From the Department of Microbiology, Tumor and Cell Biology
Karolinska Institutet, Stockholm, Sweden

GENETIC AND FUNCTIONAL ANALYSIS OF THE ADAPTIVE IMMUNE RESPONSE

Marco Mandolesi

Karolinska Institutet

Stockholm 2023

All previously published papers were reproduced with permission from the publisher.
Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2023

ISBN 978-91-8016-923-3

Genetic and functional analysis of the adaptive immune response

Thesis for Doctoral Degree (Ph.D.)

By

Marco Mandolesi

The thesis will be defended in public at Eva & Georg Klein Hall, Biomedicum, Solnavägen 9, Solna, Stockholm, Friday, February 24th, 2022 at 9:00 a.m.

Principal Supervisor:
Professor Gunilla B. Karlsson Hedestam
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Co-supervisor(s):
Assistant professor Benjamin Murrell, PhD
Karolinska Institutet
Department of Microbiology, Tumor and Cell Biology

Opponent:
Chief Robert A. Seder, MD, PhD
National Institutes of Health
National Institute of Allergy and Infectious Diseases
Vaccine Research Center

Examination Board:
Professor Anna-Lena Spetz
Stockholm Universitet
Department of Molecular Biosciences
The Wenner-Gren Institute

Associate professor Fredrik Wermeling
Karolinska Institutet
Department of Medicine Solna
Center for Molecular Medicine

Professor Johan K. Sandberg
Karolinska Institutet
Department of Medicine Huddinge
Center for Infectious Medicine

To all my loved ones. Especially to my dear grandmother "*Nonna Settimia*".

"L'arte non può mai uguagliare la ricchezza della natura."
Giacomo Leopardi

Popular science summary of the thesis

The immune system can be divided in two arms: the innate immune system and the adaptive immune system. The innate immune system is commonly known as the first line of defense against pathogens or other foreign elements, acting in a generalized manner to eliminate the threat as quickly as possible. The adaptive immune system is responsible for the production of a specialized reaction that targets the threat and generates immunological memory.

The adaptive immune response employs two main types of white blood cells: B cells and T cells. Each of these populations performs distinct tasks in the generation of pathogen-specific responses. All B and T cells produce a unique surface molecule called B cell receptor (for B cell) or T cell receptor (for T cells). When B cell receptors are secreted, they are called antibodies. In both cases, these receptors are produced by recombining genes called V, D and J genes. B cells and T cells have different sets of V, D and J genes that are encoded from different genomic locations.

The overall sets of V, D and J gene variants are different for each individual and can be considered akin to an immunological fingerprint. We call this the “germline repertoire”. As of today, we still do not the full extent of V, D and J gene variations among humans and other species. To address this question, our group developed a software called IgDiscover, to identify individual germline repertoires using modern sequencing technologies.

B cells and T cells can undergo a series of changes depending on the imminent threat. Generally, only those B cells and T cells with complementary receptors against the threat get activated. This results in the multiplication of these cells in a process called clonal expansion. Furthermore, B cells undergo additional B cell receptor modifications to improve their recognition of the threat. Furthermore, some B cells and T cells will re-program themselves to persist over long period of time. These cells make up our immunological memory, which is responsible for rapid recall responses upon subsequent encounters of the same threat.

The study of these receptors is of paramount importance to understand our adaptive immune response to infections and vaccination. In this thesis I aimed to contribute to what in the future will be a comprehensive understanding of the adaptive immune system.

In **paper I** of this thesis, I explored the T cell receptor germline gene repertoire in humans from different parts of the world. My colleagues and I described the diversity between individuals, and we discovered that some of the gene variants are population specific. Furthermore, we discovered that some currently living humans inherited some small parts of their T cell receptor germline repertoire from ancient extinct hominids.

