From Medical Biochemistry and Biophysics
Karolinska Institutet, Stockholm, Sweden

# POSITIONAL CLONING OF POLYMORPHIC LOCI THAT CONTROL AUTOREACTIVE T CELLS

Mike Aoun

Stockholm 2021

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021
© Mike Aoun, 2021
ISBN 978-91-8016-286-9

# Positional cloning of polymorphic loci that control autoreactive T cells
## THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

**Mike Aoun**

The thesis will be defended in public at Eva & Georg Klein, Biomedicum 3rd floor, Karolinska institutet, Solna, Friday the 3rd of September 2021 at 13:00.

*Principal Supervisor:*
Dr. Liselotte Bäckdahl, PhD
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Medical Inflammation Research

*Co-supervisor(s):*
Professor Rikard Holmdahl, MD, PhD
Karolinska Institutet
Department of Medical Biochemistry and Biophysics
Division of Medical Inflammation Research

*Opponent:*
Regents and Distinguished McKnight Professor
Marc Jenkins, PhD
University of Minnesota
Department of Microbiology and Immunology
Director of the Center for Immunology (CFI)
Faculty of Microbiology, Immunology and Cancer Biology

*Examination Board:*
Professor Martin Rottenberg, PhD
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Professor Gunilla Karlsson Hedestam, PhD
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Professor Mikael Sigvardsson, PhD
Lund University & Linköping University
Department of Biomedical and Clinical Sciences at Linköping University
Division of Molecular Hematology at Lund University

There are moments in moist love when heaven is jealous of what we on Earth can do.

Hafiz

*Entangled life*

# POPULAR SCIENCE SUMMARY OF THE THESIS

The immune system evolved over the years to serve its purpose in protecting us from pathogens. This phenomenon mimics a country (body) that recruits soldiers or policemen (immune cells) to defend the nation against invading armies (pathogens).

On a normal day, the human immune system resembles the city of Stockholm during '*Midsommar*', where the very few patrolling policemen tidily circulate the city to maintain an optimal homeostatic flow.

However, this is not always the case. Just like us, our immune system is not perfect. Errors in this complex system can occur, and if not rectified properly, will lead to what we call 'autoimmunity'. In a sense, autoimmunity represents a civil war where your immune cells attack your own body parts. To date, more than 100 autoimmune diseases have been identified, and their nomenclature depends on the body part inflicted.

When the joints are affected, we call it rheumatoid arthritis.
When the salivary glands are the target, we call it Sjogren's syndrome.
When the brain and spinal cord are attacked, we call it multiple sclerosis.
When the kidneys are touched, we call it system lupus erythematosus.

We still do not understand exactly how, why this occurs and what are the main triggers for this is crucial if we want to cure these diseases.

Yet, we are certain about the interplay between genetic and environmental factors that destabilize the immune system.

Although humans share 99.9% of their genetic makeup confirmed by the analysis of the Human Genome Project in 2003, the remaining 0.1% difference could hold the answer for why some people are susceptible to a disease while others are not.

This genetic variation is due to switches in nucleotides, the 4 letters (A, T, C, G) that represent the building blocks of our DNA, between one individual and other, hence the name single nucleotide polymorphism (SNP). In this thesis, I try to discuss some genes and SNPs within genes that regulate RA by heavily influencing components of the immune system.

# ABSTRACT

Autoimmune diseases are expensive diseases, both financially and socially, because they are in most cases life long and require a long-lasting treatment. They can also be life threatening, as they are among top 10 causes of death in women and therefore, they represent a serious clinical problem and require careful medical management. There is an autoimmune disease for every system or organ and the one discussed mainly in this thesis is rheumatoid arthritis (RA).

RA is the most common autoimmune disease and primarily affects the joints. It is more prevalent in women than man (3:1) and known to have a strong genetic component for its predisposition. Animal models offer an attractive strategy to study the underlying arthritis-associated genes under a controlled environment. Forward genetics is an approach to dissect genes that regulate a desired trait, from phenotype to genotype. However, this method often fails to pinpoint a single gene but identifies sized genetic fragments that harbor several point mutations inside or outside the genes. These single nucleotide polymorphisms or SNPs in short could have a dramatic effect on the transcription or translation of the gene. Once the arthritis-associated SNP is identified and the gene of interest is positionally cloned, the search for a mechanism of action commences.

All the studies included in this thesis have forward genetics as a common denominator except the last manuscript. In **Study I**, we positionally cloned the arthritis regulating gene in congenic rats that derive from a cross between an arthritis susceptible (DA) and resistant (E3) inbred strains. We found that a SNP in the second exon of *clec4b* caused the introduction of a stop codon and lead to the abrogation of the coded protein Dcar. Arthritis susceptibility increased drastically because of this loss of function mutation and influenced profoundly T cells which are the major driving force of induced arthritis models in rats. In **Study II**, a natural polymorphism in the promoter of *Vdr* was cloned in congenic mice and was found to be overexpressed selectively in T cells without any calcemic abnormalities. Consequently, antigen-specific T cell responses were enhanced, and autoimmunity was worsened. In **Study III**, we set out to investigate the interaction between the two most significant arthritis QTL in rats that were previously positionally cloned, *Ncf1* and *Clec4b*. We found that Dcar and Ncf1 regulate in concert arthritis severity and their expression on neutrophils influenced their capacity to produce reactive oxygen species. In **Study IV**, we ventured to answer one of immunology's longstanding question; *are antigen-specific B cells positively selected?* We show that indeed collagen type-II (Col2) specific B cells are positively selected and could be detected in transgenic mice, wild type mice and inbred rats. Moreover, we cloned the B cell receptor from human blood donor, expressed the antibody and validated its specificity to Col2. Moreover, we demonstrate that these Col2-specific B cells inhibit arthritis by inducing Col2-specific regulatory T cells.

The data illustrated in this thesis confirm in part, the importance of animal models in deciphering disease regulating genes and highlight the implication of antigen specific B cells in a prominent autoimmune disease.

# LIST OF SCIENTIFIC PAPERS

I. Backdahl, L.*, **Aoun, M.***, Norin, U., Holmdahl, R. Identification of Clec4b as a novel regulator of bystander activation of auto-reactive T cells and autoimmune disease. *PLoS Genet.* **16**, 1008788-1008808 (2020). doi:10.1371/journal.pgen.1008788

II. Fernandez Lahore, G., Raposo, B., Lagerquist, M., Ohlsson, C., Sabatier, P., Xu, B., **Aoun, M.**, James, J., Cai, X., Zubarev, R. A., Nandakumar, K. S., Holmdahl, R. Vitamin D3 receptor polymorphisms regulate T cells and T cell-dependent inflammatory diseases. *Proc. Natl. Acad. Sci.* **117**, 24986-24997 (2020). doi:10.1073/pnas.2001966117.

III. **Aoun, M.**, Cai, X., Xu, Bingze., Bonner, M., He, Y., Frenandez Lahore, G., Bäckdahl, L., Holmdahl, R. Glycan activation of Clec4b induces reactive oxygen species and protects against neutrophilia and arthritis [Manuscript].

IV. **Aoun, M.***, Saxena, A.*, Coelho, A., Krämer, A., Sabatier, P., Beusch, C., Lönnblom, E., Nguyen, N., Zhang, J., He, Y., Li, T., He, C., Fernandez Lahore, G., Xu, B., Viljanen, J., Rorbach, J., Kihlberg, J., Kastbom, A., Sjöwall, C., Gjertsson, I., Zubarev, R., Burkhardt, H., Holmdahl, R. Bone marrow-selected antigen specific regulatory B cells [Manuscript].

* These authors contributed equally

# CONTENTS

1 INTRODUCTION........................................................................................5
  1.1 Darwinian Immunology ........................................................................5
  1.2 The innate immune system....................................................................5
    1.2.1 History bits ....................................................................................5
    1.2.2 Introduction to PRRs........................................................................6
    1.2.3 Stratification of PRRs .....................................................................6
    1.2.4 Evolution of CLR............................................................................6
    1.2.5 Ligands and signaling .....................................................................7
    1.2.6 Phagocytosis and NOX2..................................................................7
    1.2.7 Reactive species and cellular signaling ............................................9
  1.3 The adaptive immune system................................................................9
    1.3.1 History bits .....................................................................................9
    1.3.2 Evolution of the adaptive immune system .......................................9
    1.3.3 Lymphocytes ..................................................................................10
    1.3.4 Central tolerance .............................................................................10
    1.3.5 Peripheral tolerance.........................................................................11
  1.4 Autoimmune and Inflammatory Diseases (AID) .....................................12
    1.4.1 Genetic predisposition to AID ..........................................................13
    1.4.2 Rheumatoid Arthritis........................................................................16
    1.4.3 Genetic predisposition to RA............................................................16
  1.5 Rodents as model organism ...................................................................17
    1.5.1 Autoantigen immunization ...............................................................18
    1.5.2 Transfer models...............................................................................18
    1.5.3 Environmental/adjuvant triggers ......................................................19
2 RESEARCH AIMS......................................................................................21
3 MATERIALS AND METHODS ..................................................................23
  3.1 Arthritis induction .................................................................................23
  3.2 Carrageenan air-pouch model ................................................................23
  3.3 Flow cytometry.......................................................................................23
  3.4 Extracellular and intracellular ROS detection .........................................24
  3.5 Magnetic bead-based enrichment ..........................................................24
  3.6 C1-B cell fluorescence-activated cell sorting...........................................25
  3.7 *Ex vivo* expansion of human B cells ...................................................25
  3.8 Ethical consideration ..............................................................................25
4 RESULTS...................................................................................................27
  4.1 Study I: Identification of Clec4b as a novel regulator of bystander activation of auto-reactive T cells and autoimmune disease.............................27
  4.2 Study II: Vitamin D3 receptor polymorphisms regulate T cells and T cell-dependent inflammatory diseases .........................................................28
  4.3 Study III: Glycan activation of Clec4b induces reactive oxygen species protecting against neutrophilia and arthritis ..........................................28

4.4 STUDY IV:Bone marrow-selected antigen specific regulatory B cells............29
5 CONCLUSIONS........................................................................................31
6 ACKNOWLEDGEMENTS.........................................................................33
7 REFERENCES.........................................................................................37

# LIST OF ABBREVIATIONS

| | |
|---|---|
| MACS | Macrophages |
| DCs | Dendritic Cells |
| PMN | Polymorphonuclear |
| PRRs | Pathogen Recognition Receptors |
| PAMPs | Pathogen Associated Molecular Patterns |
| TLRs | Toll-like Receptors |
| CLRs | C-type Lectin Receptors |
| NLRs | Nucleotide binding domain or NOD-like Receptors |
| RLRs | RIG-I-like Receptors |
| ALRs | AIM2-like Receptors |
| Dectin-2 | Dendritic cell-associated C-type lectin-2 |
| MCL | Macrophages C-type Lectin |
| Mincle | Macrophage Inducible C-type Lectin |
| Dcar | Dendritic Cell immunoActivating Receptor |
| BDCA-2 | Blood Dendritic Cell Antigen 2 |
| ITAM/ITIM | Immunoreceptor Tyrosine-based Activator/Inhibitor Motif |
| NOX2 | Nicotinamide adenine dinucleotide phosphate (NADPH) OXidase 2 |
| ROS | Reactive Oxygen Species |
| Rag1/2 | Recombination Activating Gene 1 and 2 |
| MHC I/II | Major Histocompatibility Complex 1 and 2 |
| SNP | Single Nucleotide Polymorphism |
| Mtb | Mycobacterium Tuberculosis |
| Col2 | Collagen type-II |
| CIA | Collagen Induced Arthritis |
| ACPA | Anti-Citrullinated Protein Antibody |

# 1 INTRODUCTION

## 1.1 DARWINIAN IMMUNOLOGY

According to historical data, seamen on board of HMS *Terror* and *Erebus* suffered from an obscure disease dubbed as '*consumption'* and died shortly after venturing out into the north pole in 1845. That common evil wreaked havoc among men and women but no one knew what caused it[1]. 13 years later '*On the Origin of Species*' was published by Charles Darwin after his intense voyage on the HMS *Beagle*. Shortly after this monumental book that describes natural selection as a mechanism for evolution, infectious agents were discovered and the Darwinian struggle – human *vs* microbes – was established[2]. To fully grasp this perspective, immunology as a field surfaced at the turn of the 19th century as the science of host defense. The announcement of Rober Koch in March of 1882 shook the world as he discovered the *consumption* bacterium, which happened to be the *tubercle bacillus*[3]. A big bang in the field of immunology followed and an established view on the immune system was concurred. The fight between Paul Ehrlich and Elie Metchnikoff in the beginning of 20th century manifested into two competing immunological defense mechanisms (1) antibodies of the humoral system[4] and (2) eerie amoebae-like cells known as phagocytes[5]. Furthermore, Burnet's description of lymphocytes that are selected within the body for purging against self-reactive clones and expanding the much-needed pathogen-reactive cells is yet another testament for Darwin's ideas and fingerprints. Today, we divide the immune system into two distinct yet intertwined fractions: the innate and the adaptive immune system.

