Pathogen-mediated selection in the immune system of rodents

Exploring selection targets, functional effects and trade-offs

Nandakumar, Mridula

2024

Pathogen-mediated selection in the immune system of rodents

Exploring selection targets, functional effects and trade-offs

MRIDULA NANDAKUMAR
DEPARTMENT OF BIOLOGY | FACULTY OF SCIENCE | LUND UNIVERSITY

Pathogen-mediated selection in the immune system of rodents

Exploring selection targets, functional effects and trade-offs

Mridula Nandakumar

LUND
UNIVERSITY

DOCTORAL DISSERTATION

Doctoral dissertation for the degree of Doctor of Philosophy (PhD) at the Faculty of Science at Lund University to be publicly defended on 23rd of April at 09.00 in Hörsalen A213, Department of Biology, Sölvegatan 35, Lund, Sweden

Faculty opponent
Prof. Aida Andrés

Department of Genetics, Evolution and Environment, University College London, United Kingdom

Organization: LUND UNIVERSITY

Document name: Doctoral Dissertation **Date of issue:** 2024-04-23

Author(s): Mridula Nandakumar **Sponsoring organization:**

Title and subtitle: Pathogen-mediated selection in the immune system of rodents: exploring selection targets, functional effects and trade-offs

Abstract:

Pathogens cause disease and play an important role in shaping evolution of the host immune system. They create pressure on host immunity to evolve in numerous ways, most commonly by increasing divergence between species (positive selection) or increasing polymorphisms within a population (balancing selection). Especially with balancing selection, trade-offs between different traits, for example responses to different pathogens, are essential. Across five papers, questions related to what immune genes are under selection, how this translates to an effect on the immune response and what trade-offs occur, are addressed using rodents as study system. Paper I utilised genomes from 30 rodent species to identify signatures of positive selection in immune genes. In general, function of immune genes was a significant determinant for signs of positive selection. This effect was significant even after accounting for potential confounding factors like gene expression and protein-protein interactions. In Paper II, the focus is on a local population of bank voles in Sweden, to look for signatures of balancing selection in the complement system – a branch of innate immunity. One complement gene, *FCNA*, was found to be under strong balancing selection. In Paper III, *FCNA* polymorphism was linked to associations with natural infections of *Borrelia afzelii*, a common pathogen for bank voles. Papers IV and V look at how the immune response of bank voles of various genotypes differ on stimulation with *B. afzelii* and the human pathogen *Streptococcus pyogenes*, captured with transcriptome sequencing of spleen cells. In Paper IV, the analysis is focused on various genotypes of *TLR2*, an immune gene under balancing selection in bank voles and associated with infection prevalence of *B. afzelii* in the wild. A stimulation-specific effect of *TLR2* on immune response was found, where the magnitude of immune response to *B. afzelii*, but not *S. pyogenes*, depends on *TLR2* expression level in a *TLR2* genotype-specific way. In Paper V, trade-offs at the *cis*-regulatory level between the response to *B. afzelii* and *S. pyogenes* was tested by searching for polymorphisms where the alleles are expressed differently to these two stimulations. Abundant *cis*-regulatory variation for responses to the two bacteria was found, but there was no evidence for trade-offs. In summary, this work pushes our knowledge of what immune genes can be expected to be under pathogen-mediated selection, as heretofore understudied categories of immune function showed signs of selection. A novel basis – the combination of genotype and expression – was uncovered for functional effects of polymorphic genes. Finally, there was no evidence for trade-offs between responses to diferent pathogens. Investigating the nature of trade-offs in the immune system further would be necessary towards understanding the causes and consequences of pathogen-mediated selection.

Key words: pathogen-mediated selection, bank voles, *Myodes glareolus*, immune system, balancing selection, immune response, rodents, positive selection, transcriptomics

Classification system and/or index terms (if any) Supplementary bibliographical information

Language: English **ISSN and key title**:

ISBN: 978-91-8039-994-4 (print)

978-91-8039-995-1 (pdf)

Recipient's notes **Number of pages**: 55

Price Security classification

Signature Date 2024-04-23

Pathogen-mediated selection in the immune system of rodents

Exploring selection targets, functional effects and trade-offs

Mridula Nandakumar

LUND
UNIVERSITY

Coverphoto by Mridula Nandakumar

Back cover is bank vole TLR2 protein structure (both haplotypes) generated by AlphaFold, superimposed over mouse Tlr2 crystal structure

‘

Faculty of Science

Department of Biology

ISBN 978-91-8039-994-4 (print)

978-91-8039-995-1 (pdf)

Printed in Sweden by Media-Tryck, Lund University

Lund 2024

“The old that is strong does not wither,
Deep roots are not reached by the frost.”

--J.R.R. Tolkien, The Fellowship of the Ring

Table of Contents

Abstract

A small proportion of microbial diversity in the world – the pathogens – causes disease and plays an important role in shaping evolution of the host, especially the immune system. Pathogens create pressure on host immunity to evolve in numerous ways, most commonly by increasing divergence between species (positive selection) or increasing polymorphisms within a population (balancing selection). Many evolutionary processes contribute to these two modes of pathogen-mediated selection, however, trade-offs between different traits, for example responses to different pathogens, are intrinsic to balancing selection, a major focus of this thesis. Across five papers, questions related to which immune genes are under selection, how this translates to an effect on the immune response and what trade-offs occur, are addressed using rodents as study system.

Paper I utilises genomes from 30 rodent species across the world to identify signatures of positive selection in the genes of the immune system and tests if immune function is an important determinant for a gene to be under positive selection. In general, all functional categories of immune genes except intracellular signalling genes were significantly enriched for signs of positive selection, as compared to a control gene set. This effect was significant even after accounting for potential confounding factors like gene expression levels and protein-protein interactions.

Paper II focuses on a local population of bank voles in Sweden to look for signatures of balancing selection in the complement system – a branch of innate immunity that is among the first responders to an infection. One complement gene, *FCNA*, was found to be under strong balancing selection. Paper III tests for the functional effects of *FCNA* polymorphism using association studies, where polymorphisms in this gene were found to be associated with natural infections of *Borrelia afzelii*, a common pathogen for bank voles.

Papers IV and V look at how the immune response of bank voles of various genotypes differ on stimulation with *B. afzelii* and the human pathogen *Streptococcus pyogenes*, captured with transcriptome sequencing of spleen cells. In Paper IV, the analysis delves into how various genotypes of *TLR2*, an immune gene under balancing selection in bank voles, contributes to genotype-dependent infection prevalence of *B. afzelii* in the wild. Results indicated that there was stimulation-specific effect of *TLR2* on the immune response, where the magnitude of immune response to *B. afzelii,* but not *S. pyogenes,* depended on *TLR2* expression level in a *TLR2* genotype-specific way. In Paper V, trade-offs at the *cis*-regulatory level between the response to *B. afzelii* and *S. pyogenes* was tested by searching for polymorphisms where the alleles are expressed differently to these two stimulations. Abundant *cis*-regulatory variation for responses to the two bacteria was found, but there was no evidence for trade-offs.

In summary, this body of work pushes our knowledge of which types of immune genes can be expected to be under pathogen-mediated selection, as heretofore understudied categories of immune function show signs of pathogen-mediated selection. It also uncovers an exciting and novel basis – combination of genotype and expression – for functional effects of polymorphic genes. Finally, there was no evidence for trade-offs between responses to different pathogens. Investigating the nature of trade-offs in the immune system further would be necessary towards understanding the causes and consequences of pathogen-mediated selection.

Popular summary

Pathogens – including bacteria, viruses, fungi and parasites, cause disease in their host. Dealing with infections can be extremely damaging, and many hosts die or suffer long-term consequences like reduced fertility. Over time, the host immune system evolves to deal with pathogens better, either by improved pathogen detection or by responding to the infection better. How this evolution that is brought about by pathogens is driven (*i.e.,* pathogen-mediated selection), is a source of ongoing interest for many biologists. This can be studied in many ways and here we look for signatures of evolution in the genes responsible for the immune system. The research in this thesis utilises rodents to study how pathogens have shaped the evolution of the immune system. Rodents are extremely suited to ask questions related to host immunity evolution because they harbour numerous pathogens – amongst the highest in mammals– and so would have had to undergone drastic changes to live with these pathogens. Many questions in the thesis are specifically focused on a local population of bank voles in Sweden, an abundant wild rodent present throughout Europe. Bank voles naturally harbour *Borrelia afzelii,* a bacterium transmitted by ticks, that causes Lyme disease in humans. Using this system, questions regarding how the population has evolved to respond to *B. afzelii* infections can be tested.

This work explores evolution of the host immune system in three ways. First, my colleagues and I identify which immune genes evolve in response to pathogens in their environment. Just like the species diversity of many animals, pathogen diversity also differs across the world. This can have strong effects on host evolution, as only individuals with the immune gene variants that provide most benefit against local pathogens will survive and reproduce. As a result, subsequent generations of the host will carry the same profile of immune gene variants – a process called positive selection. This can drive increased genetic differences between species. The first paper of the thesis looks for immune genes that show signs of positive selection using many rodent species from around the world. Many immune genes show signs of positive selection and an important determinant for which immune genes are under positive selection is their function. On the other hand, hosts can also evolve by increasing the genetic diversity of their immune genes, so that they can detect a broad range of pathogens, thereby balancing the response to multiple infections. This is enabled by a process called balancing selection, which increases genetic diversity within a species, producing many different variants of the same gene in the population. The second paper of the thesis searches for immune genes with signs of balancing selection in a part of the immune system that acts as the first line of defence against infections, called the complement system. Here, one gene, called *FCNA*, is under balancing selection.

The second aim of this work is to understand how evolution of the immune genes affects the immune response and resistance against infections. In the third chapter,

genetic variants of the gene *FCNA* are checked for associations with natural infections of *B. afzelii* in wild bank voles, which showed that some variants of this gene are indeed better than other variants at providing resistance to *B. afzelii* infections. This study demonstrates that evolution of immune genes can have real-life consequences against infections. Papers IV and V utilise data from a large laboratory experiment, where immune cells of wild bank voles are challenged with two pathogens (*B. afzelii* and *Streptococcus pyogenes*). This would test how evolution of immune genes influences the immune response by capturing the amount of gene expression in response to the two pathogens. Genes that respond to the pathogens are largely similar. However, my colleagues and I uncovered an important effect of a gene under balancing selection called *TLR2,* that has many genetic variants in the bank vole. The analyses showed that the extent of the immune response to the pathogens depends on what genetic variant of *TLR2* is present and to what extent *TLR2* is expressed. This is specific to only *B. afzelii* response, a native pathogen to bank voles, but not to the novel pathogen *S. pyogenes*, indicating that how evolution of the immune genes affects the immune response can be dependent on if the host is adapted to the pathogen or not.

The final aim was to understand what causes balancing selection in the first place. Balancing selection is a response to diversity in the environment, as the immune system must juggle cost and benefits to multiple stressors. With this juggling act, genetic variants of immune genes that are produced as a result of balancing selection would be beneficial in one context but detrimental in another, resulting in a trade-off. The potential context for this trade-off act can be the response to multiple pathogens. Using gene expression data in response to *B. afzelii* and *S. pyogenes,* variants of immune genes were checked to see if they were expressed differently based on the pathogen. If the same genetic variant of immune gene responded well to one pathogen but poorly to another, it would indicate that the evolution of the immune gene variants was due to trade-off in the response to multiple pathogens. However, no effect of the same genetic variant being beneficial against one pathogen but detrimental to another was observed. This tells us that other juggling acts, perhaps to balance the response between infection and autoimmunity for example, is more important in contributing to the evolution of the immune system.

In conclusion, this works shows that numerous immune genes have evolved as a response to selection from pathogens and evolving with pathogens has strong effects on how the host immune system responds to infections.

List of Papers

Paper I

Mridula Nandakumar, Max Lundberg, Fredric Carlsson, Lars Råberg. Positive selection on mammalian immune genes – disentangling effects of gene function from expression and protein-protein interactions ***(Manuscript)***

Paper II

Mridula Nandakumar, Max Lundberg, Fredric Carlsson, Lars Råberg. 2023. Balancing selection on the complement system of a wild rodent. *BMC Ecology and Evolution.* 23, 21. DOI: 10.1186/s12862-023-02122-0 ***(Published)***

Paper III

Elin Laike Åsberg, **Mridula Nandakumar**, Lars Råberg. Polymorphism in the complement system gene *FCNA* is associated with *Borrelia* infection in a wild rodent ***(Manuscript)***

Paper IV

Mridula Nandakumar, Mehrnaz Nouri, Christine Valfridsson, Fredric Carlsson, Lars Råberg. Unravelling the effect of balanced polymorphisms in bank vole *TLR2* on immune responses using *ex vivo* assays ***(Manuscript)***

Paper V

Mridula Nandakumar, Max Lundberg, Fredric Carlsson, Lars Råberg. *Cis*-regulatory variation for immune responses to two bacterial pathogens in a wild rodent: testing for trade-offs ***(Manuscript)***.

Author contributions

Paper I

LR, **MN** and FC designed the study. LR and **MN** performed fieldwork and **MN** performed RNA extractions. **MN**, LR and ML performed all analyses. LR and **MN** wrote the manuscript, with input from all authors.

Paper II

LR, **MN**, FC and ML designed the study. LR conducted fieldwork and labwork. ML performed whole-genome resequencing analyses. **MN** performed Sanger sequencing, selection analyses and wrote the first draft of the manuscript. All authors contributed to revising the manuscript and approved the final version.

Paper III

LR and **MN** designed the study. ELÅ and **MN** performed genotyping. LR performed field work and qPCR. ELÅ and LR analyzed the data. LR and **MN** wrote the manuscript

Paper IV

LR, **MN** and FC designed the study. LR performed fieldwork. **MN** performed lab work with contributions from MNo and CV. **MN** analysed the data and wrote the first draft of the manuscript with input from all authors.

Paper V

LR and **MN** designed the study. LR conducted fieldwork. ML performed whole-genome resequencing analyses. **MN** carried out lab work and performed ASE analysis. LR and **MN** wrote the manuscript with input from all authors.