In **paper II and III**, I used a macaque model to analyse the B cell response after vaccination with the Spike protein of SARS-CoV-2, the causative pathogen of COVID-19. In **paper II**, we found that after using an engineered Spike protein as a vaccine we can elicit potent antibody responses that can neutralize the virus. This response is specific and lasts for a long time after vaccination. In **paper III**, we were concerned about the emergence of different SARS-CoV-2 variants and boosted macaques with proteins engineered from these variants. We described how this produced potent neutralizing antibodies against different variants, resulting in broader protection.

In **paper IV**, I analyzed the B cell response of two of the macaques from Paper II in greater detail. To do so, I determined their B cell receptor V, D and J germline repertoires. I then isolated B cells that reacted against the SARS-CoV-2 Spike and identified Spike-specific antibody families characterized by a specific V, D and J gene composition. Finally, I traced the presence of each family in different lymphoid tissues and at different time points following the immunizations., and I identified a set of potent and broad antibodies with especially interesting features.

Riassunto generale della tesi

Il sistema immunitario è generalmente diviso in due componenti principali: il sistema immunitario innato e il sistema immunitario adattivo. Il sistema immunitario innato, agendo in modo generalizzato per eliminare il pericolo il più rapidamente possibile, rappresentala la prima linea di difesa contro patogeni e altri elementi estranei. Il sistema immunitario adattivo genera invece una risposta immunitaria specializzata contro ciò che minaccia la nostra salute.

La risposta immunitaria adattativa è principalmente composta da due popolazioni cellulari: i linfociti B e i linfociti T. Nonostante queste popolazioni svolgano compiti distinti, hanno in comune il fatto di generare cellule estremamente specializzate nel riconoscere le caratteristiche strutturali specifiche di ogni patogeno. Ogni volta che un linfocita B o un linfocita T viene stimolato dall' incontro con un patogeno, produce sulla propria superficie un recettore. I recettori dei linfociti B, anche noti come B cell receptors (BCRs), sono proteine transmembrana ma con il potenziale di essere secreti, prendendo il nome di anticorpi. I recettori dei linfociti T o T cell receptors (TCRs) svolgono invece la loro funzione sempre come proteina transmembrana. In entrambi i casi, questi recettori vengono prodotti ricombinando un gene da ciascuno di tre set di geni disponibili: i geni V, i geni D e i geni J. È importante evidenziare che i geni V, D e J dei BCRs e quelli dei TCRs sono posizionati in punti diversi del genoma.

L' insieme dei geni V, D e J è diverso per ogni individuo e può essere in qualche modo considerato un' "impronta digitale" immunologica. Chiamiamo questa impronta "repertorio germinale". Al momento, non siamo ancora a conoscenza di quanto variabili siano i geni V, D e J tra gli esseri umani e altre specie. Il nostro laboratorio ha sviluppato un software, chiamato IgDiscover, per identificare il repertorio germinale di ogni individuo utilizzando moderne tecnologie di sequenziamento.

I linfociti che incontrato un patogeno (o parte di esso) complementare al proprio recettore si attivano. L' attivazione innesca la moltiplicazione di queste cellule in un processo chiamato "espansione clonale". I linfociti attivati in organi linfoidi periferici (come linfonodi, tonsille, etc.) prendono parte alla formazione di microstrutture temporanee chiamate "centri germinativi". Qui, il recettore dei linfociti B si modifica e matura per acquisire più specificità e potenza. Ciò si traduce in anticorpi più efficenti.

Al termine di questi processi, alcuni linfociti B e T vengono riprogrammati per sopravvivere a lungo. Queste cellule formano la nostra memoria immunologica, che è ciò che ci permette di prevenire efficacemente la comparsa di sintomi gravi in caso di una seconda infezione con lo stesso patogeno.

Lo studio di questi recettori è di fondamentale importanza per comprendere i meccanismi immunologici e per studiare l'effetto delle infezioni da diversi patogeni o relativivaccini. Il lavoro svolto per questa tesi è finalizzato a contribuire verso ciò che in futuro sarà una comprensione globale della risposta immunitaria adattiva.