## 1.2 THE INNATE IMMUNE SYSTEM

### 1.2.1 History bits

The observed mechanisms of amoeba-like structure devouring microbes led the Russian Zoologist Elie Metchnikoff to the theory of phagocytosis. The process was not so alien as unicellular organisms undergo phagocytosis to acquire food and as evolution dictates the process morphed into a mechanism of tissue maintenance and remodeling as well as host defense[6] – another pillar or Darwinian immunology. Today, the internalization of diverse particulate components – including dead cells, pathogens, and environmental debris[6] – by cells of the innate immune system known as phagocytes is a key mechanism of innate immunity.

### 1.2.2 Introduction to PRRs

Exactly how cells of the innate immune system, consisting of macrophages, DCs and PMN cells, discriminate between self and pathogen is based on myriad of surface receptors. Receptors that recognize pathogens (PRRs) are germline encoded and vital in keeping homeostasis by constant bacterial clearance. For any given pathogen, a combination of conserved microbial motifs, denoted as PAMPs, form the basis of pathogen recognition by PRRs. The relevance of PRRs become evident when the sheer number of bacteria we host is taken into consideration. In fact, latest estimations of the total number of bacteria in a 70kg ‘reference man’ to a staggering 38 trillion (that is 1 with 12 zeros)[7]. Although the vast majority of these bacteria enjoy a commensal relationship with our body, it is hard to accurately estimate the noxious pathogens that bombard us constantly from our environment. For that PRRs evolved into a myriad of receptors, while a common denominator lies in their ability to induce a strong transcriptional response, which in turn orchestrate the entire immune system[8–10].

### 1.2.3 Stratification of PRRs

PRRs can be stratified into 5 major families consisting of TLRs, CLRs, NLRs, RLRs and ALRs. CLRs and TLRs mainly survey for the presence of extracellular ligands due to their presence on the cell surface or on endocytic compartments whereas the remaining families of receptors are typically found in the cytoplasm primarily recognizing intracellular ligands. While the award winner TLRs are the most studied type of receptors, CLRs come next forming a conglomerate of at least 17 subgroups of receptors segregated into different subfamilies based on their phylogeny and domain organization[11,12].

### 1.2.4 Evolution of CLR

Carbohydrate-binding receptors have ascended early during evolution and were found in yeast species (*S. cerevisiae*), nematode (*C. elegans*) and fruit fly (*D. melanogaster*). Conversely, distinct evolutionary patterns were observed when comparing rodent to human immune system[13]. CLRs that bind endogenous ligands and responsible for clearance of glycoproteins were conserved across species, examples of one-to-one orthologues include the mannose receptor and scavenger receptor C-type lectin. On the contrary, receptors that bind exogenous ligand, usually of pathogen nature, were forced to undergo dramatic changes due to the evolutionary pressure, leading to the absence of common orthologues between species[13]. For instance, Dectin-2 and MCL, which are localized adjacent to one another, were likely the result of gene duplication of Mincle as the latter seem to be highly conserved across the species[14]. Of note, Dectin-2 in the rat is a pseudogene, at least in the inbred strains listed in this thesis, while being a functional receptor in humans. Another example for the absence of a direct orthologue between rodents and humans is Dcar. Interestingly, the mouse possesses two isoforms (mDcar and mDcar1) while only one Dcar protein can be found in the rat. Even though Dcar is absent in humans, the closest orthologue is BDCA-2, which is a marker for plasmacytoid cells.

### 1.2.5 Ligands and signaling

CLR nomenclature stems from their ability to bind sugar residues in a $Ca^{2+}$-dependent (or independent in some cases) manner via conserved motifs in the carbohydrate recognition domain. The EPN (Glu-Pro-Asn) and the QPD (Gln-Pro-Asp) motifs determine the binding affinity towards mannose-type and/or galactose-type residues, respectively[15]. Substituting EPN with QPD motif, through site-directed mutagenesis, resulted in a switch in sugar specificity (from mannose to galactose). Nevertheless, specificity to sugar moieties does not exclusively depend on these two motifs. For example, an amino acid replacement in the region situated close to the EPN or QPD motifs can also modify sugar specificity[16]. Pathogen recognition by CLRs is well documented[17] yet ligand specificity remains a technical challenge due to the weak ligand affinity and restricted accessibility to advanced screening method[18].

CLRs like Mincle, Dcar, Dectin-2 and BDCA-2 lack an intracellular signaling domain, therefore they are linked to an ITAM-bearing Fc receptor $\gamma$ subunit. Phosphorylation of the ITAM-tyrosine(s) by different kinases, mainly the spleen tyrosine kinase (Syk), elicits a cascade of adaptor proteins that gets recruited and an avalanche of transcriptional regulations proceeds. CLRs downstream pathways are not set in stone as different reports demonstrate the versatility of the signaling machineries[19–21]. In summary, once triggered by an endogenous or exogenous stimulus, CLRs induce a complex signaling pathway to mount an immediate immune response that dictates the fate of the adaptive immune response.

### 1.2.6 Phagocytosis and NOX2

Oxidative burst is dubbed as the release of short-lived, oxygen/sulfur/chlorine/nitrogen-derived molecules [*i.e.*, hydroxide ($^{\bullet}OH$), Peroxide ($H_2O_2$), superoxide anion ($O_2^-$) and hypochlorous acid (HOCl)] by specialized cells called phagocytes[22]. Monocytes/macrophages and neutrophils were primarily identified to produce ROS, after phagocytosis of the invading pathogen, to eliminate the microorganism and thereby serving as a mediator of inflammation[23]. Even though ROS can originate mainly from two sources namely mitochondria (by-product of the electron transport chain, ETC) and NADPH oxidase, we will discuss the latter in this review. To date, the family of NADPH oxidase termed as NOX comprises seven members of membrane bound proteins, NOX 1 through 5, DUOX 1 and 2. In fact, these enzymes share the ability to transfer electrons through plasma membranes ultimately producing ROS while they differ in their expression pattern and adaptor molecules needed to perform their function[24]. For instance, NOX2 complex possess a transmembrane catalytic core formed by the heterodimerization of glycoproteins (CYBB) and (CYBA) whereas the cytosolic domain requires coupling of accessory components like P47phox (alias NCF1), P67phox (alias NCF2), P40phox (alias NCF4), and RAC[25]. Loss of function mutation in any of these components can result in defective ROS production, which is the case of chronic granulomatous disease (CGD) patients that exhibit hyper-inflammatory status with a significant increase in morbidity due to infections[26].

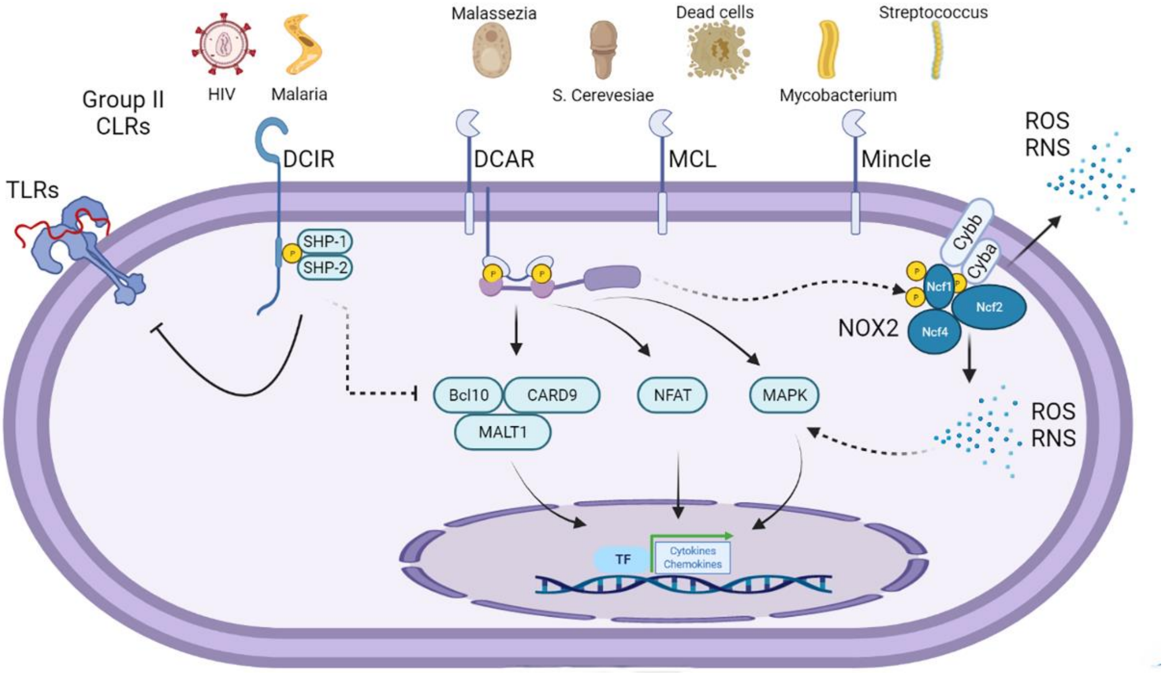
Malassezia
Dead cells
Streptococcus
HIV
Malaria
S. Cerevesiae
Mycobacterium
Group II CLRs
DCIR
DCAR
MCL
Mincle
ROS
RNS
TLRs
SHP-1
SHP-2
P
Cybb
Cyba
Ncf1
Ncf2
Ncf4
NOX2
Bcl10
CARD9
MALT1
NFAT
MAPK
ROS
RNS
TF
Cytokines
Chemokines

**Figure 1 | Group II CLRs, their known ligands and signaling pathways.** Dcir (all 4 isoforms) contain an ITIM in its intracellular domain while all other members of the group II CLRs (Dcar, MCL and Mincle) are bound to an ITAM-bearing Fc receptor γ-chain. Upon binding its cognate ligand Dcir signaling pathway involves the recruitment of spleen tyrosine phosphatase (SHP), which downregulates TLR-induced cytokine production and inhibits the formation of Card9-Malt1-Bcl10 (CMB) signalosome. Alternatively, Dcar/MCL/Mincle signal through the spleen tyrosine kinase (SYK), which is required for the assembly of the CMB signalosome, the activation of the mitogen-activated protein kinase (MAPK), the initiation of the calcineurin-NFAT pathway and the induction of the ROS producing NOX2 complex. All these proteins regulate transcription factors (TF) in the nucleus that dramatic shape the immune response in terms of cytokine/chemokine production and ROS release.

### 1.2.7 Reactive species and cellular signaling

21000 to 42000 thiols in the cysteine, methionine and selenocysteine proteome, including receptors, adaptor molecules, transporters, enzymes, heat shock proteins, ribonucleoproteins, growth factors[27] and transcription factors[28], can be oxidized by reactive species[29,30]. Iron-sulfur (Fe–S) clusters in metalloproteins that regulate ETC in the mitochondria are also sensitive to redox changes and add another layer in redox signaling diversity[31]. Due to the critical role of reactive species in influencing cellular signaling, a sensitive antioxidant system has evolved to maintain healthy homeostasis, also known as ‘oxidative eustress’. While supraphysiological level of reactive species is often denoted as ‘oxidative distress’ and therefore is countered by sets of enzymes that catalyze the reduction of these reactive species, for example, the peroxidases that is crucial for the removal of $H_2O_2$[32]. Probably the prime example to illustrate this tight regulation is the Nrf2-Keap1 system. Under physiologic condition Nrf2 and Keap1 form a complex in the cytosol, leading to the ubiquitination of Nrf2 followed by degradation by the proteasome. However, under oxidative distress the complex is dissociated because of the conformation changes in Keap1 after Cys151 oxidation (disulfide formation), this drives the release of Nrf2 that is subsequently translocated to the nucleus. Once there, Nrf2 binds to the antioxidant response elements (ARE) and promotes the expression of various detoxification, antioxidant and anti-inflammatory genes[33].

## 1.3 THE ADAPTIVE IMMUNE SYSTEM

### 1.3.1 History bits

The earliest report about a constituent of the adaptive immune system originates from a paper by Emil von Behring and Kitasato Shibasaburo. While working together in 1980, they demonstrated that serum transfer from diphtheria immunized animals to animals suffering from it, could cure the recipient animals[34]. Sadly, only one of the two received the Nobel Prize in 1901. This landmark publication was carried further by Paul Ehrlich (considered as godfather of immunology) who proposed the first model for antibody-pathogen binding[35]. Since then, the theory behind the proposed model never ceased to evolve to this day.

### 1.3.2 Evolution of the adaptive immune system

Unlike the components of the innate immune system (PRRs) that can be found throughout the animal kingdom, the adaptive immune system evolved some 400 million years ago and remained remarkably uniform over the years. The answer to how this system evolved lies in 2 macroevolutionary incidents (1) transposon molecular domestication and the (2) whole genome duplication. These events provided enough raw material and innovation to fuel the emergence of the adaptive immune system[36].

Genes that are associated with the adaptive immune system such as RAG1 and RAG2 emerged in jawed vertebrates, while invertebrates possessed a RAG1-like and RAG2-like genes[37]. The presence of RAG precursor genes in the genomes of ancient sea urchins and amphioxi support the model of transposon molecular domestication[38].

