analysis, we determine the percentage of each VSA structure amongst the 25 most successful parasite strains and compare this to the percentage of that structure amongst all strains. For the *structure* $\rightarrow$ *success* analysis, the percentage of successful strains amongst strains of each structure, again separately for each prevalence measure, is compared to the all-strain average success. Successful being defined as having a higher between-host success than the third quartile of all strains.

Test for significance are performed on the combined results of three simulations with default parameter settings and different seeds of the random number generator, using Pearson's Chi-square tests.

3 Results

3.1 General model behaviour and validation

The results shown below were found to be robust to changes in the seed of the random number generator and variations in the numbers of hosts, parasite strains, total VSA and VSA per strain. All results shown originate from one representative simulation with the default parameter settings. Where relevant, sensitivity to parameter settings is mentioned in the text.

The imposed age-dependent mortality probability ensures a stable host demography with an age-distribution characterised by a high proportion of children, similar to the ones observed in high malaria transmission areas. To validate the model, we have simulated one of the agglutination studies performed by Bull and colleagues in a high transmission area in Kenya [11, 12, 13]. Our result of an

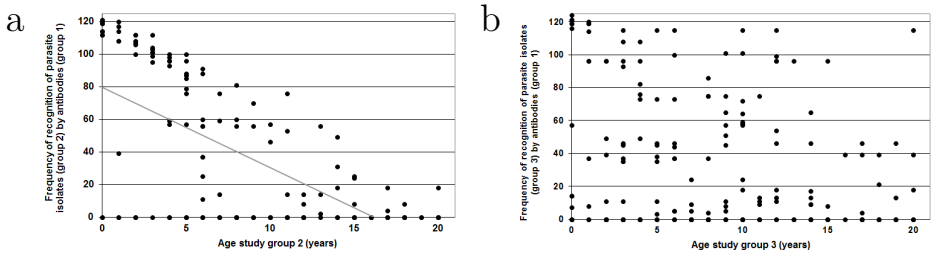


Figure 1: a) The frequency with which parasite isolates of study group 2 are recognised by antibodies from study group 1, against the age of group 2 hosts. In gray is a linear regression line. b) The frequency with which parasite isolates of study group 3 are recognised by antibodies from study group 1, against the age of group 3 hosts.

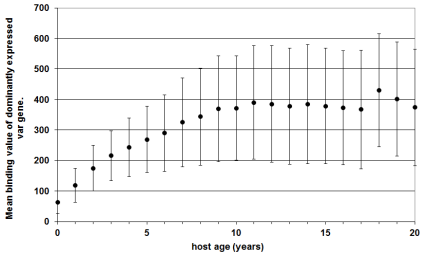


Figure 2: The average binding value of the dominantly expressed *var* genes plotted against the age of the host. Error bars denote the standard deviation.

agglutination assay between study groups three years apart is shown in Figure 1a. This graph, just like the original (see figure 2 in [12]), shows a negative correlation between the age of the host a parasite is isolated from and how often this parasite is recognised. We also find that younger children preferentially express the strongest VSA (see Figure 2). Such a hierarchy in the expression of VSA is in agreement with findings of Buckee et al. [9].

We also simulated agglutination assays between study groups many more years apart. As the time-frame between the study groups becomes larger, the negative correlation, as seen in Figure 1a becomes weaker and eventually disappears. Figure 1b shows the result from an agglutination assay between study groups that are 30 years apart (see Table 4 in Supplementary Material).

To test if stronger VSA also persist longer (i.e. are more 'common') we plot the 150 year-average of the four VSA prevalences (as defined in section 2.7) against the VSA strength. For the *Expression* prevalence we find that the stronger VSA are more often dominantly expressed (Figure 3a); *Immune* prevalence against the

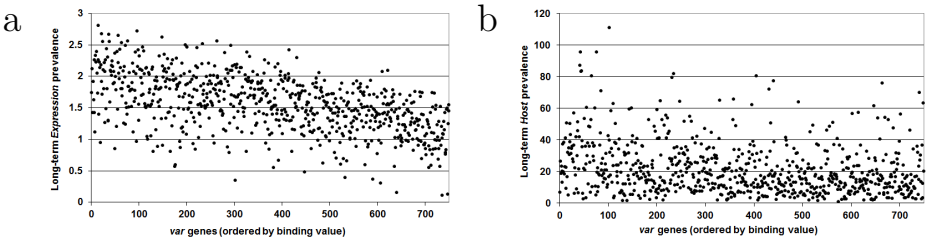


Figure 3: a) The long-term *Expression* prevalence of the *var* genes plotted against the *var* genes ordered by their binding value. The lower the binding value the stronger the *var* genes. b) The long-term *Host* prevalence of the *var* genes plotted against the *var* genes ordered by their binding value. The lower the binding value the stronger the *var* genes.

stronger VSA is therefore also higher (not shown). However, we find no correlation between the strength of a VSA and its long-term *Host* (Figure 3b) or *Parasite* prevalence (not shown).

Parasite dynamics, within and between-hosts

Although hosts are usually infected by more than one parasite strain, infection is generally dominated by the parasite strain that expresses the strongest VSA. Due to the build up of antibodies and co-infections by new strains, the strongest VSA expressed changes continuously. This results in within host dynamics characterised by periodical changes in the dominance of VSA expression and strain infection, analogous to the antigenically distinct waves of parasitemia observed in patients [46] (as shown in Supplementary Material Figure 7a and 7b). It is important to note here that the analogy only applies to the changes in dominance and not to the parasitemia since we do not keep track of the absolute number of infected rbc.

The pattern of the fluctuations in antigenic and strain dominance changes with the age of the host. In young hosts the difference between dominantly and subdominantly infecting strains is very pronounced displaying abrupt transitions between one dominant strain and the next. In older hosts, transitions between dominant strains are less pronounced with longer periods where more than one of the co-infecting parasite strains infects a considerable proportion of the rbc.

In the model, multiplicity (the number of co-infecting parasite strains) decreases with age, approaching zero in the elderly hosts. This is an artifact of the model due to the limited number of VSA, to which the elderly hosts will have built up an almost complete antibody repertoire. Indeed, simulations with twice as many VSA (1500) show that multiplicity remains in the elderly hosts despite the large size of their anti-VSA repertoire (results not shown).

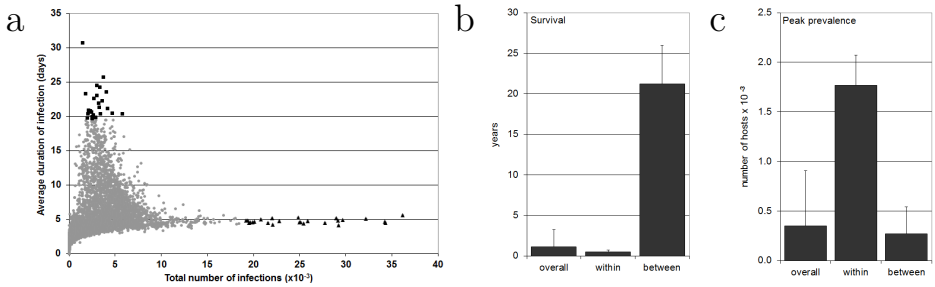


Figure 4: a) The within-host dynamics plotted against the between-host dynamics, each dot is one parasite strain. The 25 most successful within-host strains (squares) and between-host strains (triangles) are marked. For the whole set of parasite strains, the 25 most successful within-host and the 25 most successful between-host strains, the average days of parasite strain survival (b) and the average peak prevalence (c) are shown. Error bars depict standard deviations.

The between-host dynamics is characterised by semi-overlapping waves of a few strains that reach high host prevalence, combined with a continuously low background host prevalence of the remaining strains (as shown in Supplementary Material Figure 7c). To better understand the distinct difference in parasite strain behaviour, we look at the relation between the between-host success and within-host success (see section 2.7 for definitions) as shown in Figure 4a. We find that most parasite strains are not very successful either within or between-hosts. Also, the figure shows that parasite strains seem to be restricted in the sense that they are either good within-host competitors or good between-host competitors, but not both. We mark the 25 most successful parasite strains within-hosts (black squares) and between-hosts (black triangles) to compare their survival time and peak host prevalence with that of all parasite strains in Figures 4b and 4c. The parasite strains that are most successful between-host also survive much longer than the average parasite strain. The average survival time of the most successful between-host strains is 22 years. Results from simulations with different parameter settings show that the survival of the strains is positively related to the number of loci per parasite, the number of hosts and the number of parasites strains. The survival of the strains decreases with increased numbers of VSA. Parasite strains that are most successful within-host have a much higher peak in host prevalence than the average parasite strain. These are the parasite strains that sweep through the host population causing short-term outbreaks after which they go extinct. The most successful between-host parasite strains do not in general infect many hosts at the same time.

Table 1: The percentage of successful (top 25% of the between-host success for all strains) strains with a *strong-rare*, *common-rare* and *rare var* gene structure with respect to prevalence measures, *Expression*, *Immune*, *Host* and *Parasite*. Percentages shown are an average from 3 simulations with default parameter settings but different seeds for the random number generator. Asterisks denote a p-value $<$ 0.001 for Pearson's chi-square test, results without asterisks have a p-value $>$ 0.05. The average size of the subsets is also depicted.

Structure → Success						
Prevalence	**strong-rare**		**common-rare**		**rare**	
measure	successful	size	successful	size	successful	size
All strains	25	6543	25	6543	25	6543
Expression	38.8*	36.7	39.4*	83.3	49.3*	51.7
Immune	40.0	0.3	27.0	24.0	45.8	4
Host	39.4*	151.3	8.5*	69.3	45.9*	284.0
Parasite	57.2*	29.3	54.4*	38.7	71.8*	51.7

3.2 Parasite VSA structure

To examine whether under the proposed selection pressures, a *strong-rare* VSA structure is indeed the most evolutionarily advantageous, we compare the performance of strains with a *common-rare*, *strong-rare var* and *rare* VSA structure to the all strain average as described in section 2.7. The analysis is performed on the results from three simulations with different seeds for the random number generator under default parameter settings, as shown in Table 1 and 2.

The results of the *structure* → *success* analysis depicted in Table 1 show that both strains with a *rare* and a *strong-rare* structure were significantly more successful than the all-strain average. For the *Immune* prevalence measure the difference was not significant. The strains with a *rare* VSA structure were on average most successful, in particular when measured by *Parasite* prevalence. *Common-rare* structured strains were significantly more successful when measured by *Expression* and *Parasite* prevalence, but significantly less successful when measured by *Host* prevalence.

The results of the *success* → *structure* analysis depicted in Table 2 show that among the 25 most successful strains, the percentage of *rare* structured strains was the highest, in particular *rare* with respect to *Host* prevalence. The percentage of *rare* structured and *strong-rare* structured strains among the 25 most successful strains was significantly higher than among all parasite strains, except for *Immune* prevalence. In all simulations very few strains existed with a *rare* structure when measured by *Immune* prevalence, none of them among the 25 most successful strains. *Common-rare* structured strains measured by *Parasite* prevalence were significantly over-represented amongst the 25 most successful strains.

Table 2: Of the whole set of parasite strains (all) and of the 25 most successful between-host parasite strains (successful), the percentage of strains with a *strong-rare*, *common-rare* and *rare var* gene structure is shown with respect to prevalence measures, *Expression*, *Immune*, *Host* and *Parasite*. Percentages are an average from 3 simulations with default parameter settings but different seeds for the random number generator. Asterisks denote a p-value < 0.001 for Pearson's chi-square test, results without asterisks have a p-value > 0.05.

Success → Structure						
Prevalence	**strong-rare**		**common-rare**		**rare**	
measure	all	successful	all	successful	all	successful
Expression	0.67	9.33*	1.27	1.33	0.79	13.33*
Immune	0.03	0.00	0.45	0.00	0.06	0.00
Host	2.66	28.00*	0.99	0.00	4.34	66.67*
Parasite	0.71	17.33*	0.60	5.33*	0.79	21.33*

Robustness of the VSA structure results

We found our results on the parasite VSA structure to be robust for a variety of changes to the default model: i) Increasing the number of VSA to 1500, so that hosts never reach full immunity; ii) Making the VSA un-continuous in their binding value, i.e. VSA with binding value 1-5 and with value 100-844 to mimic the possibility that due to physical constraints there exists only a limited set of VSA that have optimal binding interactions with the molecules on endothelial cells; iii) Only the strain with the dominantly expressed VSA is allowed to transmit, reflecting that a significant amount of parasitemia is needed for transmission; iv) Only strains that have infected a host for more than five days can be transmitted, reflecting the time it takes to form gametocytes (the sexual parasite stage taken up by mosquitoes to transmit the parasite) [21].

4 Discussion

4.1 Emerging VSA structures.

We have built a model that captures some of the important aspects of the complex infection biology of *P. falciparum* involved in VSA dynamics, including parasite multi-level competition and host immune evasion. Also the two opposing immune selection pressures on VSA i.e., selection for optimal growth in young children and selection for rare VSA in older hosts, are both present in the model. Analysis on the results of simulations show that a *strong-rare* VSA structure is not the best performing structure despite the presence of the selection pressures held responsible for the existence of such a *strong-rare* structure. Instead, we find that parasites with a VSA structure of only *rare* VSA, in particular *rare* with respect to their presence in the host and parasite population, perform the best (Table 1 and 2).

To validate that the VSA dynamics in the model are realistic we show that findings from one of the largest serological studies on parasite recognition [11], are found in our simulations namely that there exist more commonly recognised VSA that are mainly expressed in young children (Figure 1a). Furthermore our model supports the frequently proposed idea that these commonly recognised VSA have optimal cytoadherant properties, i.e. are the stronger VSA (Figure 2) [33, 35, 14].

The benefit of expressing a strong VSA is that it allows strains to out-compete other, co-infecting, strains. However, because of this dominance they become more *common* in the hosts' antibody repertoires. Therefore strong VSA can eventually only be used to infect very young children. *Rare* VSA on the other hand, can be used to infect hosts of all ages, also the very young hosts (although only briefly because they get out-competed). Thus, it is in the benefit of a strain to carry as many *rare var* genes as possible. Previous studies had already predicted that strains with non-overlapping antigenic units can evolve in the face of immune selection pressure, even in the presence of recombination because hybrid strains will be recognised by a greater number of hosts [29, 44, 14]. The *parasite* prevalence in our model is a measure for the overlap of VSA in the parasite population. Our results confirm that strains with the least overlapping repertoires perform better.

4.2 Alternative hypotheses for a partitioned VSA structure.

For *P. falciparum* to maintain a partitioned *var* gene repertoire across all parasite strains at the cost of loosing vital strain diversity [48] and probably also at an energy cost, this partitioning must convey an essential benefit to the parasite's survival. Our model shows that safe-guarding optimal functionality for a trait, on its own, is not enough to explain this partitioning. Below, we discuss the feasibility of several alternative mechanisms to explain a partitioning in the parasites' *var* gene repertoire.

There exists the possibility that *var* gene grouping occurs due to molecular constraints rather then immune selection. Some *var* gene recombinations are likely to lead to non-functional proteins which limits recombination possibilities regardless of immune selection. This mechanism could explain why a recombination grouping in the *var* gene pool is observed, and it could also explain a proportional representation of these groups in individual *var* gene repertoires if the groups of viable recombinations are not linked to the antigenicity of the surface proteins coded by them.

In this study we chose to use differences in adhesion strength to distinguish between the VSA. There are more differences between the surface proteins coded by *var* genes, such as differences in the type of receptors they bind which results in

sequestration of infected rbc in different tissues (i.e. the brain or the placenta). Some *var* genes allow infected rbc to form 'rosettes' which plays an important role in clinical disease [43]. We argue that any other *var* gene trait that influences the within-host competition of the parasite strains, instead of binding strength, would have given the same qualitative results.

Another possibility is that the trait of the *var* genes conveys a between-host competitive advantage to the parasite expressing them (such as increased transmission). This could infer enough extra benefit to the parasite that it pays off for parasites to maintain *var* genes optimal in this trait despite the loss in diversity. Even for the trait used in our model this could be possible. If stronger *var* genes infer greater net growth rates they can cause higher parasitemia and probably a higher rate of transmission.

We model only one VSA per binding value and therefore only one antibody is needed to clear all infected rbc that have that binding strength. In reality, we expect that sometimes the adhesion site for tissue receptors is different to the adhesion site for antibodies. This opens up the possibility for *var* genes to recombine so that they keep the tissue binding site intact (maintaining functionality), while changing the antibody site to avoid recognition. A grouping in the *var* gene pool could be explained by this, but it does not explain why each strain carries representatives from each group in the pool. Instead we would expect strains to specialise in carrying *var* genes from one particular recombination group, as previously predicted [30, 44].

At least some of the anti-VSA antibodies are probably not long-lasting [6]. We expect that in the high transmission areas, as simulated in the model host antibody responses are continuously boosted by frequent re-infections and therefore effectively long-lasting. This idea is supported by the observation that protection from disease increases with age [43].

In our model, the order in which VSA are expressed within-host depends on the binding value of the VSA. Such a coupling between trait and order of expression is not unrealistic, because if infected rbc expressing strong VSA out-compete rbc expressing weaker VSA, an order of dominant expression emerges naturally. This coupling strongly affects our results because it makes antibodies against the strong VSA also more common. In a scenario where the order of dominant VSA expression would be random, anti-VSA antibodies would be equally common for all VSA and parasites carrying a strong VSA would have more benefit than in the current scenario. However, as soon as the expression of a VSA from one strain leads to competition with a strain expressing another VSA, antibody prevalence would become skewed again.

The host's immune system uses a myriad of other, non-*var* gene specific, mechanisms as a response to malaria infection. In the model these mechanisms are

jointly represented by only allowing those strains to survive that can maintain a minimum relative prevalence within the host. How these mechanisms would affect the VSA dynamics if we could model them explicitly is unclear.

In many malaria high-transmission areas there exist dry and rainy seasons. During the dry season, parasite transmission is low by reduced abundance of the mosquito vector. Parasite strains have to survive this low-transmission period inside one host, which they are thought to do by switching their VSA. Most hosts will already have immunity against the strong VSA, thus parasites will benefit from carrying *rare* VSA such that they have enough alternative VSA to switch to. We predict that including seasonality into the model will only increase the success of *rare* structured parasites.

Although we model a realistic host demography which includes high child mortality largely caused by malaria, we do not directly model host mortality due to malaria disease. We expect no qualitative difference in our results when modelling malaria morbidity directly because most malaria deaths occur amongst young children that have high parasitemia. We argue that if the infecting strain would not have been fatal, this strain would soon have been out-competed by one of the constantly infecting new strains anyway. Therefore the effect of loss in transmission by the 'killing' strain is minimal. Thus, for the VSA dynamics knowing which strain was responsible for a death is not important, only the frequency of host replacement by a non-immune new-born host is relevant. To model a realistic frequency we imposed the host demography from a high-transmission malaria-endemic region.

4.3 Parasite infection dynamics

Our model gives us insight into parasite strain dynamics that is difficult to obtain from experimental research. In the literature some (groups of) *var* genes or VSA are often referred to as 'common' and it is not always clear in which context this group is more common. In the model we defined four measures to quantify the VSA prevalence (or 'commonness'), i.e. expression prevalence, immune prevalence, host prevalence and parasite prevalence. We found that when VSA are more common according to one prevalence measure, the VSA are not necessarily also more 'common' according to another prevalence measure (Figure 3). The reason for this discrepancy is the order in VSA expression where the stronger VSA are expressed earlier. Therefore it is important to distinguish between more than one VSA prevalence measure since changing one prevalence measure (e.g. the *immune* prevalence with a vaccination program) could have a different, unexpected impact on another VSA prevalence measure.

An interesting observation from the within-host dynamics is that we found that the infection patterns of parasite strains change with host age. This occurs because the dominantly expressed VSA in younger hosts tend to be stronger than

the VSA in older hosts. The assumed difference in net growth rate between competing rbc that both express a stronger VSA is much larger than the difference in net growth rate between rbc that both express weaker VSA. As a result, in young hosts infection is very much dominated by the strain expressing the strongest VSA, while in older hosts co-infecting strains have longer transient periods were both strains have a relatively high within-host prevalence.

In infected individuals, multiplicity (the number of co-infecting parasite strains) has been found to increase with age which is curious in view of the fact that the number of anti-malaria antibodies also increases with age [50]. In the model, we find multiplicity to decrease with age. However, if we only consider parasites that infect at least ten percent of the infected rbc, we find multiplicity to increases with age, due to the changes in within-host infection patterns as described above. If we can assume that in samples of infected individuals only the most prevalent parasite strains are represented, changing infection patterns can explain why it is found that multiplicity increases with age.

4.4 Conclusion

In conclusion our model shows that the advantage for parasites to have a partitioned *var* gene repertoire is not straight forward. Merely assuming that a trade-off between *var* gene functionality versus parasite diversity can cause this partitioning is not sufficient. Our study presents a novel way for malaria researchers to look at the influence of different immune pressures on the parasite, crucial to help predict the outcome of intervention methods that effect host immunity. More studies are needed to explore the validity of alternative mechanisms explaining the partitioning of the parasites *var* gene repertoires. Insights obtained from our model are relevant in any host-pathogen system where the pathogen uses VSA to escape host immunity, like *Trypanosoma brucei*, *Giardia lamblia* and *Borellia* species.