Nell' **articolo I** di questa tesi, ho esplorato il repertorio germinale del recettore dei linfociti T in individui provenienti da diverse parti del mondo. Io e i miei colleghi abbiamo descritto la variabilità interindividuale del repertorio germinale e abbiamo scoperto che alcune delle varianti sono specifiche di certe popolazioni. Inoltre, abbiamo scoperto che alcuni individui attualmente viventi hanno ereditato alcune piccole parti del loro repertorio germinale da antichi ominidi estinti, come i Neanderthal.

Negli **articoli II e III**, ho utilizzato il modello animale del macaco per esaminare la risposta dei linfociti B dopo aver amministrato un vaccino basato sulla proteina Spike del SARS-CoV-2, il virus che causa il COVID-19. Nell' **articolo II**, abbiamo scoperto che utilizzando una spike protein ingegnerizzata, è possibile indurre una potente risposta anticorpale che può neutralizzare il virus. Questa risposta è specifica e può neutralizzare il virus per un lungo periodo di tempo dopo la vaccinazione. Nell' **articolo III**, ci siam occupati di diverse varianti del SARS-CoV-2 e abbiamo deciso di osservare come varia la risposta anticorpale se i macachi vengono immunizzati con proteine provenienti da queste varianti invece che dal virus ancestrale. Abbiamo descritto come il booster produce anticorpi neutralizzanti potenti contro le diverse varianti, generando un effetto protettivo molto ampio.

Nell' **articolo IV**, ho deciso di analizzare in maggiore dettaglio la risposta dei linfociti B di due macachi dell' articolo II. Per fare ciò, ho identificato il repertorio germinale V, D e J di ciascuna scimmia. Ho in seguito estratto e selezionato i linfociti B che presentavano una reazione specifica contro la proteina Spike di SARS-CoV-2. Il sequenziamento di queste cellule ha fornito informazioni sul loro recettore e ha permesso l' identificazione di cosiddette famiglie clonali, ovvero tutti quei linfociti B provenienti da un unico progenitore che è andato in contro al processo di espansione clonale precedentemente descritto. Ogni famiglia di linfociti è caratterizzata da una composizione specifica di geni V, D e J. Infine, ho tracciato i membri di ogni famiglia in diverse parti del corpo, come il midollo, la milza e vari linfonodi. Con i protocolli e algoritmi disponibili nel nostro laboratio, abbiamo potuto osservare come le diverse famiglie di linfociti B si diffondessero e si evolvessero dando luogo ad anticorpi con caratteristiche diverse.

Abstract

The cells of the adaptive immune system rely on somatic recombination of V, D and J genes to obtain a vast range of specificities. T cells use four chains encoded by three genomic loci, to produce alpha/beta or gamma/delta T cell receptors (TCRs). B cell receptors (BCRs) are encoded by three loci, a single heavy chain and two light chain loci, kappa and lambda. Because these genomic regions are highly polymorphic, the germline TCR and BCR repertoires are individual, shaping the elicited response against infections and vaccines in each person. In this thesis, I used a set of specialized tools and analysis pipelines to explore the adaptive immune receptors at the genomic and functional levels in both humans and non-human primates (NHP), highlighting the benefits of integrating individualized TCR/BCR repertoire analysis with functional studies to understand adaptive immune responses.

In **paper I**, we sequence expressed TCR repertoires of 45 individuals from four human populations: African, East Asian, South Asian, and European. Analysis of these repertoires with the germline gene inference tool, IgDiscover, identified 175 novel V and J alleles, most of which were characterized by codon changes or non-functional variants. The germline TCR repertoires were highly diverse between individuals, with some of the novel alleles identified only in specific populations. Furthermore, we report three introgressed regions inherited from *Homo neanderthalensis*. One of these regions includes a novel variant of TRGV4, frequent in Eurasians populations, which display altered reactivity to the ligand butyrophilin-like molecule 3 (BTNL3).