Susumu Ohno proposed this theory in 1970 and described that the genome of vertebrates suffered 2 rounds of whole genome duplication. As a result, a big bang of orthologues and paralogues arose that were essential in forming the adaptive immune system. For example, invertebrates harbor only 1 MHC-like (proto-MHC) region, which lacks MHC class I and II genes, while modern day human possesses 4 paralogons of the MHC genes on chromosome 1, 6, 9 and 19. This could only be explained by a duplication event that occurred in the past – predicted to be at the stage of fish or amphibians[39,40].

### 1.3.3 Lymphocytes

Unlike the immediate response of the innate system (minutes to hours), the adaptive system respond in a rather slower pace (days to weeks) but results in an antigen specific response as well as an immunological memory (T and B lymphocytes)[41]. The specificity of lymphocytes is produced by recombination processes *via* RAG genes, described earlier, which randomly assemble unique and diverse TCR and BCR. The repertoire generated through this process contains receptors that will bind self-proteins resulting autoreactive T and B cells. Therefore, mechanisms of tolerance evolved to keep this procedure in check and purge self-reactive cells, hence the name ‘negative selection’.

### 1.3.4 Central tolerance

T cells emigrate from the bone marrow to the thymus for antigenic education. During thymic development, many T cells that bear a randomly rearranged receptor will die by neglect because these receptors are unable to bind MHC molecules. Whereas receptors that show adequate binding to the individual’s MHC molecule, which is loaded with a self-peptide, will be positively enriched to undergo another round of selection. The reason for the second round of selection is due to the emergence of autoreactive T cells with different affinity to self-antigens during this step. Upon reaching the medulla, 3 mechanisms dictate the fate of exiting T cells. The main one being *clonal deletion*, as T cells with high affinity to self-antigens undergo suicide although the detailed molecular pathway for that process remain enigmatic. *Anergy* is the second mechanism in which T cells become functionally inactive and alive yet in a hyporesponsive state meaning that T cells cannot respond to antigens under adequate conditions. The last mechanism is the process of *receptor editing* which allows the rearrangement of surface receptors following initial binding in hopes of replacing autoreactive receptor with an innocuous one. On the T cell side, 2 crucial transcription factors govern these mechanisms, the autoimmune regulator (AIRE) and fork-head box P3 (Foxp3). Ablation or mutations in any of these genes lead to aggressive autoimmune diseases that target many organs[42]. T cells however are only half of the story as B cells represent the other part of the adaptive immune system. Likewise, mechanisms of tolerance arose to govern B cell selection which occurs in the bone marrow. Just like T cells, B cells can develop autoreactive clones and secrete self-reactive antibodies that are used as markers to diagnose autoimmune disease. Thus, *clonal deletion* and *receptor editing* are two well described mechanisms to counter the abundance of autoreactive B cells in the normal repertoire[43]. However, positively selected autoreactive B cells are still nebulous, and a potential regulatory function could explain their persistence under normal conditions.

### 1.3.5 Peripheral tolerance

Low avidity yet self-reactive T cells can escape central tolerance and can populate the periphery and thymic deletion of autoreactive T cells have been shown to be only 60 to 70% efficient. Therefore, peripheral checkpoints are in place to critically regulate the activity and pathogenicity of these dangerous T cell clones. Examples of T-cell peripheral checkpoints are quiescence, ignorance, anergy, exhaustion, senescence, and cell death[44]. *Quiescence* is a process restricted to naïve T cells to halt their expansion at the steady-state level. This is accomplished by different transcription factors that reduce the basal metabolic program and ensure a low rate of translation in naïve T cells. *Ignorance* is the least understood mechanism but can be briefly described as an autoreactive T cell that fail to induce autoimmunity even in the presence of its cognate self-antigen[44]. *Anergy* is the result of imbalanced TCR stimulation in T cells leading to a state of hypo responsiveness, however this is a reversible state and could be dependent on antigen availability[45]. It is difficult to distinguish *Exhaustion* from *Anergy* but the hallmark phenotypic changes of exhausted T cells are decreased cytokine production, elevated levels of inhibitory surface molecules, epigenetic and metabolic signatures and failed ability to mimic memory T cell-state of quiescence[44]. *Senescence* is defined when a cell reaches its end replicative stage or is subjected to different stressors and correlates with ageing[46]. The evolutionary reason behind such mechanism argue that it protects from T cell lymphoma. Even though peripheral induction of *cell death* equals that of central tolerance, the exact mechanisms are slightly overlooked. Nevertheless, studies show that it is primarily mediated by FAS and triggering caspase 8 after restimulation of T effector cells, a process that parallels the contraction phase after a productive immune challenge[47,48].

## 1.4 AUTOIMMUNE AND INFLAMMATORY DISEASES (AID)

Earliest description of an AID dates back to the Greek philosopher Hippocrates, meaning that this is an ancient disorder and not new[26]. In fact, *'Horror Autotoxicus'* or the horror of self-toxicity was a term coined by Paul Ehrlich at the beginning of the 20th century to describe autoimmune diseases[4]. Afterwards, it was for Sir Burnet's "Clonal Selection Theory" that contributed to a clearer notion regarding the selection of antigen-specific T and B cells and their clonal expansion when encountered with foreign antigens[49]. Yet, the system is far from perfect and in some cases 'forbidden clones' can be generated with a reactivity towards self-antigens leading to disturbance in homeostasis and causing tissue damage – hallmark of autoimmune diseases. However, it is now clear that AID form an umbrella that encompasses numerous diseases namely, RA, SLE, MS, T1D and should be addressed collectively instead of individually. In fact, prevalence rates of AID ranges from 5 to 500

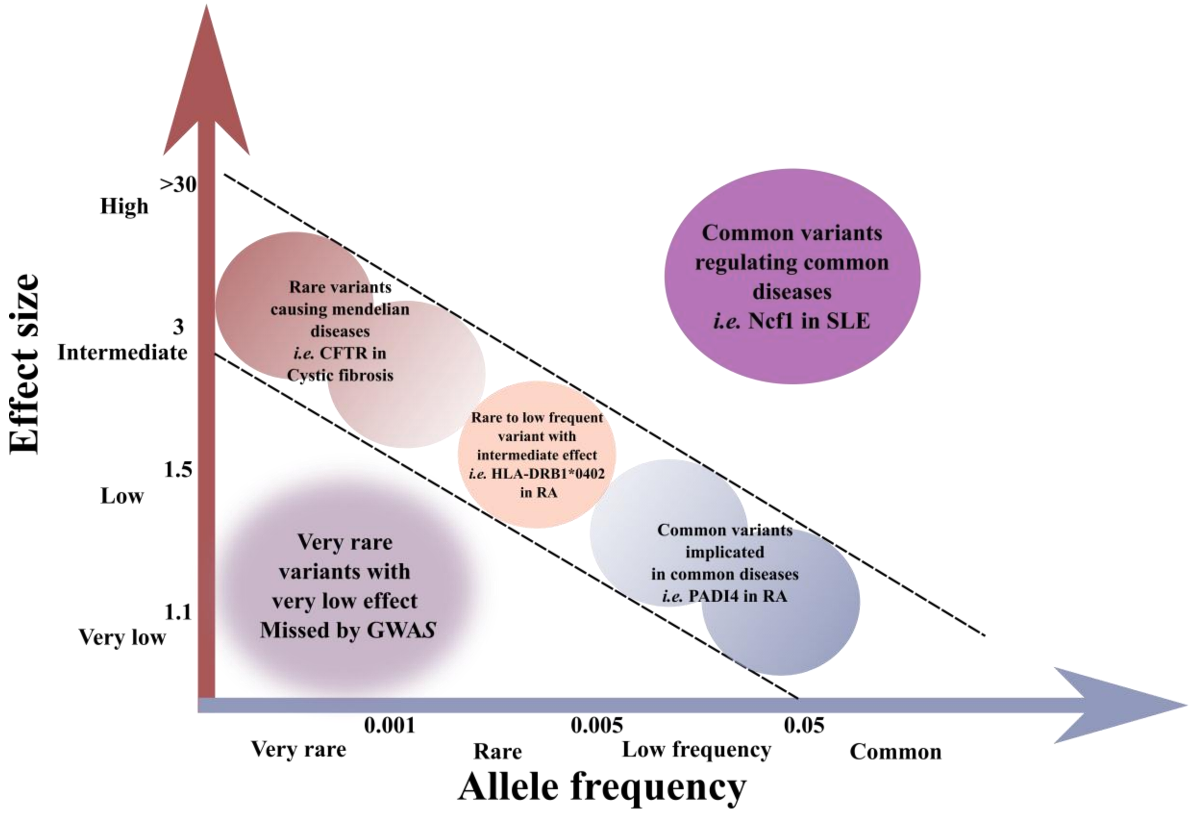
Effect size
High
>30
3
Intermediate
1.5
Low
1.1
Very low
Rare variants causing mendelian diseases i.e. CFTR in Cystic fibrosis
Common variants regulating common diseases i.e. Ncf1 in SLE
Rare to low frequent variant with intermediate effect i.e. HLA-DRB1*0402 in RA
Common variants implicated in common diseases i.e. PADI4 in RA
Very rare variants with very low effect Missed by GWAS
0.001
0.005
0.05
Very rare
Rare
Low frequency
Common
Allele frequency

**Figure 2 | Practicability of finding genetic variants by effect size (odds ratio) and allele frequency.** GWAS outshine in identifying common variants implicated in common diseases and most of these variants have modest to low effect size**.** Variants with low to rare frequency are not efficiently captured by GWA SNP-arrays, and once the frequency drops below 0.005 it becomes difficult to pick up the association by GWAS unless the effect size is very large (Mendelian diseases). Adapted from ref. 71

per 100,000, which poses a mortality burden among the general population, and a critical socio-economical pressure (estimated to be over 30€ billion/year in Europe alone) on the society as a whole[50].

### 1.4.1 Genetic predisposition to AID

Autoimmune diseases are complex diseases due to the numerous factors contributing to their aetiology. In fact, it is postulated that the combination of genetic factors, environmental and epigenetic manifestation sets off the development of AID[51]. Studies on monozygotic twins identified a high concordance rate (75-83%) for some AID namely, Coeliac Disease (CD), whilst relatively low concordance rate (12-30%) for other AID like RA, MS, and SLE[52]. However, these studies might not be completely reflective of the hereditary status of AID as they contain certain limitations including, cohort size, subjective questionnaire, lack of sufficient follow-up[52]. Contrariwise, studies done on Finnish and UK monozygotic twins identified a 60% heritability to RA, which raises the question of how accurate these numbers are as the discrepancies are quite large (12 to 60%)[53]. In order to scavenge the genome for genetic association to AID, methods like genome-wide association studies (GWAS) and linkage mapping offer the possibility to pinpoint causal variants.

#### 1.4.1.1 Genome wide association studies

Completion of the human genome in 2003 led to major scientific advances in the field of medical genetics and especially the expansion of GWAS[54]. Briefly, GWAS screen thousands of unrelated patients and controls in search for a common SNPs that underlie a specific autoimmune disease[55]. Usually, a GWAS is carried out by analyzing hundreds of thousands of SNPs in a cohort of patients vs healthy controls followed by the determination of disease causing variants[56]. As of 2020, more than 4500 GWAS have been published with more than 55 000 single nucleotide variants identified to be associated to approximately 5000 traits, and these numbers will continue to soar[57]. Applying GWAS often lead to the discovery of unknown function and/or relevance and follow-up studies catalyzes the identification of *bona fid*e mechanisms regulating traits. For instance, GWAS in Chron's disease identified SNPs in 2 autophagy related-genes (*ATG16L1*[58] and *IRGM*[59]). A follow-up study elegantly demonstrated that these variants enhance the degradation of *ATG16L1*[60] and deregulation of *IRGM*[61], respectively and thus influencing autophagy. Consequently, the link between autophagy and Crohn's disease was established. Moreover, GWAS results can be employed as a classification criterion for diseases. For example, in the case of maturity-onset diabetes of the young (MODY) and type-1 diabetes, GWAS corroborated the use of C-reactive protein as a useful biomarker for the diagnosis of diabetes subtypes[62,63]. GWAS can tap into ethnic heterogeneity and its association to complex traits as some risk loci show variable frequency and/or effect size depending on ethnicity. The association of HLA-DRB1 alleles to RA is a prime example as this association (the strongest to date) exhibit ethical variation depending on the HLA-subtype (HLA-DRB1*0401 in Caucasians and HLA-DRB1*0405 in Asians)[64]. Beyond genetic identification, GWAS provide useful applications in the field of reconstruction of population history[65], ancestral determination[66], paternity examination[67], forensic investigations[68] and copy number variants (CNV)[69] just to name a few.

Nevertheless, GWAS suffer from several limitations. Obviously, the need to adopt high level of scrutiny when it comes to statistical analysis to account for multiple testing (Bonferroni correction) comes at a price. Subsequently, the threshold for significance ($P < 5 \times 10^{-8}$) becomes harder to reach for certain SNPs and the most common strategy to overcome this limitation is by increasing the sample size[70]. In addition, the majority of GWAS identified SNPs are outside the coding sequence of a gene, therefore functional characterization of intergenic SNPs remain a very challenging task[55]. It also fails to spot rare variants of minor effect[71]. Lastly, the reference SNP-arrays used in GWAS studies lack the diversity needed to understand the genetic variation in under-represented populations and efforts in establishing reference arrays for diverse populations is crucial[72].