Lars Råberg (LR), Mridula Nandakumar (MN), Fredric Carlsson (FC), Max Lundberg (ML), Elin Laike Åsberg (ELÅ), Mehrnaz Nouri (MNo), Christine Valfridsson (CV)

Abbreviations

ASE	Allele specific expression
FCNA	Ficolin A
GWAS	Genome Wide Association Studies
HBB	Haemoglobin subunit beta
NFκB	Nuclear Factor κ Subunit B
PAMP	Pathogen Associated Molecular Pattern
PRR	Pattern Recognition Receptor
TLR	Toll like receptor 2
MHC	Major Histocompatibility Complex

Pathogen-mediated selection

Microbes are ubiquitous in any environment. While microbial diversity is astronomical, it is only a tiny fraction of this group that cause disease – in humans, one in a billion microbial species is estimated to be pathogenic (Woolhouse & Gowtage-Sequeria, 2005; Microbiology by numbers, 2011; Balloux & van Dorp, 2017). Yet it is this fraction of microbial pathogens, together with parasites (henceforth collectively called pathogens), that affects host fitness negatively and has a significant effect on host evolution. A classical illustration for the sheer impact pathogens can have on host evolution is the case of myxoma virus infections in rabbits. A poxvirus, myxoma is found naturally in American rabbits and causes very little damage to the host. When it was introduced in the 1950s to naïve rabbit populations in Europe and Australia as a population control measure, infections caused dramatic declines in rabbit populations, with a 99% mortality rate killing millions of individuals (Fenner & Ratcliffe, 1965). Continued exposure to the virus caused rabbits to evolve over time resulting in populations that were considerably more resistant and showed milder symptoms. Recent genetic analyses have revealed rapid changes in allele frequencies of a number of immune genes (Alves *et al.*, 2019). Together with many more instances of pathogens influencing host evolution (*e.g.* malaria infections in birds and humans (van Riper III *et al.*, 1986; Kwiatkowski, 2005)*, Mycoplasma* infections in American house finches (Hochachka & Dhondt, 2000)), the idea that pathogens serve as one of the strongest sources of selection pressure, and shape immune response to infection has become prevalent (Fumagalli *et al.*, 2009b, 2011). How does selection from pathogens influence evolution of the host? What changes do host populations undergo to accommodate living with its pathogens? What evolutionary processes underlie host adaptation to pathogens and what consequences can it entail? These are some of the broad questions discussed in this thesis.

A quick guide to the vertebrate immune system

In the context of pathogen-mediated selection, most changes to genetic diversity in the host can be expected to take place in the genes of the immune system, as it is the primary point of interaction between the host and pathogen. The vertebrate immune

system consists of two branches, the innate and the adaptive arms of immunity, involving multiple immune organs and cell types (Figure 1).

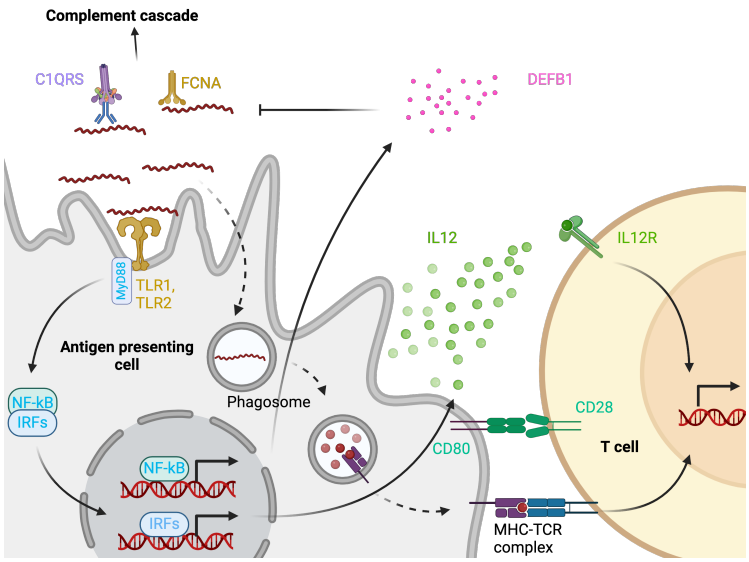


Figure 1. A simplified overview of the immune response.
Pathogens are detected with the help of innate immune receptors like toll-like receptor (TLR) 2, triggering a cascade that results in upregulation of many genes via transcription factors like nuclear factor κ subunit B (NFκB). This results in synthesis of cytokines like interleukin (IL) 12, that, in combination with antigen presentation by major histocompatibility (MHC) molecules and costimulation by cluster of differentiation (CD) 80, activates effector cells like T-cells. Detection by innate receptors also lead to secretion of effector proteins like the antimicrobial peptide defensin beta 1 (DEFB1) that lyses the pathogen. Pathogens can also be detected via the innate mechanism of the complement system, through recognition via ficolin A (FCNA) or complement components C1QRS. In the adaptive arm of immunity, phagocytosis of a pathogen can result in antigen presentation through MHC molecules and recognition through T-cell receptor (TCR). Created with BioRender.com

Innate immunity is present in various forms across different kingdoms and phyla and serves as the first barrier against infections (Kimbrell & Beutler, 2001; Fujita, 2002; Buchmann, 2014). The innate immune system spans a wide variety of immunological components, from physical barriers such as the skin and mucosal membranes to secreted serum proteins of the complement system and immune cells that recognise microbes in a relatively non-specific manner. Generally, incoming infections encounter the soluble proteins of the complement system, which can initiate a signalling cascade by three different ways, leading to *(i)* inflammation via the recruitment of innate immune cells, *(ii)* opsonisation of the invading pathogen, and *(iii)* direct lysis of the invading organism via formation of the membrane attack complex (Ricklin *et al.*, 2010). The recognition of pathogens can also be driven by a group of intracellular and membrane-bound receptors called pattern-recognition receptors (PRRs). PRRs recognise conserved motifs present in microbes, called

pathogen-associated molecular patterns (PAMPs; Akira *et al.*, 2006). On recognising the ligand, PRRs initiate a signalling cascade, involving many adaptor proteins and transcription factors, ultimately resulting in synthesis of cytokines and chemokines causing an inflammatory response (Takeuchi & Akira, 2010). These cues also promote the activation of the adaptive immune response, which requires several days to weeks to mount a targeted response against a specific pathogen (Palm & Medzhitov, 2009). The adaptive immune response is only present in vertebrates and is (in jawed vertebrates) characterised by responses produced by T-cells and B-cells (Cooper & Alder, 2006; Flajnik & Kasahara, 2010; Boehm & Swann, 2014). Antigens from the pathogens are presented with the help of major histocompatibility (MHC) molecules that result in effector responses including components of both innate and adaptive immunity (Medzhitov, 2007).

Modes of pathogen-mediated selection

Various approaches exist to measure the extent of selection in populations. For instance, selection in response to infections can be understood by investigating how variation in phenotypic traits such as immune response relate to fitness (Råberg & Stjernman, 2003). Alternatively, selection can be inferred by analysing genetic/genomic data (Sironi *et al.*, 2015). In any case, selection requires variation to act on. Variation in gene sequences can be generated by numerous processes, most importantly through mutations due to errors in sequence replication during meiosis. Much of this nucleotide variation is selectively neutral and can vary in frequency just by chance from genetic drift known as the neutral theory of molecular evolution (Kimura, 1968). If any new variants reduce host fitness, it is purged from the population by purifying selection (Figure 2). Based on deviations from the neutral theory, selection can be detected. The mode of selection can be affected by factors unrelated to pathogen pressure. For example, many genes are evolutionarily constrained due to involvement in many physiological functions (pleiotropy), breadth and depth of gene expression and longer gene length (Zhang & Yang, 2015; Soni & Eyre-Walker, 2022; Williams *et al.*, 2023). The following section will focus on detecting selection from genetic/genomic data.

Positive selection

Most studies attempting to understand adaptive response to changes in the environment focus on positive selection, which drives species level differences in adaptation. Positive selection underlies directional selection at the phenotypic level, which causes divergence between species. Signatures of positive selection can be detected where the advantageous beneficial allele dominates the population frequency and eventually reaches fixation (Figure 2). Classical methods that detect

positive selection capitalise on quantifying the amount of non-synonymous to synonymous nucleotide changes in coding sequences, using sequence information from multiple species (Hejase *et al.*, 2020; Charlesworth & Jensen, 2021). Other footprints left by positive selection include reduced genetic diversity in the regions surrounding the selected polymorphisms, as neutral variants that are in linkage disequilibrium would also be swept to fixation – a phenomenon called genetic hitchhiking (Kim & Nielsen, 2004). Contemporary methods that detect ongoing/recent positive selection make use of such signals, observable by using intraspecific (*i.e.* within a species) data, where regions under selection are present as long haplotypes with low genetic diversity due to lack of recombination (Voight *et al.*, 2006).

With the availability of multiple vertebrate genomes, many studies have focused on understanding patterns to regions of the genomes that are under positive selection. These studies reveal that sites in both coding regions and non-coding regions such as introns, regulatory regions and transposable elements are subject to positive selection (Kousathanas et al., 2011; Nellåker et al., 2012; Smith et al., 2013). Amongst the primary targets of positive selection are genes responsible for immunity, olfaction, pheromone detection, spermatogenesis, diet and metabolism (Voight et al., 2006; Kosiol et al., 2008; Staubach et al., 2012; Harris & Munshi-South, 2017; van der Lee et al., 2017; Shultz & Sackton, 2019; Roycroft et al., 2021).

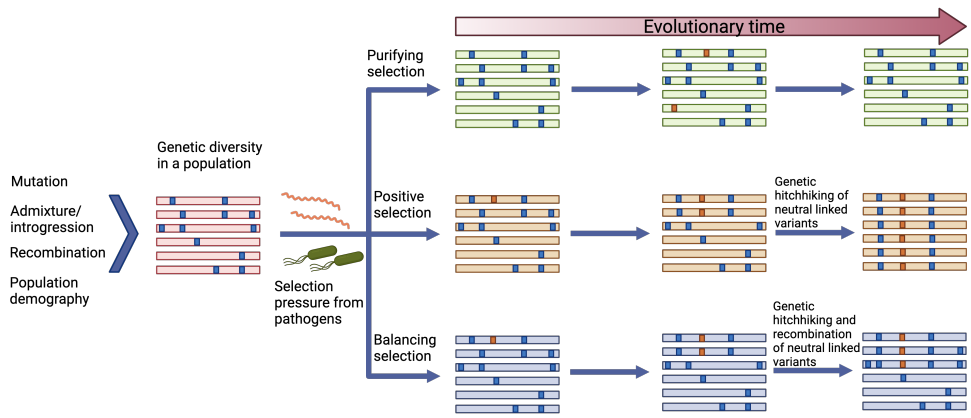


Figure 2. Modes of pathogen-mediated selection.
Genetic variation can be brought about by various processes, providing scope for pathogen-mediated selection. Different modes of selection can affect level of variation, detectable using genetic/genomic data. Each row represents sequence information from one individual in the population. Dark squares indicate existing mutations, orange squares indicate new mutations for selection to act on. Figure inspired by (Quintana-Murci & Clark, 2013; Fumagalli & Sironi, 2014), created with BioRender.com

Balancing selection

While positive selection focuses on interspecific (*i.e.,* between different species) divergence, intraspecific variation in adaptation can be determined by balancing selection, which helps maintain genetic diversity within a population and can result in diversifying selection at the phenotypic level. Affected loci are characterised by high genetic diversity, with multiple alleles present in a population simultaneously at intermediary frequencies.

Balancing selection has been well studied in certain contexts, with the *R* genes in plants drawing considerable attention (Van der Hoorn *et al.*, 2002). Other examples include self-incompatibility genes in plants, complementary sex determination genes in honey bees (Cho *et al.*, 2006; Lechner *et al.*, 2014) and blood group antigens in mammals (Ferrer-Admetlla *et al.*, 2009; Fumagalli *et al.*, 2009a; Linnenbrink *et al.*, 2011). Nonetheless, establishing balancing selection is a non-trivial endeavour, as only strong selection events can be detected. However, this effort has been greatly facilitated by newly developed methods to detect balancing selection at a genome-scale (DeGiorgio *et al.*, 2014; Siewert & Voight, 2017, 2020; Bitarello *et al.*, 2018). These tests use different signals of balancing selection and have enabled large-scale detection of balancing selection with the availability of population genomics data. Some of the regions identified to be under balancing selection include immunity genes, keratin genes and cell membrane transporters (Andrés *et al.*, 2009). On average, exonic regions encoding proteins sequences are enriched for signatures of balancing selection (Bitarello *et al.*, 2018); however such signatures can also be found in non-coding regions, as exemplified by promoter region polymorphisms in *DEFB1* that is subject to balancing selection (Cagliani *et al.*, 2008).

Targets of pathogen-mediated selection

Selection in the immune system

Much of the focus in understanding sequence evolution of immunity genes have focused on positive selection. Extensive studies across different systems identify a number of genes in the vertebrate immune system such as those encoding PRRs, cytokines, chemokines and immune receptors, along with signalling genes involved in immunity, to be under positive selection (Kosiol *et al.*, 2008; van der Lee *et al.*, 2017) and indeed this pattern is conserved across divergent vertebrate clades (van der Lee *et al.*, 2017; Shultz & Sackton, 2019). In the case of balancing selection, genes involved in antigen presentation, primarily the major histocompatibility complex (MHC) locus in vertebrates, has been of special interest due to the incredible variability observed in these loci (Hedrick, 2002; Prugnolle *et al.*, 2005).

Other genes identified that show signatures of balancing selection include viral interacting genes such as rodent 2'-5' oligoadenylate synthetase 1B (*Oas1b;* Ferguson *et al.*, 2008), cytokines and cytokine receptors like the interleukin (IL) 7 receptor (*IL7R*) and *IL18* receptor antagonist (*IL18RAP*) in humans, *Il2* and *Il1b* in rodents (Fumagalli *et al.*, 2009b; Turner *et al.*, 2012), anti-microbial peptides like *DEFB1* in humans (Cagliani *et al.*, 2008), various PRRs in rodents (Lundberg *et al.*, 2020) and other genes involved in antigen presentation like *ERAP2* and *TAP2* in humans (Andrés *et al.*, 2010; Cagliani *et al.*, 2011).

Broad patterns to immune genes under selection

Many immune genes are subject to pathogen-mediated selection, and closer analysis of their function indicates many parallels to which type of genes are favoured by selection. Principally, host genes that show signatures of pathogen-driven selection can be categorised into four groups (Quintana-Murci & Clark, 2013; Sironi *et al.*, 2015). First, genes that are involved in pathogen detection and presentation, such as the PRRs and MHC, may evolve by increasing genetic diversity to enrich pathogen detection and presentation capabilities (Sackton *et al.*, 2007; Vinkler *et al.*, 2014; Velová *et al.*, 2018; Lundberg *et al.*, 2020; Radwan *et al.*, 2020). Second, host genes that are either manipulated by pathogens to evade immunity, such as the complement regulator Factor H (*CFH*), or used to gain access to host cells such as cluster of differentiation (*CD)* 4, are sites of host-pathogen conflict and subject to selection (Cagliani *et al.*, 2016; Russell *et al.*, 2021). Third, host genes that interact with pathogens such as immune effectors like antimicrobial peptides, complement proteins, etc., show signatures of pathogen-mediated selection (Semple *et al.*, 2005; Cagliani *et al.*, 2008, 2016). Finally, genes that are not involved in immunity directly, but help mitigate infection or have fitness consequences are known to be under selection. Instances of such pattern includes the haemoglobin subunit beta (*HBB*) and glucose-6-phosphate 1-dehydrogenase (*G6PD*) variants in protecting against malaria, and galactoside alpha-(1,2)-fucosyltransferase 2 (*FUT2*) polymorphisms in protecting against norovirus in humans (Kwiatkowski, 2005; Ferrer-Admetlla *et al.*, 2009).