In **papers II and III**, we analyzed the humoral immune response in NHPs elicited by a series of immunizations with SARS-CoV-2 Spike-derived subunit proteins. In **paper II**, we observed detectable neutralization titers after priming with ancestral spike (S) protein with very high antibody titers obtained after boosting. The immunization regimen resulted in durable neutralization titers as well as S-specific memory B cells. In **paper III**, we used a heterotypic boosting strategy with beta-derived receptor binding domain (RBD) to broaden the response to circulating SARS-CoV-2 variants. The boost elicits potent and protective cross-neutralizing humoral immune responses.

In **paper IV**, we analyzed multi-compartmental longitudinal samples from two macaques used in **paper II**. We combined single cell and next generation sequencing (NGS) of BCR repertoires to characterize S-specific antibodies and S-specific B cell lineages elicited by immunizations with ancestral SARS-CoV-2 S proteins. Lineage tracing analysis identified persistent antibody lineages that were present after priming and were widely disseminated in blood, bone marrow (BM), spleen and different lymph nodes (LN), including a broadly neutralizing RBD-binding lineage. Through structural cryo-EM studies, we showed that this antibody achieved cross-neutralization by targeting conserved RBD residues with crucial interactions through its heavy chain CDR3 (HCDR3).

List of scientific papers

I. Corcoran M., Chernyshev M.*, **Mandolesi M.***, Narang S., Kaduk M., Ye K., Sundling C., Färnert A., Kreslavsky T., Bernhardsson C., Larena M., Jakobsson M. and Karlsson Hedestam G.B.

"Archaic humans have contributed to large-scale variation in modern human T cell receptor genes"

Immunity, in press. *Equal Contribution.

II. **Mandolesi M.***, Sheward D.J.*, Hanke L., Ma J., Pushparaj P., Perez Vidakovics L., Kim C., Àdori M., Lenart K., Loré K., Castro Dopico X., Coquet J.M., McInerney G.M., Karlsson Hedestam G.B.*, Murrell B.*

"SARS-CoV-2 protein subunit vaccination of mice and rhesus macaques elicits potent and durable neutralizing antibody responses"

Cell Reports Medicine. 2021 Apr 20;2(4):100252. *Equal Contribution.

III. Sheward D.J.*, **Mandolesi M.***, Urgard E.*, Kim C., Hanke L., Perez Vidakovics L., Pankow A., Smith N.L., Castro Dopico X, McInerney G.M., Coquet J.M., Karlsson Hedestam G.B.*, Murrell B.*

"Beta RBD boost broadens antibody-mediated protection against SARS-CoV-2 variants in animal models"

Cell Reports Medicine. 2021 Nov 16;2(11) *Equal Contribution.

IV. **Mandolesi M.**, Das H., Kim C., de Vries L., Dopico X.C., Chernyshev M., Hanke L., Fischbach J., Sungyong K., Àdori M., Stålmarck A., Coquet J.M., McInerney G.M., Sheward D.J., Corcoran M., Hällberg M.B., Karlsson Hedestam G.B.*, Murrell B.*.

"Broadly neutralizing antibodies in the rhesus macaque B cell repertoire after immunization with ancestral SARS-CoV-2 spike protein"

Manuscript. *Equal Contribution.

CONTENTS

List of abbreviations

1KGP	1000 Genome Project
ACE2	Angiotensin-converting enzyme 2
AIRR	Adaptive immune receptor repertoire
AIRR-seq	Adaptive immune receptor repertoire sequencing
APC	Antigen presenting cell
axLN	Axillary lymph node
BCR	B cell receptor
BM	Bone marrow
bnAb	Broadly neutralizing antibody
BTNL	Butyrophilin like
CDR	Complementarity determining region
CLP	Common lymphoid progenitor
COVID-19	Coronavirus diseases 2019
DNA	Deoxyribonucleic acid
EM	Electron microscopy
FDC	Follicular dendritic cell
FR	Framework region
GC	Germinal center
HV4	Hypervariable region 4
Ig	Immunoglobulin
IGH	Immunoglobulin heavy chain
IGK	Immunoglobulin kappa chain
IGL	Immunoglobulin lambda
IMGT	ImMunoGeneTics information system
LN	Lymph node
mAb	Monoclonal antibody
MAIT	Mucosal associated invariant T cell
MALT	Mucosa associated lymphoid tissue
mesLN	Mesenteric lymph node
nAb	Neutralizing antibody
NGS	Next generation sequencing
NHP	Non-human primate
NKT	Natural killer T cell
ORF	Open reading frame
PBMCs	Peripheral blood mononuclear cells