#### 1.4.1.2 Linkage mapping

Genetic mapping of disease-regulating genes began in the 80s, where linkage analysis using polymorphic genetic markers [restriction fragment length polymorphisms (RFLP), microsatellites or short tandem repeats (STR)] as a tool to conduct forward genetics[73]. The basic principle is that during meiotic recombination, markers that are positioned distant from each other are less likely to be inherited together (linkage equilibrium), whilst markers that are inherited more often than expected are in linkage disequilibrium (LD). To determine the linkage between traits and markers, the logarithm of odds (LOD) score is calculated. A positive, typically above 3, LOD score suggests a significant linkage, yet the threshold of significance varies between different experiments[74]. Linkage analysis emerged as a powerful tool for the discoveries of causative genes behind monogenic Mendelian human diseases, like cystic fibrosis[75]. This type of analysis requires a pedigree of related individuals accompanied by a trait. However, this approach proved to be limited in the study of complex traits in humans and for this reason it is mainly conducted in animals.

This strategy entails crossing an inbred resistant and a susceptible strain for a trait, to find genetic loci that are genetically linked to disease phenotypes in an F2 cross between the two homozygous strains. The loci identified by linkage analysis can then be introgressed into a fixed genetic background (backcrossing to one of the parental strains for many generations), thereby producing a congenic strain. Reducing the congenic fragment (typically in the cM to high Mb range) requires lengthy backcrossing strategies that often employs markers (*i.e.*, microsatellites) as a compass. Each generation should be checked to ensure the presence of the QTL within the fragment. The main pitfall of this technique is the ginormous amount of genotyping needed meaning more hands-on work and that often lead to increased technical mistakes[76]. Another variable is the dependence on good models that mimic the human diseases, hereby complex autoimmune diseases. While it is true that genetic mapping in animals is often conducted under regulated environmental conditions to increase the power of the mapping and limit the variables, it excludes the contribution of environmental factors that can be pivotal in human pathologies[77]. Obviously, allelic

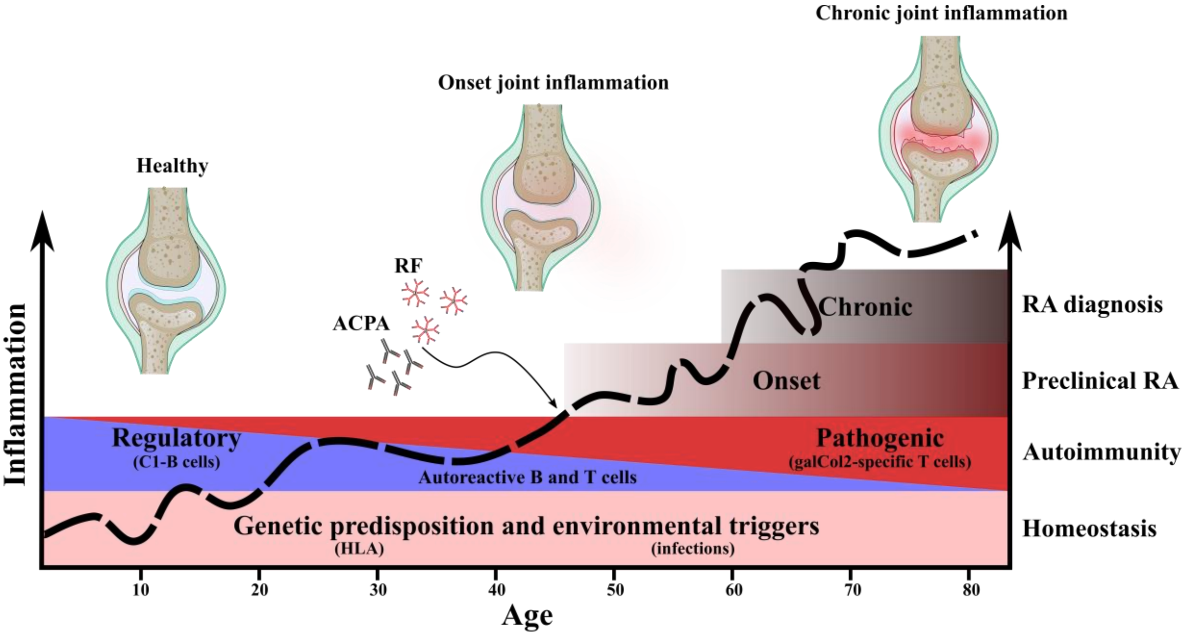
Chronic joint inflammation
Onset joint inflammation
Healthy
RF
ACPA
Inflammation
Chronic
Onset
RA diagnosis
Preclinical RA
Regulatory
(C1-B cells)
Pathogenic
(galCol2-specific T cells)
Autoimmunity
Autoreactive B and T cells
Genetic predisposition and environmental triggers
(HLA)
(infections)
Homeostasis
10
20
30
40
50
60
70
80
Age

**Figure 3 | The development of rheumatoid arthritis (inflammation) over individual's lifespan (age).** RA progression can be divided into 4 different phases (right axis). Homeostasis (healthy joint) is the lag period during which the individual's genetic susceptibility (HLA-subtype) and environmental setup (infections/smoking) influences the tight balance in the immune system (autoreactive B and T cells). Autoimmunity is characterized when the breach in tolerance occurs and the equilibrium between regulatory and pathogenic response is skewed to the point of no return. Anti-citrullinated protein antibodies (ACPA) and rheumatoid factors (RF) emerge at this step and precedes the 3rd phase of RA progression. Preclinical RA is when the individual starts experiencing symptoms (pain, stiffness, weakness, dry mouth) and it coincides with onset of joint inflammation. The last phase is the clinician diagnosis upon fulfillment of established criteria and soon after (if left untreated) the disease develops into chronic joint inflammation leading to irreversible deformities.

diversity is very restricted and the sheer amount of recombination that occurs somehow limits the mapping[78].

Lastly, results from GWAS and genetic mapping have contributed significantly to understanding genetics of complex traits, not only in AID but also in the field of agriculture. Unfortunately, many of the variants identified confer very little increments in risk and fail to fully explain familial clustering. This spiked the question of missing heritability.

### 1.4.2 Rheumatoid Arthritis

RA is a chronic autoimmune disease, the most prevalent of all AID, touching more than 1% of the general population. Clinical manifestation implicated in RA account for articular chronic inflammation that leads to permanent destruction of the articular cartilage on the long run[79]. To this date, the disease aetiology remains a mystery, nevertheless recent advances describe a multifaceted interaction between the environment on one side and genetic predisposition as well as epigenetic changes on the other side[80]. Clinically speaking, RA can be roughly divided into 3 stages, (1) the autoimmune priming in healthy people followed by (2) joint inflammatory symptoms (onset) and finally (3) the chronic inflammation/treatment[81]. The fact that the majority of GWAS studies pinpoint a strong association of HLA-genes to RA susceptibility, and the prevalence of anti-citrullinated protein antibody (ACPA) in RA patients argues for a pivotal role of T and B cells in RA pathology. What exactly unlocks joint reactive T cells and what is the role of ACPA in RA pathology are questions to be addressed for a better treatment and diagnosis.

### 1.4.3 Genetic predisposition to RA

#### 1.4.3.1 HLA loci

First reports on the association between HLA genes and AID date from the 70s[82]. HLA genes reside on chromosome 6 in humans and are divided into 3 classes: class I (A, B and C), class II (DP, DM, DO, DQ, DR) and class III (complement system)[83]. To date, the strongest association to RA (and T1D) exist in the HLA class II region, specifically in the DR alleles encoded by *HLA-DRA*, *-DRB1 –DRB3, 4* and *5* alleles. Class II are heterodimeric receptors and have an extracellular, transmembrane, and intracellular portion while unlike TCR and Ig that undergo somatic mutations and rearrangements, HLA maintain its diversity though; (1) polygenicity, (2) codominance and (3) polymorphism. In fact, this region is immensely polymorphic as it is estimated that DRB1 locus has >1700 alleles. It is established now that certain alleles within the DRB1 locus namely, *DRB1*04:01, *01:01 and *10:01* are associated, largely, in ACPA+ than in ACPA- RA. Interestingly, all these three alleles possess a shared epitope (SE) at amino acid position 70-74 in β1 domain (**QKRAA** at *DRB1*04:01*, **QRRAA** at *DRB1*01:01* and **04:04*, and **RRRAA** at *DRB1*10:01*)[84,85].

#### 1.4.3.2 Non-HLA loci

It is true that the strongest association to RA lies within the HLA locus; however, it is not the only one. Studies on Swedish EIRA cohort found at least three additional HLA unrelated genes to be associated with seropositive RA including *PTPN22, PADI4, CDK6.* Of note, *PTPN22* has been associated to other AIDs namely, SLE, Graves' disease, T1D, among others. Recent GWAS data identified *SNPs* in *TRAF1* gene to be associated with RA, specifically to the increased radiological development. Additionally, *STAT4, CD40* and *CCR6* were identified to have modest associations with RA [86–88].

More and more SNPs are constantly being discovered to have associations with RA, which is predicted due to the advances in regression models and statistical analysis, the use of more controlled and homogeneous cohorts and the technological improvements in SNPs hunting. Therefore, the need for meta-analysis studies is of utmost importance to reproduce or exclude the variants that may or may not have clinical association to RA.

Despite the effort invested in understanding the pathophysiological mechanism, a lot remain unclear about the pathways contributing to the development of RA, which is a major obstacle for drug development[89]. Fortunately, some animal models recreate different aspects of the complex human diseases and their usage depends on the question asked[90]. While they allow a functional dissection of the disease, animal models can also be used for drug discovery and evaluation. Current animal models provide fertile ground for insightful breakthroughs, yet no model shares the exact pathophysiology with the human disease[91]. In short, different models can be used to investigate different cell specific function and this, in fact has become a pre-requisite for any FDA or EMA approval[89].

## 1.5 RODENTS AS MODEL ORGANISM

Animal models are the keystone of basic research. The availability in genetic manipulation due to technological advancements (*i.e.,* CRISPR) and the economic feasibility to breed and house inbred mice strains promoted an outbreak of knowledge in immunology. Despite the differences that exist between rodent and human immunology (*i.e.,* leukocyte subsets, antibody subtypes, and innate immunity components), mice still offer tremendous advantages to abandon[92]. Animal models for autoimmunity, especially RA, diverge into two main groups, the induced and spontaneous models. The former, as the name suggests, is induced artificially while the latter is a spontaneous manifestation due to genetic modifications. The models discussed in this thesis fall into the induced section which can be subsequently categorized into three groups, 1) autoantigen immunization, 2) transfer models and 3) environmental/adjuvant triggers.

### 1.5.1 Autoantigen immunization

Col2 is the major component of the articular joint and intradermal injection of heterologous or autologous Col2 protein emulsified in an adjuvant provokes an aggressive autoimmune response that targets primarily the joints. The adaptive immune system is heavily involved in initiating the disease pathogenesis. It has been shown that anti-Col2 antibodies produced by B cells can bind the joint and induce arthritis[93]. Furthermore, mice fail to develop CIA in the absence of B cells[94]. T cells are also implicated in this model in two main facets, 1) providing T cell help for a potent production of anti-Col2 antibodies[95] and 2) sustaining joint inflammation by stimulating other cells and this has been demonstrated by the amelioration of arthritis post T cell blockade or functional inhibition[96].

A cross between the K/BxN TCR transgenic mice line and the non-obese diabetic (NOD) expressing the H-2A$^{g7}$ result in a spontaneous joint disorder that shares many clinical features of human RA. The targeted protein in this model was discovered to be the ubiquitous glucose-6-phosphate isomerase (GPI), an enzyme critical for glycolysis[97]. Intriguingly, immunization with the heterologous full GPI protein was found to be dependent, in part, on the H-2A$^{q}$. Additionally, mapping of arthritogenic epitopes on the full GPI protein identified a 14 amino acid peptide (GPI$_{325\text{-}339}$) to be a strong inducer of arthritis through B and T cells but without any antibody effector function[98].

### 1.5.2 Transfer models

Even though the role of antibodies in CIA pathogenesis is well documented, characterizing the arthritogenic anti-Col2 monoclonal antibodies was needed to provide a model to study the effect of antibodies and its mechanistic repercussion on arthritis[99]. Arthritis models consist of two separated phases, the priming and acute phase. The former is the early timeframe by which T and B cells gets primed by APC in secondary lymphoid organs while the latter corresponds to the active destruction of the articular joints. Collagen antibodies induced arthritis (CAIA) is a model that recreates the acute phase of arthritis triggered by the injected antibodies and their effect on the innate immune system (strong association to non-MHC genes)[100]. Indeed, a single clone of anti-Col2 antibody does induce a mild arthritis but a cocktail of 4 purified monoclonal antibodies with defined epitopes (C1, J1, D3 and U1) on the Col2 proteins dramatically increases the incidence and severity of arthritis[93]. Although the exact mechanism is yet to be elaborated but the accepted hypothesis argues for the immune complex formation on the surface of joint cartilage that leads to a destructive inflammation together with the contribution of other innate receptors (TLR2) and signals (ROS)[101].