It should be noted, however, that certain classes of immune genes might also be more likely to show higher genetic variability due to relaxation in evolutionary constraints, rather than balancing selection. Mostly, this is linked to the function (*i.e.,* essentiality and redundancy) of the gene (Quintana-Murci & Clark, 2013). For example, functions performed by cell surface PRRs are often redundant, and these genes show high variability, with up to 16% of individuals harbouring missense mutations (Barreiro *et al.*, 2009). On the other hand, cytosolic PRRs that recognise cytosolic PAMPs such as nucleic acids show reduced variability as microbial nucleic acids are generally conserved; these receptor are likely indispensable for immune function as no missense mutations are tolerated (Barreiro *et al.*, 2009;

Quintana-Murci & Clark, 2013). Likewise, adaptor genes such as myeloid differentiation primary response protein (*MYD88*) that is common to the signalling cascade triggered by TLRs show very little variability in their coding sequences and are thought to evolve under strong purifying selection (Nakajima *et al.*, 2008). This raises an important question of which immune genes are under selection due to their function, as opposed to factors unrelated to pathogen pressure.

Approaches to study functional effects of variants under selection

Selection tests at the genetic level only offer insights into the mode and targets of selection but provide very little information as to why some variants are favourable. Determining a functional basis for how a gene under selection operates is necessary to understand why there is selection on the gene in the first place. Here, two approaches to study the functional effects of selected genes are outlined.

Association studies

One approach to understand the functional context in which selection is relevant is to use association studies. With the advent of affordable sequencing methods, large-scale associations of variants with disease phenotypes can be performed using genome-wide association studies (GWAS). Using thousands of samples integrated with selection scans, insights into the functional relevance of the selected variants can be inferred, as demonstrated by *HBB* allele association with malaria (Band *et al.*, 2019). However, this approach is successfully applicable only in very specific situations and can only pick up associations when the selection is strong, of intermediary time scale, has monogenic basis for resistance and in endemic infections (Karlsson *et al.*, 2014). Further complications arise when there are multiple pathogen strains that are locally adapted, existence of population genetic structure, varying disease phenotype definitions, etc., as exemplified by the low power of GWAS in tuberculosis studies (Stein, 2011; Farhat *et al.*, 2019). In such cases, it is often more suitable to perform associations using candidate gene approaches, which are less resource intense than GWAS, as only hundreds of samples genotyped at a few loci are necessary. It has been observed however, that candidate gene approaches are generally poorly replicated and often fail to be picked up in GWAS (Siontis *et al.*, 2010). In both cases, further functional characterisation of the allele under selection using a more targeted approach is required to pinpoint how variants affect function and phenotypes (Mitchell-Olds *et al.*, 2007; Karlsson *et al.*, 2014).

Molecular studies

The functional advantage provided by variants is determined by their genomic location, either affecting interaction capabilities of proteins or regulating gene expression (Dean & Thornton, 2007; Mitchell-Olds *et al.*, 2007). Variants in coding regions can alter structural properties of proteins, potentially modifying interaction capabilities with ligands or other proteins. Such modifications are thought to be common in PRR and MHC, which are characterised by increased genetic diversity in ligand binding or protein interaction domains (Ohto *et al.*, 2012; Vinkler *et al.*, 2014; van der Lee *et al.*, 2017; Velová *et al.*, 2018; Arora *et al.*, 2020). Alternatively, genes under selection can also be variable in in non-coding regions. This can translate into dissimilarities in how the different variants influence gene transcription and stimulus response.

Regardless of whether a variant under selection is in coding or non-coding regions, comparative studies employing individuals with different genotypes can be used to check for the effects of different alleles on phenotypes such as expression response, or survival outcome to different infections. In an immune context, large scale transcriptome data in response to infection stimulation *in vitro/ex vivo* has been the focus of many recent studies that have attempted to understand how selection from pathogens has shaped immune response of different human populations. These studies use various approaches including expression Quantitative Trait Loci (eQTL; effect of a variant on total gene expression), allele-specific expression (ASE; expression of different alleles at variant locus) and alternate splicing (Andrés *et al.*, 2010; Grossman *et al.*, 2013; Nédélec *et al.*, 2016; Quach *et al.*, 2016; Sams *et al.*, 2016; Harrison *et al.*, 2019; Rotival *et al.*, 2019; Klunk *et al.*, 2022). Such functional characterisation has mainly been performed in the case of humans within the context of positive selection, but similar studies in other vertebrate systems or in the case of balancing selection are rare.

Evolutionary processes driving selection

Just like their hosts, pathogens too are constantly evolving to adapt to various changes in their environment, including changes in the host. This in turn renews selection pressure for host evolution, producing a dynamic cycle of adaptations in both host and pathogen. Based on the patterns of the adaptation dynamics, the mode of selection is determined. Evolutionary processes driving positive selection is often straightforward and is an outcome of response from a single source of strong selection pressure; either novel pathogens or novel pathogen strategies (corresponding to arms race scenario of host-pathogen coevolution) imposes strong selection pressure in a consistent direction. On the other hand, balancing selection is more complicated, as it is brought about by several (often overlapping) sources

of selection pressures that maintain genetic diversity in the host. In general, there are numerous mechanisms that drive balancing selection and it is often hard to distinguish between the different processes (Spurgin & Richardson, 2010; Fijarczyk & Babik, 2015). Firstly, earlier studies of balancing selection often favoured heterozygote advantage (overdominance) as a mechanistic explanation for balancing selection, whereby heterozygous individuals are more fit than homozygotes and are therefore favoured under selection (Hedrick, 2012). In this case of allelic overdominance, more heterozygotes will be found in the population than homozygotes. Secondly, fluctuations in the selective force due to variability in the environment over space and time, such that different alleles are advantageous depending on the environmental context can also lead to an increase in genetic diversity (Bergland *et al.*, 2014; Abdul-Rahman *et al.*, 2021). Finally balancing selection can be brought about by negative frequency-dependent selection, whereby alleles that occur in high frequencies are disadvantageous due to continuous pathogen adaptation to escape the most common host genotype (Gigord *et al.*, 2001; Fitzpatrick *et al.*, 2007). This leads to an increase in the frequency of the alternative allele. Once the selected allele becomes common in the population, the pathogen adapts to it making it disadvantageous and decreases in frequency; this leads to a cyclical dynamic in the allele frequency, where no selected allele reaches fixation but rather fluctuates in frequency with time. This is characteristic of host-pathogen coevolution following the so called 'Red Queen' dynamics. Which of these evolutionary forces is the main cause of balancing selection has been difficult to establish (Radwan *et al.*, 2020).

Irrespective of the mechanism, trade-offs with a genetic basis are inherent to evolutionary processes maintaining balancing selection and can manifest at different levels. Firstly, there can be trade-offs between infection and physiology. This is exemplified in the case of the haemoglobin gene *HBB*, showing a trade-off between resistance to malaria parasite and anaemia due to abnormal red blood cells. Secondly, a trade-off between infectious disease and autoimmunity/ immunopathology has been shown to exist, as in the case of the antiviral gene *IFIH1*. This viral sensing gene is thought to be under balancing selection and variants show protective effects against viral infections at the cost of increased self-recognition causing autoimmunity (Fumagalli *et al.*, 2010; Gorman *et al.*, 2017). Finally, polymorphisms under balancing selection can provide trade-offs to different pathogens, either to different strains of the same pathogen or pathogens of various types. This has been documented in invertebrates, exemplified by the plankton *Daphnia magna* and susceptibility to different strains of the pathogen *Pasteuria ramosa* (Bento *et al.*, 2017; Bourgeois *et al.*, 2021). Trade-off to different infections associated with a balanced loci in vertebrates has been demonstrated to only a limited extent (Råberg, 2023). To what extent relative trade-offs play a role in maintaining balancing selection is not well understood.

A

B

C

Fun Work Picture 1. How to identify bank voles: a comprehensive field guide.
Despite being an abundant rodent, it can be quite tricky to identify bank voles. The unsuspecting field biologist, sleepily checking traps at 6 am for bank voles, must stay alert for potential bank vole mimics. A) A sneaky blue tit trying to desperately pass as a bank vole. It should perhaps invest in a costume. B) More like it, but the blue tit has accidentally bought a harvest mouse costume. C) Finally the masquerade is complete, a bank vole! Photo: Lars Råberg (A & C).

Studying pathogen-mediated selection in rodents

Rodents provide a fantastic system to study pathogen-mediated selection. Not only are they the most speciose group of mammals present in large numbers, they also harbour a broad range of pathogens, a substantial portion of which are zoonotic (Mollentze & Streicker, 2020). All of this has driven an increased interest in using rodents as model systems for understanding host adaptation to various pathogens. Coupled with functional information from the house mouse (*Mus musculus*), valuable insights into how the immune system evolves in response to pathogen pressure can be understood.

The *Myodes glareolus- Borrelia afzelii* nexus

The primary focus of this thesis utilises the wild rodent *Myodes glareolus* (bank voles) present in the mixed deciduous forests around Revingehed, 20 km east of Lund, in southern Sweden. The bank vole is found abundantly throughout Europe and some parts of Asia (Stenseth, 1985). Many different pathogens infect bank voles, including viruses, helminths and bacteria (Johansson *et al.*, 2008; Andersson & Råberg, 2011; Clough & Råberg, 2014). A focal pathogen for this species are the bacterial spirochaetes of the *Borrelia burgdorferi sensu lato* complex, comprising of 22 genospecies (Waindok *et al.*, 2017). One among the *Borrelia burgdorferi sensu lato* complex present in Europe is *B. afzelii,* for which rodents are the preferred reservoir host (Piesman & Gern, 2004). This extracellular pathogen (atypical Gram-negative) scavenges nutrients from its host and is found at a prevalence of ca. 25% in our bank vole population (Andersson *et al.*, 2013). *B. afzelii is a* tick-borne pathogen, transmitted by *Ixodes ricinus* in Sweden (Wilhelmsson *et al.*, 2016). The tick lifecycle requires three hosts for completion, and feed on single host during each life stage. It is during the larval-nymph life stage transition that they transmit *B. afzelii* to different rodent hosts. Within the tick, the bacteria grow in the gut, migrate to the salivary gland, from where they are released into the blood stream of a new vertebrate host as ticks feed (Kurokawa *et al.*, 2020). Adult ticks feed on larger mammals like deer, which is an incompetent host for *B. afzellii*. Humans too serve as dead end hosts to *B. afzelii* following a tick bite. If

untreated, symptoms of Lyme disease sets in, and may lead to colonisation of internal tissues including joints, heart and the nervous system (Stanek *et al.*, 2012). In humans, *Borrelia* infections can also recur, while most rodent hosts are known to be chronically infected with *B. afzelii* (Gern *et al.*, 1994; Jacquet *et al.*, 2016).

Using bank voles to study pathogen-mediated selection

The system of *B. afzelii* infections in bank voles offer a good framework to explore pathogen-mediated selection and its effect on host adaptation. Bank voles and the *Borrelia burgdorferi sensu lato* complex have co-existed for long periods of time. While no estimate exists for the evolutionary origin of *B. afzelii* in Eurasia, it can be speculated to have circulated in natural populations for fairly long based on phylogenetic analyses and estimates of origin of other *Borrelia* species; in America, *Borrelia burgdorferi* origins predates the Last Glacial Maximum *i.e.,* > 20,000 years ago (Qiu & Martin, 2014; Walter *et al.*, 2017). Thus, it is possible that bank voles and its spirochaete pathogen have accumulated changes that aid in adaptation to each other. Indeed, studies have shown that *Borrelia burgdorferi sensu lato* display complex and dramatic shifts in gene expression patterns to evade host immunity as the bacteria switches between the vertebrate and invertebrate host (Kurokawa *et al.*, 2020). A prime example is the switch in outer surface protein (Osp) expression, such as OspC upregulation in the vertebrate host (Grimm *et al.*, 2004). *ospC* is highly variable and is used to characterise the different *B. afzelii* strains. Within the bank vole population at Revingehed, seven common *B. afzelii* strains have been identified (Andersson *et al.*, 2013).

Successful infections are characterised by the pathogen's ability to overcome the host immune defences. Indeed, *Borrelia* is adept at escaping both innate and adaptive immunity with the help of numerous proteins. For example, evading the complement system of the innate immunity is critical to host colonisation, and complement escape is achieved in a number of ways (Lin *et al.*, 2020). Typically, complement regulating acquired surface proteins (CRASPs), a group of borrelial proteins that help the spirochaete evade the complement system, are produced. In another instance, the adaptive immune system is evaded with the help of the VMP-like sequence (*vls*) locus, which undergoes recombination to randomly combine and produce multiple variants of the VlsE protein (Bankhead & Chaconas, 2007). These processes, along with the high diversity observed in certain genes such as *ospC* suggests that the bacteria have acquired numerous adaptations to establish infections in the host in response to host immunity.

Bank voles too demonstrate adaptation to the spirochaete in numerous ways. In infected bank voles, while *B. afzelii* infections are often disseminated to other tissues such as joints and heart, no signs of pathology (such as inflammation) were observed

in chronically infected voles (Zhong *et al.*, 2019). Using transcriptome sequencing from spleens of infected bank voles, it was found that the general pattern of infection response was downregulation of immunological pathways during the chronic phase of infection, including *IL6* signalling and complement system (Zhong *et al.*, 2020).

Bank voles also display high levels of genetic variability within the study population. A previous study investigated the targets and patterns of innate immune genes under balancing selection using whole genome resequencing methods (Lundberg *et al.*, 2020). The study identified that PRRs were more likely to be targets of balancing selection. Furthermore, PRRs recognising components of bacterial and fungal cell surfaces showed an increased likelihood of balancing selection compared to PRRs recognizing features of pathogen nucleic acids. Polymorphism at one of the genes identified to be under balancing selection, the PRR toll-like receptor (TLR) 2 recognising lipoproteins, is known to be associated with the prevalence of *B. afzelii* (Tschirren *et al.*, 2013). Nonetheless, the evolutionary pressures leading to balancing selection at *TLR2* is currently unknown.

Pathogen-mediated selection necessitates a fitness cost to infection, as otherwise both infected and uninfected individuals are equally fit. Host fitness can be evaluated in many ways, most commonly using mortality and reproductive success measures. Studies of fitness costs for bank voles infected with *B. afzelii*, while limited, show that infected individuals have no increased mortality but are rather penalised by lower reproductive success (Cayol *et al.*, 2018). This shows potential for pathogen-mediated selection to act in this system.

Fun Work Picture 2. Sweden is famous for its state of the art research facilities.
Here is Lars Råberg using a tree stump as a dissection table.