PCR	Polymerase chain reaction
PDB	Protein data bank
perLN	Periaortic lymph node
RBD	Receptor binding domain
RBM	Receptor binding motif
RNA	Ribonucleic acid
S	Spike
SARS-CoV-2	Severe acute respiratory syndrome coronavirus 2
SHM	Somatic hypermutation
TCR	T cell receptor
Tfh	Follicular T helper cell
Th	T helper cell
TRA	T cell receptor alpha
TRB	T cell receptor beta
TRD	T cell receptor delta
Treg	Regulatory T cell
TRG	T cell receptor gamma
WHO	World health organization

Introduction

The adaptive immune system is a crucial component of vertebrates' defenses against infections. It is a highly specialized system that is capable of learning and remembering past exposures to pathogens, a feature known as immunological memory. This allows the body to mount more effective responses to future encounters with the same pathogen. Vaccines aim to stimulate the formation of immunological memory without the adverse effect of a natural infection.

The adaptive immune system is made up of a variety of cells, tissues, and organs that work together to identify and eliminate foreign invaders. The protagonists of the adaptive immune system are T lymphocytes and B lymphocytes, also known as T cells and B cells. T cells and B cells undergo different developmental pathways to make up subsets of cells with specialized functions. As a population of cells, they have the potential to recognize specific pathogens thanks to the expression of their antigen receptors. T cells express T cell receptors (TCRs) while B cells express B cell receptors (BCRs). Once a B cell has been activated, it can differentiate into a plasma cell that produces BCR in a secretory format, known as antibody. The set of BCRs and TCRs expressed in an individual at given time is called the "immune repertoire".

In this thesis, I combined classic immunoassays and bioinformatic tools to study the adaptive immune receptor repertoires (AIRR) in humans and macaques. The approach I used relied primarily in the identification of individuals germline repertoires which were used as a foundation for downstream analysis. Throughout this thesis, I will introduce basic concepts behind the study of TCR and BCR repertoires and elaborate on the technologies employed in the field. I will also place these concepts and information in the context of their relevance for the examination of human populations diversity (for TCRs) and the examination of vaccine elicited B cell response (for BCRs).

I hope that the content of this thesis demonstrates the importance of generating and incorporating individualized dataset for the study of adaptive immunity.

1 Literature review

1.1 Immunology

1.1.1 Adaptive immunity

To maintain proper functionality, organisms have developed different defense mechanisms against exogenous threats. These defense mechanisms evolved into a complex network of molecular and cellular interactions that we call the immune system[1–3]. During an immune response, different functions work in synch to eliminate and/or expel the pathogen to avoid extensive damage to the host and re-establish homeostasis.

Overall, the immune system can be divided in two main categories: the innate immune system and the adaptive immune system. The components of the innate immune system are activated first, providing general protection against pathogens. The innate immune system is also tasked to activate the cells of the adaptive immune system, initiating a cascade of events[4–6]. As the name suggests, the goal of the adaptive immune system is to provide pathogen-specific customized responses. In general, cells of the adaptive immune system undergo different rounds of proliferation and selection to improve their reactivity. Therefore, the adaptive immune response requires some time before being effective.

The exclusive characteristic of adaptive immunity is the generation of immunological memory[7,8]. T cells and B cells, which make up two arms of the adaptive immune system, express antigen-specific receptors that are critical for their survival[9,10]. The loci encoding the genes for TCRs and BCRs share similar genomic structures and undergo similar events of gene rearrangement during early lymphocyte development to produce one unique functional receptor for each T cell or B cell. These mechanisms are in place to ensure that each cell expresses a unique receptor[11–13]. The collection of TCRs and BCRs make up the AIRR, which is unique to a given time point and individual[14,15]. It is important to note that, despite the fact that the process of gene rearrangement is similar between TCRs and BCRs, they recognize cognate antigen in remarkably different manners.