The spontaneous K/BxN transgenic mice did not only give rise to the GPI-induced arthritis but also to the serum transfer system that is robust, rapid, and highly reproducible[102]. Like CAIA, it enables the study of the effector mechanism as arthritis develops within days, even in the absence of lymphocytes in the recipient mice. Anti-GPI antibodies were recognized to be the causal agent as purified anti-GPI immunoglobulins from arthritic K/BxN mice was enough to induce joint-specific destruction. Evaluation of the role of Fc $\gamma$ receptors (Fc$\gamma$R) and components of the complement system in different KO animals identified a crucial contribution of Fc$\gamma$R III and the alternative pathway in K/BxN serum-transfer model[103].

### 1.5.3 Environmental/adjuvant triggers

The earliest benchmark model for the assessment of anti-inflammatory drugs (especially Cox inhibitors) was the Adjuvant Arthritis model (AA), which consists of injecting complete Freund’s adjuvant (CFA) to rats[104]. Typically, the rats develop joint specific inflammation, chronic at times depending on the strain; 10 to 12 days post injection. The fact that the disease onset is so rapid presented itself as a strong advantage to pharma industry, even though the model lacks scientific accuracy when compared to the human RA. To date the exact aetiology of AA is not very clear, yet the immunodominant antigen was found to be the heat shock protein 65K (HSP65) peptide derived from the Mtb that induces a joint specific T cell responses[105].

Most inbred rat strains develop RA-like disease after the injection of incomplete Freund’s adjuvant or Pristane (saturated alkane) which lacks the heat-killed Mtb[106]. These models are termed as Oil-induced arthritis (OIA) and Pristane-induced arthritis (PIA), respectively. Several disease modifying anti-rheumatic drugs (DMARDs) and corticosteroids were validated for efficacy in PIA rats before reaching the clinic as gold standard treatment for RA[107]. For instance, administration of Methotrexate, Cyclosporine A, Dexamethasone and Etanercept in rats pre and post onset of PIA significantly reduced arthritis severity[108]. This highlights the similarities in PIA pathology compared to RA however, no gender differences were detected which contradicts the discrepancy in gender prevalence found in human RA. Furthermore, post-onset administration of $\alpha$-IL17 antibody intravenously to rats significantly reduces arthritis severity and neutrophil infiltration in PIA, which hints for a role of IL17 producing T cells in maintaining inflammation and in recruiting neutrophils to the joints[108]. What really stands out in this model is the ability to adoptively transfer PIA using pristane primed CD4 T cells. In fact, it has been elegantly demonstrated that T cells ,derived from joint-draining lymph nodes of pristane immunized rats, induce a chronic relapsing arthritis that is highly associated with IL-17 production and neutralization of that cytokine dampened the severity of PIA[109].

# 2 RESEARCH AIMS

Applying forward genetics to map QTL has proven to be the optimal strategy to unravel arthritis-regulating genes in an unbiased way. After decades of breeding and tedious genetics work, several QTL were identified in rodents and positional cloning proceeded. The **overall aim** of this thesis is to pinpoint disease-regulating genes and explore potential mechanisms that define their function. This is true for the first 3 studies, while the last manuscript drifts away from the forward genetics approach and dives into the realm of regulatory B cells and their role as regulatory mediators of inflammation.

More tailored aims for each study are mentioned below:

- **Study I**, the focus was to positionally clone and functionally characterize the arthritis-regulating gene within a 200kb chromosomal interval defining the *Pia7/Cia13/Oia2* QTL using congenic rats.

- **Study II**, the goal was to identify the gene responsible for susceptibility to T cell-driven autoimmune models within a 500kb chromosomal fragment on the *Cia37* locus in mice.

- **Study III**, in this manuscript we set to investigate the interaction between the two most significant arthritis QTL in rats.

- **Study IV**, we questioned the curent dogma concerning the selection of antigen-specific B cells and their role in controlling antigen-specific autoimmunity.

# 3 MATERIALS AND METHODS

## 3.1 ARTHRITIS INDUCTION

For PIA induction, a single dose of 150µl pristane (2,6,10,14-tetramethylpentadecane) was injected intradermally. For OIA induction, a single dose of 150µl mineral oil (incomplete Freund's adjuvant /IFA/) was injected intradermally. PIA and OIA are rat restricted models used in **Study I** and **III**.

For CIA induction in rats of **Study I** and **III**, a single dose of 100µg of rat Col2 in 100µl of a 1:1 emulsion with IFA and PBS was injected intradermally. For CIA induction in **Study II** and **IV**, mice were immunized at day 0 with 100µg of rat Col2 in 100µl of a 1:1 emulsion in CFA and PBS intradermally. Mice were boosted at day 35 with 50µg of rat Col2 in 50µl of a 1:1 elusion in IFA and PBS.

All used rodents were older than 8 weeks old. All experiments were performed with littermate rats and mice that are age-matched, distributed within the cages and blindly evaluated by the investigator. The rats and mice were daily inspected which involves monitoring the limbs for arthritis development by a previously described macroscopic scoring system[108]. Briefly, 1 point was given for each individual swollen and erythematic toe and up to 5 points for an inflamed ankle (15 points in total per paw). The scoring was carried out every other day for 20-30 days after disease induction. No points were given to deformed paws that did not exhibit signs of edema.

## 3.2 CARRAGEENAN AIR-POUCH MODEL

Air pouches were formed in 10- to 12- weeks old rats according to the described protocol[110] and was only used in **Study III**. Briefly, 5 to 6 ml of sterile air was injected in the region between the scapulae of the animals. 3 days later, 2 to 3 ml of air was reinjected in the same spot to reconstitute the pouch structure. Induction of the inflammation was carried out by injecting at day 6, 1ml of 1% λ-Carrageenan solution and rats were sacrificed shortly after. Inflammatory exudate as well as air pouch cells were harvested using PBS supplemented with EDTA. Cells were processed by flow cytometry and cytokine concentrations were determined by sandwich ELISA.

## 3.3 FLOW CYTOMETRY

Single cells suspensions were carried out in ice-cold fluorescence-activated cell sorting buffer ($Ca2^+$-and $Mg2^+$-free Dulbecco's PBS supplemented with 1% fetal calf serum, 2 mmol/L EDTA) and blocking was presumed using 2.4G2/BD Fc Block (CD32) antibody. Subsequently, cells were stained with a saturating concentration of monoclonal antibodies on 96-well V-bottom polypropylene plates. After extracellular staining, cells were incubated in fixation/permeabilization buffer and washed with permeabilization buffer before intracellular staining. Both live/dead stain (Invitrogen, Carlsbad, CA) and the forward scatter *vs*. side scatter plot were used to include only live/singlets. A SORP BD LSR II analytic flow cytometer was used for acquisition, and the data were analyzed with FlowJo software version 8.8.6 (Tree Star, Ashland, OR). For intracellular cytokine staining, cells were stimulated in vitro with PMA 10 ng/mL, ionomycin 1 µg/mL, and brefeldin A 10

µg/mL for 4 to 6 h at 37 °C prior to fixation, permeabilization, and staining. This same staining procedure was carried out in **Study I**, **II**, **III** and **IV**.
The protocol I **Study III** for the ligand detection using bead-based assay was based on a previously described protocol with minor changes[111]. Briefly, compensation beads were coated with anti-histidine mAb. Subsequently, Dcar-EPN, Dcar-QPD and a decoy protein was used to couple it to the beads *via* the polyhistidine-tag and the whole complex was incubated with FITC-tagged Zymosan bioparticles at different concentration as illustrated in Fig. 6A. Beads were gated, and non-singlets were excluded using FSC-A *vs*. SSC-A followed by determination of geometric mean fluorescence intensity of FITC channel.

## 3.4 EXTRACELLULAR AND INTRACELLULAR ROS DETECTION

For flow cytometry-based detection of intracellular ROS in **Study III**, blood and BM cells were incubated with dihydrorhodamine (DHR) 123 and stimulated with PMA, Fmlp, Zymosan and absence of stimuli. Geometric mean fluorescence intensity of rhodamine 123 was corrected for background fluorescence (dimethyl sulfoxide) and fold increase was determined. For chemiluminescence based detection of extracellular ROS, BM cells were plated in duplicates and stimulated PMA, Fmlp, Zymosan and absence of stimulus. Next, extracellular ROS was measured chemiluminescence assay in HBSS$^{++}$ containing isoluminol/ HRPII. Data output was measured in relative light units (RLU), and total RLU integrated over time was plotted.

## 3.5 MAGNETIC BEAD-BASED ENRICHMENT

**Study IV** required sorting of antigen-specific B and T cells, we therefore resorted to magnetic enrichment for flow cytometric analysis and downstream assays.

### *Col2-specific T cells*

The PE-enrichment method was adopted from a previously published protocol[112]. Briefly, spleens and iLNs were mashed and pooled, and single-cell suspensions were prepared from immunized and naïve animals. Cells were stained with Aq-galCol2-PE tetramer. Next, cells were washed and incubated with MACS anti-R-Phycoerythrin. Cells were washed again, loaded onto LS columns, and washed with MACS buffer. Tetramer-specific T cells were eluted by flushing 5 mL of MACS buffer using the plunger. Collected cells were stained for surface markers and analyzed by Attune NxT (ThermoFisher) or LSR II (BD).

### *C1-specific B cells*

Typically, five spleens were mashed, and pooled, and single-cell suspensions were prepared from naïve animals. Cells were stained with C1-PE tetramer, washed, and incubated with MACS anti-R-phycoerythrin. Cells were washed again and loaded onto LS columns and washed MACS buffer. C1 B cells were eluted by flushing MACS buffer using the plunger. Collected cells were counted, stained for surface markers, and analyzed by Attune NxT (ThermoFisher) or LSR II (BD).

## 3.6 C1-B CELL FLUORESCENCE-ACTIVATED CELL SORTING

Epitope-specific B cells in **Study IV** were sorted from naïve and immunized mice. Briefly, single-cell suspensions were pelleted, stained with TCRβ and CD11b biotinylated antibodies and then washed extensively after incubation. Cells were suspended in MACS buffer and incubated with anti-biotin microbeads followed by a washing step. Next, cells were passed through LD columns and washed twice with MACS buffer. Collected unbound cells were resuspended in PBS and stained for CD19/B220, C1-tetramer, and live/dead dye to be further sorted *via* BD FACSAria III. Cells were gated for live, CD19$^+$B220$^+$ and sorted depending on C1-tetramer staining. 0.5 to 1 million pure (>90% purity) C1$^-$ B cells and C1$^+$ B cells were collected for downstream application.

For human single-cell sorting, whole blood (100-400 mL) from healthy blood donors was purchased from the transfusion unit at Karolinska Hospital (Solna, Stockholm) and peripheral blood mononuclear cells were isolated *via* density gradient centrifugation using Ficoll-Paque Plus in SepMate$^{TM}$-50 tubes. Cells were counted and stained with C1-PE tetramer. Next, cells were washed twice with MACS buffer and anti-PE beads were added followed by a washing step. Cells were passed through LS columns according to manufacturer’s protocol. Collected C1-B cells were stained with dead cells discriminator dye and anti-human antibodies, CD3, CD14, CD19, GPO-tetramer-APC. C1-tetramer$^+$ B cells (Live/CD3$^-$CD14$^-$/CD19$^+$/GPO-tetramer$^-$), GPO-tetramer$^+$ cells (Live/CD3$_-$CD14$_-$/CD19$^+$/C1-tetramer$^-$) and random B cells (Live/CD3$^-$CD14$^-$/CD19$^+$/GPO and C1-tetramer$^-$) and single-cell sorted into 96 or 384 well plates *via* BD FACSAria III.

## 3.7 *EX VIVO* EXPANSION OF HUMAN B CELLS

Complete B-cell medium was prepared using RPMI supplemented with ultra-low IgG fetal bovine serum, MEM non-essential amino acids, sodium pyruvate, β-mercaptoethanol, and penicillin/streptomycin. We supplemented the complete medium with recombinant human IL-21, recombinant human IL-2 , human holo transferrin, and ODN 2006 (5′-tcgtcgttttgtcgttttgtcgtt-3′ with a nuclease-resistant phosphorothioate backbone). C1-B cells in **Study IV** were enriched magnetically from HD PBMC and single-cell sorted into TC-treated 96 well U-bottom plates prefilled with 1 x 10$^4$ CD40L-expressing irradiated fibroblasts (50 Gy dose) in complete B-cell medium. Cells were cultured for a determined period. Culture supernatants were carefully aspirated and screened for anti-C1, anti-GPO, total IgG, and IgM antibodies.

## 3.8 ETHICAL CONSIDERATION

The guidelines supporting the humane use of animals in basic research are known as the 3Rs. Replace or avoid the use of animals when possible. Reduce the number of experimental animals to obtain comparable level of information. Refine the method to minimize pain, distress, suffering and improve animal welfare. Implementing these guiding principles ensures an optimal experimental design and statistical evaluation, transparency in conducting animal research, standardized husbandry at different facilities and promote reproducibility.