References

[1] Salim Abdullah, Kubaje Adazu, Honorati Masanja, Diadier Diallo, Abraham Hodgson, Edith Ilboudo-Sanogo, Ariel Nhacolo, Seth Owusu-Agyei, Ricardo Thompson, Thomas Smith, and Fred N Binka. Patterns of age-specific mortality in children in endemic areas of sub-saharan africa. *Am J Trop Med Hyg*, 77(6 Suppl): 99–105, 2007.

[2] Kubaje Adazu, Kim A Lindblade, Daniel H Rosen, Frank Odhiambo, Peter Ofware, James Kwach, Anna M Van Eijk, Kevin M Decock, Pauli Amornkul, Diana Karanja, John M Vulule, and Laurence Slutsker. Health and demographic surveillance in rural western kenya: a platform for evaluating interventions to reduce morbidity and mortality from infectious diseases. *Am J Trop Med Hyg*, 73(6):1151–1158, 2005.

[3] Alyssa E Barry, Aleksandra Leliwa-Sytek, Livingston Tavul, Heather Imrie, Florence Migot-Nabias, Stuart M Brown, Gilean A V McVean, and Karen P Day. Population genomics of the immune evasion (var) genes of plasmodium falciparum. *PLoS Pathog*, 3(3):e34, 2007. doi: 10.1371/journal.ppat.0030034.

[4] D. I. Baruch, B. L. Pasloske, H. B. Singh, X. Bi, X. C. Ma, M. Feldman, T. F. Taraschi, and R. J. Howard. Cloning the p. falciparum gene encoding pfemp1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell*, 82(1):77–87, 1995.

[5] Ayaga A Bawah and Fred N Binka. How many years of life could be saved if malaria were eliminated from a hyperendemic area of northern ghana? *Am J Trop Med Hyg*, 77(6 Suppl):145–152, 2007.

[6] James G Beeson, Faith H A Osier, and Christian R Engwerda. Recent insights into humoral and cellular immune responses against malaria. *Trends Parasitol*, 24(12):578–584, 2008. doi: 10.1016/j.pt.2008.08.008.

[7] R. Bødker, J. Akida, D. Shayo, W. Kisinza, H. A. Msangeni, E. M. Pedersen, and S. W. Lindsay. Relationship between altitude and intensity of malaria transmission in the usambara mountains, tanzania. *J Med Entomol*, 40 (5):706–717, 2003.

[8] P. Borst, W. Bitter, R. McCulloch, F. Van Leeuwen, and G. Rudenko. Antigenic variation in malaria. *Cell*, 82(1):1–4, 1995.

[9] Caroline O Buckee, Peter C Bull, and Sunetra Gupta. Inferring malaria parasite population structure from serological networks. *Proc R Soc B*, 276(1656):477–485, 2009. doi: 10.1098/rspb.2008.1122.

[10] P. C. Bull, B. S. Lowe, M. Kortok, C. S. Molyneux, C. I. Newbold, and K. Marsh. Parasite antigens on the infected red cell surface are targets for naturally acquired immunity to malaria. *Nat Med*, 4(3):358–360, 1998.

[11] P. C. Bull, B. S. Lowe, M. Kortok, and K. Marsh. Antibody recognition of plasmodium falciparum erythrocyte surface antigens in kenya: evidence for rare and prevalent variants. *Infect Immun*, 67(2): 733–739, 1999.

[12] P.C. Bull, M. Kortok, O. Kai, F. Ndungu, A. Ross, B.S. Lowe, C.I. Newbold, and K. Marsh. Plasmodium falciparum-infected erythrocytes: agglutination by diverse Kenyan plasma is associated with severe disease and young host age. *J Infect Dis*, 182(1):252–9, 2000.

[13] Peter C Bull, Matthew Berriman, Sue Kyes, Michael A Quail, Neil Hall, Moses M Kortok, Kevin Marsh, and Chris I Newbold. Plasmodium falciparum variant surface antigen expression patterns during malaria. *PLoS Pathog*, 1(3):e26, 2005. doi: 10.1371/journal.ppat.0010026.

[14] Peter C Bull, Caroline O Buckee, Sue Kyes, Moses M Kortok, Vandana Thathy, Bernard Guyah, Jos A Stoute, Chris I Newbold, and Kevin Marsh. Plasmodium falciparum antigenic variation. mapping mosaic var gene sequences onto a network of shared, highly polymorphic sequence blocks. *Mol Microbiol*, 68(6):

1519–1534, 2008. doi: 10.1111/j.1365-2958.2008.06248.x.

[15] Q. Chen, V. Fernandez, A. Sundstrm, M. Schlichtherle, S. Datta, P. Hagblom, and M. Wahlgren. Developmental selection of var gene expression in plasmodium falciparum. *Nature*, 394(6691):392–395, 1998. doi: 10.1038/28660.

[16] Qijun Chen. The naturally acquired immunity in severe malaria and its implication for a pfemp-1 based vaccine. *Microbes Infect*, 9(6):777–783, 2007. doi: 10.1016/j.micinf.2007.02.009.

[17] Thanat Chookajorn, Ron Dzikowski, Matthias Frank, Felomena Li, Alisha Z Jiwani, Daniel L Hartl, and Kirk W Deitsch. Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci U S A*, 104(3):899–902, 2007. doi: 10.1073/pnas.0609084103.

[18] W. E. Collins and G. M. Jeffery. A retrospective examination of the patterns of recrudescence in patients infected with plasmodium falciparum. *Am J Trop Med Hyg*, 61(1 Suppl):44–48, 1999.

[19] W. E. Collins and G. M. Jeffery. A retrospective examination of sporozoite- and trophozoite-induced infections with plasmodium falciparum: development of parasitologic and clinical immunity during primary infection. *Am J Trop Med Hyg*, 61(1 Suppl):4–19, 1999.

[20] Madeleine Dahlbck, Thomas Lavstsen, Ali Salanti, Lars Hviid, David E Arnot, Thor G Theander, and Morten A Nielsen. Changes in var gene mrna levels during erythrocytic development in two phenotypically distinct plasmodium falciparum parasites. *Malar J*, 6:78, 2007. doi: 10.1186/1475-2875-6-78.

[21] Matthew W A Dixon, Joanne Thompson, Donald L Gardiner, and Katharine R Trenholme. Sex in plasmodium: a sign of commitment. *Trends Parasitol*, 24(4):168–175, 2008. doi: 10.1016/j.pt.2008.01.004.

[22] Rémy Durand, Frédéric Ariey, Sandrine Cojean, Arnaud Fontanet, Louise Ranaivo, Lanto-Alisoa Ranarivelo, Jeanne Aimée Vonimpaisomihanta, Didier Menard, Virginio Pietra, Jacques Le Bras, David Modiano, and Milijaona Randrianarivelojosia. Analysis of circulating populations of plasmodium falciparum in mild and severe malaria in two different epidemiological patterns in madagascar. *Trop Med Int Health*, 13(11):1392–1399, 2008. doi: 10.1111/j.1365-3156.2008.02156.x.

[23] Ron Dzikowski, Thomas J Templeton, and Kirk Deitsch. Variant antigen gene expression in malaria. *Cell Microbiol*, 8 (9):1371–1381, 2006. doi: 10.1111/j.1462-5822.2006.00760.x.

[24] Marcelo U Ferreira, Martine Zilversmit, and Gerhard Wunderlic. Origins and evolution of antigenic diversity in malaria parasites. *Curr Mol Med*, 7(6):588–602, 2007.

[25] Matthias Frank and Kirk Deitsch. Activation, silencing and mutually exclusive expression within the var gene family of plasmodium falciparum. *Int J Parasitol*, 36(9):975–985, 2006. doi: 10.1016/j.ijpara.2006.05.007.

[26] Matthias Frank, Laura Kirkman, Daniel Costantini, Sohini Sanyal, Catherine Lavazec, Thomas J Templeton, and Kirk W Deitsch. Frequent recombination events generate diversity within the multi-copy variant antigen gene families of plasmodium falciparum. *Int J Parasitol*, 38(10):1099–1109, 2008. doi: 10.1016/j.ijpara.2008.01.010.

[27] Malcolm J Gardner, Neil Hall, Eula Fung, Owen White, Matthew Berriman, Richard W Hyman, Jane M Carlton, Arnab Pain, Karen E Nelson, Sharen Bowman, Ian T Paulsen, Keith James, Jonathan A Eisen, Kim Rutherford, Steven L Salzberg, Alister Craig, Sue Kyes, Man-Suen Chan, Vishvanath Nene, Shamira J Shallom, Bernard Suh, Jeremy Peterson, Sam Angiuoli, Mihaela Pertea, Jonathan Allen, Jeremy Selengut, Daniel Haft, Michael W Mather, Akhil B Vaidya, David M A Martin, Alan H Fairlamb, Martin J Fraunholz, David S Roos, Stuart A Ralph, Geoffrey I McFadden, Leda M Cummings, G. Mani Subramanian, Chris Mungall, J. Craig

Venter, Daniel J Carucci, Stephen L Hoffman, Chris Newbold, Ronald W Davis, Claire M Fraser, and Bart Barrell. Genome sequence of the human malaria parasite plasmodium falciparum. *Nature*, 419(6906):498–511, 2002. doi: 10.1038/nature01097.

[28] H. A. Giha, T. Staalsoe, D. Dodoo, C. Roper, G. M. Satti, D. E. Arnot, L. Hviid, and T. G. Theander. Antibodies to variable plasmodium falciparum-infected erythrocyte surface antigens are associated with protection from novel malaria infections. *Immunol Lett*, 71(2): 117–126, 2000.

[29] S. Gupta and R. M. Anderson. Population structure of pathogens: the role of immune selection. *Parasitol Today*, 15(12): 497–501, 1999.

[30] S. Gupta, M. C. Maiden, I. M. Feavers, S. Nee, R. M. May, and R. M. Anderson. The maintenance of strain structure in populations of recombining infectious agents. *Nat Med*, 2(4):437–442, 1996.

[31] S. Gupta, R. W. Snow, C. A. Donnelly, K. Marsh, and C. Newbold. Immunity to non-cerebral severe malaria is acquired after one or two infections. *Nat Med*, 5 (3):340–343, 1999. doi: 10.1038/6560.

[32] M. Ho and N. J. White. Molecular mechanisms of cytoadherence in malaria. *Am J Physiol*, 276(6 Pt 1):C1231–C1242, 1999.

[33] Anja T R Jensen, Pamela Magistrado, Sarah Sharp, Louise Joergensen, Thomas Lavstsen, Antonella Chiucchiuini, Ali Salanti, Lasse S Vestergaard, John P Lusingu, Rob Hermsen, Robert Sauerwein, Jesper Christensen, Morten A Nielsen, Lars Hviid, Colin Sutherland, Trine Staalsoe, and Thor G Theander. Plasmodium falciparum associated with severe childhood malaria preferentially expresses pfemp1 encoded by group a var genes. *J Exp Med*, 199(9):1179–1190, 2004. doi: 10.1084/jem.20040274.

[34] Mirjam Kaestli, Ian A Cockburn, Alfred Corts, Kay Baea, J. Alexandra Rowe, and Hans-Peter Beck. Virulence of malaria is associated with differential expression of plasmodium falciparum var gene subgroups in a case-control study. *J Infect Dis*, 193(11):1567–1574, 2006. doi: 10.1086/503776.

[35] Karin Kirchgatter and Hernando A Del Portillo. Clinical and molecular aspects of severe malaria. *An Acad Bras Cienc*, 77(3):455–475, 2005. doi: /S0001-37652005000300008.

[36] Susan M Kraemer and Joseph D Smith. Evidence for the importance of genetic structuring to the structural and functional specialization of the plasmodium falciparum var gene family. *Mol Microbiol*, 50(5):1527–1538, 2003.

[37] Susan M Kraemer and Joseph D Smith. A family affair: var genes, pfemp1 binding, and malaria disease. *Curr Opin Microbiol*, 9(4):374–380, 2006. doi: 10.1016/j.mib.2006.06.006.

[38] Susan M Kraemer, Sue A Kyes, Gautam Aggarwal, Amy L Springer, Siri O Nelson, Zoe Christodoulou, Leia M Smith, Wendy Wang, Emily Levin, Christopher I Newbold, Peter J Myler, and Joseph D Smith. Patterns of gene recombination shape var gene repertoires in plasmodium falciparum: comparisons of geographically diverse isolates. *BMC Genomics*, 8: 45, 2007. doi: 10.1186/1471-2164-8-45.

[39] Sue A Kyes, Susan M Kraemer, and Joseph D Smith. Antigenic variation in plasmodium falciparum: gene organization and regulation of the var multigene family. *Eukaryot Cell*, 6(9):1511–1520, 2007. doi: 10.1128/EC.00173-07.

[40] Jean Langhorne, Francis M Ndungu, Anne-Marit Sponaas, and Kevin Marsh. Immunity to malaria: more questions than answers. *Nat Immunol*, 9(7):725–732, 2008. doi: 10.1038/ni.f.205.

[41] Thomas Lavstsen, Ali Salanti, Anja T R Jensen, David E Arnot, and Thor G Theander. Sub-grouping of plasmodium falciparum 3d7 var genes based on sequence analysis of coding and non-coding regions. *Malar J*, 2:27, 2003. doi: 10.1186/1475-2875-2-27.

[42] Thomas Lavstsen, Pamela Magistrado, Cornelus C Hermsen, Ali Salanti, Anja

T R Jensen, Robert Sauerwein, Lars Hviid, Thor G Theander, and Trine Staalsoe. Expression of plasmodium falciparum erythrocyte membrane protein 1 in experimentally infected humans. *Malar J*, 4(1):21, 2005. doi: 10.1186/1475-2875-4-21.

[43] Kevin Marsh and Robert W Snow. Host-parasite interaction and morbidity in malaria endemic areas. *Phil Trans R Soc Lond B Biol Sci*, 352(1359):1385–1394, 1997. doi: 10.1098/rstb.1997.0124.

[44] F. E. McKenzie, M. U. Ferreira, J. K. Baird, G. Snounou, and W. H. Bossert. Meiotic recombination, cross-reactivity, and persistence in plasmodium falciparum. *Evolution*, 55(7):1299–1307, 2001.

[45] L. H. Miller, M. F. Good, and G. Milon. Malaria pathogenesis. *Science*, 264(5167): 1878–1883, 1994.

[46] Louis H Miller, Dror I Baruch, Kevin Marsh, and Ogobara K Doumbo. The pathogenic basis of malaria. *Nature*, 415(6872):673–679, 2002. doi: 10.1038/415673a.

[47] Noa D Pasternak and Ron Dzikowski. Pfemp1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite plasmodium falciparum. *Int J Biochem Cell Biol*, 41(7):1463–1466, 2009. doi: 10.1016/j.biocel.2008.12.012.

[48] Mario Recker, Nimalan Arinaminpathy, and Caroline O Buckee. The effects of a partitioned var gene repertoire of plasmodium falciparum on antigenic diversity and the acquisition of clinical immunity. *Malar J*, 7:18, 2008. doi: 10.1186/1475-2875-7-18.

[49] Matthias Rottmann, Thomas Lavstsen, Joseph Paschal Mugasa, Mirjam Kaestli, Anja T R Jensen, Dania Mller, Thor Theander, and Hans-Peter Beck. Differential expression of var gene groups is associated with morbidity caused by plasmodium falciparum infection in tanzanian children. *Infect Immun*, 74(7):3904–3911, 2006. doi: 10.1128/IAI.02073-05.

[50] T. Smith, H. P. Beck, A. Kitua, S. Mwankusye, I. Felger, N. Fraser-Hurt, A. Irion, P. Alonso, T. Teuscher, and M. Tanner. Age dependence of the multiplicity of plasmodium falciparum infections and of other malariological indices in an area of high endemicity. *Trans R Soc Trop Med Hyg*, 93 Suppl 1:15–20, 1999.

[51] X. Z. Su, V. M. Heatwole, S. P. Wertheimer, F. Guinet, J. A. Herrfeldt, D. S. Peterson, J. A. Ravetch, and T. E. Wellems. The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of plasmodium falciparum-infected erythrocytes. *Cell*, 82(1):89–100, 1995.

[52] H. M. Taylor, S. A. Kyes, and C. I. Newbold. Var gene diversity in plasmodium falciparum is generated by frequent recombination events. *Mol Biochem Parasitol*, 110(2):391–397, 2000.

[53] Thomas J Templeton. The varieties of gene amplification, diversification and hypervariability in the human malaria parasite, plasmodium falciparum. *Mol Biochem Parasitol*, 166(2):109–116, 2009. doi: 10.1016/j.molbiopara.2009.04.003.

[54] Lasse S Vestergaard, John P Lusingu, Morten A Nielsen, Bruno P Mmbando, Daniel Dodoo, Bartholomew D Akanmori, Michael Alifrangis, Ib C Bygbjerg, Martha M Lemnge, Trine Staalsoe, Lars Hviid, and Thor G Theander. Differences in human antibody reactivity to plasmodium falciparum variant surface antigens are dependent on age and malaria transmission intensity in north-eastern tanzania. *Infect Immun*, 76(6): 2706–2714, 2008. doi: 10.1128/IAI.01401-06.

Supplementary Material

The model structure

The model is written in the C++ programming language. The main actors in our model are hosts, parasite strains and *var* genes, all individually labelled by a unique ID number.

The host population

At initiation of the simulations 2000 hosts are created, each with age zero, and without antibodies. Once every year (365 time steps) either the host's age is incremented by one or it dies. The number of hosts in the population remains constant. When a host dies it is replaced by an uninfected newborn host with a new ID number, age zero and no antibodies. To obtain a realistic host demography we have imposed age-dependent mortality probabilities from a demographic study in a high-transmission malaria area (see Figure 6) characterised by high infant mortality, for a large part caused by malaria [1, 2, 5].

The parasite population

The parasite population consists of a set of 50 parasite strains, where a strain is defined as a combination of *var* genes with a unique parasite strain ID number. Each strain is assumed to have 10 loci for *var* genes which at initialisation of a simulation are chosen at random from the *var* gene pool without allowing duplicate *var* genes within a strain.

The *var* genes

There exists a fixed set of 750 unique *var* genes in the model. The ID number of a *var* gene is denoted 'binding value' and assumed to be inversely correlated with the ability of the rbc expressing this *var* gene to cytoadhere. This avoids splenic clearance of the infected rbc, thereby determining the within-host growth rate of the infecting parasites. The lower its binding value, the better the cytoadherent properties of the *var* gene, i.e the 'stronger' the *var* gene.

Infection, immunity and transmission

Infection

In the model, hosts have a fixed default daily probability of becoming infected of 0.679, in agreement with infective biting rates measured in hyperendemic areas [7]. Because we are interested in the dynamics of the proteins that are expressed on the cell surface of infected red blood cells we focuss only on the asexual multiplication of the parasite that takes place in the general bloodstream.
Parasite proliferation in the earlier liver-stage of the parasite is taken into account

by assuming that a large number of new parasites enter the general bloodstream and infect rbc, each infected cell expressing only one of the 10 *var* genes at a time. We assume that initially every *var* gene is expressed with equal probability, i.e. in 10% of the infected rbc. This assumption is supported by a study of Lavstsen et al. [42] showing that first generation asexual parasites transcribe all *var*-genes at roughly similar levels.

Inside rbc, parasite multiplication takes approximately 48 hours, after which the cell ruptures and a new generation of parasites is released into the blood stream infecting new rbc [46, 24]. In the model, intracellular infection is updated once every two days by calculating the new proportions with which infected rbc express each *var* gene. This calculation is based on the assumptions that i.) epigenetic mechanisms within the parasites ensure that new generations of infected rbc remember the *var* gene that was expressed by their parental rbc [17]; and ii.) the adhesion properties of individual *var* genes determine the growth rate of the parasites expressing them on the surface of an rbc (as suggested in [11]). As a result within only a few generations of intra-cellular parasites the *var* gene with the best cytoadhering properties will be expressed on the majority of infected rbc and dominate the infection. Details on calculating the within-host dynamics are given below.

Within-host dynamics

The first generation of parasite strains, that enter the bloodstream from the liver, express the *var* genes they carry with an equal relative prevalence of 0.03 on the surface of the rbc they infect (see Table 5). To obtain the new relative prevalence at which the *var* genes of the parasite are expressed by the rbc it infects we first determine the factor, x, that normalises the weighted sum of the frequencies to one. Frequencies are weighted by the inverse of the ID of the *var* genes i.e., with the strength of the *var* genes as follows:

$$\sum_{l=1}^{n} \frac{1}{ID_l} \cdot p_l^{old} \cdot x = 1 \tag{1}$$

Here, l, are the *var* genes, n is the total number of *var* gene loci per strain and p^{old} is the current relative prevalence of the *var* genes. Then the new relative prevalence for each *var* gene is calculated by:

$$p_l^{new} = \frac{1}{ID_l} \cdot p_l^{old} \cdot x \tag{2}$$

Once every other day, for every new generation of parasites a new normalisation factor is determined and the new relative prevalences of the expressed *var* genes are calculated, an example with a strain of 4 *var* gene loci is shown in table 5.