Aims of the thesis

Using rodents in general, and the bank vole-*Borrelia* study system in particular, questions related pathogen-mediated selection, mainly focused on innate immunity, are explored along the following themes:

1) Identifying targets and functional patterns of genes under pathogen-mediated selection (Papers I, II):
    a) What is the extent of nucleotide divergence in the immune genes across different rodent species? (Paper I)
    b) Are certain types of immune genes more likely to be positively selected? If so, is that a result of immune function as such, or other factors, like gene expression and protein-protein interaction? (Paper I)
    c) What is the extent of balancing selection in a key immune pathway, the complement system, in bank voles? (Paper II)
    d) Do we see patterns to complement genes under balancing selection? (Paper II)
2) Understanding the functional effects of genetic polymorphism on immune responses (Papers III, IV, V)
    a) Is there an association for a complement gene under balancing selection with natural infections of *B. afzelii* in bank voles? (Paper III)
    b) What mechanism underlies balancing selection at a pattern-recognition receptor? (Paper IV)
    c) Is there an effect of *cis*-regulatory variation on immune response to diverse pathogen stimuli? (Paper V)
3) What evolutionary processes underlie responses to diverse pathogens?
    a) Are there trade-offs in the *ex vivo* spleen expression profiles from bank voles with different TLR2 genotypes in response to diverse bacterial stimuli? (Paper IV)
    b) Are there trade-offs for allelic expression of genetic variants in response to diverse bacterial stimuli? (Paper V)

Contributions of the thesis towards understanding pathogen-mediated selection

Targets and functional patterns of genes under pathogen-mediated selection (Papers I, II)

Searching for genes under selection (Papers I and II)

The first two chapters mainly relies on the vast information generated through genomic sequencing and methods developed therein to systematically look for signatures of selection in the immune system. In both papers, a set of control genes are used to compare estimates against, to ensure that signatures of selection are not attributable to neutral processes and population demographics.

Estimates of divergence and positive selection

Rodents are the most species-rich mammalian clade, with two families – Cricetidae and Muridae, making up a large chunk of this diversity. Recent adaptive radiations to varied habitats across multiple continents (Fabre *et al.*, 2012) has presumably exposed these newly formed clades to a diversity of pathogens and adaptation to these events are expected to be reflected in their gene sequences. In paper I, my co-authors and I use 30 publicly available genomes across these two families to look for signatures of positive selection, using the classical method PAML (Yang, 2007).

Divergence, an estimate of species-level differences in sequence information, is defined as the ratio (ω) of the rate of non-synonymous substitutions (dN) to the rate of synonymous substitutions (dS). A higher rate of non-synonymous substitutions inflates ω to >1 and is indicative of positive selection. Put simply, PAML runs two models to infer positive selection: *(i)* a model assuming neutral evolutionary processes, *i.e*., the gene is not under selection with ω=1, and *(ii)* a model assuming positive selection, *i.e.,* ω> 1. A gene is considered to be under selection if the fit of the second model (positive selection) is better than the first (neutrality). Using this approach, it was found that immune genes in general had higher ω than control

genes. This extended to positive selection as well, with 217 of 821 genes were under positive selection amongst immune genes, as opposed to 79 out of 902 in control gene set. These results are in line with a bevy of previous studies that find an enrichment of immune genes under positive selection (*e.g.* (Kosiol *et al.*, 2008; Shultz & Sackton, 2019)).

Estimates of polymorphisms and balancing selection

In Paper II, two methods to detect balancing selection in the complement system of our study population of bank voles were used, aimed at targeting long-term balancing selection events. The first, the Hudson- Kreitman-Aguadé (HKA) test, a classical method for determining balancing selection, is based on the idea that under neutral evolution new mutations should affect polymorphisms (within species) and divergence (between species) equally and these estimates should therefore be proportionate (Hudson *et al.*, 1987). By using intra-specific and inter-specific sequence information, deviations from this expectation can be tested and if true, would indicate an excess of polymorphisms, representative of balancing selection. To this end, we used genome sequences from 31 bank voles, generated as part of an earlier study, along with *Mus musculus* sequences, to find estimates of both polymorphisms and balancing selection (Lundberg *et al.*, 2020).

To compensate for the limitations of the HKA test, a second method of balancing selection was used; BetaScan2 is a genomic scan capable of identifying targets of balancing selection in both coding and non-coding regions (Siewert & Voight, 2020). If a SNP is under balancing selection, it accumulates neutral variants around it that would also show similar frequencies as the balanced SNP due to linkage disequilibrium (LD) and genetic hitchhiking. This is captured with the β statistic, that can be normalised with the mutation rate (β_{std}) and represented with the maximum value per gene ($\beta_{std.max}$) (Siewert & Voight, 2020). By estimating $\beta_{std.max}$ for a set of control genes, outliers for $\beta_{std.max}$ (in our case > 95th percentile) would indicate balancing selection.

With this approach, a complement gene ficolin A (*FCNA*), was found to be under balancing selection by both HKA and BetaScan2. *FCNA* is a pattern recognition molecule that is part of the lectin pathway of the complement system. It recognises acetylation marks on carbohydrates such as N-acetyl glucosamine (GlcNAc), a key subunit of peptidoglycan layer of bacterial cell walls. Haplotype network of a region near a high nucleotide diversity window (π), showed that there were two main haplotype groups, defined by polymorphisms that were in strong linkage disequilibrium with each other (Figure 3). The strongest localised signature of balancing selection was found to be in exons 6-8, which encodes the fibrinogen domain of the protein, known to directly interact with PAMPs. Together, all the data consistently point to a classical notion of an immune gene under balancing selection – one that recognises pathogenic components (PAMPs) and shows high genetic diversity in the regions interacting with the pathogen.

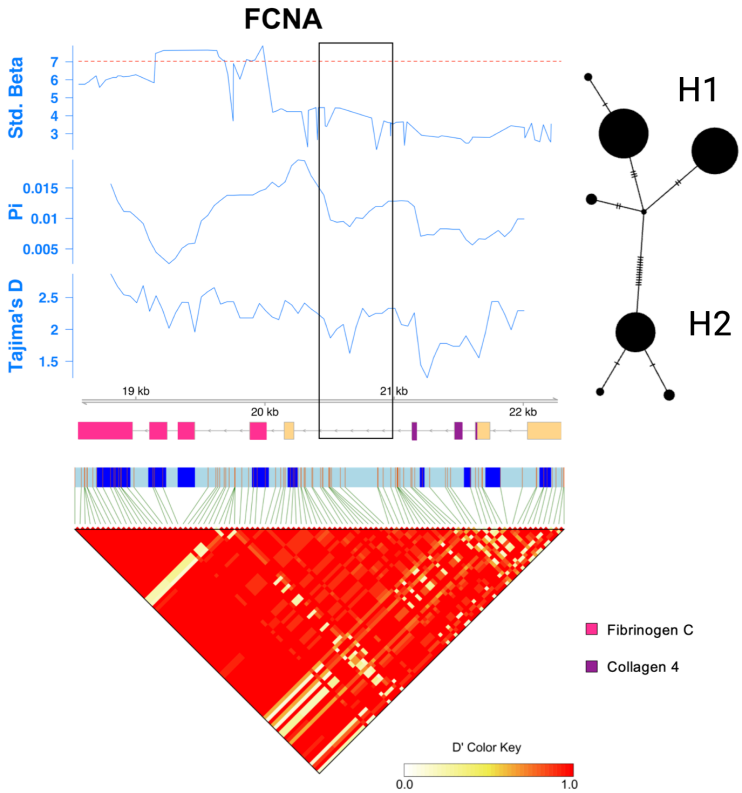


Figure 3. Sliding window analysis of β_{std}, π, and Tajima's D, haplotype network, and linkage disequilibrium (LD) plot for FCNA.
Red dashed line indicates the 95th percentile of $\beta_{std.max}$ of control genes. The haplotype network was constructed from the region marked with black square. Gene structure below sliding window plots show protein domains as predicted by Pfam. The LD plot was constructed for the whole gene using D', with LD blocks shown. D' refers to the normalised values of the coefficient of LD, with high values indicating strong LD. Gene structure corresponding to the LD plots show introns in light blue and exons in dark blue. Two major haplotype clusters – H1 and H2, were identified from haplotype networks.

Functional patterns to immune genes under selection (Papers I and II)

Positive selection

To understand if pathogen-mediated selection is more likely to act on immune genes with certain functions than others, my colleagues and I rely on literature survey and manual curation of immune gene function. Online resources for gene function such as Gene Ontology (GO) terms, while useful, have several, often overlapping, functions associated with each gene which makes it difficult to categorise immune

function. Additionally, what constitutes an immune gene is often debatable. In Paper I, a list of genes was curated to be immunity related if they occurred in at least two of four online functional databases (Panther, KEGG, GO, Reactome). For these genes a literature survey was undertaken to place them in one of six categories (Figure 4): *(i)* PRRs involved in sensing PAMPs (*e.g., TLR1* and *TLR2*), *(ii)* intracellular signalling proteins that transmit signals within a cell (*e.g., MYD88*, *NFκB*), *(iii)* cytokines, chemokines and their receptors that are involved in transmitting signal to other cells (*e.g., IL12*, *IL12R*), *(iv)* other cell-surface proteins involved in receptor and co-stimulatory function (*e.g., CD80*, *CD28*), *(v)* effector proteins such as anti-microbial peptides and anti-viral restriction factors (e.g., *DEFB1*), *(vi)* extracellular proteases and their inhibitors involved primarily in the complement and coagulation pathways (*e.g., C1Q*). Differences between these categories in their evolutionary patterns were tested using estimates of positive selection, after removing genes that have poor one-to-one orthology like *MHC* and *TCR*.

When looking for functional enrichment of genes under selection it was found that, with the exception of intracellular signalling, all other gene categories show a higher proportion of genes under positive selection compared to non-immune control genes (Figure 4). This effect was significant even when controlling for other factors that might affect selection (gene expression level, tissue specificity of expression and gene length). While genes that interact with pathogens such as PRRs and effectors are expected to be enriched for signals of positive selection, it is surprising that other categories such as extracellular proteases, cytokines and cell-surface proteins are also enriched. This suggests that these genes are potentially involved in other antagonistic interactions with pathogens, such as being targets of pathogen evasion, as illustrated with many complement proteins (Cagliani *et al.*, 2016).

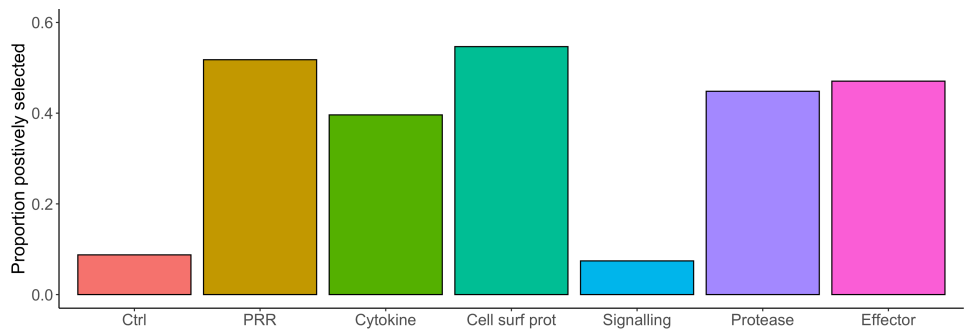


Figure 4. The proportion of genes with signatures of positive selection in six different categories of immune genes and non-immune control genes.
With the exception of intracellular signalling genes, all other categories of immune function had a significantly higher proportion of genes under positive selection as compared to non-immune controls.

Balancing selection

It is well known that many pathogens evade the complement system by manipulating the host immune system, placing these genes under presumably strong selection pressure (Lambris et al., 2008; Cagliani et al., 2016). Using this rationale, we expect that complement genes targeted by pathogen evasion strategies should make them more likely to be "targets of balancing selection". A curated list of complement genes that were targets of pathogen evasion was used to check if patterns of balancing selection (measured as $\beta_{std.max}$) differed between targets of balancing selection vs other complement genes (Figure 5). Complement genes in general displayed higher values of $\beta_{std.max}$ than a set of control genes (Mann-Whitney U test: $p = 0.014$). However, there was no difference between the two complement gene categories. Thus, in the case of complement genes, direct interactions with pathogens alone does not place genes at higher likelihood for balancing selection.

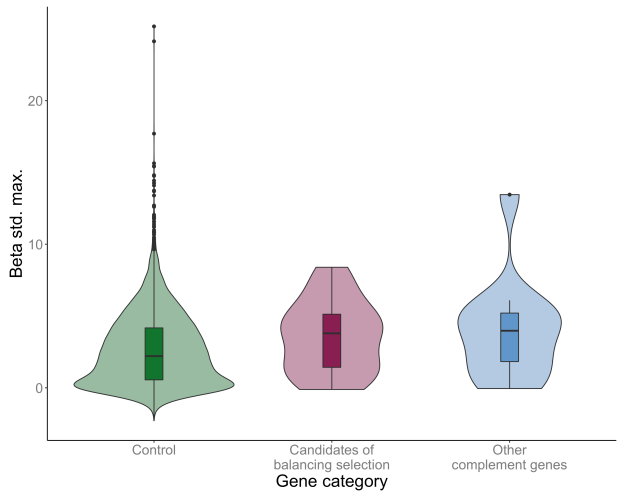


Figure 5. Patterns to complement genes with signs of balancing selection.
Violin plots of $\beta_{std.max}$ for non-complement control genes (n = 8465), complement genes that are candidates of balancing selection *(i.e.,* involved in pathogen recognition or are targets of immune evasion; n = 25), and other complement genes (n= 12).

Understanding functional effects of polymorphisms (Papers III, IV, V)

The final three chapters delve into functional effects of polymorphisms under balancing selection using both association studies with infection and *ex vivo* assays to investigate how polymorphisms influence immune responses.

Association study (Paper III)

Paper III presents an investigation into the functional effect of the complement gene *FCNA,* found to be under balancing selection in Paper II, via association with natural infections of *B. afzelii*. Hundreds of adult bank voles (n=281), caught between 2010-2013, were genotyped at this locus, for which corresponding *B. afzelii* infection information measured with qPCR was available. We then tested if infection prevalence and *FCNA* genotype were associated.

Individuals with two copies of the H1 haplotype were more resistant to *B. afzelii* infections (Figure 6; infection prevalence of 31.6%), than the homozygous H2 (infection prevalence 58%), while heterozygous individuals (H1/H2) showed intermediary infection prevalence (44%). This indicates that selection from *B. afzelii* infections is one factor leading to balancing selection at the *FCNA* locus.

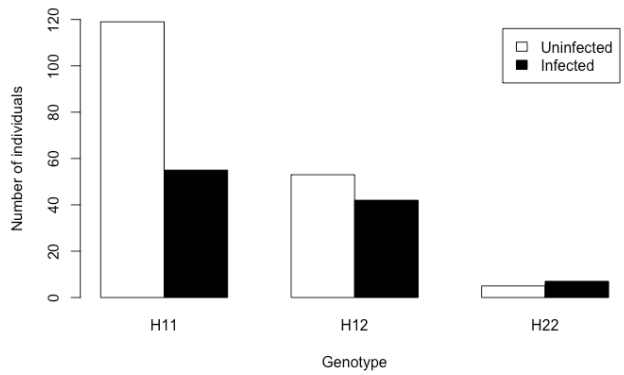


Figure 6. Association between *FCNA* genotype and *Borrelia afzelii* infection in adult bank voles. Haplotype groups: H11 are individuals homozygous for haplotype group 1, H12 are heterozygous, and H22 are homozygous for haplotype group 2. There was signficant association of *FCNA* genotype with infection prevalence.

Ex vivo assays (Papers IV, V)

Ex vivo assays offer an alternative to *in vivo* infections to isolate functional effects of genetic polymorphisms. Data for Papers IV and V comes from *ex vivo* manipulations of the spleen, a key immune organ involved in immune response against infections, to understand the functional effects of polymorphisms on immune response. The spleen contains all immune cells necessary for responding to infection and would therefore be representative of a typical immune reaction.