*That said, do we really need months of negotiation and political lobbying to accept or decline an ethical permit application?*

Ethical permit can be considered as the rulebook of animal handling and practice. They are specifically tailored for each field to serve scientific purposes. If you lack that holy document, working on animals is strictly prohibited. Renewing the ethical permit is a tedious process that requires endless forms and pretentious signatures. Problems will most likely arise along this procedure and solving them is even more time consuming. Infinite meetings, legislations, rebuttals, and agreements that ultimately lead to a butchered version of the starting ethical application with ridiculous changes that affect the overall experimental procedures.

We definitely need regulations, but they also need to be backed by a strong rational and not by the fallacies of collective belief. Most of ethical beliefs are fictional, a projection of our limbic system that do have advantages and disadvantages. We need to keep in mind that It was not the scientists that led the war, it was not researchers that killed people, instead it is a decision taken by the powerful minority that are often exempt from any sort of regulation (presidents and governmental officials).

Lastly, instead of wasting so much time and effort on tightening the belly chain, we should loosen the restrictions and revisit our ethical considerations. Ethics is an evolving practice rather than a static process.

# 4 RESULTS

## 4.1 STUDY I: IDENTIFICATION OF CLEC4B AS A NOVEL REGULATOR OF BYSTANDER ACTIVATION OF AUTO-REACTIVE T CELLS AND AUTOIMMUNE DISEASE

The dark agouti (DA) rat is an extraordinarily inbred strain due to their susceptibility to different inflammatory and autoimmune diseases[106]. The arthritis permeable phenotype can be mapped to loci on distinct chromosomes that control some aspects of the disease like incidence, severity, and antibody level. The mapped chromosomal fragments or quantitative trait loci (QTL) come in variable sizes (from thousand to million bases) and harbor polymorphic genes. Identification and positional cloning of disease-causing genes have been proven to be a difficult task but not impossible. *PIA7* is one of the most prominent rat QTL that was previously identified to regulate CIA, OIA and even EAE in a cross between DA and an arthritis resistant inbred strain (E3)[113]. Located on chromosome 4, *PIA7* harbored a cluster of genes coding for the group II of CLRs including *Clec4a* isoforms (*Clec4a-Clec4a3*), *Clec4b*, *Clec4d*, *Clec4e*, *Clec4n* (pseudogene). After sequencing of the complex, 2 main coding polymorphisms were identified in DA rats and were linked to the disease phenotype observed.

The first was a nonsense mutation in the second exon of *Clec4b* leading to a stop codon and the second was a non-synonymous SNP in the transmembrane domain of *Clec4e*. The latter was dismissed because of a cross between DA and another inbred strain (BN), that shares the same *Clec4e* SNP but differs in the *Clec4b* region, did reproduce the QTL. The *Clec4b*$^{\mathrm{DA}}$ variant became the most probable candidate for disease regulation as it is a natural knock-out, the 'cleanest' genetically unmodified knock-out. Indeed, we presumed to characterize the function of this receptor **Study I** to elucidate its role in autoimmunity.

Even with the scarce antibody panels designed for the rat, we could isolate CD4$^{+}$ antigen presenting cells (APC) and demonstrate that they express of Clec4b at a steady state level. Moreover, this population expanded in the spleens of DA rats with the highest expansion at day 3 and 4 when compared to the congenic rat bearing the protected *Clec4b*$^{E3}$ variant. *In vivo* recall assay of T cell adoptive transfer revealed the regulatory function of *Clec4b* over the proliferation, activation, and IL17 production in transferred CD4$^{+}$ T.

## 4.2 STUDY II: VITAMIN D3 RECEPTOR POLYMORPHISMS REGULATE T CELLS AND T CELL-DEPENDENT INFLAMMATORY DISEASES

CIA is an extensively used model that closely mimic RA. Like RA, CIA is a complex, polygenic disease that depends on T and B cell responses. Previous QTL were mapped following crosses between arthritis susceptible and resistant mice inbred strains and subsequent underlying polymorphisms were identified[114–116]. **Study II** focuses on positionally cloning the disease-regulating gene within *CIA37* mouse QTL previously mapped on chromosome 15. The sizable *CIA37* locus (54 centimorgan) from the DBA/1J was introgressed into H-2A$^q$ bearing C57BL/10 mice and through selective breeding the fragment was reduced to 500kb.

This subcongenic strain retained the disease phenotype observed previously in the parental strain but was demonstrated to solely regulate T cell mediated autoimmune models. Within the fragment 8 polymorphic genes were identified and sequencing revealed 5 coding non-synonymous SNPs. None of the coding SNPs had any functional or expressional implication and were therefore dismissed. Thereafter, the focus was then shifted towards noncoding SNPs. Expression analysis of the fragment-containing genes identified *Vdr* as a potential candidate as SNPs within its promoter region regulated its expression mainly in T cells. Remarkably, the overexpression of *Vdr* in the sub-congenic strain had no repercussion on bone density and calcium as was previously described in knock-out animals. This could be explained by the restricted overexpression of *Vdr* in T cells.

The increased susceptibility in the *Vdr* overexpressed sub-congenic strain was attributed to an enhanced T cell response which ultimately resulted in an increased inflammation. Proteomic analysis targeting VDR signaling uncovered a proinflammatory phenotype in activated T cells. Yet, non-physiological supplementation with vitamin D3 *via* osmotic pumps reversed the phenotype. Altogether, we show the anti-inflammatory properties of this hormone when used in excess, we demonstrate that physiological levels of VDR and vitamin D3 are necessary to mount an efficient immune response.

## 4.3 STUDY III: GLYCAN ACTIVATION OF CLEC4B INDUCES REACTIVE OXYGEN SPECIES PROTECTING AGAINST NEUTROPHILIA AND ARTHRITIS

Even when different QTL are elucidated and disease-regulated genes are positionally cloned, it remains crucial to investigate the potential interaction between them. In fact, the two most significant arthritis-regulating genes in DA rats have been found to be *Ncf1* and *Clec4b*[117]. The former is a vital member of the NADPH oxidase 2 and is imperative for a physiological phagocyte-associated oxidative burst[118], while the latter was the focus of **Study I**. In an attempt to reproduce the human scenario of polygenicity, we proceeded to examine the interaction of both genes (under a controlled environment) in models of autoimmunity and inflammation as depicted in **Study III**.

To qualitatively measure this interaction, we bred 4 groups of rats on the DA background that harbor the E3 variant of *Ncf1* or *Clec4b,* a double congenic rat containing both variants and a DA control. Next, we assessed PIA severity, incidence, and oxidative burst capabilities in all subcongenic groups and uncovered an epistatic relationship between *Clec4b* and *Ncf1*. Neutrophils express high levels of Dcar (protein encoded by *Clec4b*) at the steady state and their dramatic infiltration to lymphoid organs (spleen and draining lymph nodes) during the first days of PIA and OIA correlated with the loss of *Clec4b*. Because conventional arthritis models are governed by intricate interaction between both arms of the immune system, we resorted to an acute model of inflammation where only the innate immune cells are implicated. The neutrophil, cytokine, and chemokine phenotype replicated in the carrageenan air pouch was solely regulated by the expression of *Clec4b*.

Phagocytes undergo oxidative burst when subjected to innate stimuli such as components of the yeast cell wall (Zymosan). Interestingly, the neutrophils that contained E3 variants of *Clec4b* and *Ncf1* exhibited the highest ROS production when Zymosan was used to stimulated neutrophils. Therefore, we hypothesized that Zymosan-induced ROS production could be mediated by Dcar. To test this, we recombinantly expressed Dcar and assessed its binding affinity to Zymosan using a bead-based method. Indeed, Dcar binding to Zymosan was mostly dependent on a tripeptide motif within the carbohydrate recognition domain. Taken together, we conclude that *Clec4b* signaling leads to ROS production which inhibits neutrophilia and arthritis.

## 4.4 STUDY IV:BONE MARROW-SELECTED ANTIGEN SPECIFIC REGULATORY B CELLS

Unlike the constant evolution in the field of T cell selection, B cell selection has been dominated by the dogma established by seminal studies at the end of the 20th century[43]. The accepted theory is that autoreactive B cells are functionally impaired or deleted in the bone marrow. The models used to demonstrate these mechanisms of selection are elegant yet suffer from couple of shortcomings. One is the nature of the antigen used which is many cases is not truly ‘self-antigens’ but rather neo-antigens (hen egg lysozyme) in combination with a transgenic B cell cognate receptor[119]. The second is the systemic availability and the nature of the antigen utilized (soluble/cellular)[120].

Over decades our lab has extensively studied mouse models for rheumatoid arthritis, the most common autoimmune disease. This resulted in identifying highly conserved B cell-restricted epitopes (C1-epitope) on the injected antigen, collagen type-II (Col2). Thus, utilizing tetramers loaded with one of the Col2-specific peptides, we followed the fate of epitope-specific B cells, termed C1-specific B cells, in the presence of a physiologically expressed self-antigen (expressed in the bone marrow and articular joints). Surprisingly, C1-specific B cells were not deleted nor anergic as they were tracked throughout B cell ontogeny. Based on our observation that C1-specific B cells remained functionally active and essential for regulating arthritis, we dubbed them regulatory B cells. Due to the scarcity of C1-specific B cells, we employed VDJ knock-in mice to further analyze and characterize these cells by proteomics, metabolic assays, and flow cytometric analyses.

We were able to translate our findings to human by isolating the C1-specific B cells from healthy blood circulating cells using the same tetramer used in mice. Then, we expressed 3 monoclonal anti-Col2 antibodies from their recombined immunoglobulin VDJ sequences and validated the binding and specificity of these C1-specific antibodies. These data corroborate the notion that antigen-specific regulatory B cells populate the physiological repertoire in all species analyzed (mouse, rat and human).

# 5 CONCLUSIONS

A general conclusion can be drawn from **Study I, II** and **III**, collectively. Despite the focus on T cells in **Study I** and **II**, the modulation of the innate compartment described in **Study III** depicts a more comprehensive story regarding the implication of both arms of the immune system in mediating autoimmunity. The role of T cells in autoimmune arthritis models is evident in T cell deficient mice (RAG$^{-/-}$ mice) and autoreactive T cells (Col2-specific) can be readily isolated in mice as well as humans. **Study I** and **II** describe the influence of different genes in the priming of autoreactive T cells, in a bystander or intrinsic way. While **Study III** shed some light on the mechanisms that could explain the effector phase of arthritis.

B cells were the focal point of **Study IV** for their indispensable role in inducing arthritis. Autoreactive B cells in RA are poorly defined because of the lack of a relevant autoantigen on the one hand and the mechanisms of B cell selection on the other hand. We identified a subset of B cells, specific for a physiologically expressed protein (Col2), to be positively selected for regulatory reasons. The main role of these autoreactive B cells is to uptake, prime regulatory T cells and ultimately inhibit arthritis in an antigen-specific manner.

# 6 ACKNOWLEDGEMENTS

I would like to thank different people I got to rub elbows with over the years at **MIR**, starting with our king **Rikard**, thank you the financial, scientific, and motivational support you provided over the years. I value the freedom you bestowed upon me in different projects and for trusting my decisions (some of them were idiotic, to say the least).

**Liselotte**, I need to apologize for being a clown at the start and for driving you insane at times. However, our relationship evolved over the years, and I learned a great deal of genetics from you. I am also grateful for accepting me as a PhD student and for being on my side when it mattered. **Angel**, I can not think of a better nickname, you have been my angel during this whole period, you fulfilled the much-needed ‘mother’ duty at MIR and continue to do so. **Bingze**, the only true biochemist around, thank you for your help in expressing the proteins that were crucial for this thesis. Your role at MIR is vital, the pillar that holds this lab together.

I would like to extend my gratitude to the previous and current members of MIR that molded my scientific knowledge and altered my personal views (hopefully for the better).

**Clara**, you made the first years of my PhD bearable, we formed a nasty trio with Gonzo, I will never forget the hours we spent discussing controversial topics and I am still waiting for the podcast we three are destined to have. **Amit**, my favorite immunology course at KI was organized by you. I learnt a big deal from you both professionally and personally to the point that listening to your ranting became my preferred activity. **Danielle**, we shared a lot of things in common even our tiny office space. I enjoyed your company, positive mentality, and cheerful attitude very much. I will never forget the very first lunch I had at KI, it was you Amit and I sitting on the bench outside JJ under the soothing sun of august! **Ulrika**, ‘another pea in the pot’ sums it all, I guess. Same brain different sex. That China trip will go down in my history book for all the fun and adventures we had. **Johnny** the hardest working fellow in the lab, I am forever grateful for handing me your beautiful apartment in Sollentuna and for sharing the hardcore data you collected at every lab meeting. **Bruno**, in case you don’t remember you taught me how to perform splenectomy and the art of stitching. I thank you for the remarks and the help you provided during my early PhD phase. **Min**, the master of ROS detection, thank you for teaching me the tips and tricks of a successful burst experiment. **Anthony**, you were the last Rat fellow, and that fan noise still haunts me to this day.

**Gonzalo**, I still remember how well dressed you were on your first day at MIR, now you come in pink shorts and plain white t-shirts (what a transformation). I am very lucky that we spent 7 years together, we shared insane experiences and you will always have a special place in my heart. I love you man and I wish you all the best! One of the smartest, most eloquent, and funniest people I met by far. **Erik**, or as I always called you the Brad Pitt of MIR, I thank you dearly for all the books you shared with me, all the topics we covered in our post-lunch break and all the wine bottles I sneakily stole from your stash! **Ana**, without you the projects would have lasted me another 2 to 3 years. As I always said, you are a working horse, super organized and ultra-friendly. You will have a bright future I have no doubt, just keep this rhythm. We will continue to work closely on different projects, so bear with me.