Immunity

When the proportion of an rbc population expressing a specific *var* gene crosses a threshold value (0.3), the immune system of the infected host starts to build a specific antibody response against this expressed *var* gene. Once an antibody response is initiated it will continue to build up even if the proportion of cells expressing the *var* gene drops below the threshold again. Antibodies become effective 22 days after initiation of the specific response (as suggested in [18, 19]. The effect of the presence of specific antibodies is that the relative prevalence of cells expressing the *var* gene under consideration is permanently set to zero. Because of this, the infection will switch to being dominated by the next best *var* gene expressing rbc, eliciting a new immune response. This cycle results in periodical changes of dominant *var* gene expression, analogous to the antigenically distinct waves of parasitemia observed in patients [46]. When a host has built up immunity against all *var* genes that a particular parasite can express, that parasite is cleared from the host (Figure 5a).

Transmission

The mosquitoes that transmit the parasite between hosts are not explicitly included in the model. Instead we reduce the complexity of the model by letting the human hosts become infected by parasite strains proportionally to $1 + strain\,prevalence$ in the human host population. We thereby implicitly assume that the prevalence of the parasite strains in the host population is proportional to the prevalence of the parasite strains in the mosquito population, i.e. all hosts have the same probability of infecting a mosquito regardless of which parasite is infecting them. We also assume that all infecting parasites are equally transmissible, regardless of parasitemia. By adding one to the strain prevalence, parasites that have a prevalence of zero, still have a small probability of infecting a host. This is important because it allows new parasites that enter the parasite population through mutation or recombination (see below) to infect a host.

Co-infections

It is common for malaria patients to be co-infected by more than one parasite strain [50, 22]. In the model we also allow this to happen. Infection can occur daily, while the within-host generation cycle is updated only once every other day. If the newly infecting parasite is out of phase with the resident parasite, it is synchronised for simplicity. After co-infection, calculation of the new proportions of expressed *var* genes continues as if the host were infected by one parasite strain with twice the number of *var* genes.
When parasites are released into the blood stream after rupture of the infected red blood cells they are extremely vulnerable to attacks by the immune system. Thus sufficient infected rbc from the previous generation must avoid splenic clearance to ensure that enough parasites are released into the blood stream simultaneously, such that sufficiently many will survive to find and infect new rbc.

When multiple parasite strains co-infect a host, the dominant *var* gene expression of one of the strains might lower the proportion of rbc infected by other parasite strains so much that they are cleared from the host. We have set a minimal proportion of rbc that must express at least one of the *var* genes of a parasite for the parasite to survive into the next generation. If none of the *var* genes of a parasite strain are expressed above this threshold, the parasite strain is out-competed and cleared from the host (Figure 5b).

Recombination

Malaria parasites are known to undergo frequent meiotic recombination of their *var*-gene sequences [52, 3]. This recombination takes place during sexual reproduction of the parasite inside the mosquitoes. The mosquito life-stage of the parasites is not explicitly included in the model; instead, we set a fixed daily recombination event probability of 0.12. When a recombination event happens, the parasite strain infecting the least hosts (typically none) is set to go extinct and is replaced by a new strain with a new ID number. The *var* genes of the new parasite strain are selected from the set of all 750 *var* genes with a (within-host) prevalence-weighted probability of: $1 + var\,prevalence$, in much the same way as the selection of parasites for infection. The reasoning behind this is that recombination in mosquitoes is more likely to take place between *var* genes with a higher prevalence in the population, simply because more mosquitoes will be infected by them. The small probability for new *var* genes, with zero prevalence, to form part of the new parasite, reflects the introduction of a new *var* gene into the parasite population through mutation which happens at a lower rate than recombination [24].

Extended description of model results

When parasites carry equally *rare* VSA there is still no difference in parasite success for strains carrying a stronger or weaker VSA (not shown). This is because carrying a strong or a weak VSA each has its own small advantage. To illustrate, we consider two identical parasite strains, where one of the strains carries one stronger VSA than the other but of equal 'rarity'. Then, in a few hosts, the strain with the strong VSA will have the advantage of infecting a host longer because it is less likely to be out-competed by other strains. On the other hand, in a few hosts, the strain carrying the weaker VSA will be able to co-infect without being the dominant strain thus hiding from the immune system which will allow it to infect the same host for a second time using the same VSA.

From the model we learn that parasite strains balance between not infecting too many hosts to avoid depletion of susceptibles and at the same time infecting enough hosts to survive. This balance is very fragile, if a new parasite strain en-

ters the population through recombination, sharing one or a few of its VSA with a resident parasite strain it disturbs the infection balance of the resident strain. The greatest problem that parasite strains seem to face is that they have to maintain this infection balance in an ever changing (immune) environment while using the same VSA. Replacing their 'VSA-coat' through continuous recombination events is crucial for parasite survival since no combination of VSA in a strain was able to survive throughout the 150 years in any of the simulations.

Results from simulations with different parameter settings show that the survival of the strains is positively related to the number of loci per parasite, the number of hosts and the number of parasites strains. The survival of the strains decreases with increased numbers of VSA. At first, it seems counter-intuitive that parasite strains can survive longer when there are more competitors around and that they survive shorter when there are more VSA to choose from. The reason for this result has to do with the proportion of VSA from the VSA pool that are actually present in the parasite strain (and host) population. When the VSA-pool becomes larger without an increase in the number of different strains, it means that more VSA are not seen by the host population for longer stretches of time. When, by recombination, they enter the strain and host population again there is no immunity for these VSA and they can cause a sweep through the population. These sweeps disturb the balance (as described above) of other strains, causing them to go extinct. Therefore the more VSA from the VSA-pool that are represented in the parasite strain and host population, the longer strains tend to survive.

Table 3: Default parameter settings

parameter	value	tests	unit	description
Hosts	2000	5000	numbers	the number of hosts
Parasite strains	50	100	numbers	the number of parasite strains
Var genes	750	1500	numbers	the number of unique *var* genes
Var gene loci	10	15	per parasite	the number of different *var* gene loci
Recombination probability	0.12		per day	How often a new parasite strain arises in the population
Infection probability	0.679		per day	daily probability of a host to become infected [54]
Infection frequency	0.03			relative prevalence with which all *var* gene of a newly infecting parasite strains are expressed within a host upon co-infection
Survival frequency	0.05			minimal relative within host prevalence needed for parasite strain survival
Immune frequency	0.3			minimal relative within host expression prevalence needed to trigger immune response against a VSA
Building time	22		days	time taken for the immune system to build (enough) effective antibodies

Table 4: Details of the groups of hosts that are sampled at different time points during simulation.

Patient group	Selection year(s)	Number of patients	Average age (years)
1	150	119	4.1
2	153-154	184	4.0
3	180-182	242	3.6
4	210	138	3.9
5	212	108	3.9
6	240	112	3.8
7	242	94	3.6

Table 5: Example of the within-host relative prevalence of *var*-expression by rbc infected with a strain that carries 4 different *var* genes, binding values of each gene are in boldface.

		2	**15**	**20**	**25**
day	X	prevalences			
0	50.76	0.03	0.03	0.03	0.03
2	2.54	0.76	0.10	0.08	0.06
4	2.06	0.97	0.02	0.01	<0.01
6	2.00	>0.99	<0.01	<0.01	<0.01

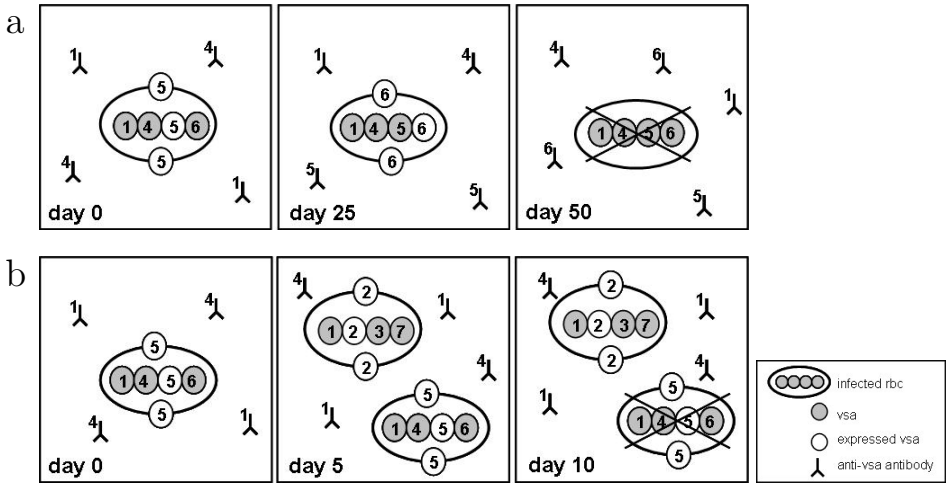


Figure 5: a) A parasite infecting an rbc, expresses its strongest *var* gene for which the host has no antibodies (day 0). The host has built up antibodies against the expressed *var* gene, the parasite switches to expressing its next strongest *var* gene (day 25). The host has antibodies against all of the *var* genes of the parasite, it is cleared from the host (day 50). b) A parasite infecting an rbc, expresses its strongest *var* gene for which the host has no antibodies (day 0). The host is co-infected with another parasite that also expresses its strongest *var* gene for which the host has no antibodies (day 5). The expressed *var* gene of one of the parasites is so much stronger than the *var* gene expressed by the other parasite that it out-competes the other parasite which is cleared from the host (day 20).

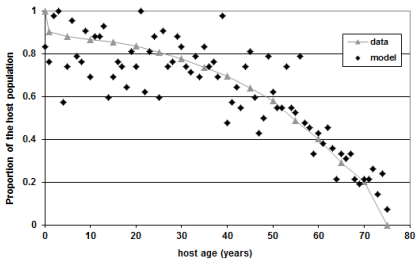


Figure 6: Imposed host demography: Age profile of host population at year 15 of simulation, line is produced using mortality data from a high malaria transmission area [5].

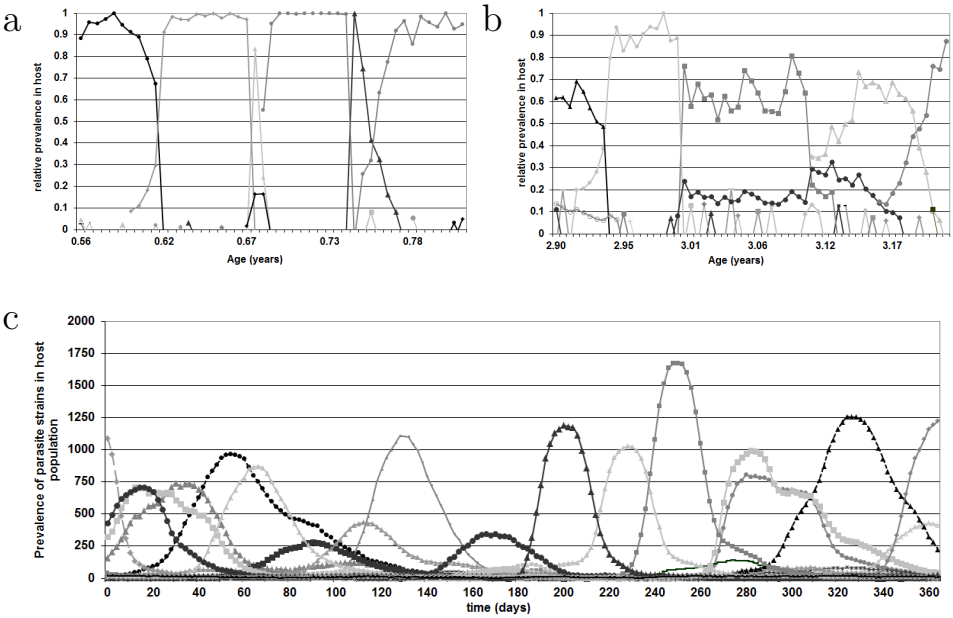


Figure 7: Parasite dynamics, within and between-hosts: Representative sample of within-host parasite dynamics in a young host ≈ 9 months (a) and in an older host ≈ 3 years (b), showing the relative prevalence of co-infecting parasite strains. Each line denotes a different parasite strain. c) Representative sample of between-host dynamics showing the prevalence of circulating parasite strains in the host population, each line denotes a different parasite strain.

CHAPTER 6

How selection forces dictate the variant surface antigens used by malaria parasites.

M. Severins, D. Klinkenberg, and J.A.P Heesterbeek

Published in:
Journal of the Royal Society Interface (2011) **9**(67):246-60

Abstract

Red blood cells infected by the malaria parasite *Plasmodium falciparum*, express variant surface antigens (VSA) that evade host immunity and allow the parasites to persist in the human population. There exist many different VSA and the differential expression of these VSA is associated with the virulence (damage to the host) of the parasites. The aim of this study is to unravel the differences in the effect key selection forces have on parasites expressing different VSA such that we can better understand how VSA enable the parasites to adapt to changes in their environment (like control measures) and how this may impact the virulence of the circulating parasites.

To this end we have built an individual-based model that captures the main selective forces on malaria parasites namely, parasite competition, host immunity, host death and mosquito abundance at both the within and between-host level. VSA are defined by the net growth rates they infer to the parasites and the model keeps track of the expression of, and antibody build-up against, each VSA in all hosts. Our results show an ordered acquisition of VSA-specific antibodies with host age which causes a dichotomy between the more virulent VSA that reach high parasitemias but are restricted to young relatively un-immune hosts, and less virulent VSA that do not reach such high parasitemias but can infect a wider range of hosts. The out-come of a change in the parasite's environment in terms of parasite virulence depends on the exact balance between the selection forces which sets the limiting factor for parasite survival. Parasites will evolve towards expressing more virulent VSA when the limiting factor for parasite survival is the within-host parasite growth and the parasites are able to minimise this limitation by expressing more virulent VSA.

1 Introduction

Malaria parasites are subject to a diverse range of selection pressures. Key selection pressures, as reviewed in Mackinnon & Marsh 2010 [46], are parasite competition, host immunity, host death, mosquito abundance and human interventions (such as the use of antimalarial drugs). Malaria parasites have a complex life cycle, with a variety of life stages within two different hosts, that allows them to adapt and survive when changes occur in these selection pressures. It is due to this adaptability that malaria still causes the death of over one million children each year despite continuous efforts of prevention, control and eradication [46]. One trait that parasites can adapt that is of key importance to public health is the parasites' virulence i.e., the harm caused to the hosts. Theoretical studies have shown that some seemingly sensible measures for control, for example vaccines that lower the growth rate of the parasites, can cause the parasites to increase their virulence to such extent that host mortality even increases [32]. Understanding how parasites adapt to changes in their environment is therefore important to develop safe and effective control measures [49].

Parasite virulence is considered to be an unavoidable consequence of host exploitation required for parasite reproduction [3, 52]. Parasites cannot exploit their hosts limitlessly in order to increase their reproduction and transmission because increased virulence also reduces the host life expectancy and consequently can reduce parasite transmission. Classical virulence theory predicts that, due to this trade-off between costs and benefits of virulence, parasites evolve toward intermediate levels of virulence [3, 52, 27, 62]. Virulence evolution is influenced by several host-pathogen interactions that need to be taken into account when predicting how a pathogen's virulence will be adapted in response to changes in their environment. Examples of these are infections that allow for multiple- or superinfection [31, 33, 8], infections where immunity to the pathogen is built up gradually [48, 2], and infections where there exists a reciprocal interaction between the within-host and between-host dynamics [54]. Understanding and predicting virulence adaptation in malaria is difficult because the parasite's life cycle is complex and all of these host-pathogen interactions apply to malaria infections.

A substantial part of the parasites adaptability and survival in humans arises from the parasite's so-called variant surface antigens (VSA). The best-studied VSA in malaria, and the ones we will focus on in this paper, are *Plasmodium falciparum* erythrocyte membrane proteins 1 (PfEMP1's)[34, 9, 22]. These VSA are expressed on the surface of infected red blood cells (rbc) and the immune system builds effective antibody responses against them [14, 59]. In addition to being strong antigens, these VSA have cythoadhesive properties and depending on which VSA is expressed, infected rbc can adhere to different host tissues obstructing local blood flow, which is an important virulence determinant of infection [51, 9, 58, 61]. For example, VSA have been associated with various life-threatening clinical manifestations of the disease such as cerebral malaria, pregnancy malaria, and the

formation of so-called rbc 'rosettes' [56, 58, 66, 38, 61].

Each parasite carries approximately 60 genes coding for different VSA [6, 69] of which only one is expressed at a time [17, 25]. When an antibody response against a particular VSA has grown strong, the small number of parasites that express a different VSA have a benefit allowing growth of parasites expressing another VSA. The ongoing VSA changes evade immune recognition and allow for persistent and frequent (re)infections. In endemic areas with high exposure to infection, individuals gradually build up a repertoire of antibodies against a large set of these VSA [59, 42, 70]. In concurrence to the build-up of antibodies against PfEMP1 VSA, individuals become resistant, first to severe malaria, then to mild malaria and eventually to all clinical malaria [18, 50]. The number of different VSA in the entire parasite population is unknown, but presumably very large [28] which is why infections remain common even at old age. Due to these VSA that potentially form an important link between the parasites' virulence and host immunity, understanding virulence adaptation for malaria parasites is a major challenge.

To increase our understanding of virulence adaptation in *P. falciparum* infections we have developed an individual-based computational model that includes the key selective forces on malaria parasites at both the within and between-host level and explicitly takes feedback between these levels into account. The model keeps track of parasitemia, VSA expression and immunity within all individual hosts of the population. We make no other assumption on the differences between the VSA other than that parasites expressing different VSA have different net growth rates. This assumption is based on the argument that parasites expressing VSA with stronger cythoadhesive power are better at avoiding clearance by the spleen [16, 35]. Under this assumption we find that the model yields realistic infection dynamics and reproduces key features of the epidemiological characteristics of *P. falciparum* malaria.

Which VSA a parasite expresses determines how parasites perceive the pressure of selection on them. For example, the pressure of immune selection on a parasite expressing a VSA for which the host has formed antibodies will be different than on a parasites expressing a VSA for which the host has no antibodies. Also, the pressure of selection through competition between a parasite expressing a very virulent VSA and a parasite expressing a very mild VSA will be different. The first of two aims of this study is to unravel these differences in the effect key selection forces have on parasites expressing different VSA such that we can better understand how VSA enable *P. falciparum* parasites to adapt to changes in their environment. During a real-life infection parasites switch between VSA which makes it very difficult to gauge the relative effect of each of the parasite's VSA on for example, its within-host infection duration, transmission success and its survival in general. Because parasites express only one VSA at a time we can, as a first approximation for this problem, limit the parasites to carrying only one VSA

while still keeping the environment in which the parasites experience competition and interact as diverse as in a scenario where parasites are able to switch VSA. This allows us to assess the effect on a parasite of expressing one specific VSA relative to the expression of the other VSA.

The second aim of this study is to understand how alterations in the parasites' living conditions may impact the virulence of the circulating parasites. The implementation of measures to prevent and control malaria such as drugs, vaccination and extensive use of bed nets, impose a strong selection pressure on the parasites. Due to the wide variety of VSA and existing genetic recombination mechanisms that generate both new VSA and parasites with new sets of VSA, it is likely that virulence adaptation will affect the population-wide repertoire of VSA available to individual parasites. To understand the effect of different alterations in the parasite's environment on their virulence and host mortality we run, under varying parameter and model conditions, a wide range of simulations with each parasite having a repertoire of five VSA from a much larger set of possible VSA.

2 Model and methods

The model functions through keeping track of the individual parasitemia of every host specifying how many rbc are infected by each infecting parasite and which VSA these rbc express. For every host, the model also keeps track of the host's antibody repertoire against specific VSA. At the between-host level the model keeps track of the prevalence and transmission success of every individual parasite and in addition, the degree of host immunity against every parasite at host population level. To accomplish this the model is equipped with actors, processes, variables, parameters, and iterators. The actors in our model, are VSA, parasites and hosts. The actors carry variables, the most important of these being parasite prevalence, parasitemia, host antibodies and parasite within-host growth rate, which are under the influence of regularly executed model processes. The model processes are transmission, parasite within-host growth and competition, malaria caused host mortality, specific host immune response, parasite extinction and recombination, and host demography. The model parameters are used to set the context of the model environment and remain constant during simulation. Iterators are used to identify or denote a specific instance of an actor. An overview of the model structure is given in Figure 1.

The numbers of actors remain constant during simulation. The size of the host population is 1000. A parasite consist of a set of five different VSA chosen at random from a pool of 1500 unique VSA. We introduce the term parasite 'type' to denote parasites with a specific subset of VSA. These parasite types bear no relation to possible 'real-life' strains. We allow 75 different parasite types to circulate in the model at a time.