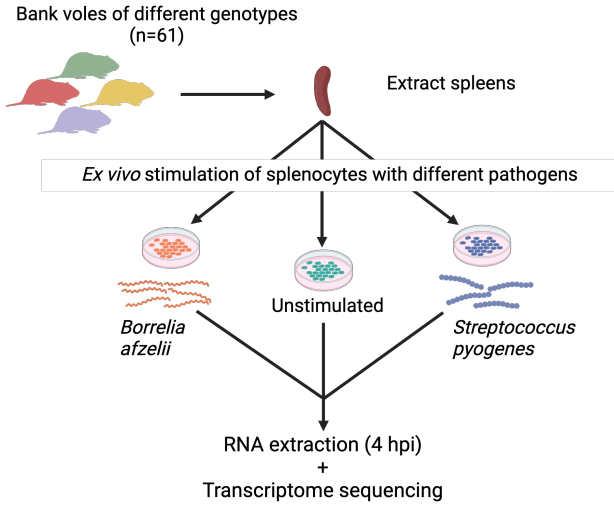


Figure 7. Experimental design of *ex vivo* assays to test functional effects of pathogen-mediated selection on immune response. Created with BioRender.com.

For Papers IV and V, wild bank voles were used to prepare primary cell cultures of spleens, called splenocytes, which are then stimulated under laboratory conditions with two bacterial pathogens (Figure 7); a natural bank vole pathogen (*B. afzelii*) and a human-specific pathogen (*Streptococcus pyogenes*), novel to bank voles. A feature of *S. pyogenes* (group A streptococcus) and other pathogenic *Streptoccus spp* is β-haemolytic activity – the ability to lyse host cells by the synthesis of bacterial toxin. This ability has been detected in unidentified Streptococcal species isolated from the microbiota of bank vole upper-respiratory tracts in our local population (K. Wollein Waldetoft, pers. comm.). The presence of other potentially pathogenic Streptococcal species indicates that *S. pyogenes* would be a relevant model pathogen in bank voles. Furthermore, *S. pyogenes* is an extracellular Gram-positive bacterium containing lipoprotein PAMPs that will bind *TLR2*, activating similar immune pathways as *B. afzelii*. After stimulation, RNA from the cell cultures was extracted and the transcriptome sequenced to capture the gene expression response in its entirety. Together, this experimental set-up offers a good framework to test various aspects of how selection on immune genes influences immune response.

Functional effects of polymorphisms in TLR2 (Paper IV)

TLR2 is a gene under balancing selection in bank voles with two major haplotype clusters, called c1 and c2 (Tschirren et al., 2013; Lundberg et al., 2020). Similar to *FCNA*, the different genotypes of *TLR2* are associated with infection of *B. afzelii* in the wild. Bank voles homozygous for the c1 haplotype have a higher prevalence of

B. afzelii than homozygous c2, with the heterozygote c1/c2 showing similar prevalence to c2/c2 (Figure 8).

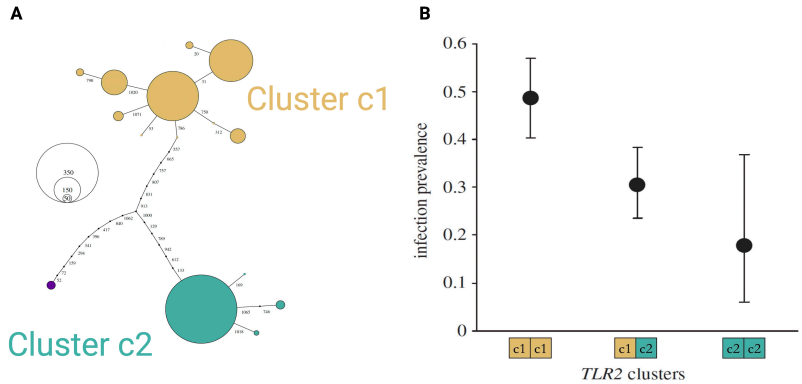


Figure 8. Polymorphism of *TLR2* and its association with *Borrelia afzelii* infections.
A) Haplotype network of *TLR2* in bank voles shows two major haplotypes at intermediate frequencies, with five non-synonymous SNPs distinguishing the two clusters B) Association of different *TLR2* genotypes with natural infection of *B. afzelii* infections in bank voles. Figure modified from (Tschirren *et al.*, 2013).

A key aim with the *ex vivo* set-up is to understand the immunological basis that underlies the observed differences in *B. afzelii* infection prevalence between the different *TLR2* genotypes in bank voles. TLR2, together with its co-receptor TLR1 or TLR6, recognise lipoproteins present in bacterial cells wall. In the coding sequence, five non-synonymous SNPs differentiate the two *TLR2* haplotypes in bank voles. One of these polymorphism (T276A) is located in the ligand binding region of TLR2, although this site does not interact directly with ligands in crystal structures of mouse and human TLR2 heterodimers (Figure 9; (Jin et al., 2007; Kang et al., 2009). The other polymorphisms are found outside the ligand binding and dimerization regions, and instead closer to the transmembrane and intracellular domains of TLR2 (Figure 9). Regardless of the causative SNP(s), any effect of *TLR2* would be observed in downstream signalling of the *TLR2* pathway.

Network theory for gene expression suggests that genes under a common regulation (such as genes in the same pathway) would show correlation in their expression. Network analysis can be used to find downstream genes that are influenced by *TLR2* activation by identifying genes whose expression is correlated with *TLR2*. Even if different stimulation conditions do not show differences in the global expression pattern between different *TLR2* genotypes, the magnitude of gene expression for genes correlated with *TLR2* expression could differ between *TLR2* genotypes. In paper IV, the *ex vivo* set-up was used to first see how stimulation induced by *B. afzelii* and *S. pyogenes* via different *TLR2* genotypes affected the global immune response (differential gene expression analysis with DESeq2 (Love et al., 2014)).

Then, the effect of different genotypes on the magnitude of gene co-expression with *TLR2* was tested using network analysis implemented in WGCNA (Langfelder & Horvath, 2008).

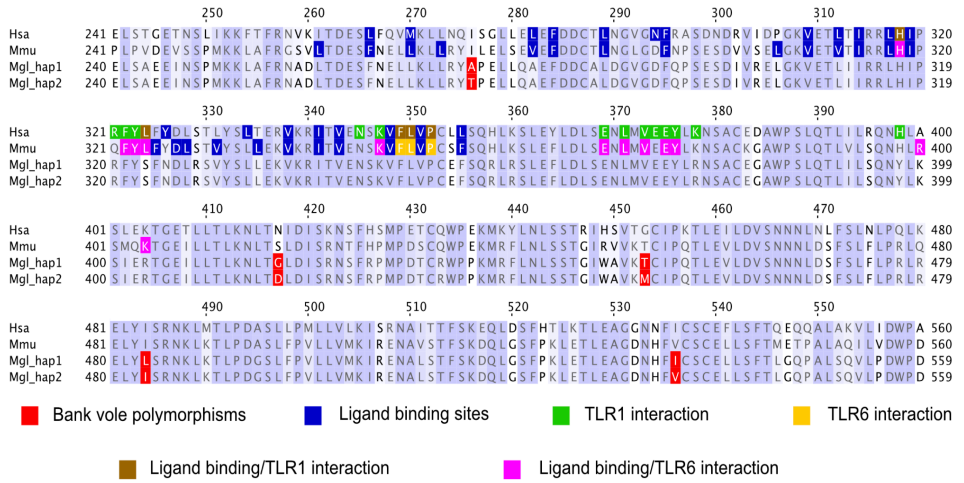


Figure 9: Multiple protein sequence alignment of TLR2 from various species.
The major bank vole TLR2 haplotypes c1 and c2 (Mgl_hap1 and Mgl_hap2 respectively) are aligned with human (Hsa) and mouse (Mmu) orthologs. Sites that interact with ligands (PAMPs) or TLR1 or TLR6 heterodimerization are taken from crystal structure of human/mice TLR2 with TLR1 or TLR6 (Jin et al., 2007; Kang et al., 2009). Polymorphic sites that separate the two haplotypes are highlighted in red.

Differential gene expression analysis revealed that the response to *B. afzelii* and *S. pyogenes* activated largely similar genes and pathways across various TLR2 genotypes. Network analysis however indicated a potential mode of action for how *TLR2* can exert an effect on the immune response. A module of co-expressed genes containing *TLR2*, along with many chemokines and cytokines, was identified to be strongly positively correlated with *B. afzelii* stimulation. Most genes in the module were also differentially expressed. Further analysis of this module to understand if its expression co-varied with *TLR2* genotype and expression was performed with linear mixed effect modelling. The first principal component of this module (excluding *TLR2*) was modelled as an effect of condition, *TLR2* expression, *TLR2* genotype and their interactions, with individual vole as random effect (Figure 10). There was a significant effect of all three factors interacting ($F=5.37$, $p=0.023$). Isolating this effect further yielded significant results for TLR2 genotype and expression interaction in *B. afzelii*-stimualtion alone ($F_{1,51}= 5.45$, $p= 0.023$). Similar analyses for two members of this module upregulated in response to TLR stimulation – cytokines *IL6* and tumour necrosis factor (*TNF)*, yielded similar results (interaction between *TLR2* genotype and expression for *IL6*: $F_{1,51}=4.37$, $p= 0.042$; *TNF*: $F_{1,51}=8.18$ $p=0.006$). Our results indicate that downstream signalling of

cytokines and chemokines is dependent more strongly on *TLR2* expression for c2/c2 genotype than the c1/c1 genotype, although the average expression of module genes was the same for all genotypes. This expression pattern correlated with the infection prevalence of *B. afzelii* in wild populations, with the more susceptible genotype c1/c1 having a weaker association between expression of the module genes and *TLR2* expression, in comparison to the more resistant genotype c2/c2. This study unravels a novel layer to how genotypes affect observed phenotypes as the expression of downstream genes is dependent on *TLR2* genotype for a given *TLR2* expression level.

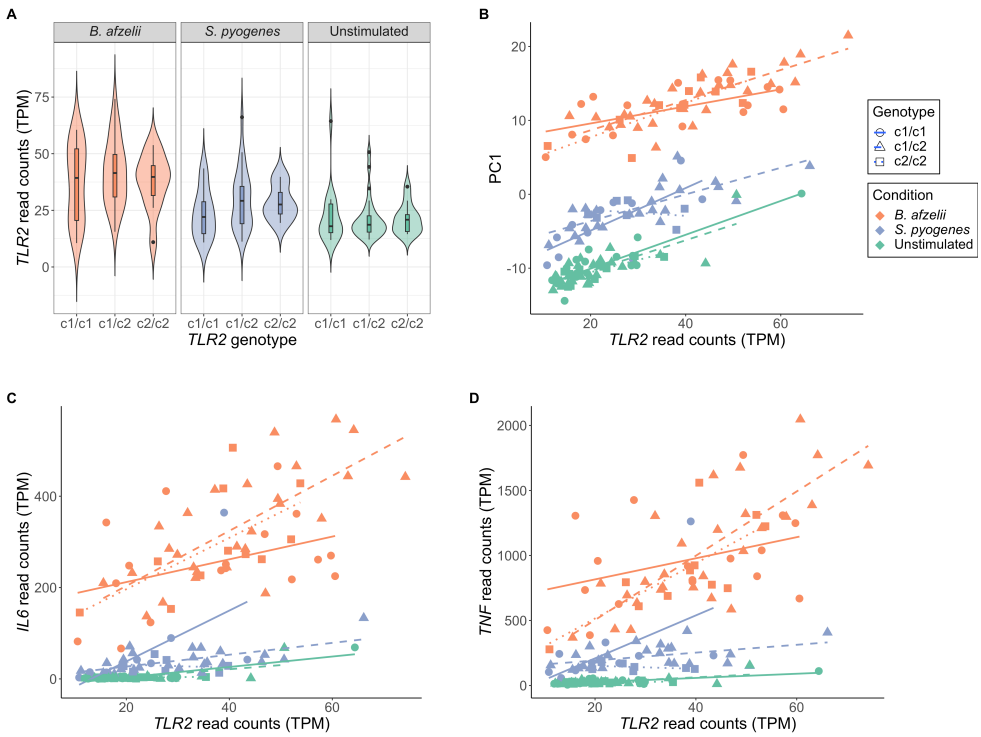


Figure 10. Association of *TLR2* genotypes with expression of a module of genes associated with response *to B. afzelii*.
A) Expression levels of *TLR2* for each genotype and condition. B) Effect of condition, *TLR2* genotype and *TLR2* expression on module 5 expression (first principal component). C) Effect of condition, *TLR2* genotype and *TLR2* expression on *IL6* expression. D) Effect of condition, *TLR2* genotype and *TLR2* expression on *TNF* expression.

Cis-regulatory variation in immune responses (Paper V)

Variants in non-coding regions can affect immune response via transcriptional regulation in two ways. *Cis*-acting variants regulate gene expression in nearby genes and are present as, for example, promoter or enhancer region SNPs. *Trans*-acting

variants control gene expression through polymorphisms in genes encoding, for example, transcription factors. The *ex vivo* data was used to search for differences in immune response driven by *cis*-regulatory variants as measured by allele-specific expression (ASE). *Cis*-regulatory variants are often linked with polymorphisms in coding regions of the gene (Figure 11). Allele-specific expression leverages this information by looking at how different exonic variants at a locus are expressed. In heterozygous individuals, when there is no allelic imbalance (*i.e.,* no ASE), both alleles (for biallelic SNPs) are expressed to the same extent. However, if alleles of a *cis*-regulatory variant are linked to variants in the coding region, one allele will be expressed more than the other (*i.e.,* ASE). By measuring the relative expression of two alleles at an exonic SNP, ASE can be used to measure the effect of *cis*-regulatory variants. ASE analysis can be performed by calling variants from transcriptome data itself and. Thus, with this strategy, the effect of *cis*-regulatory variants can be estimated, without it being genotyped, although it comes with the caveat that only variants that are expressed would be captured.