**James** ‘Black’ for being a strong independent feminist and politically correct in every aspect, a true professional. **Yibo** ‘Mr. Doomer’ for music and gaming being our glue, I enjoyed every cigarette break with you. **Alex**, your German efficiency, deep knowledge in cloning and critical mind were crucial for the B cell project and I thank you repeatedly for that. **Michael**, for sharing mutual interest in politics. **Taotao**, for the histology section you provided, they were very valuable, I am still very impressed by the evolution of your English language! **Vilma**, for your fantastic photography that I always enjoy watching, I still have the printed photo you generously offered me hanging on my bedroom wall. **Laura** and **Rajan**, great addition to the lab as you can attract enough capital to maintain us afloat. **Outi**, for replying to all the questions about the Ncf1 mice even when you were not working in the lab. **Tamas**, for your enthusiasm regarding iNOS and ROS. **Weiwei** and **Lei**, for the Luminex run. **Huqiao**, for sharing the Ncf1 passion and projects. **Zongwei**, for being the ‘Future’ of this lab. **Changrong**, for your help in modeling and docking. The Palestro mafia, **Carlos** and **Kristina** and **Louisa** for keeping my animals happy and healthy, for being good friends underground, where things are cold and dark. Send my hugs to **Oliver** as well.

To my friends at Biomedicum including, **Frank**, for agreeing to help me out during the teaching expeditions, cover my ass when I was away. A genuine philanthropist. **Pierre**, your ‘JAQUES’ body has been the constant source for my gym motivation. I appreciate your kindness (at least towards me) and your exquisite culinary skills. **Christian** ‘KIKI’, the computer genius. A big thank you for the long and draining MS data curation. You have a bright future ahead of you dude. **Amir**, for discussing many topics during our 2 cigarettes break and this has culminated into a beautiful collaboration between our labs. Your readiness to contribute to projects is what I appreciate the most. **Jing**, for trying hard to help me with the metabolic assays. This is just the beginning buddy; we shall work on strengthening this collaboration.

To our collaborators in China **Jing**, **Liesu** and **Wenhua**, thank you for the input and for always being helpful.

To the old member of **TIM** at Huddinge that planted the PhD seed in me.

**Isabelle**, the Lab coat police! Thank you for grilling me during my halftime and for accepting this role on a very short notice and without any hesitation. **Thomas**, for discussing music, for playing basketball in Solna, for drinking relentlessly during Huddinge era and for your cool-headed attitude. You grew to become a solid family man and I wish you and Anneli all the best. Also, Goodluck with your new work and your international missions. **Adi**, I was not motivated much when I arrived at Huddinge but under your supervision, I grew to love science and appreciate clinical work. I miss you man and I hope you are still doing great science in Stanford.

All my students that helped me during the years, especially **Philip Welin-berger** and **Natalie Lager** for being my first students and good training material, **Thijmen** for your friendship, **Essam Barmada** for the genotyping (it finally worked), and lastly **Xiaojie** for working with me for almost 2 years, you have been a tremendous help and robust impetus for me personally and for Mincle specifically.

**Keka** for pushing me to wrap up my work and finish my PhD. This would have taken another year or two without you. Thank you for listening to my boring Lab related stories and proofreading my manuscripts.

Lastly, to my **Friends** and **Family**, for constant support and encouragement.

# 7 REFERENCES

1. Simmons, D. *The terror*. *Back Bay Books* (Little, Brown, 2007).
2. Travis, J. On the origin of the immune system. *Science (80-. ).* **324**, 580–582 (2009).
3. Sakula, A. Robert Koch: Centenary of the Discovery of the Tubercle Bacillus, 1882. *Can. Vet. J.* **24**, 127–131 (1983).
4. Valent, P. *et al.* Paul Ehrlich (1854-1915) and His Contributions to the Foundation and Birth of Translational Medicine. *J. Innate Immun.* **8**, 111–120 (2016).
5. Metchnikoff, E. Elie Metchnikoff (1845-1916), advocate of phagocytosis. *JAMA J. Am. Med. Assoc.* **203**, 139–141 (1968).
6. Underhill, D. M. & Goodridge, H. S. Information processing during phagocytosis. *Nat. Rev. Immunol.* **12**, 492–502 (2012).
7. Sender, R., Fuchs, S. & Milo, R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* **14**, 1–14 (2016).
8. Brubaker, S. W., Bonham, K. S., Zanoni, I. & Kagan, J. C. *Innate Immune Pattern Recognition: A Cell Biological Perspective*. *Annual Review of Immunology* vol. 33 (2015).
9. Takeuchi, O. & Akira, S. Pattern Recognition Receptors and Inflammation. *Cell* **140**, 805–820 (2010).
10. Jang, J. H. *et al.* An Overview of Pathogen Recognition Receptors for Innate Immunity in Dental Pulp. *Mediators Inflamm.* **2015**, (2015).
11. Brown, G. D., Willment, J. A. & Whitehead, L. C-type lectins in immunity and homeostasis. *Nature Reviews Immunology* vol. 1 1–16 (2018).
12. Vidya, M. K. *et al.* Toll-like receptors: Significance, ligands, signaling pathways, and functions in mammals. *Int. Rev. Immunol.* **37**, 20–36 (2018).
13. Drickamer, K. & Taylor, M. E. Recent insights into structures and functions of C-type lectins in the immune system. *Curr. Opin. Struct. Biol.* **34**, 26–34 (2015).
14. Miyake, Y. *et al.* C-type lectin MCL is an FcRγ-coupled receptor that mediates the adjuvanticity of mycobacterial cord factor. *Immunity* **38**, 1050–62 (2013).
15. Furukawa, A. *et al.* Structural analysis for glycolipid recognition by the C-type lectins Mincle and MCL. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 17438–17443 (2013).
16. Moriuchi, H. *et al.* Mannose-recognition mutant of the galactose/N-acetylgalactosamine-specific C-type lectin CEL-I engineered by site-directed mutagenesis. *Biochim. Biophys. Acta - Gen. Subj.* **1850**, 1457–1465 (2015).
17. Geijtenbeek, T. B. H. & Gringhuis, S. I. Signalling through C-type lectin receptors: Shaping immune responses. *Nat. Rev. Immunol.* **9**, 465–479 (2009).
18. Lee, R. T. *et al.* Survey of immune-related, mannose/fucose-binding C-type lectin receptors reveals widely divergent sugar-binding specificities. *Glycobiology* **21**, 512–520 (2011).
19. Ostrop, J. & Lang, R. Contact, Collaboration, and Conflict: Signal Integration of Syk-Coupled C-Type Lectin Receptors. *J. Immunol.* **198**, 1403–1414 (2017).
20. Martínez-López, M. *et al.* Microbiota Sensing by Mincle-Syk Axis in Dendritic Cells Regulates Interleukin-17 and -22 Production and Promotes Intestinal Barrier Integrity. *Immunity* **50**, 446–461.e9 (2019).
21. Deerhake, M. E. *et al.* Dectin-1 limits autoimmune neuroinflammation and promotes myeloid cell-astrocyte crosstalk via Card9-independent expression of Oncostatin M. *Immunity* **54**, 484–498.e8 (2021).
22. El-Benna, J., Dang, P. M. C., Gougerot-Pocidalo, M. A., Marie, J. C. & Braut-Boucher, F. p47phox, the phagocyte NADPH oxidase/NOX2 organizer: Structure, phosphorylation and implication in diseases. *Exp. Mol. Med.* **41**, 217–225 (2009).
23. Levine, A. P. & Segal, A. W. The NADPH Oxidase and Microbial Killing by Neutrophils, With a Particular Emphasis on the Proposed Antimicrobial Role of Myeloperoxidase within the Phagocytic Vacuole. *Microbiol. Spectr.* **4**, 1–14 (2016).
24. Zhong, J. *et al.* Association of NOX2 subunits genetic variants with autoimmune diseases. *Free Radic. Biol. Med.* 1–9 (2018) doi:10.1016/j.freeradbiomed.2018.03.005.
25. Sareila, O., Kelkka, T., Pizzolla, A., Hultqvist, M. & Holmdahl, R. NOX2 Complex–Derived ROS as Immune Regulators. *Antioxid. Redox Signal.* **15**, 2197–2208 (2011).
26. Arnold, D. E. & Heimall, J. R. A Review of Chronic Granulomatous Disease. *Adv. Ther.* **34**, 2543–2557 (2017).
27. Behring, J. B. *et al.* Spatial and temporal alterations in protein structure by EGF regulate cryptic cysteine oxidation. *Sci. Signal.* **13**, (2020).
28. Reynaert, N. L. *et al.* Dynamic redox control of NF-κB through glutaredoxin-regulated S-glutathionylation of inhibitory κB kinase β. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 13086–13091 (2006).
29. Jones, D. P. Radical-free biology of oxidative stress. *Am. J. Physiol. - Cell Physiol.* **295**, (2008).
30. Go, Y. M. & Jones, D. P. The redox proteome. *J. Biol. Chem.* **288**, 26512–26520 (2013).
31. Hidalgo, E. & Demple, B. An iron-sulfur center essential for transcriptional activation by the redox-sensing SoxR protein. *EMBO J.* **13**, 138–146 (1994).

32. Sies, H. & Jones, D. P. Reactive oxygen species (ROS) as pleiotropic physiological signalling agents. *Nat. Rev. Mol. Cell Biol.* **21**, 363–383 (2020).
33. Silva, M. de F. *et al.* The Keap1/Nrf2-ARE pathway as a pharmacological target for chalcones. *Molecules* **23**, 1–22 (2018).
34. Behring, E. Von & Kitasato, S. Uber das Zustandekommen Der Diphtherie- Immunitat Und der Tetanus-Immunitat Bei Thieren. *Dtsch Med Wochenschr* **49**, 1113–1121 (1980).
35. Davies, D. R. & Chacko, S. Antibody structure. *Acc. Chem. Res.* **26**, 421–427 (1993).
36. Flajnik, M. F. & Kasahara, M. Origin and evolution of the adaptive immune system: Genetic events and selective pressures. *Nat. Rev. Genet.* **11**, 47–59 (2010).
37. Fugmann, S. D., Messier, C., Novack, L. A., Andrew Cameron, R. & Rast, J. P. An ancient evolutionary origin of the Rag1/2 gene locus. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 3728–3733 (2006).
38. Zhang, Y. *et al.* Transposon molecular domestication and the evolution of the RAG recombinase. *Nature* **569**, 79–84 (2019).
39. Kaufman, J. Unfinished Business: Evolution of the MHC and the Adaptive Immune System of Jawed Vertebrates. *Annu. Rev. Immunol.* **36**, 383–409 (2018).
40. Ohno, S. Gene duplication and the uniqueness of vertebrate genomes circa 1970-1999. *Semin. Cell Dev. Biol.* **10**, 517–522 (1999).
41. Yatim, K. M. & Lakkis, F. G. A brief journey through the immune system. *Clin. J. Am. Soc. Nephrol.* **10**, 1274–1281 (2015).
42. Hogquist, K. A., Baldwin, T. A. & Jameson, S. C. Central tolerance: Learning self-control in the thymus. *Nat. Rev. Immunol.* **5**, 772–782 (2005).
43. Nemazee, D. Mechanisms of central tolerance for B cells. *Nat. Rev. Immunol.* **17**, 281–294 (2017).
44. ElTanbouly, M. A. & Noelle, R. J. Rethinking peripheral T cell tolerance: checkpoints across a T cell's journey. *Nat. Rev. Immunol.* **21**, 257–267 (2021).
45. Pape, K. A., Merica, R., Mondino, A., Khoruts, A. & Jenkins, M. K. Direct evidence that functionally impaired CD4+ T cells persist in vivo following induction of peripheral tolerance. *J. Immunol.* **160**, 4719–4729 (1998).
46. Van Deursen, J. M. The role of senescent cells in ageing. *Nature* **509**, 439–446 (2014).
47. Dhein, J., Walczak, H., Bäumlert, C., Debatint, K. M. & Krammer, P. H. Autocrine T-cell suicide mediated by APO-1/(Fas/CD95). *Nature* **373**, 438–441 (1995).
48. Krammer, P. H., Arnold, R. & Lavrik, I. N. Life and death in peripheral T cells. *Nat. Rev. Immunol.* **7**, 532–542 (2007).
49. Frappier, A. Clonal Selection Theory of Acquired Immunity. *American Journal of Public Health and the Nations Health* vol. 51 488–488 (1961).
50. Cho, J. H. & Feldman, M. Heterogeneity of autoimmune diseases: Pathophysiologic insights from genetics and implications for new therapies. *Nat. Med.* **21**, 730–738 (2015).
51. Alzabin, S. & Venables, P. J. Etiology of autoimmune disease: Past, present and future. *Expert Rev. Clin. Immunol.* **8**, 111–113 (2012).
52. Bogdanos, D. P. *et al.* Twin studies in autoimmune disease: Genetics, gender and environment. *J. Autoimmun.* **38**, J156–J169 (2012).
53. MacGregor, A. J. *et al.* Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum.* **43**, 30–37 (2000).
54. Hebbring, S. J. The challenges, advantages and future of phenome-wide association studies. *Immunology* **141**, 157–165 (2014).
55. Hu, X. & Daly, M. What have we learned from six years of GWAS in autoimmune diseases, and what is next? *Curr. Opin. Immunol.* **24**, 571–575 (2012).
56. Kochi, Y. Genetics of autoimmune diseases: Perspectives from genome-wide association studies. *Int. Immunol.* **28**, 155–161 (2016).
57. Loos, R. J. F. 15 Years of Genome-Wide Association Studies and No Signs of Slowing Down. *Nat. Commun.* **11**, 10–12 (2020).
58. Hampe, J. *et al.* A genome-wide association scan of nonsynonymous SNPs identifies a susceptibility variant for Crohn disease in ATG16L1. *Nat. Genet.* **39**, 207–211 (2007).
59. Burton, P. R. *et al.* Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
60. Murthy, A. *et al.* A Crohn's disease variant in Atg16l1 enhances its degradation by caspase 3. *Nature* **506**, 456–462 (2014).
61. Brest, P. *et al.* A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nat. Genet.* **43**, 242–245 (2011).
62. Thanabalasingham, G. *et al.* A large multi-centre European study validates high-sensitivity C-reactive protein (hsCRP) as a clinical biomarker for the diagnosis of diabetes subtypes. *Diabetologia* **54**, 2801–2810 (2011).
63. Ridker, P. M. *et al.* Loci Related to Metabolic-Syndrome Pathways Including LEPR,HNF1A, IL6R, and GCKR Associate with Plasma C-Reactive Protein: The Women's Genome Health Study. *Am. J. Hum.*