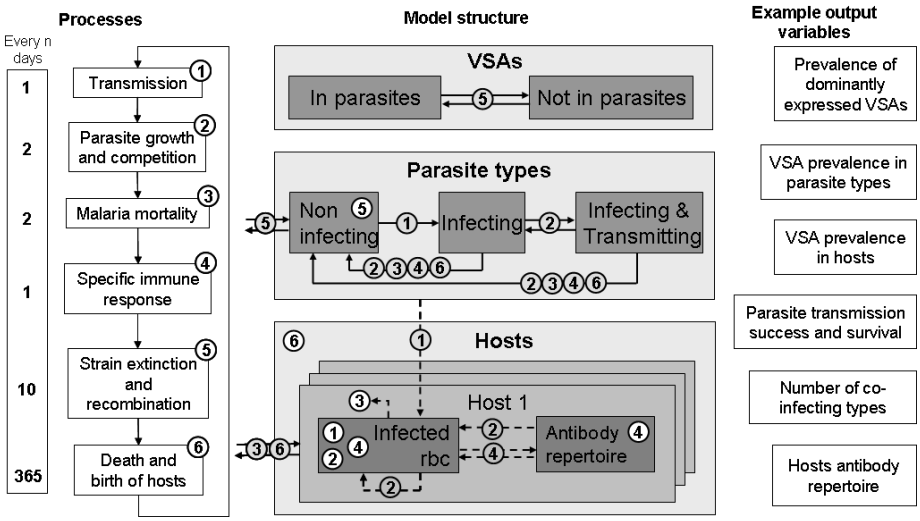


Figure 1: Model Processes: The main model processes are shown inside the numbered boxes. Within each time step (day) the processes are executed in numerical order although not every process is executed each day. On the left hand side is given how often every process takes place. Model structure: Model actors are visualised with boxes. Numbered white process circles indicate on which actors the process directly acts, shaded process circles are linked to arrows. Arrows indicate either actor displacement by a process (solid) or how a process is influenced by actors or their variables (dashed).

The VSA in the pool are ranked and their rank reflects how strong their binding is to the receptors lining the host's blood vessels. We assume that this difference in binding strength causes differences in net growth rate because stronger binding VSA are better at avoiding splenic clearance of the infected rbc on which they are expressed [35]. We therefore refer to high ranking VSA as 'strong' and lower ranking VSA as 'weak'.

We have coded the model in the C++ programming language. Model simulations run with time steps t, that are equivalent to days in real life. During a time step the processes are executed in a fixed order although not every process is executed each day, as shown on the left-hand side of Figure 1. Below we describe the model processes in detail, using lower case notation for parameters and capitals for the variables. We use iterator i to denote the ith host in the host population, iterator n to denote the nth type in the parasite type population, and iterator v to denote the vth VSA in the VSA pool which is equal to its rank or 'strength'. We use iterator j to denote the jth parasite type infecting a particular host and iterator k to denote the expression of the kth VSA of an infecting parasite type.

2.1 Between-host infection and transmission

In real life, a parasite can transmit to mosquitoes when a proportion of the infecting parasites differentiates into gametocytes, the life stage of the parasite which is taken up by the mosquitoes [23]. Although the exact relation between parasitemia and gametocyte abundance is unclear [53] there is a robust relationship between parasitemia and transmission [65, 24].

The model's transmission process is executed every day and involves two processes, first, host to mosquito transmission and second, mosquito to host transmission. During the first process a pool of parasites is created that represents the parasites that have been taken up by mosquitoes. During the second step, hosts are infected with a parasite from the pool and at the end of this process the pool is deleted.

Every parasite infecting a host has a probability to be added to the pool. $S_{ij}(t)$ is the probability of the jth parasite infecting host i to be included in the pool on day t. To take into account the time it takes parasites to form gametocytes we impose a delay between infection and admission to the pool, regulated by parameter d. Before d days of infection, $S_{ij}(t) = 0$. After d days of infection the probability of being added to the pool depends on the parasitemia of the parasite. The parasitemia of a parasite in a host is calculated by summing all the infected rbc expressing one of the VSA of the parasite $\sum_k P_{ijk}(t)$ (see section Within-host growth for details). We also introduce parameter p_0, between zero and one, that reflects how much gametocyte formation depends on parasitemia. When, $p_0 = 0$ (default), transmission is fully dependent on parasitemia (implying that a fixed

proportion of parasites differentiates into gametocytes), and when $p_0 = 1$, transmission is fully independent of parasitemia (implying that there are always enough parasites that differentiate into gametocytes to infect a mosquito regardless of parasitemia). The probability to be added to the pool for each infecting parasite that has formed gametocytes thus becomes: $S_{ij}(t) = min(1, \sum_k P_{ijk}(t) + p_0)$. At the end of the first process, a pool has been created of parasites that are chosen through a random process but where weighed by their relative prevalence in the parasite population, i.e. they proportionally represent the parasite types in the host population. We assume that the mosquito population is large enough such that stochastic variation in the proportion of types transmitted can be neglected.

During the second process, each host is infected with probability l, the so-called transmission intensity or force of infection. The infecting parasite is randomly selected from all the parasites in the pool. To allow new parasite types (see section Parasite recombination) to enter the host population, there exists a very small probability (set at 0.0001) that a host is not infected with one of the parasites from the pool but with one of the 75 parasite types at random.

In some malaria endemic areas the number of mosquitoes drops substantially during the yearly dry season. In the simulations where this seasonality is taken into account we introduce a lower force of infection, l_s, during a yearly recurring season of 122 days.

2.2 Within-host parasite growth

During the blood stage of malaria the parasites infect red blood cells (rbc) in repeated two-day cycles [57, 28]. The parasites multiply inside the rbc and at the end of the cycle the infected rbc rupture, releasing a new generation of parasites into the blood stream [57]. In the model, the parasite growth and competition process explicitly simulates the within-host dynamics of the infected rbc and VSA expressed on their surface. Liver-stage parasite proliferation is implicitly taken into account by letting parasites enter the general bloodstream in the order of magnitude in which they are released from the liver which is approximately 10,000 parasites (0.002 per microlitre blood)[71]. We assume that rbc homeostasis ensures that for every new generation of free parasites there is a fixed number, r, of rbc available for infection per microlitre blood. We set r, regardless of infection and host age, to 100,000 rbc which, is approximately 20% of rbc in a microlitre of blood. This 20% lies in the order of magnitude of the observed maximum parasitemia per microlitre blood in patients [21, 71]. In the model parasitemia is defined as the infected fraction of the r available rbc.

Each parasite infecting an rbc expresses only one of its VSA and epigenetic mechanisms within the parasites ensure that new generations of parasites, originating from that infected rbc, remember the VSA that was expressed by their parental

parasite [19]. Supported by findings of Lavstsen et al. 2005 [44] we assume that the first generation of new parasites in a host express each one of the VSA they carry on an equal fraction of the rbc they infect.

Let $P_{ijk}(t)$ be the fraction of r rbc that is infected and expressing VSA k, of parasite type j inside host i at time t. This fraction is updated every other day. The parasitemia caused by parasite type j, is defined as the sum of the fractions corresponding to all the VSA of that type, $\sum_k P_{ijk}(t)$. The total host parasitemia is defined as the sum of the parasitemias of all parasite types infecting host i: $P_{tot,i}(t) = \sum_{j,k} P_{ijk}(t)$.

Each $P_{ijk}(t)$ infected rbc fraction has its own per generation multiplication, or growth, factor $G_{ijk}(t)$. The growth factor sets the average net number of free parasites (merozoites) released in host i per ruptured rbc expressing VSA k, at the end of an infection cycle. The core of the growth factor is defined as: $G_{ijk}(t) = c + \frac{c_v}{v_k}$. It is composed from: c, an intrinsic, non-VSA associated, growth component; c_v, a VSA-associated growth component; and v_k which is how the kth VSA of parasite j ranks in the overall VSA pool (see start of section 2). The VSA rank influences the growth factor in such a way that higher ranking VSA allow for relatively fast growth ('strong' VSA) and lower ranking VSA allow for relatively slower growth ('weak' VSA). In addition, the growth factor is influenced by the host's non-specific immune response, regulated by parameter m_n, and cross-immunity through the hosts VSA-specific antibody repertoire ($A_i(t)$), regulated by parameter m_c, explained in detail in section Host immunity. The resulting growth factor is:

$$G_{ijk}(t) = c + \frac{c_v}{v_k + A_i(t)m_c}(1 - P_{tot,i}(t)m_n).$$

In each host i, the total number of free parasites (merozoites) per microlitre blood at the end of an infection cycle is obtained by multiplying the fraction of rbc infected with parasite j expressing VSA k, with the growth factor of that fraction, then adding these products for all VSA of all infecting types, and finally multiplying by the total number of rbc r: $r \sum_{j,k} P_{ijk}(t)G_{ijk}(t)$. These free parasites compete with each other for infection of the available rbc in the new infection cycle. The mean number of free parasites, $F_i(t)$, that compete with each other for one rbc is given by: $F_i(t) = \sum_{j,k} P_{ijk}(t)G_{ijk}(t)$. We assume that the number of parasites each rbc encounters is Poisson-distributed with mean $F_i(t)$. That means that a proportion $1 - e^{-F_i(t)}$ of the r rbc will be infected in the next cycle. The fraction of rbc in host i infected by the host's jth infecting parasite expressing its kth VSA in the new infection cycle then becomes:

$$P_{ijk}(t+2) = \frac{P_{ijk}(t)G_{ijk}(t)}{F_i(t)}\left(1 - e^{-F_i(t)}\right)$$

A very small fraction ($1x10^{-4}$ %) of the most dominantly expressed VSA of an infecting parasite is allowed to switch its VSA expression evenly to the other VSA of the parasite. We assume that each infecting parasite type j, infecting a fraction smaller then 0.013 of the r rbc available for infection per microlitre blood (i.e. if $\sum_k P_{ijk}(t+2) < 0.013$), does not survive and is cleared from the host.

2.3 Host immunity

The model implements VSA-specific host antibodies, cross-immunity between these antibodies and non-specific, or general, immunity. Maternal immunity which is thought to protect infants in the first few months of life is ignored in our model [64, 36] (see Discussion).

Specific host immunity of host i against VSA v is denoted by $A_{iv}(t) \in \{0, 1\}$. If $A_{iv}(t) = 0$, the host is 'susceptible' for that VSA whereas if $A_{iv} = 1$, the host is immune. New-born hosts are assumed to be susceptible to all VSA: $A_{iv}(0) = 0$ for all v. The total size of the antibody repertoire of host i at time t is denoted by $A_i(t) = \sum_v A_{iv}(t)$. An infected host i is triggered to start to build a specific antibody response against the kth VSA of one of its infecting parasite types j from the moment that $\frac{P_{ijk}(t)}{\sum_{j,k} P_{ijk}(t)} > m_a$. This relative antibody threshold value, m_a, ensures that hosts only mount specific immunity to dominantly expressed VSA. It takes time to build an immune response and therefore antibodies become effective ($A_{iv}(t) = 1$) m_d days after the specific response is triggered (default $m_d = 18$ days which is within the range suggested by Collins & Jeffery 1999 [20, 21]). We assume that antibody binding to a VSA clears all the infected rbc expressing that VSA and as a result the fraction of available rbc infected and expressing this VSA is set to zero from that moment. When a host has built up effective antibodies against all the VSA of a parasite type, the type can no longer survive in the host and is cleared. Because we assume life-long antibody immunity, only parasites that carry at least one VSA for which the host has no antibodies can infect the host.

We implement cross-immunity between VSA antibodies [26, 43] by letting the growth factors of the different VSA-expressing rbc fractions, $G_{ijk}(t)$, be affected by the total size of the host's antibody repertoire, $A_i(t)$. In the simulations that are used to study the effect of cross-immunity we use a parameter, m_c, to regulate the degree in which the host's antibody repertoire influences the core of the parasite's net growth factor: $G_{ijk}(t) = c + \frac{c_v}{v_k + A_i(t)\, m_c}$. When m_c is zero there is

no cross-immunity and with increasing values of m_c the effect of cross-immunity becomes larger.

Non-specific immunity, or general general immunity, is assumed to effect the net growth of the parasites proportional to the host's total parasitemia. This reduction in net growth is implemented like: $G_{ijk}(t) = c + \frac{c_v}{v_k + A_i(t)\, m_c}(1 - P_{tot,i}(t) m_n)$. Here parameter m_n (default 0.7) sets the strength of general immunity. The smaller m_n is, the smaller the strength of general immunity. When $m_n = 0$, there is no general immunity.

2.4 Malaria mortality and host demography

In the model there is malaria-induced host mortality (executed after every parasite growth process) and non-malaria related mortality executed yearly. The probability of dying from malaria depends on the parasitemia in the host and is scaled by a factor f_m. We set factor f_m to 0.0001 such that malaria mortality is slightly higher than in the demographic study by Bawah & Binka 2007 [7] to simulate a host population without any medical intervention (see Figure 6 in ESM). Thus the daily probability of dying from malaria of a host i is: $P_{tot,i}(t) f_m$. The probability of dying of other causes than malaria is age-dependent, and adopted from the same demographic study in a high-transmission malaria area that specifies mortality rates due to malaria and non-malaria related causes [7]. When a host dies it is replaced by an uninfected newborn host, with age zero and no antibodies.

2.5 Parasite recombination

Parasite types can end up not infecting any hosts due to host immunity, parasite competition or, when at low prevalence already, due to the model's stochasticity. Every 10 days all parasite types that are not infecting any hosts are removed i.e., we assume they go extinct. These types are replaced by new parasites types which we assume arise from parasite recombination. In real-life malaria parasites are known to undergo frequent recombination of their VSA set [68, 5, 30]. Most recombination takes place during sexual reproduction of the parasite inside the mosquitoes. The new parasite types in our model are assigned a set of 5VSA drawn at random from the pool of VSA. This means not only that parasite types with new combinations of VSA enter the host population but also that VSA's which where not represented in one of the parasite types previously, and thus unseen by the hosts, can enter the host population.

2.6 Model simulations

Default simulation settings

Simulations last the equivalent of 200 years. Host are initially assigned a random age between zero and 75, and initially have no antibodies. The first 100

Table 1: Default model parameter values and the variations in these parameters simulated.

parameter	default value	variations			
r	100,000				
l	0.7	0.1	0.2	0.3	0.9
l_s	-	0.5	0.1	0.01	0
p_0	0	0.5	0.25	0.1	0.01
d	10	0	5	15	25
f_m	0.0001	$5\text{x}10^{-5}$	$1.25\text{x}10^{-4}$	$1.5\text{x}10^{-4}$	$2.5\text{x}10^{-4}$
c	1	0.1	0.25	0.5	0.75
c_v	500	100	250	750	1500
m_c	1	0	0.1	0.5	1.5
m_a	0.25	0.1	0.2	0.5	0.6
m_d	18	7	14	21	28
m_n	0.7	0	0.1	0.5	0.9

years of each simulation are regarded as a transient period for which no output is recorded to allow for the host demography and antibody repertoires to stabilise.

The default parameter values for the simulations is given in Table 1. In the default simulations there is no seasonality. The force of infection, l, is derived from infective mosquito biting rates measured in high-transmission areas [10]. For the first aim of the study we perform simulations where the parasites carry only one VSA per parasite (1VSA simulation) and for the second aim of the study we perform simulations where the parasites carry five VSA per parasite (5VSA simulations).

Model behaviour

To determine if the model gives realistic infection dynamics we study a 5VSA simulation with default parameter settings. In this simulation we look at the daily parasitemia of the hosts of which one typical host is plotted. We sample all 1000 hosts at yearly intervals to study the average parasitemia, average number of co-infecting parasite types, average strength of the expressed VSA and, the average size of the hosts VSA-specific antibody repertoires at all host ages.

First aim simulations

As part of the first aim of our study we perform several within-host experiments to study the selection forces of parasite competition and host immunity within-host for parasites expressing different VSA. In these experiments we look at one isolated host of which we can manually set the size and repertoire of its VSA-specific antibodies and also all other immune functions. Under pre-defined static

immune conditions we allow one or two parasites, expressing a VSA of our choice to which the host has no antibodies, to grow inside the host.

For the other part of the first aim of our study we look at several life-history traits of the parasite types from a 1VSA simulation with default parameter settings. One of those traits is the yearly number of transmission events of a parasite type. A transmission event is registered when a parasite type is selected from the parasite pool, that is created during the transmission process, to infect a host. A successful transmission event is registered when a transmission event results in infection of the host which happens when the parasite type selected for infection carries at least one VSA to which the host has no antibodies. The life-history trait 'host immunity' of a parasite type is defined as the average number of hosts that have VSA antibodies against one of its VSA during the existence of the type. We distinguish between two ways an infecting parasite can be cleared from the host namely, clearance through the build-up of specific VSA antibodies ('immune clearance') or clearance because parasitemia drops below the threshold level ('competition clearance'). The latter mostly occurs through competition with co-infecting parasites or to a lesser extend though cross-immunity. Another life-history trait we define is the fraction of all clearances of a parasite type that was due to competition clearance.

Second aim simulations

For the second aim of our study we perform many 5VSA simulations in which we change one parameter value from the default value. After each simulation we evaluate the average yearly host mortality and compare these results. We also evaluate the average rank of the VSA dominantly expressed in infants (hosts under age one) for every simulation as a proxy for the virulence of the parasites that infect the most vulnerable hosts. The parameter values used in the simulations are shown in Table 1.

3 Results & Conclusions

3.1 Model behaviour

To determine if the model gives realistic infection dynamics we study the hosts' infection burden and immunity with age. Figure 2 shows results of a 5VSA simulation with default parameter values. This same figure for a 1VSA simulation is show in the supplementary material (Figure 7). The parasitemia ($P_{tot,i}(t)$) during the lifetime of a typical host, depicted in Figure 2a, shows that parasitemia is highest in early childhood after which parasite levels drop in between occasional peaks. Even in elderly hosts infections still frequently occur albeit at much lower levels of parasitemia than in younger hosts (see Figure 2b). Because transmission is related to parasitemia, it is the youngest hosts that are the most infectious.

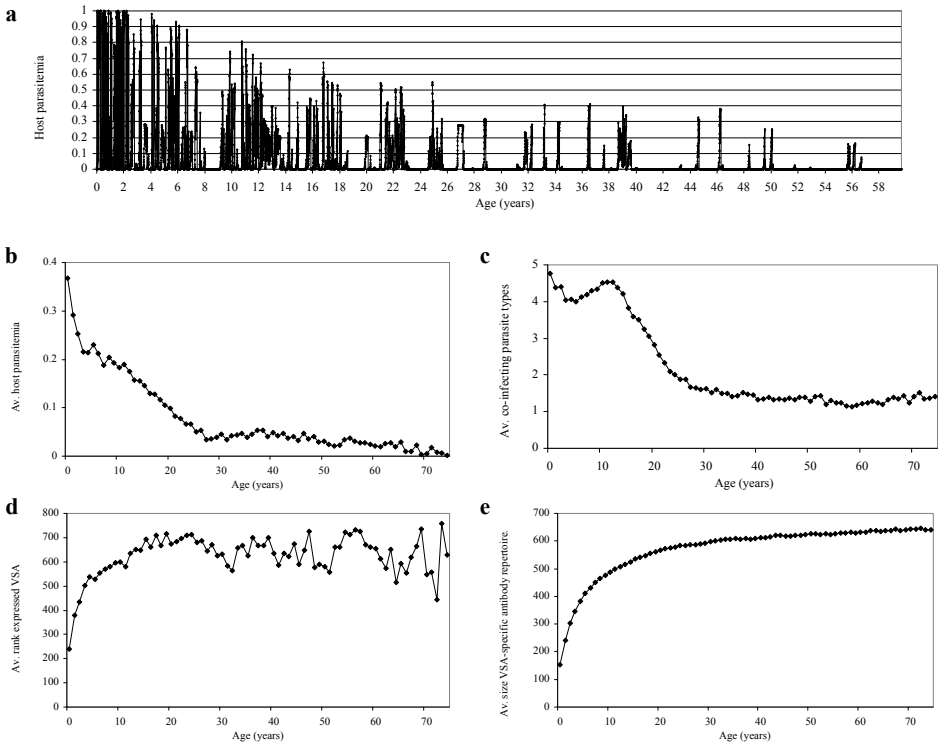


Figure 2: Results from a 5VSA simulation under default model and parameter conditions. Panel a shows the life-time parasitemia of a typical host. (Please note that a parasitemia of 1 means only 20% of rbc are infected). Panels b-d show averages by host age from yearly samples taken of all hosts during 60 years of the simulation. Panel b shows the average host parasitemia. Panel c shows the average number of co-infecting parasite types. Panel d shows the average rank of the dominantly expressed VSA. Panel e shows the average size of the hosts VSA-specific antibody repertoires.

Co-infections with multiple parasite types are common in the simulations and vary on average between one and five co-infecting types (see Figure 2c) although infections with over 10 types do occur. Despite the presence of multiple parasite types in one host, we find that infection is mostly dominated by only one VSA, especially in younger hosts. Switches in VSA dominance occur after VSA antibody build-up or through parasite competition. Because the hosts only build up specific immunity against VSA that cross a relative infection threshold, the time it takes from infection to the build up of specific antibodies varies depending on the competition between infecting parasite types and the infection history of the host.