While both T cells and B cells originate from common lymphoid progenitors (CLP) located in the BM, they undergo early developmental phases in different compartments. T cell formation takes place in the thymus while B cells are produced in the BM. Cells with autoreactive or non-functional receptors are eliminated or rendered non-responsive. This process is called “central tolerance” and ensures that only cells with appropriate receptors are expressed in any given functional repertoire[16,17].

Naïve cells that pass the central tolerance selection migrate to different secondary lymphoid organs such as spleen, lymph nodes (LNs) or mucosa-associated lymphoid tissues (MALT). Here, encounter with their cognate antigen triggers further differentiation

of T cells and B cells. At this stage, reactive cells undergo rounds of proliferation and selection to gain specific effector functions. Again, cells with autoreactive or non-functional receptors are eliminated or rendered non-responsive in a process called "peripheral tolerance"[16,18–25].

During these events, a subset of the activated cells develops a long-living phenotype and are retained in the system as components of immunological memory[7,8,26].

1.1.2 Lymphocytes biology

1.1.2.1 T cells

T cells originate from CLPs that undergo differentiation and selection in the thymus. Early development begins at the outer cortex of the thymus. Here, T cells attempts to generate functional TCRs will define commitment to either a αβ or a γδ lineage (Figure 1)[27]. Generally, αβ T cells are responsible for recognizing antigens that are presented by major histocompatibility complex (MHC) molecules while γδ T cells recognize antigens independent of classical MHC molecules[28,29].