*Genet.* **82**, 1185–1192 (2008).
64. Drongelen, V. van & Holoshitz, J. HLA-Disease Associations in Rheumatoid Arthritis. *Rheum Dis Clin North Am* **43**, 363–376 (2017).
65. Lazaridis, I. *et al.* Ancient human genomes suggest three ancestral populations for present-day Europeans. *Nature* **513**, 409–413 (2014).
66. Price, A. L. *et al.* Discerning the ancestry of European Americans in genetic association studies. *PLoS Genet.* **4**, 0009-0017 (2008).
67. Kerr, S. M. *et al.* Pedigree and genotyping quality analyses of over 10,000 DNA samples from the Generation Scotland: Scottish Family Health Study. *BMC Med. Genet.* **14**, (2013).
68. Homer, N. *et al.* Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet.* **4**, (2008).
69. Monzon, F. A. *et al.* Whole genome SNP arrays as a potential diagnostic tool for the detection of characteristic chromosomal aberrations in renal epithelial tumors. *Mod. Pathol.* **21**, 599–608 (2008).
70. Korte, A. & Farlow, A. The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* **9**, 1 (2013).
71. Manolio, T. A. *et al.* Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
72. Bond, K. M., McCarthy, M. M., Rubin, J. B. & Swanson, K. R. Molecular omics resources should require sex annotation: a call for action. *Nat. Methods* **18**, 585–588 (2021).
73. Altshuler, D., Daly, M. J. & Lander, E. S. Genetic mapping in human disease. *Science (80-. ).* **322**, 881–888 (2008).
74. Pulst, S. M. Genetic linkage analysis. *Arch. Neurol.* **56**, 667–672 (1999).
75. Kerem, B. S. Identification of the cystic fibrosis gene: genetic analysis. *Trends Genet.* **5**, 363 (1989).
76. Ahlqvist, E., Hultqvist, M. & Holmdahl, R. The value of animal models in predicting genetic susceptibility to complex diseases such as rheumatoid arthritis. *Arthritis Res. Ther.* **11**, 1–10 (2009).
77. Ahlqvist, E., Hultqvist, M. & Holmdahl, R. The value of animal models in predicting genetic susceptibility to complex diseases such as rheumatoid arthritis. *Arthritis Res. Ther.* **11**, 1–10 (2009).
78. Ott, J., Wang, J. & Leal, S. M. Genetic linkage analysis in the age of whole-genome sequencing. *Nat. Rev. Genet.* **16**, 275–284 (2015).
79. Firestein, G. S. & McInnes, I. B. Immunopathogenesis of Rheumatoid Arthritis. *Immunity* **46**, 183–196 (2017).
80. Angelotti, F. *et al.* One year in review 2017: the pathogenesis of rheumatoid arthritis. *Clin. Exp. Rheumatol.* **35**, 368–378 (2017).
81. Holmdahl, R., Malmström, V. & Burkhardt, H. Autoimmune priming, tissue attack and chronic inflammation - The three stages of rheumatoid arthritis. *Eur. J. Immunol.* **44**, 1593–1599 (2014).
82. Vladutiu, A. O. & Rose, N. R. Autoimmune murine thyroiditis relation to histocompatibility (H-2) type. *Science (80-. ).* **174**, 1137–1139 (1971).
83. Howell, W. M., Carter, V. & Clark, B. The HLA system: Immunobiology, HLA typing, antibody screening and crossmatching techniques. *J. Clin. Pathol.* **63**, 387–390 (2010).
84. Anderson, K. M. *et al.* A Molecular Analysis of the Shared Epitope Hypothesis: Binding of Arthritogenic Peptides to DRB1*04 Alleles. *Arthritis Rheumatol.* **68**, 1627–1636 (2016).
85. Gregersen, P. K., Silver, J. & Winchester, R. J. The Shared Epitope Hypothesis. *Arthritis Rheum.* **30**, 1205–1212 (1987).
86. Cheng, P. *et al.* Association between CCR6 and rheumatoid arthritis: A meta-analysis. *Int. J. Clin. Exp. Med.* **8**, 5388–5396 (2015).
87. Van Der Linden, M. P. M. *et al.* Association of a single-nucleotide polymorphism in CD40 with the rate of joint destruction in rheumatoid arthritis. *Arthritis Rheum.* **60**, 2242–2247 (2009).
88. El-Gabalawy, H. S. *et al.* Non-HLA genes modulate the risk of rheumatoid arthritis associated with HLA-DRB1 in a susceptible North American native population. *Genes Immun.* **12**, 568–574 (2011).
89. Kollias, G. *et al.* Animal models for arthritis: Innovative tools for prevention and treatment. *Ann. Rheum. Dis.* **70**, 1357–1362 (2011).
90. Vincent, T. L. *et al.* Mapping pathogenesis of arthritis through small animal models. *Rheumatol. (United Kingdom)* **51**, 1931–1941 (2012).
91. Bevaart, L., Vervoordeldonk, M. J. & Tak, P. P. Evaluation of therapeutic targets in animal models of arthritis: How does it relate to rheumatoid arthritis? *Arthritis Rheum.* **62**, 2192–2205 (2010).
92. Tao, L. & Reese, T. A. Making Mouse Models That Reflect Human Immune Responses. *Trends Immunol.* **38**, 181–193 (2017).
93. Nandakumar, K. S. & Holmdahl, R. Efficient promotion of collagen antibody induced arthritis (CAIA) using four monoclonal antibodies specific for the major epitopes recognized in both collagen induced arthritis and rheumatoid arthritis. *J. Immunol. Methods* **304**, 126–136 (2005).
94. Svensson, L., Jirholt, J., Holmdahl, R. & Jansson, L. B cell-deficient mice do not develop type II collagen-induced arthritis (CIA). *Clin. Exp. Immunol.* **111**, 521–526 (1998).
95. Corthay, A., Johansson, Å., Vestberg, M. & Holmdahl, R. Collagen-induced arthritis development requires αβ T cells but not γδ T cells: Studies with T cell-deficient (TCR mutant) mice. *Int. Immunol.* **11**,

1065–1073 (1999).

96. Goldschmidt, T. J. & Holmdahl, R. Anti-T cell receptor antibody treatment of rats with established autologous collagen-induced arthritis: suppression of arthritis without reduction of anti-type II collagen autoantibody levels. *Eur. J. Immunol.* **21**, 1327–1330 (1991).
97. Matsumoto, I. & Staub, A. Arthritis provoked by linked T and B cell recognition of a glycolytic enzyme. *Science (80-. ).* **286**, 1732–1735 (1999).
98. Pizzolla, A., Wing, K. & Holmdahl, R. A glucose-6-phosphate isomerase peptide induces T and B cell-dependent chronic arthritis in C57bl/10 mice: Arthritis without reactive oxygen species and complement. *Am. J. Pathol.* **183**, 1144–1155 (2013).
99. Nandakumar, K. S. & Holmdahl, R. Antibody-induced arthritis: disease mechanisms and genes involved at the effector phase of arthritis. *Arthritis Res. Ther.* **8**, 223 (2006).
100. Nandakumar, K. S., Svensson, L. & Holmdahl, R. Collagen Type II-Specific Monoclonal Antibody-Induced Arthritis in Mice: Description of the Disease and the Influence of Age, Sex, and Genes. *Am. J. Pathol.* **163**, 1827–1837 (2003).
101. Kelkka, T., Hultqvist, M., Nandakumar, K. S. & Holmdahl, R. Enhancement of antibody-induced arthritis via toll-like receptor 2 stimulation is regulated by granulocyte reactive oxygen species. *Am. J. Pathol.* **181**, 141–150 (2012).
102. Maccioni, M. *et al.* Arthritogenic monoclonal antibodies from K/BxN mice. *J. Exp. Med.* **195**, 1071–1077 (2002).
103. Ji, H. *et al.* Arthritis critically dependent on innate immune system players. *Immunity* **16**, 157–168 (2002).
104. Anderson, G. D. *et al.* Selective inhibition of cyclooxygenase (COX)-2 reverses inflammation and expression of COX-2 and interleukin 6 in rat adjuvant arthritis. *J. Clin. Invest.* **97**, 2672–2679 (1996).
105. Eden, W. Van *et al.* Cloning of the mycobacterial epitope recognized by T lymphocytes in adjuvant arthritis. *Nature* **331**, 171–173 (1988).
106. Rintisch, C. & Holmdahl, R. DA rats from two colonies differ genetically and in their arthritis susceptibility. *Mamm. Genome* **19**, 420–428 (2008).
107. Faisal, R., Ahmad, N., Fahad, Y. S. & Chiragh, S. Anti-Arthritic Effect Of Thymoquinone In Comparison With Methotrexate On Pristane Induced Arthritis In Female Sprague Dawley Rats. *J. Ayub Med. Coll. Abbottabad* **30**, 3–7 (2018).
108. Tuncel, J. *et al.* Animal models of rheumatoid arthritis (I): Pristane-induced arthritis in the rat. *PLoS One* **11**, 1–17 (2016).
109. Tuncel, J. *et al.* Self-reactive T cells induce and perpetuate chronic relapsing arthritis. *Arthritis Res. Ther.* **22**, 1–12 (2020).
110. Duarte, D. B., Vasko, M. R. & Fehrenbacher, J. C. Models of inflammation: Carrageenan air pouch. *Curr. Protoc. Pharmacol.* **2016**, 5.6.1-5.6.9 (2016).
111. Mccoy, C. E. *Toll-Like Receptors Practice and Methods Second Edition Methods in Molecular Biology 1390*.
112. Kotov, D. I. & Jenkins, M. K. Peptide:MHCII Tetramer-Based Cell Enrichment for the Study of Epitope-Specific CD4+ T Cells. *Curr. Protoc. Immunol.* **125**, (2019).
113. Rintisch, C. *et al.* Finemapping of the arthritis QTL Pia7 reveals co-localization with Oia2 and the APLEC locus. *Genes Immun.* **11**, 239–45 (2010).
114. Vingsbo-Lundberg, C. *et al.* Genetic control of arthritis onset, severity and chronicity in a model for rheumatoid arthritis in rats. *Nat. Genet.* **20**, 401–404 (1998).
115. Ahlqvist, E., Bockermann, R. & Holmdahl, R. Fragmentation of Two Quantitative Trait Loci Controlling Collagen-Induced Arthritis Reveals a New Set of Interacting Subloci. *J. Immunol.* **178**, 3084–3090 (2007).
116. Johannesson, M. *et al.* Identification of epistasis through a partial advanced intercross reveals three arthritis loci within the Cia5 QTL in mice. *Genes Immun.* **6**, 175–185 (2005).
117. Olofsson, P. *et al.* A comparative genetic analysis between collagen-induced arthritis and pristane-induced arthritis. *Arthritis Rheum.* **48**, 2332–2342 (2003).
118. Holmdahl, R., Sareila, O., Olsson, L. M., Bäckdahl, L. & Wing, K. Ncf1 polymorphism reveals oxidative regulation of autoimmune chronic inflammation. *Immunol. Rev.* **269**, 228–247 (2016).
119. Goodnow, C. C. *et al.* Altered immunoglobulin expression and functional silencing of self-reactive B lymphocytes in transgenic mice. *Nature* **334**, 676–682 (1988).
120. Nemazee, D. A. & Bürki, K. Clonal deletion of B lymphocytes in a transgenic mouse bearing anti-MHC class I antibody genes. *Nature* **337**, 562–566 (1989).