The hosts' VSA antibody repertoire gradually increases with age as shown in Figure 2e. We find that, due to the differences in growth rate inferred by the VSA expressed by the parasites, a hierarchy in VSA expression with host age arises naturally through the infection dynamics (see Figure 2d). Younger hosts are infected with parasites expressing stronger VSA and as the hosts grow older and their VSA antibody repertoire increases they are infected with parasites expressing weaker VSA as hypothesised by Hviid 2010 [35].

3.2 First aim simulations

Within-host competition and immunity

As part of the first aim of our study i.e., understanding the effect of key selection forces on parasites expressing different VSA, we show the results of several within-host experiments (Figure 3) as described in the Model and Methods section 2.6. Figure 3a shows that, without competition, the rate at which parasitemia increases and the maximum parasitemia a parasite can reach is determined by the VSA expressed by the parasite and the amount of cross-immunity in the host ($A_i(t)$). Figure 3b shows that when two parasite types compete the parasite expressing the strongest VSA will dominate the infection and suppress the parasite expressing the weaker VSA. In the absence of specific immunity against the stronger VSA the parasite expressing this stronger VSA will eventually out-compete the parasite with a weaker VSA.

Under some circumstances parasites expressing a weaker VSA do better in the within-host competition than parasites expressing a stronger VSA. Figure 3c shows that parasites entering a host later can have a benefit over resident parasites, even if the parasite that enters the host later expresses a weaker VSA. This occurs when the parasite that infects later grows slow enough to stay below the threshold for the build up of specific immunity (i.e., it 'hides' from the immune system) without being pushed under the clearance threshold by the resident parasite. Figure 3d shows that parasites expressing weaker VSA can be better at hiding from specific immunity than parasites expressing stronger VSA.

Gametocyte formation takes 10 days and the time it takes to build up antibod-

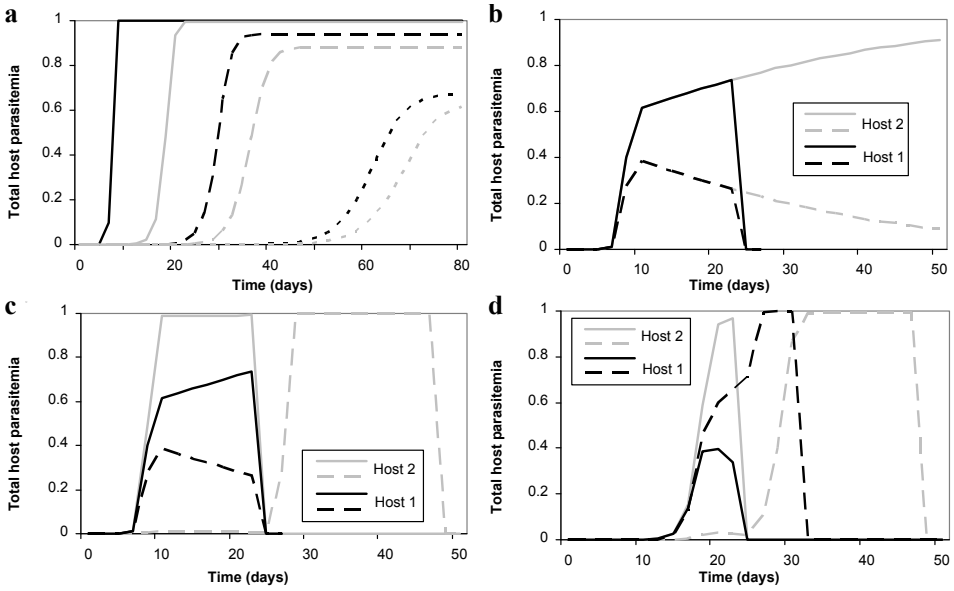


Figure 3: Results from within-host experiments. Panel a compares parasite growth in three naive hosts ($A_i(t) = 0$, black lines) and three hosts with an antibody repertoire of size 100 ($A_i(t) = 100$, grey lines) infected at day=0 with a parasite expressing a VSA with rank 5 (solid lines), 250 (dashed lines) or 750 (dotted lines). The hosts are not allowed to build any immune response during simulations. Panels b-d compare parasite growth in naive hosts that are infected with two different parasite types. In panel b both hosts are infected at day=0 with one parasite type expressing a VSA with rank 10 (solid lines) and the other VSA with rank 11 (dashed lines). Host 1 is allowed to build up VSA-specific antibodies during the simulation (black lines), host 2 is not (grey lines). In panel c the hosts are infected with a parasite type expressing a VSA with rank 10 (solid lines) and one expressing a VSA with rank 11 (dashed lines). In host 1 both parasite types start infecting the host at day=0 (black lines), in host 2 the parasite expressing the VSA with rank 11 does not super-infect until day=2 (grey lines). In panel d both hosts are infected at day=0 with a parasite type expressing a VSA with rank 100 (solid lines). At day=2, host 1 is super-infected with a parasite type expressing a VSA with rank 75 (black dashed line), host 2 is super-infected with a parasite strain expressing a VSA with rank 125 (gray dashed line).

ies from the moment the immune threshold is crossed is 18 days. Therefore any parasite that is not cleared by competition can transmit for at least 8 days. The results from the within-host experiments shows that parasites expressing strong VSA can reach high parasitemia quickly, hereby maximising transmission during the time before specific immunity clears infection. During this period they can also out-compete parasites expressing weaker VSA. Parasites expressing weaker VSA, in co-infections, are better at hiding from specific immunity which allows them to stretch the period in which they can transmit long beyond 8 days albeit at lower parasitemia and thus smaller transmission probability per day.

To further understand the effect of the key selection forces on parasites expressing different VSA we also run a full 1VSA simulation with default parameter settings and study several life-history traits of the parasites (Figure 4). The supplementary material has the same figure (Figure 8) for a 5VSA simulation which shows how difficult it is to gauge the effect on a parasite of carrying a specific VSA when the parasites carry 5 VSA because the results are diluted and distorted by the other VSA of the parasite.

Mosquito abundance

Mosquito abundance is an important selective force that acts on the boundary of the within- and between-host level. It is most important to parasites that have a low frequency of transmission because when the number of mosquitoes drops, the total number of transmission events drops and parasites that have a lower frequency of transmission will be more likely to go extinct. Although the model does not take the mosquito population explicitly into account, it does keep track of the number of transmission events for each parasite type. Figure 4a shows that parasites carrying weaker VSA experience much less transmission events than parasites carrying a stronger VSA meaning that the parasites that carry weaker VSA are more vulnerable to drops in mosquito abundance.

Competition and immunity

Confirming what we would expect from the within-host experiments, the parasites that carry strong VSA are mostly cleared from hosts by specific immunity while the parasites that carry weak VSA are more often cleared by competition (Figure 4d). We also find that at the host population level there are much more hosts that have specific immunity against stronger VSA than against weaker VSA (Figure 4c). Consequently the fraction of transmission events that result in infection i.e., a successful transmission event, is very low for parasites that carry a strong VSA (Figure 4b).

Panel 4e shows the resulting average lifetime of parasite types carrying a specific VSA. This result does not imply that there is an 'optimal' VSA for the parasites. The benefit to a parasite of expressing a certain VSA is always relative to the other

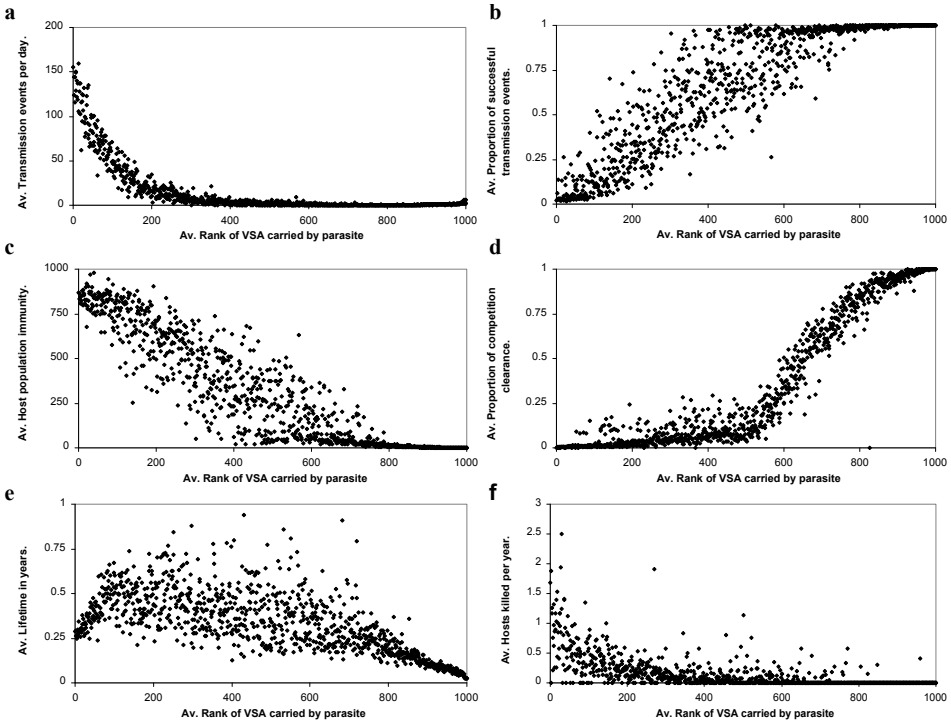


Figure 4: Results from a 1VSA simulation under default parameter conditions. Panels show averages from all parasite types that existed during the simulation (∼18,000) and carried the specified VSA. Panel a shows the daily number of transmission events i.e., how many times a day the parasite is selected to infect a host. Panel b shows the proportion of successful transmission i.e., transmission events to hosts without protecting VSA-specific immunity. Panel c shows the number of hosts with specific antibodies against the VSA carried. Panel d shows the proportion of clearances from the host due to competition (opposed to immunity clearance). Panel e shows the lifetime of the parasite types in years. Panel f shows the yearly number of hosts dying from malaria while infected by the parasite type.

VSA expressed in the host and to the VSA-specific antibodies that exist in the host and the host population at the time of expression. These other VSA expressed and the anti-VSA antibodies in the population are continually fluctuating, as such, an optimal VSA does not really exist.

Host mortality

Host death is a selective force that is experienced by the parasite types both at the within- and between-host level. Within-host, host death stops infection and transmission of all infecting parasite types; at the between-host level it removes a host from the population that was susceptible for some VSA, and immune for others. Parasites expressing strong VSA can reach the highest parasitemias, mainly in children, and therefore cause more host death, than weaker VSA (see Figure 4f). Loss of transmission due to host death is hence more important for parasites expressing a stronger VSA.

Conclusions first aim

In conclusion, the differences in the effect key selection forces have on parasites expressing different VSA result in differences in the contribution the VSA have to the survival of the parasites carrying them as follows:

Strong VSA allow for high parasitemia in the infected host which means that per day of infection the parasite can infect more mosquitoes. Strong VSA also contribute to the survival of the parasite by suppressing or even clearing other co-infecting parasites. A high proportion of the host population has specific immunity against strong VSA therefore strong VSA hardly contribute to the number of hosts the parasite can infect. Because they allow for high parasitemia strong VSA contribute to loss of infection and transmission of the parasite due to host death.

Weak VSA contribute little to transmission per day but in some cases allow for a longer infection time because they do not trigger a specific immune response so quickly. Their largest contribution to parasite survival is in terms of the number of hosts the parasite can infect since the proportion of hosts with specific immunity against the weaker VSA is very small. Weak VSA hardly contribute to loss of infection and transmission due to host death.

3.3 Second aim simulations

The second aim of this study is to understand how alterations in the parasites' living conditions may impact the virulence of the circulating parasites. For this purpose we run a wide range of 5VSA simulations with different parameter settings and look at both host mortality and the rank i.e., strength of the VSA expressed in infants (Figure 5). The strength of expressed VSA is used as a proxy for parasite virulence.

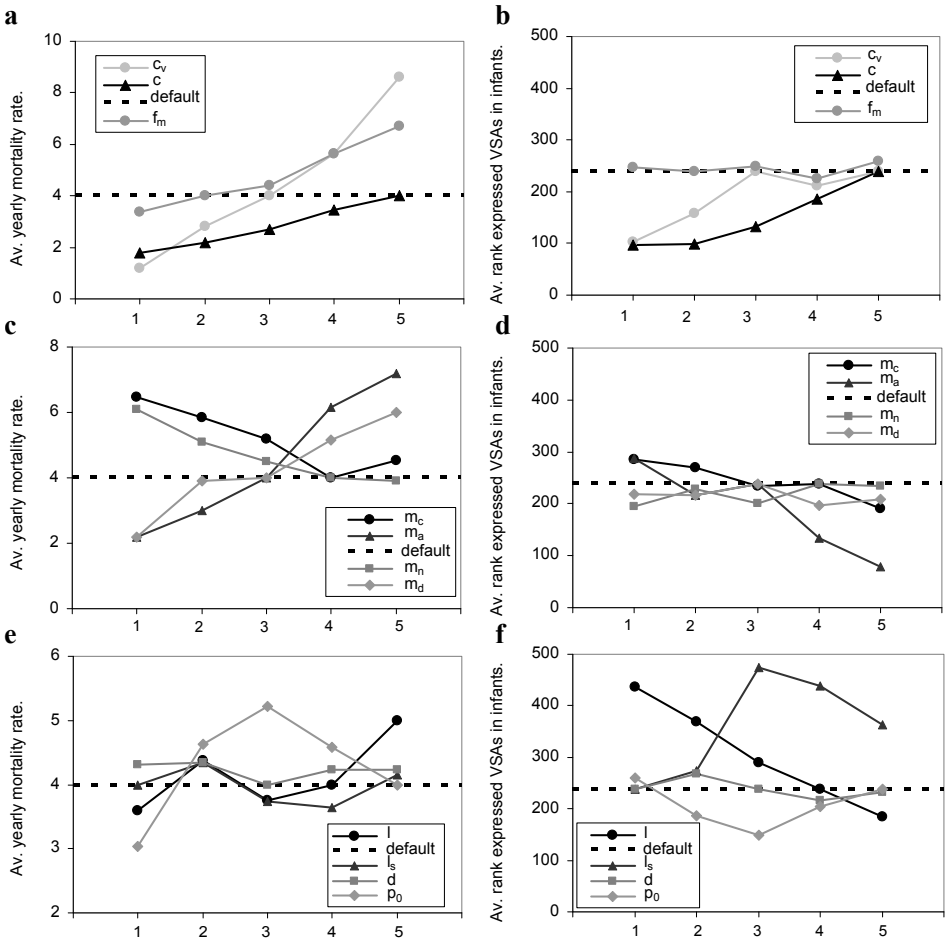


Figure 5: Results from 5VSA simulations. Effect of within- and between-host parameter variations on the yearly host mortality rate and the average dominantly expressed VSA in infants (< 1 year). Horizontal axes indicate increase in the effects of one of the variations, values are as follows (*=default value): VSA-related growth factor, c_v (1: 100, 2: 250, 3: 500*, 4: 750, 5: 1500), intrinsic growth factor, c (1: 0.1, 2: 0.25, 3: 0.5, 4: 0.75, 5: 1*), malaria mortality probability, f_m (1: $5\text{x}10^{-5}$, 2: $1\text{x}10^{-4}$*, 3: $1.25\text{x}10^{-4}$, 4: $1.5\text{x}10^{-4}$, 5: $2.5\text{x}10^{-4}$), cross-immunity, m_c (1: 0, 2: 0.1, 3: 0.5, 4: 1*, 5: 1.5), relative antibody threshold, m_a (1: 0.1, 2: 0.2, 3: 0.4*, 4: 0.5, 5: 0.6), general immunity, m_n (1: 0, 2: 0.1, 3: 0.5, 4: 0.7*, 5: 0.9), antibody building time, m_d (1: 7, 2: 14, 3: 18*, 4: 21, 5: 28), force of infection, l (1: 0.1, 2: 0.2, 3: 0.3, 4: 0.7*, 5: 0.9), seasonality, l_s (1: 0.7*, 2: 0.5, 3: 0.1, 4: 0.01, 5: 0), gametocyte formation time, d (1: 0, 2: 5, 3: 10*, 4: 15, 5: 25), parasitemia dependant transmission and, p_0 (1: 0.5, 2: 0.25, 3: 0.1, 4: 0.01, 5: 0*)

Parasite growth & host mortality

Decreasing parasite growth, wether by the intrinsic component, c, or the VSA specific component, c_v, decreases host mortality (Figure 5a) but increases the strength of the VSA expressed in infants (Figure 5b) because the parasites compensate for the loss in growth potential. Increasing the probability of malaria-induced host death increases host mortality yet has a minimal effect on the VSA expressed in infants. That the parasite population does not move towards expressing weaker VSA even though parasites expressing strong VSA are most affected by host mortality might be because these parasites are instantly compensated when a deceased host is replaced by a non-immune new host which are mostly infected by parasites expressing strong VSA.

Host immunity

A general observation from model simulations with different immune parameter values, is that any change that improves the performance of the host's immune system, lowers host mortality (Figure 5c). Even though an increase in general immunity lowers the growth potential of the parasites, they do not move towards expressing stronger VSA as we found when we decreased the growth rate of the parasites (Figure 5d). We believe this is because expressing stronger VSA cannot compensate for the loss of growth since it is the parasites expressing the stronger VSA that are most suppressed by general immunity due to its relation with host parasitemia. Changes in the time it takes to build up antibodies do not affect the strength of the VSA expressed in infants (Figure 5d). This could be because these changes are equally beneficial or detrimental to parasites expressing any VSA. With increasing cross-immunity the parasites move towards expressing stronger VSA. An increase in cross-immunity mostly effects parasites expressing weaker VSA because they are expressed in older hosts with a larger VSA antibody repertoire and hence more cross-immunity.

Transmission

Decreasing the transmission intensity, whether all year round, l, or during a yearly recurring season, l_s, decreases the strength of the VSA expressed in infants (Figure 5f). Thus at lower transmission intensity it becomes more beneficial for the parasites to express a VSA to which very few hosts have specific immunity than to express a VSA that can reach higher parasitemia. This means that the limiting factor to parasite survival is finding enough hosts without specific immunity to infect, rather than the within-host growth rate of the parasites or within-host competition. Notably, only when the transmission intensity is decreased all year round does it decrease host mortality (Figure 5e).

Changes in the time it takes to form gametocytes had no effect on either host mortality or the VSA expressed in infants (Figure 5e, Figure 5f) which suggests

that the gametocyte formation time is not the limiting factor to parasite transmission and survival. A small increase in the probability of parasite transmission independent of parasitemia, increases host mortality and the strength of the expressed VSA. This is likely to be because of an increase in transmission and we saw the same effect when the transmission intensity was increased. When the parasitemia independent transmission is increased even further, host mortality decreases again and so does the strength of the expressed VSA. We argue that this is because now all infecting parasites transmit with almost equal probability and thus there is much less need to reach high parasitemia and express strong VSA.

Conclusions second aim

The second aim of this study is to understand when and how alterations in the parasites' living conditions may impact the VSA, and thus the virulence, of the circulating parasites, in particular those expressed in infants. When the parasites' environment doesn't change the interplay of the selective forces on the parasites will reach an equilibrium at which there is an average VSA strength expressed in infants that is most optimal for parasite survival. As we learned from our first aim simulations the different VSA contribute differently to the parasites' survival. Changes in virulence due to parameter changes are an indication to what the limiting factors are to parasite survival. Parasites will evolve towards expressing stronger VSA when (one of) the limiting factor(s) for parasite survival is the within-host parasite growth and the parasites are able to minimise this limitation by expressing stronger VSA.

4 Discussion

In this study we have built an individual-based computational model to better understand how VSA enable *P. falciparum* parasites to adapt to changes in their environment and how this adaptability can affect the success of malaria prevention and control measures. Malaria prevention and control measures are aimed to reduce host morbidity and mortality. From the parasites' perspective, these measures are yet another force of selection that interacts with the existing selective forces. Because these forces are connected across the within and between-host infection level, to fully understand how these selective forces interact, it is crucial to model these forces at both levels explicitly. Our model does this and as a result it reproduces key features of the epidemiological characteristics of *P. falciparum* infections and VSA-specific immunity.

Our results show how parasites expressing different VSA perceive the pressure of the selection forces differently at both the within and between host level. We also show the effect of a wide range of changes to the parasites living conditions on host mortality and parasite virulence in infants. From the simulations we learn

that the outcome of a change in the parasite's environment in terms of parasite virulence depends on the balance between the selection forces which sets the limiting factor for parasite survival.

Due to limitations in existing data, computational restrictions and to maintain transparency, our model is a simplification of reality and not designed to mimic the exact balance between the forces of selection. In addition, this balance in reality is likely to differ between communities due to regional differences. Therefore, while our results can be used as an indication to how the virulence of parasites may change in response to certain measures they should not be interpreted as accurate predictions.

In our 5VSA default simulations the limiting factor for parasite survival is finding a non-immune host therefore, when the transmission intensity is decreased the virulence of the parasites decreases. However when the limiting factor to survival is e.g. the within-host competition between parasites, the virulence of the parasites would increase when the transmission intensity is decreased.