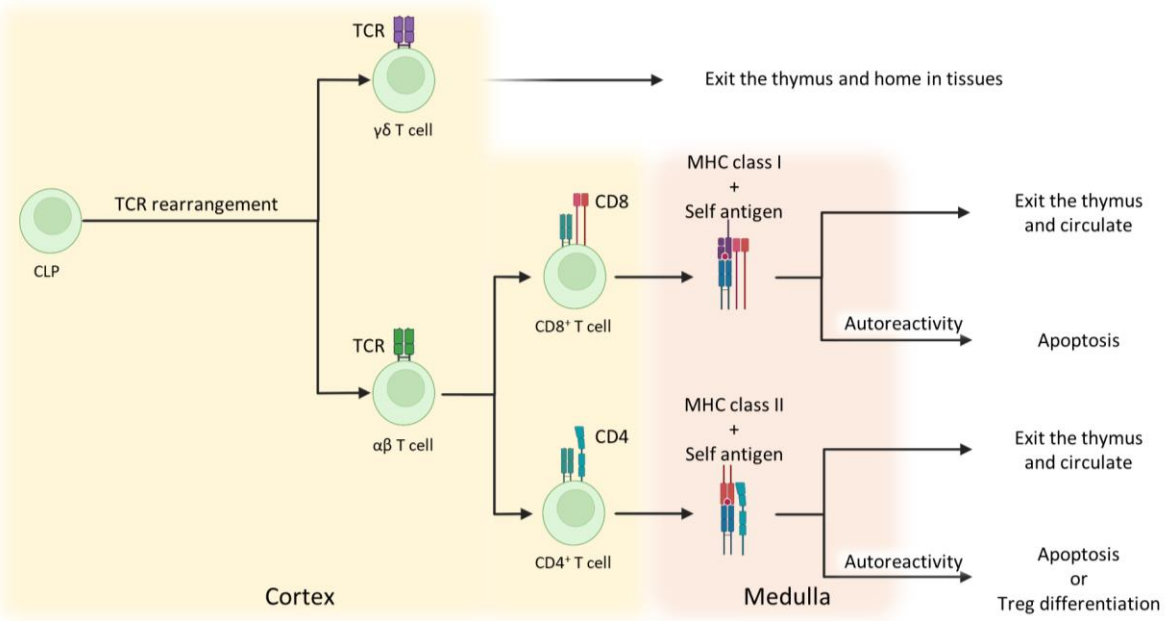


Figure 1. ***Early phase T cell development****. A simplified representation of T cell lineage commitment in the thymus. Briefly, T cells undergo initial TCR loci rearrangement events which determine αβ or γδ fate. αβ lineages further split into CD4⁺ and CD8⁺ T cells, which determine their effector function. Non-autoreactive circulate and patrol the body for possible threats.*

Recognition of MHC molecules is a crucial step for αβ T cells development, and it involves interaction with co-receptors CD4 or CD8[30]. In a process called "positive selection", αβ T cells without functional TCRs or without stable interaction with MHC molecules don't

receive survival signals and do not further differentiate. At this stage, αβ T cells commit to either a $CD4^+$ or a $CD8^+$ phenotype, which determines their effector functions. $CD4^+$ T cells recognize MHC class II molecules, while $CD8^+$ T cells recognize MHC class I molecules. Canonically, $CD4^+$ T cells modulate and regulate the immune response via cytokines, for example as T helper (Th) cells or regulatory T (Treg) cells[31]. $CD8^+$ T cells detect and eliminate infected or tumor cells via lysis, which is why they are also called cytotoxic T cells. Autoreactive $CD8^+$ and $CD4^+$ T cells receive apoptotic signals leading to cell death. This process, called "negative selection", takes place in the medulla of the thymus, and determines the thymic output of naïve T cells[17].

Naïve T cells migrate to secondary lymphoid organs and other peripheral tissues patrolling the environment for cognate antigens to become activated. In case of a virus infection, an infected cell can present foreign antigens via MHC class I (expressed by all cell types except erythrocytes) for $CD8^+$ T cell recognition leading to the elimination of the infected cell. $CD4^+$ T cell activation requires the presence of antigen presenting cells (APC), such as dendritic cells, as they are the only cells expressing MHC class II molecules[30].

In secondary lymphoid organs, activation of certain Th cells, called follicular T helper cells (Tfh), triggers the formation of specialized structure called germinal centers (GC), where Tfh cells mediate selection of mutated B cells[31].

The role of γδ T cells is not as extensively studied as that of αβ T cells. Most of the γδ T cells develop from early TCR rearrangement events before T cells start to express $CD4^+$ and $CD8^+$ molecules[27]. After functional TCR recombination, γδ T cells migrate to different tissues[32]. They do not require extra events to develop effector functions as they are already capable of reacting against antigens and other stress signals without antigen-specific selection phases[33,34]. This feature confers them with an "innate-like" role and places them in grey zone between the innate and the adaptive immune system. However, some γδ T cells react in response to infection by certain types of microorganisms and via TCR interactions and show expression of T cells memory makers[32].

The TCR repertoire varies depending on different T cells subpopulations. Canonically, αβ T cells generate repertoire various order of magnitude more diverse than γδ T cells. However, natural killer T (NKT) cells and mucosal-associated invariant (MAIT) T cells express semi-invariant αβ TCR repertoires which recognize non-peptide antigens presented by non-classical MHC class molecules[35,36]. More on their repertoire will be expanded in dedicated sections below.

1.1.2.2 B cells

B cells originate from CLP that undergo differentiation and selection in the BM (Figure 2)[16]. The most important developmental stage of B cells in the BM is the formation and

selection of a functional BCR. This process is divided in different phases, where, first, the heavy chain locus undergoes rearrangements for the formation of a pre-BCR. A pre-BCR is composed of a full heavy chain in association with a surrogate light chain. Pre-BCR-expressing B cells undergo proliferation and start to rearrange the light chain locus. The first attempts of rearrangements are performed in the kappa locus, followed by the lambda locus. Once a fully functional BCR is expressed, B cells expressing autoreactive or non-reactive receptors, undergo either receptor editing, anergy or deletion. This process is called "central tolerance"[16]. B cells with functional, non-autoreactive BCRs enter the circulation as immature naïve B cells and home to different secondary lymphoid organs. The BCR can have different isotypes, at this stage BCRs are expressed in the form of IgM and IgD.