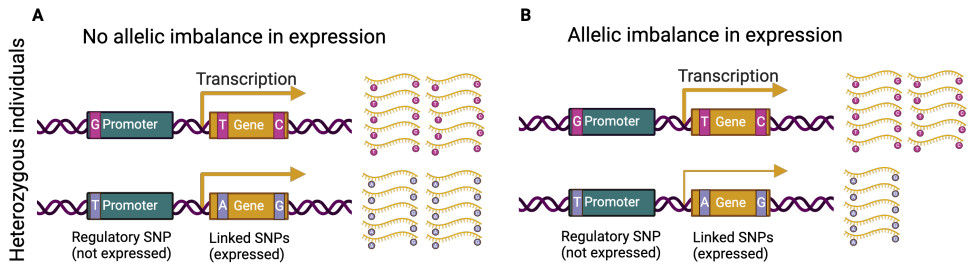


Figure 11. Capturing allele-specific expression using transcriptome information from heterozygous individuals.
A) If alleles of a *cis*-regulatory SNP have no effect on expression of linked coding-region SNPs, no differences in expression between alleles of an SNP will be detected B) If alleles of a *cis*-regulatory variant show differences in transcription ability, differences in allelic expression of linked coding-region SNPs would be observed. Figure inspired by (Fan *et al.*, 2020), created with BioRender.com

With this approach ~3k SNPs that show allelic imbalance for each condition (*B. afzelii*-stimulated, *S. pyogenes*-stimulated, unstimulated controls) were identified in the *ex vivo* data. This translates to ~1.3k genes with signatures of ASE and is quite consistent for all three conditions. While many SNPs that show ASE are specific to each condition, nearly half are shared between conditions, indicating that allelic imbalance underlies both baseline and immune response expression (Figure 12). Thus, this analysis showed abundant *cis*-regulatory variation for immune responses. We further investigated if *cis*-regulatory variants maintained by balancing selection contributed disproportionately to ASE. This was found to be true; 38.8% of genes that had signatures of balancing selection (as estimated by outliers of $\beta_{std.max}$) also showed ASE, as opposed to 28.4% of genes that showed allelic imbalance without signatures of balancing selection. This effect was significant only for *B. afzelii* and

S. pyogenes stimulations (χ^2 = 4.00, df=1, p=0.046 and χ^2 = 4.79, df=1, p=0.029, respectively) but not unstimulated (χ^2 = 1.54, df=1, p=0.21), highlighting that *cis*-regulatory variants maintained by balancing selection is an important factor to ASE in an infection context.

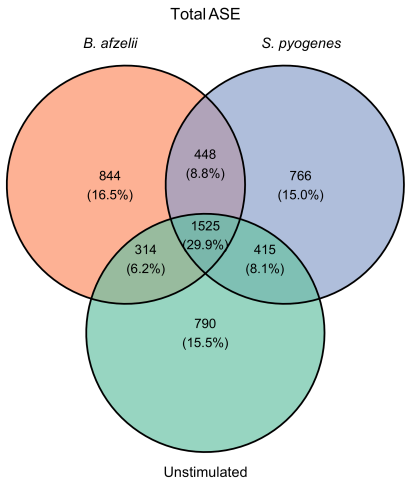


Figure 12. Analysis of allele-specific expression (ASE) in the three conditions of the *ex vivo* assays.
Venn diagram representing total number of SNPs that showed imbalances in allelic expression towards either reference or alternate allele.

Evolutionary processes driving balancing selection (Papers IV and V)

The *ex vivo* setup also provides a good opportunity to test if trade-offs for responses to different pathogens underlies maintenance of genetic polymorphisms in immune genes. Paper IV investigated how balanced polymorphisms at *TLR2* responded differently to the two infection stimulations. The results indicated that there is an effect of genotype only with regards to *B. afzelii* but not *S. pyogenes* response, indicating a specificity in pattern-recognition capability of *TLR2*. This suggests that the effect of *TLR2* genotype on immune responsiveness might be due to selection from *B. afzelii.* There is, however, as yet no indication of trade-offs between responses to different pathogens based on *TLR2* genotype.

Paper V similarly evaluated if allelic response was condition-dependent, by testing for trade-offs in response to the two bacterial stimulations. This was inquired by checking if the direction of ASE flips based on the stimulation, *i.e.,* if alleles that show biased expression towards the reference allele in *B. afzelii*-stimulated cells

would show expression bias towards alternate allele in *S. pyogenes* and vice versa. However, no such SNPs that have condition-dependent effect on ASE was found (Figure 13). This inquiry reveals that trade-offs mediated by *cis*-regulatory variation in response to different bacteria does not play an important role in maintaining genetic variation in bank voles, although it does not preclude trade-offs mediated by *cis*-regulatory variants in other contexts. These results are aligned with the rarity of findings that show *cis*-regulatory variants affecting immune response to dissimilar pathogens in other studies (Nédélec *et al.*, 2016; Quach *et al.*, 2016; Häder *et al.*, 2023).

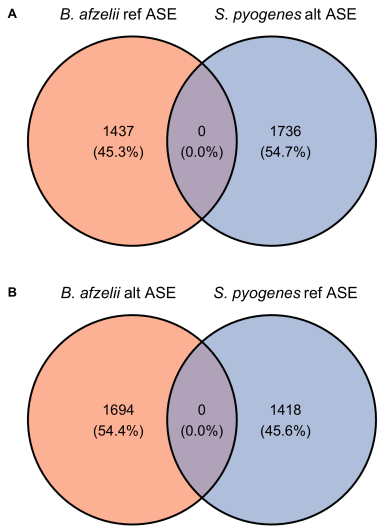


Figure 13. Testing for trade-offs at the *cis*-regulatory level by intersecting SNPs showing allele-specific experssion (ASE).
Allelic imbalance A) towards reference allele under *B. afzelii* stimulation and towards alternate allele under *S. pyogenes* stimulation, and B) towards alternate allele under *B. afzelii* stimulation and towards reference allele under *S. pyogenes* stimulation.

Concluding remarks and perspectives

This thesis explored three facets to pathogen-mediated selection in rodents. In the first, the key aim was to identify gene targets and their functional patterns to pathogen-mediated selection. Paper I provided overwhelming evidence that positive selection strongly shapes the divergence of a wide diversity of immune genes between species. Analysis in Paper II revealed a complement gene – encoding a receptor that recognises pathogenic components – with strong signatures of balancing selection. Together, Papers I and II reinforce ideas that immune genes under pathogen-mediated selection include those that interact with pathogens directly and thereby participate in host-pathogen coevolution, such as PRRs and effector proteins. However, Paper I also expands the ideas of what types of immune genes can be under selection, as it identified heretofore overlooked components of the immune system, such as cytokines and cell-surface proteins, to be highly enriched for positive selection. This work can be expanded further to understand to what extent selection on for example cytokines and cell surface receptors are driven by direct interactions with pathogens in the form of pathogen immune evasion. To get a more complete picture of how immune systems evolve, it would also be of interest to investigate if divergence in protein coding sequences correlates with divergence in gene expression and the extent of copy-number variations in immune genes (Sackton, 2019).

The second aspect of pathogen-mediated selection dealt with in this thesis was its effect on immune function. An association of *FCNA* genotypes under balancing selection with natural infections of *B. afzelii* was identified in Paper III. To supplement this functional effect, it would be of future interest to understand if balanced polymorphisms at *FCNA* mediate complement-induced killing of bacteria differently. Paper IV offers an exciting role for how different *TLR2* genotypes show different reactions to *B. afzelii* – one that involves genotype-dependent expression specific to an endemic pathogen. More experiments are required to inform this line of thought, particularly to understand how this interaction plays out as differences in the immune response and resistance to infection. These two papers jointly demonstrate how coupling functional studies to scans of selection improves our understanding of pathogen-mediated selection.

The third and final aspect of pathogen-mediated selection addressed in this thesis is in relation to evolutionary processes driving balancing selection. What evolutionary forces drive signatures of selection is often unclear, especially with respect to balancing selection. Papers IV and V tested for effects of trade-offs in immune responses to two bacterial pathogens. It is interesting to note that in both cases, no trade-offs were uncovered. There are numerous explanations for this, as we solely focused on trade-offs to two pathogens. Perhaps, the chances of finding a significant effect would be higher if we checked for trade-offs to two pathogens bank voles are adapted to? Or perhaps trade-offs are more likely to occur between more similar pathogens (*e.g.,* different strains of a pathogen), or between more distantly related pathogens (*e.g.,* bacteria vs parasites)? Alternatively, other types of trade-offs, such as between resistance to infection and autoimmunity or immunopathology might be more important. In any case, searching for this elusive evolutionary process would benefit the quest for understanding the causes of balancing selection.

It is evident from our results, as well as others, that only a subset of immune genes are subject to selection from pathogens. But, what of pathogens itself? Are all pathogens equally influential on host evolution? Surely not, as there seems to be a hierarchy to the type of pathogens that cause host evolution (Figure 14). The primary determinant for a "good pathogen" is the balance between virulence (defined in the ecological sense: cost borne by the host during infection) and transmissibility (ability of pathogen to transmit to other hosts) (Alizon *et al.*, 2009). It is only when there is an optimum balance between virulence and transmissibility that there is sufficient selection pressure on host fitness from a pathogen to influence host evolution. Within this narrow condition, the mode of selection is then influenced by for example, how a pathogen spreads through a population (epidemic or endemic). Epidemic infections spread rapidly through a naïve population and can have massive consequences on host fitness and population size. It can be imagined that such epidemics will immensely impact allele frequencies over short timescales, where fixation of alleles can occur through strong positive selection favouring resistance or founder effects following population bottleneck. This would indiscriminately alter allele frequencies of multiple genes. Pathogens that cause epidemic outbreaks can eventually become endemic, as they adapt locally, resulting in pathogenic variants that infect a steady number of individuals. Long-time evolution with an endemic pathogen allows for host-pathogen coevolution following either arms race or Red Queen dynamics (trade-off to different pathogenic variants), as antagonistic interactions involving specific loci can be expected (Woolhouse *et al.*, 2002). Endemic infections can also turn into epidemics, as pathogens evolve new variants or strategies that challenge host immunity, altering the mode of selection. If there are multiple factors that affect host fitness equally resulting in a trade-off (such as between endemic pathogens and physiology, autoimmunity or other endemic pathogens), then balancing selection through other mechanisms can be expected.

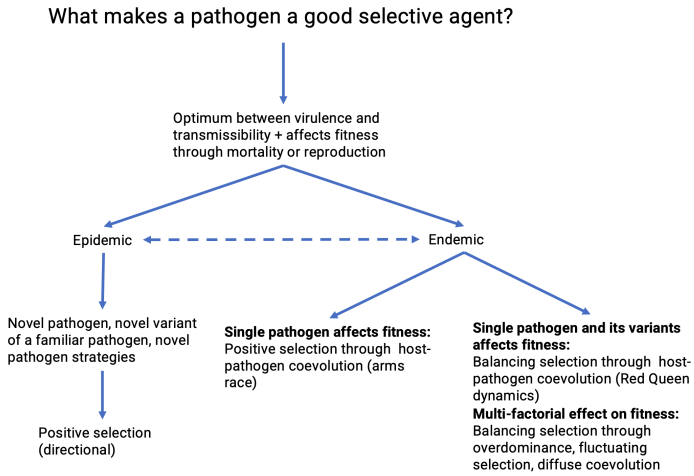


Figure 14. The heirarchy to factors affecting selection pressure from pathogens.
The mode of pathogen-mediated selection on the host is influenced by the type of pathogen.

The selection pressure imposed by a pathogen is of course an extremely dynamic process, as both hosts and pathogens evolve, and any change could alter the position of the pathogen in the hierarchy and influence the strength and mode of selection. Considering this tight-rope a pathogen must walk to impose selective pressure, it is remarkable that there are so many immune genes with signatures of positive and balancing selection, potentially coming from a small number of pathogens, underscoring a likely polygenic basis of adaptation to a single pathogen. A future course for studies of pathogen-mediated selection would be to understand which process contributes more to host evolution – is it selection from past epidemics or continuous coevolution with endemic pathogens?

References

Abdul-Rahman, F., Tranchina, D. & Gresham, D. 2021. Fluctuating Environments Maintain Genetic Diversity through Neutral Fitness Effects and Balancing Selection. *Molecular Biology and Evolution* **38**: 4362–4375.

Akira, S., Uematsu, S. & Takeuchi, O. 2006. Pathogen Recognition and Innate Immunity. *Cell* **124**: 783–801.

Alizon, S., Hurford, A., Mideo, N. & Van Baalen, M. 2009. Virulence evolution and the trade-off hypothesis: history, current state of affairs and the future. *Journal of Evolutionary Biology* **22**: 245–259.

Alves, J.M., Carneiro, M., Cheng, J.Y., Matos, A.L. de, Rahman, M.M., Loog, L., *et al.* 2019. Parallel adaptation of rabbit populations to myxoma virus. *Science* **363**: 1319–1326. American Association for the Advancement of Science.

Andersson, M. & Råberg, L. 2011. Wild Rodents and Novel Human Pathogen Candidatus Neoehrlichia mikurensis, Southern Sweden. *Emerg Infect Dis* **17**: 1716–1718.

Andersson, M., Scherman, K. & Råberg, L. 2013. Multiple-Strain Infections of Borrelia afzelii: A Role for Within-Host Interactions in the Maintenance of Antigenic Diversity? *The American Naturalist* **181**: 545–554. The University of Chicago Press.

Andrés, A.M., Dennis, M.Y., Kretzschmar, W.W., Cannons, J.L., Lee-Lin, S.-Q., Hurle, B., *et al.* 2010. Balancing Selection Maintains a Form of ERAP2 that Undergoes Nonsense-Mediated Decay and Affects Antigen Presentation. *PLOS Genetics* **6**: e1001157. Public Library of Science.

Andrés, A.M., Hubisz, M.J., Indap, A., Torgerson, D.G., Degenhardt, J.D., Boyko, A.R., *et al.* 2009. Targets of Balancing Selection in the Human Genome. *Molecular Biology and Evolution* **26**: 2755–2764.

Arora, J., Pierini, F., McLaren, P.J., Carrington, M., Fellay, J. & Lenz, T.L. 2020. HLA Heterozygote Advantage against HIV-1 Is Driven by Quantitative and Qualitative Differences in HLA Allele-Specific Peptide Presentation. *Molecular Biology and Evolution* **37**: 639–650.

Balloux, F. & van Dorp, L. 2017. Q&A: What are pathogens, and what have they done to and for us? *BMC Biol* **15**: 91.

Band, G., Le, Q.S., Clarke, G.M., Kivinen, K., Hubbart, C., Jeffreys, A.E., *et al.* 2019. Insights into malaria susceptibility using genome-wide data on 17,000 individuals from Africa, Asia and Oceania. *Nat Commun* **10**: 5732. Nature Publishing Group.

Bankhead, T. & Chaconas, G. 2007. The role of VlsE antigenic variation in the Lyme disease spirochete: persistence through a mechanism that differs from other pathogens. *Molecular Microbiology* **65**: 1547–1558.

Barreiro, L.B., Ben-Ali, M., Bouchier, C., Tichit, M., Neyrolles, O., Gicquel, B., *et al.* 2009. Evolutionary Dynamics of Human Toll-Like Receptors and Their Different Contributions to Host Defense. *PLoS Genetics* **5**: 18.

Bento, G., Routtu, J., Fields, P.D., Bourgeois, Y., Pasquier, L.D. & Ebert, D. 2017. The genetic basis of resistance and matching-allele interactions of a host-parasite system: The Daphnia magna-Pasteuria ramosa model. *PLOS Genetics* **13**: e1006596. Public Library of Science.

Bergland, A.O., Behrman, E.L., O'Brien, K.R., Schmidt, P.S. & Petrov, D.A. 2014. Genomic Evidence of Rapid and Stable Adaptive Oscillations over Seasonal Time Scales in Drosophila. *PLOS Genetics* **10**: e1004775. Public Library of Science.

Bitarello, B.D., de Filippo, C., Teixeira, J.C., Schmidt, J.M., Kleinert, P., Meyer, D., *et al.* 2018. Signatures of Long-Term Balancing Selection in Human Genomes. *Genome Biol Evol* **10**: 939–955. Oxford Academic.