There are a few parameter changes for which we can compare our results with findings from other theoretical or experimental studies. That a decrease in growth rate of the parasites leads to an increase in virulence has also been found by Gandon et al. 2001 [32] and Mackinnon et al. 2008 [47]. The virulence of circulating parasites in regions with different transmission intensities (i.e., the number of infective bites/day) has been studied and shows that with decreasing transmission intensity the proportion of severe malaria cases that are caused by malarial anaemia decrease while the proportion of cases due to cerebral malaria increase [67, 60]. Because in our model we do not specify the location in the body where the VSA preferentially adhere these findings are not easily comparable. However a decrease in malarial anaemia could indicate that there are less strong VSA to cause ongoing high parasitemia.

Below we discuss some of the simplifications and assumptions that where made in order to create the model and the implications these may have on our results.

4.1 Growth rate differences related to VSA expression

The expression of VSA give infected rbc the capacity to sequester in body tissues and reduce their passage through the spleen where infected rbc are recognised and cleared from the blood. Following Hviid 2010 [35], we assume that the strength of the binding between a VSA and its host receptor is a measure for how well infected rbc can avoid splenic clearance and consequently we envisage different VSA as conferring different net growth rates to parasites.

Classical work on VSA in African *Trypanosomas* parasites [39] shows that the observed differences in growth rates alone, between parasites expressing different

VSA, cannot account for the ordered appearance of dominant VSA during *Trypanosoma* infection as reported in the literature. This work is often referred to in the context of malaria VSA models. However, *Trypanosomas* differ from malaria parasites in many ways. For our analysis, notably, the observed differences in growth rate between *Trypanosomas* parasites expressing different VSA are very small compared to those of *P. falciparum* parasites [4]. In addition, *Trypanosomas* VSA do not have adhesion properties like the VSA of *P. falciparum* do.

In our model the differences in growth rate allow for dominant VSA expression and regular changes in this dominance due to the build-up of antibodies against the dominant VSA or competition with a newly infecting parasite. Because every host has a different history of infection, every host also has a unique order in which it builds up antibodies against the VSA. However, on the host population level we find that on average there is a generic order in dominant VSA expression and host age, where antibodies against the stronger VSA are acquired earlier in life. A recent study indeed shows that there appears to be such a generic order in which antibodies against VSA are acquired [16]. The transmission intensity does not influence this order in antibody acquisition but merely dictates its pace.

It is, however, likely that dominant VSA expression and the antigenic variation as seen in malaria infections is not dictated by differences in growth rate alone, especially in elderly hosts where the differences in growth rate between the VSA is much smaller. Various other mechanisms have been suggested in the literature [1, 29, 45]. A model by Recker et al. 2004 [63] has shown that e.g. transient cross-reactive immune responses are able to orchestrate realistic within-host antigenic variation. However this mechanism on its own does not generate a generic order of antibody acquisition at host population level which is what makes parasites expressing different VSA experience selection pressures such as competition and immunity differently. For this we believe that a feedback mechanism through the population as incorporated in our model is required.

4.2 Disease and Immunity

Malaria mortality in the model depends only on host parasitemia while in reality disease and disease outcome is more complex. For example, the location of infected rbc sequestration and the formations of rosettes by infected rbc are important factors in disease outcome [38]. Consequently, VSA that have disease exacerbating properties will cause more host mortality than other VSA that infer the same growth rate but do not have these properties.

In our model we do not take into account maternal antibodies which are thought to protect infants the first few months of life. Because of these antibodies, from the parasites' perspective, infants behave like adults for a few months after which they become non-immune again until they build up their own antibodies. Therefore, apart from a somewhat reduced infant death (which would than happen at

a slightly older age) the overall dynamics will largely stay the same.

It is not clear whether immunity against malaria, in particular VSA specific antibodies, is life-long. Here we assume that VSA-specific immunity is effectively life-long due to frequent re-infections that cause sufficient antibody boosting. If future studies find that specific VSA antibodies can not be assumed to effectively provide life-long protection this can be implemented in the model. In that case also older hosts could be infected by parasites expressing strong VSA which would change the balance of the selection forces as we present here considerably. We have assumed that the gametocytes of a parasite are cleared when the parasite is cleared from a host, however there could be a delay in gametocyte clearance as they carry different VSA and are relatively long-lived [11]. Depending on the extend of the delay this may have a significant impact on transmission.

4.3 Peripheral blood compartment

We have made a few assumptions regarding the rbc in the model. Firstly, the blood volume of the hosts in our model is equal for hosts of all ages, whereas in reality blood volume increases with age until adulthood. Secondly, all parasite types in our model are equal except for their VSA, while in reality some parasite types specialise in infecting rbc of specific ages [55]. Thirdly, in the model every within-host infection cycle, a fixed amount of new uninfected rbc becomes available for infection. In reality, at very high parasitemia new rbc production is not able to keep up the formation of new rbc, hence patients become anaemic. We argue that in all these cases parasite types have fewer rbc to infect than in the model which would imply that the within-host competition in reality is likely to be stronger than in our model.

4.4 Antigenically distinct groups of VSA

From serological and genomic data there is evidence that the immense VSA pool is organised into specific groups of VSA [13, 40, 41, 15, 12]. It is unclear how the VSA from the different groups are represented in individual parasite genomes but it seems that all parasites carry at least a few members from each group [41]. Antibodies against VSA from the so-called group A VSA seem to be more cross-reactive than antibodies against VSA from other groups [37]. Also, group A VSA have been associated with more severe disease and appear to be mostly expressed at young age [37, 66].

In our model, we only assume one difference between the VSA, namely the difference in growth rate inferred by different VSA, to fully understand the consequence of that assumption. We expect that the effect on our results, of adding antigenically distinct groups of VSA to the model, depends on the order in which antibodies against VSA from each of the groups is acquired by the hosts. In our model all VSA experience equal inhibition from cross-immunity which, on aver-

age, is also the case if antibodies against VSA from all groups are acquired at equal pace. However, if e.g. antibodies against group A VSA are acquired earlier in life, on average, VSA from this group will experience more inhibition from cross-immunity than VSA from other groups. We predict that this will enhance the dichotomy we find in our model between strong VSA that allow for high parasitemia only in young relatively un-immune hosts, and weaker VSA that allow for infection of a wider range of hosts but lower parasitemia. However, implementation of such grouping into the model is necessary to fully understand the effect of this extra level of organisation on the dynamics of the VSA.

4.5 Concluding

The first aim of this study was to unravel the differences in the effect key selection forces have on parasites expressing different VSA such that we can better understand how VSA enable *P. falciparum* parasites to adapt to changes in their environment. We find distinct differences in the perception of the key selection forces by parasites expressing strong VSA and weak VSA. With this background we continued to the second aim of our study which was to understand how alterations in the parasites' environment may impact the virulence of the circulating parasites. We find that an increase in virulence only occurs when (one of) the limiting factor(s) for parasite survival is the within-host parasite growth and a change to the environment strengthens this limiting factor. To be able to make predictions on whether implementation of control measures in a community will increase the virulence of the circulating parasites a thorough understanding is needed of the interactions and balance between the selective forces on the parasites in this community.

References

[1] Z. Agur, D. Abiri, and L. H. Van der Ploeg. Ordered appearance of antigenic variants of african trypanosomes explained in a mathematical model based on a stochastic switch process and immune-selection against putative switch intermediates. *Proc Natl Acad Sci U S A*, 86(23): 9626–9630, 1989.

[2] Samuel Alizon and Minus van Baalen. Acute or chronic? within-host models with immune dynamics, infection outcome, and parasite evolution. *Am Nat*, 172(6):E244–E256, 2008. doi: 10.1086/592404.

[3] R. M. Anderson and R. M. May. Coevolution of hosts and parasites. *Parasitology*, 85 (Pt 2):411–426, 1982.

[4] N. Aslam and C. M. Turner. The relationship of variable antigen expression and population growth rates in trypanosoma brucei. *Parasitol Res*, 78(8):661–664, 1992.

[5] Alyssa E Barry, Aleksandra Leliwa-Sytek, Livingston Tavul, Heather Imrie, Florence Migot-Nabias, Stuart M Brown, Gilean A V McVean, and Karen P Day. Population genomics of the immune evasion (var) genes of plasmodium falciparum. *PLoS Pathog*, 3(3):e34, 2007. doi: 10.1371/journal.ppat.0030034.

[6] D. I. Baruch, B. L. Pasloske, H. B. Singh, X. Bi, X. C. Ma, M. Feldman, T. F. Taraschi, and R. J. Howard. Cloning the p. falciparum gene encoding pfemp1, a malarial variant antigen and adherence receptor on the surface of parasitized human erythrocytes. *Cell*, 82(1):77–87, 1995.

[7] Ayaga A Bawah and Fred N Binka. How many years of life could be saved if malaria were eliminated from a hyper-endemic area of northern ghana? *Am J Trop Med Hyg*, 77(6 Suppl):145–152, 2007.

[8] Andrew S Bell, Jacobus C de Roode, Derek Sim, and Andrew F Read. Within-host competition in genetically diverse malaria infections: parasite virulence and competitive success. *Evolution*, 60(7): 1358–1371, 2006.

[9] B. A. Biggs, R. F. Anders, H. E. Dillon, K. M. Davern, M. Martin, C. Petersen, and G. V. Brown. Adherence of infected erythrocytes to venular endothelium selects for antigenic variants of plasmodium falciparum. *J Immunol*, 149(6):2047–2054, 1992.

[10] R. Bødker, J. Akida, D. Shayo, W. Kisinza, H. A. Msangeni, E. M. Pedersen, and S. W. Lindsay. Relationship between altitude and intensity of malaria transmission in the usambara mountains, tanzania. *J Med Entomol*, 40 (5):706–717, 2003.

[11] Teun Bousema, Lucy Okell, Seif Shekalaghe, Jamie T Griffin, Sabah Omar, Patrick Sawa, Colin Sutherland, Robert Sauerwein, Azra C Ghani, and Chris Drakeley. Revisiting the circulation time of plasmodium falciparum gametocytes: molecular detection methods to estimate the duration of gametocyte carriage and the effect of gametocytocidal drugs. *Malar J*, 9:136, 2010. doi: 10.1186/1475-2875-9-136.

[12] Caroline O Buckee, Peter C Bull, and Sunetra Gupta. Inferring malaria parasite population structure from serological networks. *Proc R Soc B*, 276(1656):477–485, 2009. doi: 10.1098/rspb.2008.1122.

[13] P. C. Bull, B. S. Lowe, M. Kortok, and K. Marsh. Antibody recognition of plasmodium falciparum erythrocyte surface antigens in kenya: evidence for rare and prevalent variants. *Infect Immun*, 67(2): 733–739, 1999.

[14] Peter C Bull and Kevin Marsh. The role of antibodies to plasmodium falciparum-infected-erythrocyte surface antigens in naturally acquired immunity to malaria. *Trends Microbiol*, 10(2):55–58, 2002.

[15] Peter C Bull, Caroline O Buckee, Sue Kyes, Moses M Kortok, Vandana Thathy, Bernard Guyah, Jos A Stoute, Chris I Newbold, and Kevin Marsh. Plasmodium falciparum antigenic variation. mapping mosaic var gene sequences onto a network of shared, highly polymorphic sequence blocks. *Mol Microbiol*, 68(6):

1519–1534, 2008. doi: 10.1111/j.1365-2958.2008.06248.x.

[16] Gerald K K Cham, Louise Turner, John Lusingu, Lasse Vestergaard, Bruno P Mmbando, Jonathan D Kurtis, Anja T R Jensen, Ali Salanti, Thomas Lavstsen, and Thor G Theander. Sequential, ordered acquisition of antibodies to plasmodium falciparum erythrocyte membrane protein 1 domains. *J Immunol*, 183 (5):3356–3363, 2009. doi: 10.4049/jimmunol.0901331.

[17] Q. Chen, V. Fernandez, A. Sundstrm, M. Schlichtherle, S. Datta, P. Hagblom, and M. Wahlgren. Developmental selection of var gene expression in plasmodium falciparum. *Nature*, 394(6691):392–395, 1998. doi: 10.1038/28660.

[18] Qijun Chen. The naturally acquired immunity in severe malaria and its implication for a pfemp-1 based vaccine. *Microbes Infect*, 9(6):777–783, 2007. doi: 10.1016/j.micinf.2007.02.009.

[19] Thanat Chookajorn, Ron Dzikowski, Matthias Frank, Felomena Li, Alisha Z Jiwani, Daniel L Hartl, and Kirk W Deitsch. Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci U S A*, 104(3):899–902, 2007. doi: 10.1073/pnas.0609084103.

[20] W. E. Collins and G. M. Jeffery. A retrospective examination of the patterns of recrudescence in patients infected with plasmodium falciparum. *Am J Trop Med Hyg*, 61(1 Suppl):44–48, 1999.

[21] W. E. Collins and G. M. Jeffery. A retrospective examination of sporozoite- and trophozoite-induced infections with plasmodium falciparum: development of parasitologic and clinical immunity during primary infection. *Am J Trop Med Hyg*, 61(1 Suppl):4–19, 1999.

[22] A. Craig and A. Scherf. Molecules on the surface of the plasmodium falciparum infected erythrocyte and their role in malaria pathogenesis and immune evasion. *Mol Biochem Parasitol*, 115(2):129–143, 2001.

[23] Matthew W A Dixon, Joanne Thompson, Donald L Gardiner, and Katharine R Trenholme. Sex in plasmodium: a sign of commitment. *Trends Parasitol*, 24(4):168–175, 2008. doi: 10.1016/j.pt.2008.01.004.

[24] Chris Drakeley, Colin Sutherland, J. Teun Bousema, Robert W Sauerwein, and Geoffrey A T Targett. The epidemiology of plasmodium falciparum gametocytes: weapons of mass dispersion. *Trends Parasitol*, 22(9):424–430, 2006. doi: 10.1016/j.pt.2006.07.001.

[25] Ron Dzikowski, Felomena Li, Borko Amulic, Andrew Eisberg, Matthias Frank, Suchit Patel, Thomas E Wellems, and Kirk W Deitsch. Mechanisms underlying mutually exclusive expression of virulence genes by malaria parasites. *EMBO Rep*, 2007. doi: 10.1038/sj.embor.7401063.

[26] Salenna R Elliott, Paul D Payne, Michael F Duffy, Timothy J Byrne, Wai-Hong Tham, Stephen J Rogerson, Graham V Brown, and Damon P Eisen. Antibody recognition of heterologous variant surface antigens after a single plasmodium falciparum infection in previously naive adults. *Am J Trop Med Hyg*, 76 (5):860–864, 2007.

[27] Paul W. Ewald. Host-parasite relations, vectors, and the evolution of disease severity. *Annual Review of Ecology and Systematics*, 14:pp. 465–485, 1983. ISSN 00664162.

[28] Marcelo U Ferreira, Martine Zilversmit, and Gerhard Wunderlic. Origins and evolution of antigenic diversity in malaria parasites. *Curr Mol Med*, 7(6):588–602, 2007.

[29] Matthias Frank and Kirk Deitsch. Activation, silencing and mutually exclusive expression within the var gene family of plasmodium falciparum. *Int J Parasitol*, 36(9):975–985, 2006. doi: 10.1016/j.ijpara.2006.05.007.

[30] Matthias Frank, Laura Kirkman, Daniel Costantini, Sohini Sanyal, Catherine Lavazec, Thomas J Templeton, and Kirk W Deitsch. Frequent recombination events generate diversity within the multi-copy variant antigen gene families of plasmodium falciparum. *Int J*

Parasitol, 38(10):1099–1109, 2008. doi: 10.1016/j.ijpara.2008.01.010.

[31] S. Gandon, V. A. Jansen, and M. van Baalen. Host life history and the evolution of parasite virulence. *Evolution*, 55 (5):1056–1062, 2001.

[32] S. Gandon, M. J. Mackinnon, S. Nee, and A. F. Read. Imperfect vaccines and the evolution of pathogen virulence. *Nature*, 414(6865):751–756, 2001. doi: 10.1038/414751a.

[33] Sylvain Gandon, Minus van Baalen, and Vincent A A Jansen. The evolution of parasite virulence, superinfection, and host resistance. *Am Nat*, 159(6):658–669, 2002. doi: 10.1086/339993.

[34] R. J. Howard, J. W. Barnwell, E. P. Rock, J. Neequaye, D. Ofori-Adjei, W. L. Maloy, J. A. Lyon, and A. Saul. Two approximately 300 kilodalton plasmodium falciparum proteins at the surface membrane of infected erythrocytes. *Mol Biochem Parasitol*, 27(2-3):207–223, 1988.

[35] Lars Hviid. The role of plasmodium falciparum variant surface antigens in protective immunity and vaccine development. *Hum Vaccin*, 6(1), 2010.

[36] Lars Hviid and Trine Staalsoe. Malaria immunity in infants: a special case of a general phenomenon? *Trends Parasitol*, 20(2):66–72, 2004.

[37] Anja T R Jensen, Pamela Magistrado, Sarah Sharp, Louise Joergensen, Thomas Lavstsen, Antonella Chiucchiuini, Ali Salanti, Lasse S Vestergaard, John P Lusingu, Rob Hermsen, Robert Sauerwein, Jesper Christensen, Morten A Nielsen, Lars Hviid, Colin Sutherland, Trine Staalsoe, and Thor G Theander. Plasmodium falciparum associated with severe childhood malaria preferentially expresses pfemp1 encoded by group a var genes. *J Exp Med*, 199(9):1179–1190, 2004. doi: 10.1084/jem.20040274.

[38] Mirjam Kaestli, Ian A Cockburn, Alfred Corts, Kay Baea, J. Alexandra Rowe, and Hans-Peter Beck. Virulence of malaria is associated with differential expression of plasmodium falciparum var gene subgroups in a case-control study. *J Infect Dis*, 193(11):1567–1574, 2006. doi: 10.1086/503776.

[39] R. J. Kosinski. Antigenic variation in trypanosomes: a computer analysis of variant order. *Parasitology*, 80(2):343–357, 1980.

[40] Susan M Kraemer and Joseph D Smith. A family affair: var genes, pfemp1 binding, and malaria disease. *Curr Opin Microbiol*, 9(4):374–380, 2006. doi: 10.1016/j.mib.2006.06.006.

[41] Susan M Kraemer, Sue A Kyes, Gautam Aggarwal, Amy L Springer, Siri O Nelson, Zoe Christodoulou, Leia M Smith, Wendy Wang, Emily Levin, Christopher I Newbold, Peter J Myler, and Joseph D Smith. Patterns of gene recombination shape var gene repertoires in plasmodium falciparum: comparisons of geographically diverse isolates. *BMC Genomics*, 8: 45, 2007. doi: 10.1186/1471-2164-8-45.

[42] Sue A Kyes, Susan M Kraemer, and Joseph D Smith. Antigenic variation in plasmodium falciparum: gene organization and regulation of the var multigene family. *Eukaryot Cell*, 6(9):1511–1520, 2007. doi: 10.1128/EC.00173-07.

[43] Jean Langhorne, Francis M Ndungu, Anne-Marit Sponaas, and Kevin Marsh. Immunity to malaria: more questions than answers. *Nat Immunol*, 9(7):725–732, 2008. doi: 10.1038/ni.f.205.

[44] Thomas Lavstsen, Pamela Magistrado, Cornelus C Hermsen, Ali Salanti, Anja T R Jensen, Robert Sauerwein, Lars Hviid, Thor G Theander, and Trine Staalsoe. Expression of plasmodium falciparum erythrocyte membrane protein 1 in experimentally infected humans. *Malar J*, 4(1):21, 2005. doi: 10.1186/1475-2875-4-21.

[45] Katrina A Lythgoe, Liam J Morrison, Andrew F Read, and J. David Barry. Parasite-intrinsic factors can explain ordered progression of trypanosome antigenic variation. *Proc Natl Acad Sci U S A*, 104(19):8095–8100, 2007. doi: 10.1073/pnas.0606206104.

[46] M. J. Mackinnon and K. Marsh. The selection landscape of malaria parasites. *Science*, 328(5980):866–871, 2010. doi: 10.1126/science.1185410.

[47] M. J. Mackinnon, S. Gandon, and A. F. Read. Virulence evolution in response to vaccination: the case of malaria. *Vaccine*, 26 Suppl 3:C42–C52, 2008.

[48] Margaret J Mackinnon and Andrew F Read. Immunity promotes virulence evolution in a malaria model. *PLoS Biol*, 2(9):E230, 2004. doi: 10.1371/journal.pbio.0020230.

[49] Margaret J Mackinnon and Andrew F Read. Virulence in malaria: an evolutionary viewpoint. *Philos Trans R Soc Lond B Biol Sci*, 359(1446):965–986, 2004. doi: 10.1098/rstb.2003.1414.

[50] Claire Mackintosh, Zoe Christodoulou, Tabitha Mwangi, Moses Kortok, Robert Pinches, Thomas Williams, Kevin Marsh, and Christopher Newbold. Acquisition of naturally occurring antibody responses to recombinant protein domains of plasmodium falciparum erythrocyte membrane protein 1. *Malar J*, 7(1):155, 2008. doi: 10.1186/1475-2875-7-155.

[51] C. Magowan, W. Wollish, L. Anderson, and J. Leech. Cytoadherence by plasmodium falciparum-infected erythrocytes is correlated with the expression of a family of variable proteins on infected erythrocytes. *J Exp Med*, 168(4):1307–1320, 1988.

[52] R. M. May and R. M. Anderson. Epidemiology and genetics in the coevolution of parasites and hosts. *Proc R Soc Lond B Biol Sci*, 219(1216):281–313, 1983.

[53] Nicole Mideo and Troy Day. On the evolution of reproductive restraint in malaria. *Proc Biol Sci*, 275(1639):1217–1224, 2008. doi: 10.1098/rspb.2007.1545.

[54] Nicole Mideo, Samuel Alizon, and Troy Day. Linking within- and between-host dynamics in the evolutionary epidemiology of infectious diseases. *Trends Ecol Evol*, 23(9):511–517, 2008. doi: 10.1016/j.tree.2008.05.009.

[55] Nicole Mideo, Victoria C. Barclay, Brian H. K. Chan, Nicholas J. Savill, Andrew F. Read, and Troy Day. Understanding and predicting strain-specific patterns of pathogenesis in the rodent malaria plasmodium chabaudi. *Am Nat*, 172(5):214–238, 2008. doi: 10.1086/591684.

[56] L. H. Miller, M. F. Good, and G. Milon. Malaria pathogenesis. *Science*, 264(5167): 1878–1883, 1994.

[57] Louis H Miller, Dror I Baruch, Kevin Marsh, and Ogobara K Doumbo. The pathogenic basis of malaria. *Nature*, 415(6872):673–679, 2002. doi: 10.1038/415673a.

[58] C. I. Newbold, A. G. Craig, S. Kyes, A. R. Berendt, R. W. Snow, N. Peshu, and K. Marsh. Pfemp1, polymorphism and pathogenesis. *Ann Trop Med Parasitol*, 91(5):551–557, 1997.

[59] Michael F Ofori, Daniel Dodoo, Trine Staalsoe, Jrgen A L Kurtzhals, Kwadwo Koram, Thor G Theander, Bartholomew D Akanmori, and Lars Hviid. Malaria-induced acquisition of antibodies to plasmodium falciparum variant surface antigens. *Infect Immun*, 70(6):2982–2988, 2002.

[60] Emelda A Okiro, Abdullah Al-Taiar, Hugh Reyburn, Richard Idro, James A Berkley, and Robert W Snow. Age patterns of severe paediatric malaria and their relationship to plasmodium falciparum transmission intensity. *Malar J*, 8:4, 2009. doi: 10.1186/1475-2875-8-4.

[61] Noa D Pasternak and Ron Dzikowski. Pfemp1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite plasmodium falciparum. *Int J Biochem Cell Biol*, 41(7):1463–1466, 2009. doi: 10.1016/j.biocel.2008.12.012.

[62] Andrew F. Read. The evolution of virulence. *Trends in Microbiology*, 2(3): 73 – 76, 1994. ISSN 0966-842X. doi: 10.1016/0966-842X(94)90537-1.

[63] Mario Recker, Sean Nee, Peter C Bull, Sam Kinyanjui, Kevin Marsh, Chris Newbold, and Sunetra Gupta. Transient

cross-reactive immune responses can orchestrate antigenic variation in malaria. *Nature*, 429(6991):555–558, 2004. doi: 10.1038/nature02486.

[64] E. M. Riley, G. E. Wagner, B. D. Akanmori, and K. A. Koram. Do maternally acquired antibodies protect infants from malaria infection? *Parasite Immunol*, 23 (2):51–59, 2001.

[65] Amanda Ross, Gerry Killeen, and Thomas Smith. Relationships between host infectivity to mosquitoes and asexual parasite density in plasmodium falciparum. *Am J Trop Med Hyg*, 75(2 Suppl): 32–37, 2006.

[66] Matthias Rottmann, Thomas Lavstsen, Joseph Paschal Mugasa, Mirjam Kaestli, Anja T R Jensen, Dania Mller, Thor Theander, and Hans-Peter Beck. Differential expression of var gene groups is associated with morbidity caused by plasmodium falciparum infection in tanzanian children. *Infect Immun*, 74(7):3904–3911, 2006. doi: 10.1128/IAI.02073-05.

[67] Robert W Snow and Kevin Marsh. The consequences of reducing transmission of plasmodium falciparum in africa. *Adv Parasitol*, 52:235–264, 2002.

[68] H. M. Taylor, S. A. Kyes, and C. I. Newbold. Var gene diversity in plasmodium falciparum is generated by frequent recombination events. *Mol Biochem Parasitol*, 110(2):391–397, 2000.

[69] Thomas J Templeton. The varieties of gene amplification, diversification and hypervariability in the human malaria parasite, plasmodium falciparum. *Mol Biochem Parasitol*, 166(2):109–116, 2009. doi: 10.1016/j.molbiopara.2009.04.003.

[70] Lasse S Vestergaard, John P Lusingu, Morten A Nielsen, Bruno P Mmbando, Daniel Dodoo, Bartholomew D Akanmori, Michael Alifrangis, Ib C Bygbjerg, Martha M Lemnge, Trine Staalsoe, Lars Hviid, and Thor G Theander. Differences in human antibody reactivity to plasmodium falciparum variant surface antigens are dependent on age and malaria transmission intensity in northeastern tanzania. *Infect Immun*, 76(6): 2706–2714, 2008. doi: 10.1128/IAI.01401-06.

[71] Christian W Wang, Cornelus C Hermsen, Robert W Sauerwein, David E Arnot, Thor G Theander, and Thomas Lavstsen. The plasmodium falciparum var gene transcription strategy at the onset of blood stage infection in a human volunteer. *Parasitol Int*, 58(4):478–480, 2009. doi: 10.1016/j.parint.2009.07.004.

Supplementary Material

Table 2: Default model parameter values and the variations in these parameters simulated.

r	Number of available rbc for infection each parasite generation/μL blood.
l	Force of infection.
l_s	Force of infection during a dry season.
p_0	Degree in which gametocyte formation depends on parasitemia.
d	Days of infection before gametocytes are formed
f_m	Scaling factor for malaria related mortality
c	Intrinsic parasite net growth rate
c_v	VSA-enabled parasite net growth rate
m_c	Degree of cross-immunity
m_a	Relative threshold for antibody trigger.
m_d	Time in days before triggered antibody response becomes effective.
m_n	Degree of general or innate immunity.

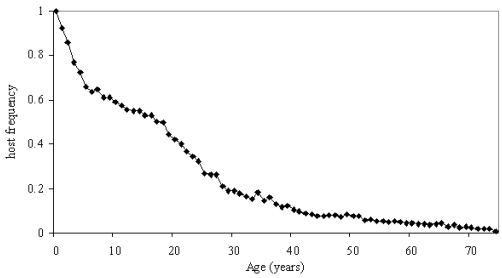


Figure 6: Host demography from a 5VSA simulation under default parameter conditions.

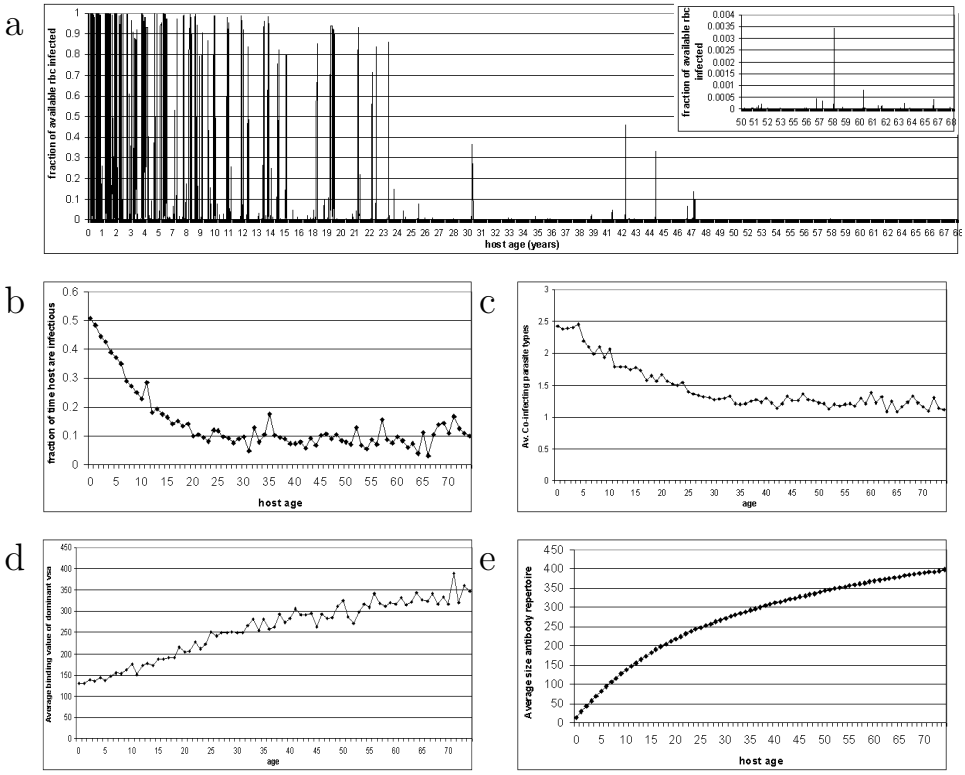


Figure 7: Results from a 1VSA simulation under default model conditions. Panel a shows the life-time parasitemia of a typical host. (Please note that a parasitemia of 1 means only 20% of rbc are infected). Panels b-d show averages by host age from yearly samples taken of all hosts during 60 years of the simulation. Panel b shows the average host parasitemia. Panel c shows the average number of co-infecting parasite types. Panel d shows the average rank of the dominantly expressed VSA. Panel e shows the average size of the hosts VSA-specific antibody repertoires.

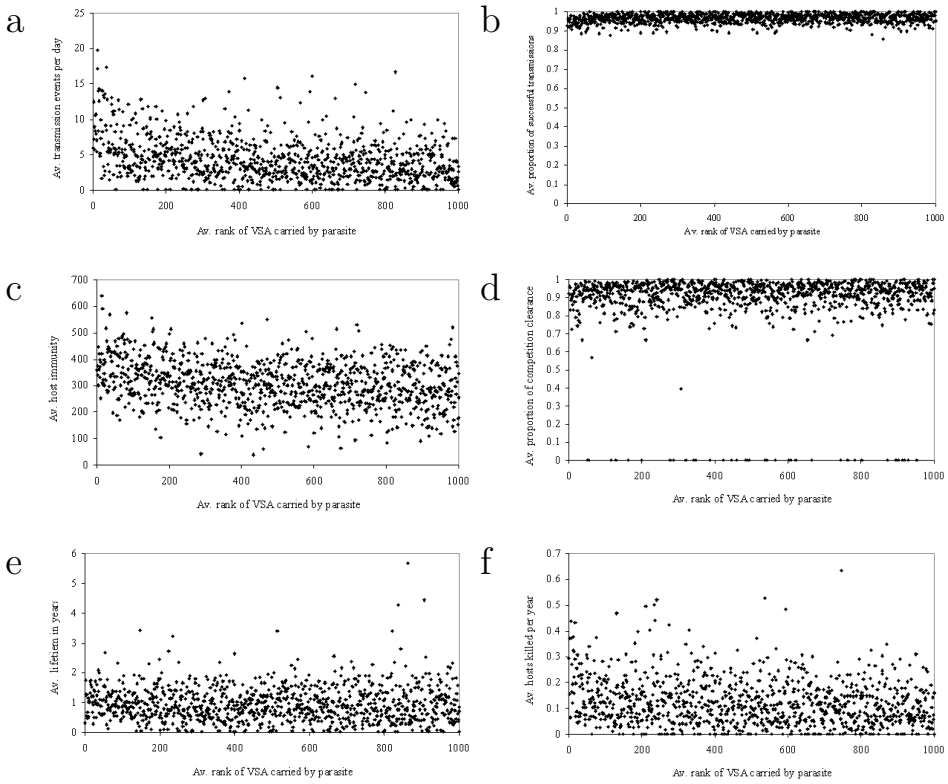


Figure 8: Results from a 5VSA simulation under default parameter conditions. Panels show averages from all parasite types that existed during the simulation and carried the specified VSA. Panel a shows the daily number of transmission events i.e., how many times a day the parasite is selected to infect a host. Panel b shows the proportion of successful transmission i.e., transmission events to hosts without protecting VSA-specific immunity. Panel c shows the number of hosts with specific antibodies against the VSA carried. Panel d shows the proportion of clearances from the host due to competition (opposed to immunity clearance). Panel e shows the lifetime of the parasite types in years. Panel f shows the yearly number of hosts dying from malaria while infected by the parasite type.

English Summary

In order to survive, infectious agents must succeed at two scales, within-host and between-host. Within-host survival involves processes such as reproduction and immune evasion, while between-host survival involves transmission and persistence in the host population. Although in all infections these processes are connected, in some infections, the processes influence each other extensively through means of feedback loops between the processes. This can lead to a large network of interactions across both infection scales. Developing and predicting the effect of potential control measures on the complex infection dynamics in these systems poses a challenge. Part of this challenge comes from the large heterogeneity often observed in these systems for example in, infection history, immunity and parasite appearance.

The goal of this PhD thesis is to increase our understanding of some of these complex within and between-host infection dynamics through the creation of mathematical and computational models that are able to capture the existing host and/or parasite heterogeneity. This goal is reached through a series of research projects that gradually build up in complexity of both the system modelled and the modelling techniques used. The models created in the projects help to explain a wide range of sometimes contradictory experimental results, are used to predict the effect of control measures and generate ideas for the development of new methods of control.

In the first project a mathematical model is built to help understand a large body of experimental results on Experimental Autoimmune Encephalomyelitis (EAE), a mouse-model for the human autoimmune disease MS. Many elements of the immune system are involved in EAE and using the model it was possible to find a limited set of cell populations and their interactions that can explain these results. The model hereby reduced the complexity of the experimental results allowing more focussed continued research and the formation of new hypotheses on treatments.

The second project looks at Paratuberculosis in cows, a chronic infection without cure caused by *Mycobacterium avium* subspecies *paratuberculosis* (MAP). A mathematical model of MAP infection is built to study the difference of Type 1 and Type 2 immune responses on disease progression. The model shows that both responses are in principle able to clear infection and that the killing rate of the Type 1 and Type 2 adaptive immune responses together is much stronger than the sum of their individual killing rates. Subsequently simulations of the model are compared to infection and immunity data from a 4-year longitudinal experimental study to better understand the Type 1 and Type 2 responses against MAP in real cows. The comparison shows that a Type 1 response which is observed to stop the initial growth of infection does not necessarily have the critical killing rate needed to control the infection. The results from this project give new insight into Type 1 and Type 2 adaptive immune responses against MAP infections which are important for the development of effective control measures

against Paratuberculosis.

In the third project a computational model is built of Coccidiosis in broiler chickens caused by the protozoan parasite *Eimeria acervulina*. The model is built to understand the progression of infection in flocks of commercial broiler chickens together with the influence of the stable cleaning intensity between subsequent flocks on this progression. The model shows a wave-like relationship between cleaning intensity and the severity of the Coccidiosis outbreaks. This relationship is caused by an increased heterogeneity in the host population (i.e. in infection history) at the peaks of the wave. The heterogeneity originates from stochastic parasite uptake combined with infection feedback loops that span the within-host and between-host scales. As such the model confirms previous experimental and computational findings, and additionally gives an explanation for the cause of this wave-like relationship.

The fourth and fifth projects study malaria in humans caused by the *Plasmodium falciparum* parasite. This parasite uses a group of variant surface antigens (VSA) to escape host immunity. These VSA play an important role in disease severity due to the ability they give infected red blood cells to adhere to different host tissues. Each parasite carries a different set of these VSA and the number of different VSA in the population is immense. During infection, parasites regularly change which VSA they express to remain undetected by the antibodies built by the immune system. In high-transmission areas people build up a large repertoire of anti-VSA antibodies during their lifetime which is thought to play an important role in the gradual protection from severe, mild and, eventually, all clinical malaria, although subclinical malaria infections remain common even at old age.

In the fourth project a computational model is built to better understand the role of these VSA on host immunity and parasite evolution. The model shows that hosts do not acquire antibodies against VSA at random but that there is a general order in which they are acquired where immunity against stronger adhering VSA is acquired earlier. The model hereby confirms recent experimental findings indicating the existence of such an order and furthermore gives an explanation for the origin of this order. In addition, the model suggests that the difference in a younger versus older host environment, at least on its own, does not give a selective advantage for parasites with a 'strong-rare' set of VSA over parasites with other, random, sets of VSA.

In the fifth project the previous model is extended to study how VSA enable the parasites to adapt to changes in their environment and to examine the effect of current and potential control measures on the sets of VSA carried by the circulating parasites. The results, amongst others, show that some seemingly sensible measures, such as reducing the reproduction rate of the parasites, could have the unwanted effect of increasing the frequency of parasites carrying stronger adhering VSA, such as those causing cerebral malaria. The results from this project

give insight into how malaria parasites are able to adapt so well to changes in their environment, as caused by drugs and other control measures, and shows the importance of this knowledge for the success of these control measures.

Nederlandse Samenvatting

Om te overleven moeten infectieuze ziektekiemen (hierna parasieten genoemd) succesvol zijn op twee niveaus, namelijk binnen de gastheren (intra-gastheer) en tussen de gastheren (inter-gastheer). Intra-gastheer overleving draait om processen zoals reproductie en het vermijden van gastheerimmuniteit, terwijl inter-gastheer overleving draait om processen zoals transmissie en persistentie in de gastheerpopulatie. Alhoewel voor alle infecties geldt dat de processen op beide niveaus met elkaar verbonden zijn, beinvloeden zij elkaar in sommige gevallen zeer sterk door middel van terugkoppelingslussen. Hierdoor ontstaat een groot netwerk aan interacties dat beide infectieniveaus overbrugt. Door de complexe infectiedynamica binnen zo'n systeem is het ontwikkelen en het voorspellen van het effect van potentiële controle- en preventiemaatregelen voor dergelijke infecties een enorme uitdaging. Deze wordt nog eens verder bemoeilijkt door de grote mate van heterogeniteit die vaak wordt geobserveerd in deze systemen zoals bijvoorbeeld in de individuele infectiegeschiedenis van elke gastheer, zijn immuniteit en het uiterlijk van de afzonderlijke parasieten.

Het doel van dit proefschrift is om de complexe intra- en inter-gastheer infectiedynamica in deze systemen beter te begrijpen met behulp van wiskundige en simulatie modellen die in staat zijn om de bestaande gastheer en/of parasiet heterogeniteit te beschrijven. Dit doel wordt bereikt aan de hand van een serie onderzoeksprojecten die geleidelijk toenemen in complexiteit van zowel het beschreven infectiesysteem als van de modeltechnieken die worden gebruikt.

In het eerste project wordt een wiskundig model ontwikkeld voor het verklaren vaneen grote hoeveelheid aan experimentele resultaten afkomstig uit onderzoek aan Experimentele Autoimmuun Encephalomyelitis (EAE), een muismodel van de humane aandoening Multiple Sclerose. Het zoeken naar een behandelmethode voor EAE wordt bemoeilijkt door het grote aantal elementen van het immuunsysteem dat actief is tijdens EAE. Met behulp van het ontwikkelde model is het mogelijk om een beperkte set van cruciale celpopulaties aan te wijzen die met hun gezamelijke interactie de onderzoeksresultaten kunnen verklaren. Het model reduceert hiermee de complexiteit van de experimentele resultaten en maakt het mogelijk om gerichter vervolgonderzoek te doen naar behandelmethoden. Ook heeft het model geleid tot de formulering van nieuwe hypothesen.

Het tweede onderzoeksproject is gericht op Paratuberculose in koeien, een chronische aandoening zonder effectieve behandeling, die wordt veroorzaakt door *Mycobacterium avium* subsoort *paratuberculosis* (MAP). Er wordt een wiskundig model van MAP infectie ontwikkeld om het verschil tussen Type 1 en Type 2 adaptieve immuunresponsen op het infectieverloop te bestuderen. Het model laat zien dat in theorie beide responsen afzonderlijk in staat zijn om de infectie onder controle te houden en te verwijderen, al duurt volledige verwijdering erg lang. Het model laat ook zien dat de snelheid waarmee de infectie verwijderd wordt wanneer beide responsen gezamenlijk actief zijn, veel hoger is dan de som van de afzonderlijke snelheden. Model simulaties worden vervolgens vergeleken

met infectie- en immuundata van een 4 jaar lange experimentele studie om meer inzicht te verwerven in de bijdrage die Type 1 en Type 2 responsen leveren aan de afweerreacties tegen MAP in de praktijk. Deze vergelijking laat zien dat een Type 1 respons in een koe die in staat is om de initiële infectiegroei te stoppen niet noodzakelijkerwijs ook de kritieke bacterie-verwijderingssnelheid heeft om de MAP infectie blijvend onder controle te houden. De resultaten uit dit onderzoek geven nieuwe inzichten in de Type 1 en Type 2 immuunresponsen tegen MAP infecties, een belangrijke basis voor de ontwikkeling van een meer doeltreffende bestrijding van Paratuberculose.