Boehm, T. & Swann, J.B. 2014. Origin and evolution of adaptive immunity. *Annu Rev Anim Biosci* **2**: 259–283.

Bourgeois, Y., Fields, P.D., Bento, G. & Ebert, D. 2021. Balancing Selection for Pathogen Resistance Reveals an Intercontinental Signature of Red Queen Coevolution. *Molecular Biology and Evolution* **38**: 4918–4933.

Buchmann, K. 2014. Evolution of Innate Immunity: Clues from Invertebrates via Fish to Mammals. *Frontiers in Immunology* **5**: 459.

Cagliani, R., Forni, D., Filippi, G., Mozzi, A., De Gioia, L., Pontremoli, C., *et al.* 2016. The mammalian complement system as an epitome of host–pathogen genetic conflicts. *Molecular Ecology* **25**: 1324–1339.

Cagliani, R., Fumagalli, M., Riva, S., Pozzoli, U., Comi, G.P., Menozzi, G., *et al.* 2008. The signature of long-standing balancing selection at the human defensin β-1 promoter. *Genome Biol* **9**: 1–11. BioMed Central.

Cagliani, R., Riva, S., Pozzoli, U., Fumagalli, M., Comi, G.P., Bresolin, N., *et al.* 2011. Balancing selection is common in the extended MHC region but most alleles with opposite risk profile for autoimmune diseases are neutrally evolving. *BMC Evolutionary Biology* **11**: 171.

Cayol, C., Giermek, A., Gomez-Chamorro, A., Hytönen, J., Kallio, E.R., Mappes, T., *et al.* 2018. Borrelia afzelii alters reproductive success in a rodent host. *Proceedings of the Royal Society B: Biological Sciences* **285**: 20181056. Royal Society.

Charlesworth, B. & Jensen, J.D. 2021. Effects of Selection at Linked Sites on Patterns of Genetic Variability. *Annual Review of Ecology, Evolution, and Systematics* **52**: 177–197.

Cho, S., Huang, Z.Y., Green, D.R., Smith, D.R. & Zhang, J. 2006. Evolution of the complementary sex-determination gene of honey bees: Balancing selection and trans-species polymorphisms. *Genome Res* **16**: 1366–1375.

Clough, D. & Råberg, L. 2014. Contrasting patterns of structural host specificity of two species of Heligmosomoides nematodes in sympatric rodents. *Parasitol Res* **113**: 4633–4639.

Cooper, M.D. & Alder, M.N. 2006. The Evolution of Adaptive Immune Systems. *Cell* **124**: 815–822. Elsevier.

Dean, A.M. & Thornton, J.W. 2007. Mechanistic approaches to the study of evolution: the functional synthesis. *Nat Rev Genet* **8**: 675–688.

DeGiorgio, M., Lohmueller, K.E. & Nielsen, R. 2014. A Model-Based Approach for Identifying Signatures of Ancient Balancing Selection in Genetic Data. *PLOS Genetics* **10**: e1004561. Public Library of Science.

Fabre, P.-H., Hautier, L., Dimitrov, D. & P Douzery, E.J. 2012. A glimpse on the pattern of rodent diversification: a phylogenetic approach. *BMC Evolutionary Biology* **12**: 88.

Fan, J., Hu, J., Xue, C., Zhang, H., Susztak, K., Reilly, M.P., *et al.* 2020. ASEP: Gene-based detection of allele-specific expression across individuals in a population by RNA sequencing. *PLoS Genet* **16**: e1008786.

Farhat, M.R., Freschi, L., Calderon, R., Ioerger, T., Snyder, M., Meehan, C.J., *et al.* 2019. GWAS for quantitative resistance phenotypes in Mycobacterium tuberculosis reveals resistance genes and regulatory regions. *Nat Commun* **10**: 2128. Nature Publishing Group.

Fenner, F. & Ratcliffe, F.N. 1965. *Myxomatosis*. Cambridge University Press.

Ferguson, W., Dvora, S., Gallo, J., Orth, A. & Boissinot, S. 2008. Long-term balancing selection at the west nile virus resistance gene, Oas1b, maintains transspecific polymorphisms in the house mouse. *Mol Biol Evol* **25**: 1609–1618.

Ferrer-Admetlla, A., Sikora, M., Laayouni, H., Esteve, A., Roubinet, F., Blancher, A., *et al.* 2009. A Natural History of FUT2 Polymorphism in Humans. *Molecular Biology and Evolution* **26**: 1993–2003.

Fijarczyk, A. & Babik, W. 2015. Detecting balancing selection in genomes: limits and prospects. *Mol Ecol* **24**: 3529–3545.

Fitzpatrick, M.J., Feder, E., Rowe, L. & Sokolowski, M.B. 2007. Maintaining a behaviour polymorphism by frequency-dependent selection on a single gene. *Nature* **447**: 210–212. Nature Publishing Group.

Flajnik, M.F. & Kasahara, M. 2010. Origin and evolution of the adaptive immune system: genetic events and selective pressures. *Nat Rev Genet* **11**: 47–59. Nature Publishing Group.

Fujita, T. 2002. Evolution of the lectin–complement pathway and its role in innate immunity. *Nat Rev Immunol* **2**: 346–353.

Fumagalli, M., Cagliani, R., Pozzoli, U., Riva, S., Comi, G.P., Menozzi, G., *et al.* 2009a. Widespread balancing selection and pathogen-driven selection at blood group antigen genes. *Genome Res.* **19**: 199–212.

Fumagalli, M., Cagliani, R., Riva, S., Pozzoli, U., Biasin, M., Piacentini, L., *et al.* 2010. Population Genetics of IFIH1: Ancient Population Structure, Local Selection, and Implications for Susceptibility to Type 1 Diabetes. *Mol Biol Evol* **27**: 2555–2566. Oxford Academic.

Fumagalli, M., Pozzoli, U., Cagliani, R., Comi, G.P., Riva, S., Clerici, M., *et al.* 2009b. Parasites represent a major selective force for interleukin genes and shape the genetic predisposition to autoimmune conditions. *Journal of Experimental Medicine* **206**: 1395–1408.

Fumagalli, M. & Sironi, M. 2014. Human genome variability, natural selection and infectious diseases. *Current Opinion in Immunology* **30**: 9–16.

Fumagalli, M., Sironi, M., Pozzoli, U., Ferrer-Admettla, A., Pattini, L. & Nielsen, R. 2011. Signatures of Environmental Genetic Adaptation Pinpoint Pathogens as the Main Selective Pressure through Human Evolution. *PLOS Genetics* **7**: e1002355. Public Library of Science.

Gern, L., Siegenthaler, M., Hu, C.M., Leuba-Garcia, S., Humair, P.F. & Moret, J. 1994. Borrelia burgdorferi in rodents (Apodemus flavicollis andA. sylvaticus): Duration and enhancement of infectivity forIxodes ricinus ticks. *Eur J Epidemiol* **10**: 75–80.

Gigord, L.D.B., Macnair, M.R. & Smithson, A. 2001. Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid Dactylorhiza sambucina (L.) Soò. *PNAS* **98**: 6253–6255. National Academy of Sciences.

Gorman, J.A., Hundhausen, C., Errett, J.S., Stone, A.E., Allenspach, E.J., Ge, Y., *et al.* 2017. The A946T variant of the RNA sensor IFIH1 mediates an interferon program that limits viral infection but increases the risk for autoimmunity. *Nature Immunology* **18**: 744–752. Nature Publishing Group.

Grimm, D., Tilly, K., Byram, R., Stewart, P.E., Krum, J.G., Bueschel, D.M., *et al.* 2004. Outer-surface protein C of the Lyme disease spirochete: A protein induced in ticks for infection of mammals. *PNAS* **101**: 3142–3147. National Academy of Sciences.

Grossman, S.R., Andersen, K.G., Shlyakhter, I., Tabrizi, S., Winnicki, S., Yen, A., *et al.* 2013. Identifying Recent Adaptations in Large-Scale Genomic Data. *Cell* **152**: 703–713.

Häder, A., Schäuble, S., Gehlen, J., Thielemann, N., Buerfent, B.C., Schüller, V., *et al.* 2023. Pathogen-specific innate immune response patterns are distinctly affected by genetic diversity. *Nat Commun* **14**: 3239. Nature Publishing Group.

Harrison, G.F., Sanz, J., Boulais, J., Mina, M.J., Grenier, J.-C., Leng, Y., *et al.* 2019. Natural selection contributed to immunological differences between hunter-gatherers and agriculturalists. *Nat Ecol Evol* **3**: 1253–1264.

Hedrick, P.W. 2002. Pathogen resistance and genetic variation at MHC loci. *Evolution* **56**: 1902–1908.

Hedrick, P.W. 2012. What is the evidence for heterozygote advantage selection? *Trends in Ecology & Evolution* **27**: 698–704.

Hejase, H.A., Dukler, N. & Siepel, A. 2020. From Summary Statistics to Gene Trees: Methods for Inferring Positive Selection. *Trends in Genetics* **36**: 243–258.

Hochachka, W.M. & Dhondt, A.A. 2000. Density-dependent decline of host abundance resulting from a new infectious disease. *Proceedings of the National Academy of Sciences* **97**: 5303–5306. Proceedings of the National Academy of Sciences.

Hudson, R.R., Kreitman, M. & Aguadé, M. 1987. A Test of Neutral Molecular Evolution Based on Nucleotide Data. *Genetics* **116**: 153–159.

Jacquet, M., Margos, G., Fingerle, V. & Voordouw, M.J. 2016. Comparison of the lifetime host-to-tick transmission between two strains of the Lyme disease pathogen Borrelia afzelii. *Parasites & Vectors* **9**: 645.

Jin, M.S., Kim, S.E., Heo, J.Y., Lee, M.E., Kim, H.M., Paik, S.-G., *et al.* 2007. Crystal Structure of the TLR1-TLR2 Heterodimer Induced by Binding of a Tri-Acylated Lipopeptide. *Cell* **130**: 1071–1082.

Johansson, P., Olsson, G.E., Low, H.-T., Bucht, G., Ahlm, C., Juto, P., *et al.* 2008. Puumala hantavirus genetic variability in an endemic region (Northern Sweden). *Infection, Genetics and Evolution* **8**: 286–296.

Kang, J.Y., Nan, X., Jin, M.S., Youn, S.-J., Ryu, Y.H., Mah, S., *et al.* 2009. Recognition of Lipopeptide Patterns by Toll-like Receptor 2-Toll-like Receptor 6 Heterodimer. *Immunity* **31**: 873–884.

Karlsson, E.K., Kwiatkowski, D.P. & Sabeti, P.C. 2014. Natural selection and infectious disease in human populations. *Nat Rev Genet* **15**: 379–393.

Kim, Y. & Nielsen, R. 2004. Linkage Disequilibrium as a Signature of Selective Sweeps. *Genetics* **167**: 1513–1524.

Kimbrell, D.A. & Beutler, B. 2001. The evolution and genetics of innate immunity. *Nat Rev Genet* **2**: 256–267.

Kimura, M. 1968. Evolutionary rate at the molecular level. *Nature* **217**: 624–626.

Klunk, J., Vilgalys, T.P., Demeure, C.E., Cheng, X., Shiratori, M., Madej, J., *et al.* 2022. Evolution of immune genes is associated with the Black Death. *Nature* 1–8. Nature Publishing Group.

Kosiol, C., Vinař, T., Fonseca, R.R. da, Hubisz, M.J., Bustamante, C.D., Nielsen, R., *et al.* 2008. Patterns of Positive Selection in Six Mammalian Genomes. *PLOS Genetics* **4**: e1000144. Public Library of Science.

Kurokawa, C., Lynn, G.E., Pedra, J.H.F., Pal, U., Narasimhan, S. & Fikrig, E. 2020. Interactions between Borrelia burgdorferi and ticks. *Nature Reviews Microbiology* 1–14.

Kwiatkowski, D.P. 2005. How Malaria Has Affected the Human Genome and What Human Genetics Can Teach Us about Malaria. *The American Journal of Human Genetics* **77**: 171–192.

Lambris, J.D., Ricklin, D. & Geisbrecht, B.V. 2008. Complement evasion by human pathogens. *Nat Rev Microbiol* **6**: 132–142. Nature Publishing Group.

Langfelder, P. & Horvath, S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**: 559.

Lechner, S., Ferretti, L., Schöning, C., Kinuthia, W., Willemsen, D. & Hasselmann, M. 2014. Nucleotide variability at its limit? Insights into the number and evolutionary dynamics of the sex-determining specificities of the honey bee Apis mellifera. *Mol Biol Evol* **31**: 272–287.

Lin, Y.-P., Diuk-Wasser, M.A., Stevenson, B. & Kraiczy, P. 2020. Complement Evasion Contributes to Lyme Borreliae–Host Associations. *Trends in Parasitology* **36**: 634–645. Elsevier.

Linnenbrink, M., Johnsen, J.M., Montero, I., Brzezinski, C.R., Harr, B. & Baines, J.F. 2011. Long-term balancing selection at the blood group-related gene B4galnt2 in the genus Mus (Rodentia; Muridae). *Mol Biol Evol* **28**: 2999–3003.

Love, M.I., Huber, W. & Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**: 550.

Lundberg, M., Zhong, X., Konrad, A., Olsen, R.-A. & Råberg, L. 2020. Balancing selection in Pattern Recognition Receptor signalling pathways is associated with gene function and pleiotropy in a wild rodent. *Molecular Ecology* **29**: 1990–2003.

Medzhitov, R. 2007. Recognition of microorganisms and activation of the immune response. *Nature* **449**: 819–826.

Microbiology by numbers. 2011. *Nat Rev Microbiol* **9**: 628–628. Nature Publishing Group.

Mitchell-Olds, T., Willis, J.H. & Goldstein, D.B. 2007. Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat Rev Genet* **8**: 845–856.

Mollentze, N. & Streicker, D.G. 2020. Viral zoonotic risk is homogenous among taxonomic orders of mammalian and avian reservoir hosts. *Proceedings of the National Academy of Sciences* **117**: 9423–9430. Proceedings of the National Academy of Sciences.

Nakajima, T., Ohtani, H., Satta, Y., Uno, Y., Akari, H., Ishida, T., *et al.* 2008. Natural selection in the TLR-related genes in the course of primate evolution. *Immunogenetics* **60**: 727–735.

Nédélec, Y., Sanz, J., Baharian, G., Szpiech, Z.A., Pacis, A., Dumaine, A., *et al.* 2016. Genetic Ancestry and Natural Selection Drive Population Differences in Immune Responses to Pathogens. *Cell* **167**: 657-669.e21. Elsevier.

Ohto, U., Yamakawa, N., Akashi-Takamura, S., Miyake, K. & Shimizu, T. 2012. Structural Analyses of Human Toll-like Receptor 4 Polymorphisms D299G and T399I*. *Journal of Biological Chemistry* **287**: 40611–40617.

Palm, N.W. & Medzhitov, R. 2009. Pattern recognition receptors and control of adaptive immunity. *Immunological Reviews* **227**: 221–233.