In het derde project wordt een computermodel ontwikkeld voor het bestuderen van Coccidiose infecties in vleeskuikens veroorzaakt door de protozoa parasiet *Eimeria acervulina*. Met het model wordt het verloop van infectieuitbraken in fokkerijen onderzocht en de invloed die de mate van stalreiniging tussen de opeenvolgende groepen vleeskuikens op dit verloop heeft. Het model laat zien dat er een op een golf lijkende relatie bestaat tussen de reinigingsintensiteit en de ernst van de uitbraak. Deze relatie komt voort uit een verhoogde heterogeniteit (o.a. in infectiegeschiedenis) in de gastheerpopulatie op de toppen van de golf. Deze heterogeniteit wordt veroorzaakt omdat kippen willekeurig Eimeria oocysten uit hun omgeving opnemen en de afzonderijke kippen hierdoor in verschillende mate geïnfecteerd raken. Er zijn stalcontaminatie niveaus waarbij de terugkoppelingslussen tussen de intra- en inter-gastheer infectieprocessen deze heterogenitiet in besmetting versterken of juist uitdempen. Het model bevestigt hiermee niet alleen eerdere experimentele en model resultaten, het geeft ook een verklaring voor de oorzaak van dit verband.

Het vierde en vijfde onderzoeksproject zijn gericht op Malaria bij de mens, veroorzaakt door de *Plasmodium falciparum* parasiet. Deze parasiet maakt gebruik van een groep van variabele oppervlakte antigenen (VOA) om het immunsysteem van de gastheer te omzeilen. Deze VOA zijn tevens medebepalend voor de ernst en dodelijkheid van malaria bij patiënten, door hun vermogen om geïnfecteerde rode bloedcellen te laten binden aan verschillende interne weefsels. Elke parasiet heeft een andere set van deze VOA tot zijn beschikking en het aantal verschillende VOA die bestaan in de gehele parasietpopulatie is immens. Tijdens een infectie wisselen parasieten geregeld de VOA die ze gebruiken om zo onherkenbaar te blijven voor de antilichamen die het immuunsysteem maakt tegen de reeds herkende VOA. In gebieden waar veel malaria voorkomt bouwen mensen gedurende hun leven een groot repertoire aan VOA antilichamen op waarvan vermoed wordt dat het verantwoordelijk is voor de geleidelijke opbouw van bescherming tegen ernstige, milde en uiteindelijk alle klinische vormen van malaria, alhoewel subklinische infecties ook op latere leeftijd blijven voorkomen.

In het vierde onderzoeksproject wordt een computermodel ontwikkeld om de rol van de VOA op de immuniteit van de gastheer en de evolutie van de parasiet beter te kunnen begrijpen. Het model laat zien dat gastheren niet willekeurig

antilichamen tegen VOA verwerven maar dat er een algemene volgorde bestaat waarbij immuniteit tegen sterker bindende VOA eerder wordt verworven. Het model bevestigt hiermee recent experimenteel onderzoek dat een volgorde in de verwerving van deze antilichamen aanduidt en geeft een verklaring voor de oorsprong van deze volgorde. Ook laat het model zien dat, in tegenstelling tot wat eerdere studies veronderstelden, het verschil in leefomgeving voor de parasieten tussen jonge gastheren (met weinig immuniteit) en oudere gastheren (met veel immuniteit), op zichzelf niet genoeg is om een selectievoordeel op te leveren voor parasieten met een set van zowel sterke als zeldzame VOA vergeleken met parasieten die een willekeurige set van VOA tot hun beschikking hebben.

In het vijfde project wordt het malaria model uitgebreid om te bestuderen hoe de VOA de parasieten in staat stellen om zich aan te passen aan veranderingen in hun omgeving en om te onderzoeken wat het gevolg kan zijn van de huidige en potentieel nieuwe controlemaatregelen op de sets van VOA onder de circulerende parasieten. De resultaten laten o.a. zien dat sommige ogenschijnlijk doeltreffende maatregelen, zoals het verlagen van de voortplanting van de parasieten, kunnen leiden tot een stijging van het percentage parasieten met sterker bindende VOA, zoals bijvoorbeeld de veroorzakers van cerebrale malaria. De resultaten uit dit project geven inzicht in het aanpassingsvermogen van malariaparasieten aan veranderingen in hun leefomgeving, zoals medicijnen en andere bestrijdingsmaatregelen, en tonen het belang aan van dit inzicht voor het slagen van deze maatregelen.

Samenvattend, kunnnen de hierboven beschreven onderzoeksprojecten en de daarin ontwikkelde modellen helpen een breed scala aan soms tegenstrijdige experimentele resultaten te verklaren, worden ze gebruikt om het effect van controlemaatregelen te voorspellen en leiden ze tot inzichten en ideeën voor de ontwikkeling van nieuwe methoden om infectieziekten te bestrijden.

Resumen en Español

Para su supervivencia, los agentes infecciosos deben sobrevivir en dos niveles, dentro y entre sus hospederos. Sobrevivir dentro del hospedero incluye procesos como reproducción y evasión del sistema inmune del hospedero, mientras que sobrevivir entre hospederos incluye procesos como transmisión y persistencia en la población de hospederos. Aunque en todas las infecciones estos procesos están conectados, en algunas infecciones se influencian mutuamente mas gracias a bucles de retroalimentación entre los procesos. En estos casos se puede formar una gran red de interacciones a través de los niveles de infección. Desarrollar y predecir el efecto de posibles medidas de control en estos sistemas con una dinámica de infección tan compleja plantean un difícil desafío. Parte de este desafío proviene de la gran heterogeneidad que a menudo se observa en estos sistemas, por ejemplo en la historia de la infección y de la inmunidad del hospedero, y en la apariencia del parásito.

El objetivo de esta tesis doctoral es el de aumentar nuestra comprensión sobre algunas de estas infecciones con dinámica compleja dentro y entre los hospederos, a través de la creación de modelos matemáticos y computacionales, capaces de capturar la heterogeneidad del hospedero y/o el parásito. Este objetivo se alcanza a través de una serie de proyectos de investigación que van aumentando tanto en la complejidad del sistema modelado como en las técnicas que se utilizan. Los modelos creados en estos proyectos ayudan a explicar una amplia gama de resultados experimentales, a veces contradictorios, y se utilizan tambien para predecir el efecto de medidas de control, ademas de generar ideas para el desarrollo de nuevos métodos de control.

En el primer proyecto se construye un modelo matemático para ayudar a comprender una extensa cantidad de resultados experimentales en encefalomielitis auto-inmune experimental (EAE), un modelo en ratones de la enfermedad autoinmune humana Sclerosis Multiple. Una multitud de elementos presentes en el sistema inmunológico se activan durante EAE, lo cual dificulta la búsqueda de elementos esenciales que puedan ser manipulados para un tratamiento. Haciendo uso de este modelo matemático fue posible encontrar un pequeño grupo de células cruciales, las cuales junto con sus interacciones, son capaces de explicar los resultados experimentales. De esta forma, el modelo reduce la complejidad de los resultados experimentales a un subconjunto de células, permitiendo así que futuras investigaciones puedan centrarse en las células mas importantes. Este proyecto también llevo a nuevas hipótesis para el tratamiento de EAE.

El segundo proyecto se dedica a paratuberculosis en vacas (ganado), una infección crónica sin cura, causada por la bacteria *Mycobacterium avium*, subespecie *paratuberculosis* (MAP). Un modelo matemático de la infección con MAP es diseñado para estudiar la diferencia entre las respuestas inmune Tipo 1 y Tipo 2 en la progresión de la infección. El modelo muestra que ambas respuestas son, en teoría, capaz de eliminar la infección y que la velocidad de eliminación, cuando ambos tipos de respuestas están activas simultaneamente es mucho mayor que la suma de las velocidades cuando están activas individualmente. Asimismo, las simulaciones

del modelo matemático se comparan con datos de infección y de inmunidad de un estudio experimental longitudinal de 4 años, para estudiar las repuestas inmunes Tipo 1 y 2 contra MAP en la práctica. La comparación muestra que una respuesta del Tipo 1, la cual puede detener el aumento inicial de la infección, no tiene necesariamente la capacidad de controlar la infección a largo plazo. Los resultados de este proyecto arrojan nuevos elementos para entender las respuestas inmunes Tipo 1 y Tipo 2 contra MAP, las cuales son importantes para el desarrollo de medidas eficaces de control contra la paratuberculosis.

En el tercer proyecto se construye un modelo computacional sobre coccidiosis en pollos de engorde causada por el parásito protozoo *Eimeria acervulina*. El modelo es diseñado para estudiar la progresión de la infección en las parvadas de pollos de engorde comerciales, asi como la influencia que tiene sobre esta progresión la intensidad en la limpieza de los galpones entre parvadas sucesivas . El modelo muestra que existe un efecto de onda entre la intensidad de la limpieza y la gravedad de los brotes de coccidiosis. Esta relación es causada por un aumento en la heterogeneidad de los hospederos (entre otros en historia de la infección) en los picos de la onda de infeccion . La heterogeneidad se origina en los pollos al contraer parásitos del suelo del galpon de forma aleatoria, obtienendo así diferentes grados de infección. De acuerdo a los niveles de contaminación inicial del galpon, los circuitos de retroalimentación entre los dos niveles de infección (dentro y entre los hospederos) pueden aumentar o silenciar esta heterogeneidad. Así, el modelo confirma previos resultados experimentales y computacionales, además de explicar la causa de esta relación ondular.

El cuarto y quinto proyecto se dedican a la malaria en los seres humanos causada por el parásito protozoo *Plasmodium falciparum*. Este parásito utiliza un grupo de antígenos de superficie variantes (ASV) para escapar a la inmunidad del hospedero. Estos ASV juegan un papel importante en la gravedad de la enfermedad, debido a la habilidad que confieren a las células rojas infectadas para adherirse a los tejidos internos del hospedero. Cada parásito lleva una combinación diferente de estos ASV, siendo la variedad de ASV en la población inmensa. Durante la infección, los parásitos cambian regularmente el ASV que expresan para no ser detectados por los anticuerpos previamente producidos contra el anterior ASV. En áreas de alta transmisión, individuos producen un gran repertorio de anticuerpos anti-ASV a lo largo de su vida. Se cree que este repertorio es, en parte, responsable de la protección que las personas adquieren primero contra la malaria grave, luego la malaria moderada y, finalmente, contra todo tipo de malaria clínica, a pesar de que infecciones subclínicas de malaria siguen siendo comunes incluso en las personas mayores.

En el cuarto proyecto se construye un modelo computacional para comprender mejor el papel de estos ASV en la inmunidad del hospedero y en la evolución del parásito. El modelo muestra que los hospederos no adquieren anticuerpos contra ASV al azar, sino atraves de un orden general en el cual la inmunidad

contra ASV con fuerte capacidad de adhesión se adquiere primero. El modelo confirma recientes resultados experimentales que indican la existencia de tal orden y, además, da una explicación para la causa de esta orden. Asimismo, el modelo sugiere que la diferencia en el entorno dentro de hospederos jóvenes y hospederos mayores, al menos por sí solo, no es suficiente ventaja selectiva para los parásitos con una combinacion de ASV 'fuerte-raros' sobre los parásitos con combinaciones aleatorias de ASV.

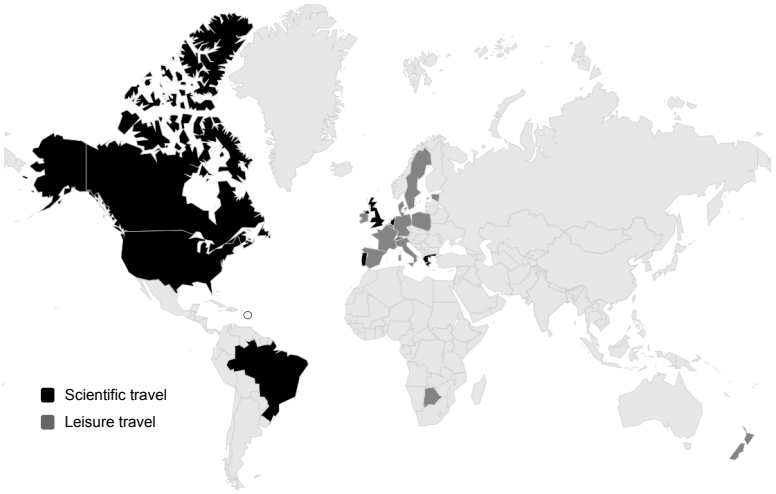
En el quinto proyecto se amplía el modelo anterior para estudiar cómo los ASV permiten a los parásitos adaptarse a los cambios en su entorno, asi como para estudiar el efecto de medidas de control actuales y potenciales sobre las combinaciones de ASV en los parásitos circulantes. Los resultados, entre otros, indican que algunas medidas aparentemente sensatas, como una reducción en la reproducción de los parásitos, podría tener el efecto indeseado de aumentar la frecuencia de los parásitos que llevan los ASV con más fuerte capacidad adherente, como por ejemplos los que causan la malaria cerebral. Los resultados de este proyecto nos dan una idea de la forma como los parásitos son capaces tan bien de adaptarse a los cambios en su entorno, causadas por ejemplo por medicinas anti-malaria y otras medidas de control, y ademas muestra la importancia de este conocimiento para el éxito de dichas medidas de control.

Thesis & Project Facts

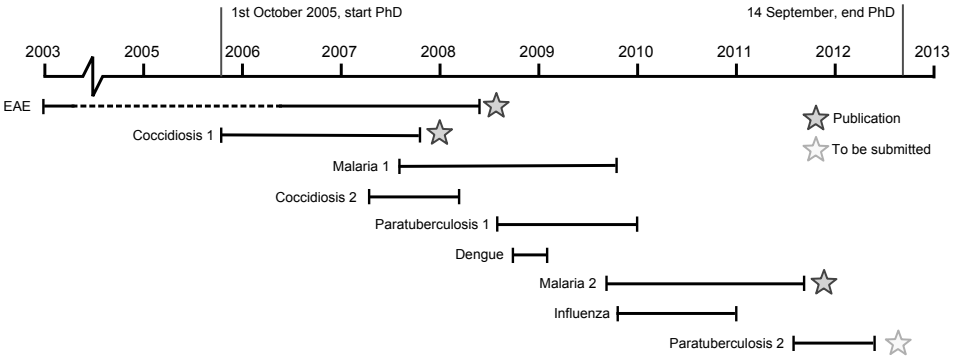
Did you know that:

- The number of words in this thesis are 40,779. It would take more than 12 hours to read it all out loud.
- Approximately 65% percent of work done during the PhD is represented in this thesis.
- The simulations in this thesis were run on 7 different computers in 5 different offices in 3 countries.
- Stephen & Maite both met each other and got engaged during the making of this thesis.
- During the making of this thesis Maite made 21 trips back and forth to Oxford.
- If instead of starting this thesis on Oct 1st 2005, Maite had got in a car and driven at 80km/h, she would now have travelled 4,877,000 km kilometers, enough to get to the moon and back 6 times!!
- There are more than 50 spelling mistakes in this thesis, can you find them all?
- In Maite's department in just one hallway there are researchers working on: chickens, pigs, cows, goats, udders, gerbils, crows, midges, sandflies, bank voles, and tiger mosquitoes.
- No harm was done to any of the studied bugs during this thesis (except coding bugs).

Travel



Research projects



Acknowledgements


AdKoets
ArjanStegeman
AntonCrombach
SanjaSelakovic
WiebrenSantema
Thanks
ArjanTholen
JoseMorata
AnnemarieBouma
SanderVanNoort
Gracias
JohnGalsworthy
LizetteMaljaars
LindaGalsworthy
MarjoleinDerks
Danke
SarahGalsworthy
LiesbethWilschut
OscarMorata
GhislaineFilot
Bedankt
DanielaCianci
JudithLópez
KirstenTenTusscher
JanSeverins
NerminkaMuslija
RocioLópez
WilmaSteeneveld
DanielVanDerPost
DonKlinkenberg
FranciscaVelkers
ClaraLópez
PieterPoorthuis
RosanneVanOudheusden
AnnekeBatenburg
MariekeJesse
ThankYou
JanvandenBroek
MaiteLópez
DieuwertjeSpekreijse
GloriaMorata
LennyHogerwerf
BrunoBonier
HansHeesterbeek
CamielVanLenteren
RichiLópez
TijsTobias
AnneriekeHeuvelink
AmparoSeverins
DiederikLiebrand
CarmenSepulveda
CarlineVanDenDool
ChristianBertholle
KristyPickwell
AnnaMontero
StephenGalsworthy
MilGracias
SusanneEisenberg
SaskiaVanDerDrift
MilanVanHoek
JanKeesVanAmerongen
NienkeHartemink
RicardoLópez
IgnacioLópez
MariskaVogelenzangDeJong
EllenMeijer
ChrisFleer
AngelinesSeverins
Obrigada
PiliLópez
GuillermoBonier
GustavoBonier
Dankjewel
MarianGroeneboom
MaartenPieterse
DankU

Curriculum Vitae

English

Maite Severins was born on June 19th, 1979 in Heerlen (The Netherlands). She attended the Bisschoppelijk College Echt high school from 1991 to 1997 and studied Biology with a masters degree in Theoretical Biology at Utrecht University from 1998 to 2005. From 2000 to 2001 she interrupted her studies to become a fulltime treasurer for the Utrechtse Biologen Vereniging. As part of her studies she spent 9 months in Xalapa (Mexico) to do field work on ant-plant interactions. After her studies, in the summer of 2005 she did an internship working on the modelling of Altzheimer's Disease at Novartis Pharma Services AG in Basel, Switzerland. In October 2005 she started her PhD studies on infection dynamics at within-host and between-host scales at the Theoretical Epidemiology group of the Farm Animal Health department (Faculty of Veterinary Medicine, Utrecht University). During this work she has been a guest scientist for several months at the Gulbenkian Institute in Lisbon (Portugal). In May 2011 she started working as a classification specialist in the Buyer Experience & Search team at eBay Inc., in London (United Kingdom).

Nederlands

Maite Severins is geboren op 19 juni 1979 in Heerlen. Zij heeft in 1997 haar VWO-diploma behaald aan het Bisschoppelijk College Echt. Hierna volgde zij een Biologie opleiding aan de Universiteit Utrecht welke zij heef afgerond met een doctoraal diploma in Theoretische Biologie in 2005. Van 2000 tot 2001 onderbrak zij haar studie om voltijd penningmeester te worden van de Utrechtse Biologen Vereniging gefinancierd door de Universiteit Utrecht. Als onderdeel van haar studie bracht ze negen maanden door in Xalapa (Mexico) om veldwerk uit te voeren aan insect-plant interacties. Na haar studie, in de zomer van 2005, heeft zij werkervaring opgedaan in het modelleren van de ziekte van Altzheimer tijdens een stage bij Novartis Pharma Services AG in Basel, Zwitserland. In oktober 2005 begon zij aan haar promotie traject 'Infectie dynamica op binnen-gastheer en tussen-gastheer niveaus' bij de vakgroep Theoretische Epidemiology van het departement Gezondheidszorg Landbouwhuisdieren (Faculteit Diergeneeskunde, Universiteit Utrecht). Tijdens haar promotie is zij enkele maanden gastonderzoeker geweest aan het Gulbenkian Institute in Lissabon (Portugal). In mei 2011 begon zij haar werk als 'classification specialist' in het Buyer Experience & Search team bij eBay Inc., in London (Verenigd Koninkrijk).

Español

Maite Severins nació el 19 de junio de 1979 en Heerlen (Países Bajos). Asistió la escuela secundaria Bisschoppelijk College Echt desde 1991 hasta 1997 y estudió Biología con una maestría en Biología Teórica en la Universidad de Utrecht desde 1998 hasta 2005. Entre 2000 y 2001 interrumpió sus estudios para tomar el puesto

de tesorera del Utrechtse Biologen Vereniging (Union de Biologos de Utrecht) financiado por la Universidad de Utrecht. Como parte de sus estudios pasó 9 meses en Xalapa (México) para a realizar trabajo de campo en el estudio de las interacciones entre plantas e insectos. Después de sus estudios, en el verano de 2005 realizó una estancia en Novartis Pharma Services AG en Basilea, Suiza modelando la enfermedad de Alzeimer. En octubre de 2005 comenzó sus estudios de doctorado sobre la dinámica de infección a nivel dentro-hospedero y entre-hospederos en el grupo de Epidemiología Teórica del Departamento de Sanidad Animales de Granja (Facultad Veterinaria, Universidad de Utrecht). Durante estos estudios ha sido invitada a hacer su investigación por varios meses en el Instituto Gulbenkian de Lisboa (Portugal). En mayo de 2011 comenzó a trabajar como especialista de clasificación parte del equipo de Búsqueda de eBay Inc., en Londres (Reino Unido).