Piesman, J. & Gern, L. 2004. Lyme borreliosis in Europe and North America. *Parasitology* **129**: S191–S220. Cambridge University Press.

Prugnolle, F., Manica, A., Charpentier, M., Guégan, J.F., Guernier, V. & Balloux, F. 2005. Pathogen-Driven Selection and Worldwide HLA Class I Diversity. *Current Biology* **15**: 1022–1027.

Qiu, W.-G. & Martin, C.L. 2014. Evolutionary genomics of Borrelia burgdorferi sensu lato: Findings, hypotheses, and the rise of hybrids. *Infection, Genetics and Evolution* **27**: 576–593.

Quach, H., Rotival, M., Pothlichet, J., Loh, Y.-H.E., Dannemann, M., Zidane, N., *et al.* 2016. Genetic Adaptation and Neandertal Admixture Shaped the Immune System of Human Populations. *Cell* **167**: 643-656.e17. Elsevier.

Quintana-Murci, L. & Clark, A.G. 2013. Population genetic tools for dissecting innate immunity in humans. *Nat Rev Immunol* **13**: 280–293.

Råberg, L. 2023. Human and pathogen genotype-by-genotype interactions in the light of coevolution theory. *PLOS Genetics* **19**: e1010685. Public Library of Science.

Råberg, L. & Stjernman, M. 2003. Natural Selection on Immune Responsiveness in Blue Tits Parus caeruleus. *Evolution* **57**: 1670–1678. [Society for the Study of Evolution, Wiley].

Radwan, J., Babik, W., Kaufman, J., Lenz, T.L. & Winternitz, J. 2020. Advances in the Evolutionary Understanding of MHC Polymorphism. *Trends in Genetics* **36**: 298–311. Elsevier.

Ricklin, D., Hajishengallis, G., Yang, K. & Lambris, J.D. 2010. Complement: a key system for immune surveillance and homeostasis. *Nature Immunology* **11**: 785–797. Nature Publishing Group.

Rotival, M., Quach, H. & Quintana-Murci, L. 2019. Defining the genetic and evolutionary architecture of alternative splicing in response to infection. *Nat Commun* **10**: 1671.

Russell, R.M., Bibollet-Ruche, F., Liu, W., Sherrill-Mix, S., Li, Y., Connell, J., *et al.* 2021. CD4 receptor diversity represents an ancient protection mechanism against primate lentiviruses. *PNAS* **118**. National Academy of Sciences.

Sackton, T.B. 2019. Comparative genomics and transcriptomics of host–pathogen interactions in insects: evolutionary insights and future directions. *Current Opinion in Insect Science* **31**: 106–113.

Sackton, T.B., Lazzaro, B.P., Schlenke, T.A., Evans, J.D., Hultmark, D. & Clark, A.G. 2007. Dynamic evolution of the innate immune system in Drosophila. *Nat Genet* **39**: 1461–1468. Nature Publishing Group.

Sams, A.J., Dumaine, A., Nédélec, Y., Yotova, V., Alfieri, C., Tanner, J.E., *et al.* 2016. Adaptively introgressed Neandertal haplotype at the OAS locus functionally impacts innate immune responses in humans. *Genome Biol* **17**: 1–15. BioMed Central.

Semple, C.A.M., Maxwell, A., Gautier, P., Kilanowski, F.M., Eastwood, H., Barran, P.E., *et al.* 2005. The complexity of selection at the major primate beta-defensin locus. *BMC Evol Biol* **5**: 32.

Shultz, A.J. & Sackton, T.B. 2019. Immune genes are hotspots of shared positive selection across birds and mammals. *Elife* **8**.

Siewert, K.M. & Voight, B.F. 2020. BetaScan2: Standardized Statistics to Detect Balancing Selection Utilizing Substitution Data. *Genome Biology and Evolution* **12**: 3873–3877.

Siewert, K.M. & Voight, B.F. 2017. Detecting Long-Term Balancing Selection Using Allele Frequency Correlation. *Mol Biol Evol* **34**: 2996–3005. Oxford Academic.

Siontis, K.C.M., Patsopoulos, N.A. & Ioannidis, J.P.A. 2010. Replication of past candidate loci for common diseases and phenotypes in 100 genome-wide association studies. *Eur J Hum Genet* **18**: 832–837. Nature Publishing Group.

Sironi, M., Cagliani, R., Forni, D. & Clerici, M. 2015. Evolutionary insights into host–pathogen interactions from mammalian sequence data. *Nat Rev Genet* **16**: 224–236.

Soni, V. & Eyre-Walker, A. 2022. Factors That Affect the Rates of Adaptive and Nonadaptive Evolution at the Gene Level in Humans and Chimpanzees. *Genome Biology and Evolution* **14**: evac028.

Spurgin, L.G. & Richardson, D.S. 2010. How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B: Biological Sciences* **277**: 979–988. Royal Society.

Stanek, G., Wormser, G.P., Gray, J. & Strle, F. 2012. Lyme borreliosis. *The Lancet* **379**: 461–473.

Stein, C.M. 2011. Genetic Epidemiology of Tuberculosis Susceptibility: Impact of Study Design. *PLOS Pathogens* **7**: e1001189. Public Library of Science.

Stenseth, N.C. 1985. Clethrionomys biology:population dynamics, dispersal, reproduction and social structure. *Annales Zoologici Fennici* **22**.

Takeuchi, O. & Akira, S. 2010. Pattern Recognition Receptors and Inflammation. *Cell* **140**: 805–820.

Tschirren, B., Andersson, M., Scherman, K., Westerdahl, H., Mittl, P.R.E. & Råberg, L. 2013. Polymorphisms at the innate immune receptor TLR2 are associated with Borrelia infection in a wild rodent population. *Proceedings of the Royal Society B: Biological Sciences* **280**: 20130364. Royal Society.

Turner, A.K., Begon, M., Jackson, J.A. & Paterson, S. 2012. Evidence for selection at cytokine loci in a natural population of field voles (Microtus agrestis). *Mol Ecol* **21**: 1632–1646.

Van der Hoorn, R.A.L., De Wit, P.J.G.M. & Joosten, M.H.A.J. 2002. Balancing selection favors guarding resistance proteins. *Trends in Plant Science* **7**: 67–71.

van der Lee, R., Wiel, L., van Dam, T.J.P. & Huynen, M.A. 2017. Genome-scale detection of positive selection in nine primates predicts human-virus evolutionary conflicts. *Nucleic Acids Res* **45**: 10634–10648.

van Riper III, C., van Riper, S.G., Goff, M.L. & Laird, M. 1986. The Epizootiology and Ecological Significance of Malaria in Hawaiian Land Birds. *Ecological Monographs* **56**: 327–344.

Velová, H., Gutowska-Ding, M.W., Burt, D.W. & Vinkler, M. 2018. Toll-Like Receptor Evolution in Birds: Gene Duplication, Pseudogenization, and Diversifying Selection. *Molecular Biology and Evolution* **35**: 2170–2184.

Vinkler, M., Bainová, H. & Bryja, J. 2014. Protein evolution of Toll-like receptors 4, 5 and 7 within Galloanserae birds. *Genet Sel Evol* **46**: 72.

Voight, B.F., Kudaravalli, S., Wen, X. & Pritchard, J.K. 2006. A Map of Recent Positive Selection in the Human Genome. *PLoS Biol* **4**: e72.

Waindok, P., Schicht, S., Fingerle, V. & Strube, C. 2017. Lyme borreliae prevalence and genospecies distribution in ticks removed from humans. *Ticks and Tick-borne Diseases* **8**: 709–714.

Walter, K.S., Carpi, G., Caccone, A. & Diuk-Wasser, M.A. 2017. Genomic insights into the ancient spread of Lyme disease across North America. *Nat Ecol Evol* **1**: 1569–1576.

Wilhelmsson, P., Fryland, L., Lindblom, P., Sjöwall, J., Ahlm, C., Berglund, J., *et al.* 2016. A prospective study on the incidence of Borrelia burgdorferi sensu lato infection after a tick bite in Sweden and on the Åland Islands, Finland (2008–2009). *Ticks and Tick-borne Diseases* **7**: 71–79.

Williams, A.M., Ngo, T.M., Figueroa, V.E. & Tate, A.T. 2023. The Effect of Developmental Pleiotropy on the Evolution of Insect Immune Genes. *Genome Biology and Evolution* **15**: evad044.

Woolhouse, M.E.J. & Gowtage-Sequeria, S. 2005. Host Range and Emerging and Reemerging Pathogens. *Emerg Infect Dis* **11**: 1842–1847.

Woolhouse, M.E.J., Webster, J.P., Domingo, E., Charlesworth, B. & Levin, B.R. 2002. Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nat Genet* **32**: 569–577.

Yang, Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution* **24**: 1586–1591.

Zhang, J. & Yang, J.-R. 2015. Determinants of the rate of protein sequence evolution. *Nat Rev Genet* **16**: 409–420.

Zhong, X., Lundberg, M. & Råberg, L. 2020. Comparison of spleen transcriptomes of two wild rodent species reveals differences in the immune response against Borrelia afzelii. *Ecology and Evolution* **10**: 6421–6434.

Zhong, X., Lundberg, M. & Råberg, L. 2021. Divergence in Coding Sequence and Expression of Different Functional Categories of Immune Genes between Two Wild Rodent Species. *Genome Biol Evol* **13**. Oxford Academic.

Zhong, X., Nouri, M. & Råberg, L. 2019. Colonization and pathology of Borrelia afzelii in its natural hosts. *Ticks and Tick-borne Diseases* **10**: 822–827.

Acknowledgements – a thousand thanks!

Just like the saying, “It takes a village to raise a child”, it has definitely taken a battalion of mentors, colleagues, friends and family to raise this academic toddler to graduation. I am filled deeply with gratitude for this outstanding experience and have a ton of people to thank, none more than you, **Lars**. When I moved from India to start my PhD, little did I expect that working with you would set the benchmark for so many things I want to be – mentor, researcher, teacher. You have been so incredibly supportive every step of the way and I have had a fantastic time doing my PhD with you, I cannot thank you enough for being all-round awesome! If Lars was water, then you are *firee* **Fredric**! It’s always been so much fun talking to you about all things science (you always bring sharp insights and new perspectives to my research), listening to your zany stories (for the n^{th} time), or teaming up against Lars (you can always count on me). Thank you for completing my supervisory dream team! **Dennis,** thank you for being my examiner and organising everything for my studies to progress! **Marjorie**, thank you for being my scientific mentor – you have always checked in and looked out for me and I really appreciate it!

My family – **Amma, Appa, Swetha, Venkat and little Vasudha** – for having so much faith and belief in me that you thought I would complete a four-year PhD in three! I really owe everything to you.

My co-authors, **Max, Mehrnaz, Christine**, for patiently teaching me things I didn’t know and things that I should have known, I am really grateful!

To other members of the extended group that are great colleagues and even better friends – **Esther, Katie, Giulia and Jaume** – it has really been an absolute joy sharing workspaces and more with you, thank you. **Juliaaa**! It’s been so much fun teaching and suffering in gym classes together! Let’s do it again! Past members – **Xiuqin and Elin M.,** thank you for getting me started in the lab and to the newest

member **Ximena,** I hope you have the PhD experience you dream of! Other colleagues that have looked out for me – **Nick, Elisa, Carl-Johan, Nat** – I've always had a blast hanging out! Other colleagues that have always lent a helping hand for all things administrative or technical – **Camilla, Agnieszka, Lena, Sara**.

The MHC group – **Helena, Emily, Anna, Hannah, Samantha** – working on similar topics, going to conferences and having weekly journal clubs together has really helped me learn so much from you (and a lot of fun was had along the way). Thank you for that!

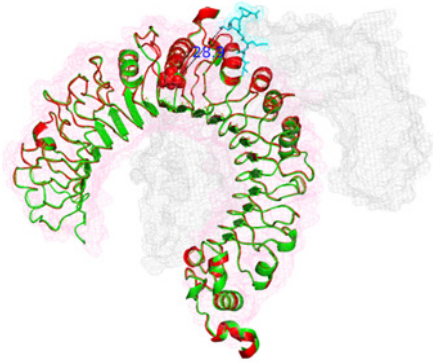
The **MEEL group and everyone at Ecology** – for always making me feel welcome and including me in your activities even though I was from the "dark side" across the stream! **Charlie**, thank you for being my half-time opponent, I learnt a lot from our discussion!

A whole bunch of brilliant fellow PhDs and colleagues that have enriched the work atmosphere and added enjoyment (and parties) to my life – **Simon, Hongkai, Karla, Elsie, Emma, Kalle, Zach, Sofie, Erica, Pedro, Jane, Dima, Andre, Daniel, Zaide, David, Qinyang, Robin and many others I am probably forgetting (sorry!)**.

I've particularly enjoyed and appreciated teaching; students have made me question my own understanding and forced me to explain things better and I am thankful for the experience. It's been even more rewarding to mentor thesis students – **Veda, Nisha, Max H, Samuel, Cornelia, Ayushi** – it was a pleasure helping you. **Elin LÅ**, thank you for your hard work and pulling the weight for Paper III!

A huge thanks to all the funding agencies and research infrastructure that have enabled this work – Swedish Research Council, Jörgen Lindström Foundation, Royal Physiographic Society, Carl-Trygger Foundation, National Bioinformatics Infrastructure Sweden, Department of Biology DNA Sequencing Facility. And finally, to all the little bank voles that have contributed to this research – thank you!

Personally, writing up the thesis, I have come to realise and appreciate just how interdisciplinary the thesis and my learning has become – integrating and testing ideas and aspects from evolution, ecology, immunology with methods both old and new. Yet with it comes a nagging feeling of being a "Jack of all trades, master of none". I hope to continue my research journey by stewing in this intersection of ideas, and hopefully someday I become "Jack of all trades, master of some".



List of papers

I. **Mridula Nandakumar**, Max Lundberg, Fredric Carlsson, Lars Råberg. Positive selection on mammalian immune genes – disentangling effects of gene function from expression and protein-protein interactions (*Manuscript*)

II. **Mridula Nandakumar,** Max Lundberg, Fredric Carlsson, Lars Råberg. 2023. Balancing selection on the complement system of a wild rodent. *BMC Ecology and Evolution*. 23, 21. DOI: 10.1186/s12862-023-02122-0 (*Published*)

III. Elin Laike Åsberg, **Mridula Nandakumar**, Lars Råberg. Polymorphism in the complement system gene *FCNA* is associated with *Borrelia* infection in a wild rodent (*Manuscript*)

IV. **Mridula Nandakumar**, Mehrnaz Nouri, Christine Valfridsson, Fredric Carlsson, Lars Råberg. Unravelling the effect of balanced polymorphisms in bank vole *TLR2* on immune responses using *ex vivo* assays (Manuscript)

V. **Mridula Nandakumar**, Max Lundberg, Fredric Carlsson, Lars Råberg. *Cis*-regulatory variation for immune responses to two bacterial pathogens in a wild rodent: testing for trade-offs (*Manuscript*).

Printed by Media-Tryck, Lund 2024 NORDIC SWAN ECOLABEL 3041 0903

LUND UNIVERSITY

Lund University
Faculty of Science
Department of Biology
ISBN 978-91-8039-994-4