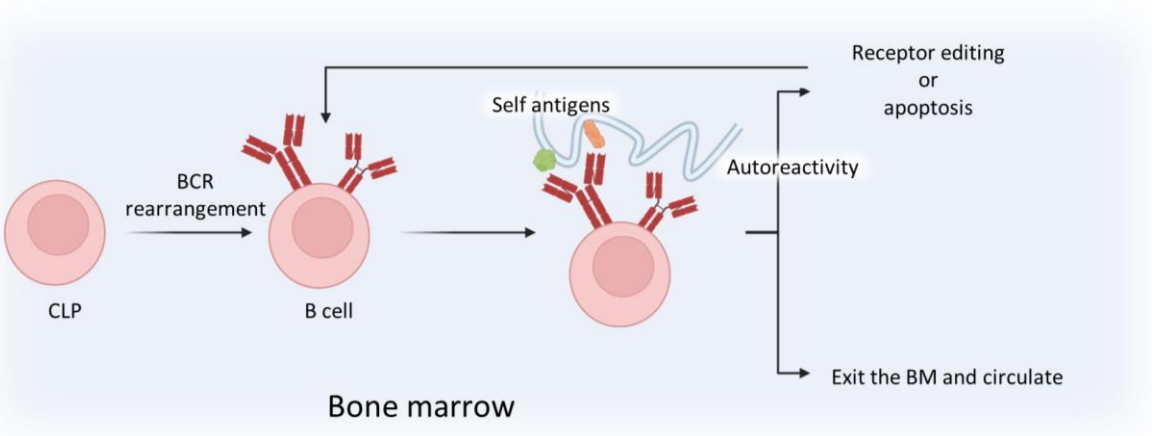


Figure 2. ***Early phase B cell development****. A simplified representation of core events determining B cell generation from the BM, Briefly, each B cell undergo BCR rearrangement events to generate a functional BCR. Non-autoreactive cells circulate to search for their cognate antigen.*

Depending on the type of exogenous stimuli, B cells can get activated with or without T cells intervention (T-dependent vs T-independent)[37]. A T-independent response is typically triggered by antigens that have repetitive structures, like polysaccharides, and lipopolysaccharides, as these antigens are capable of efficiently cross-linking with cognate BCRs. This response is characterized by a rapid production of IgM antibodies with no formation of GCs.

During a T-dependent response, initial interaction with cognate antigen triggers B cells proliferation and differentiation into plasmablasts. Plasmablasts are short-lived antibody secreting cells with the role of mounting a preliminary response against infections. Simultaneously, cognate B cells function as APCs by internalizing and processing pathogenic antigens. Interaction with cognate Tfh cells in secondary lymphoid organs triggers the formation of GCs (Figure 3)[38]. Here, B cells proliferate while their BCRs undergo rounds of somatic hypermutation (SHM) with the goal to improve BCR affinity to relevant antigens. At this stage, BCRs can change isotype resulting in IgG, IgA or IgE. This process

is called "class switching" and is dependent on the cytokine microenvironment, which is generally determined by Th cells[39]. Since SHMs are introduced in a random manner, some BCRs might acquire deleterious mutations or become autoreactive. To avoid so, B cells undergo a selection phase in an extremely selective environment. Matured BCRs compete for antigen affinity with each other by interacting with follicular dendritic cells (FDCs). The antigen is then internalized and presented to cognate follicular Tfh cells which provide survival and differentiation signals. B cell proliferation, SHM and class switching takes place at a distinct zone of the GC, called dark zone. Selection and differentiation take place in the light zone. B cells can undergo different rounds of migrations between these two zones. B cells passing selection will either differentiate into memory B cells or long-lived plasma cells, both responsible of immunological memory[26,40].

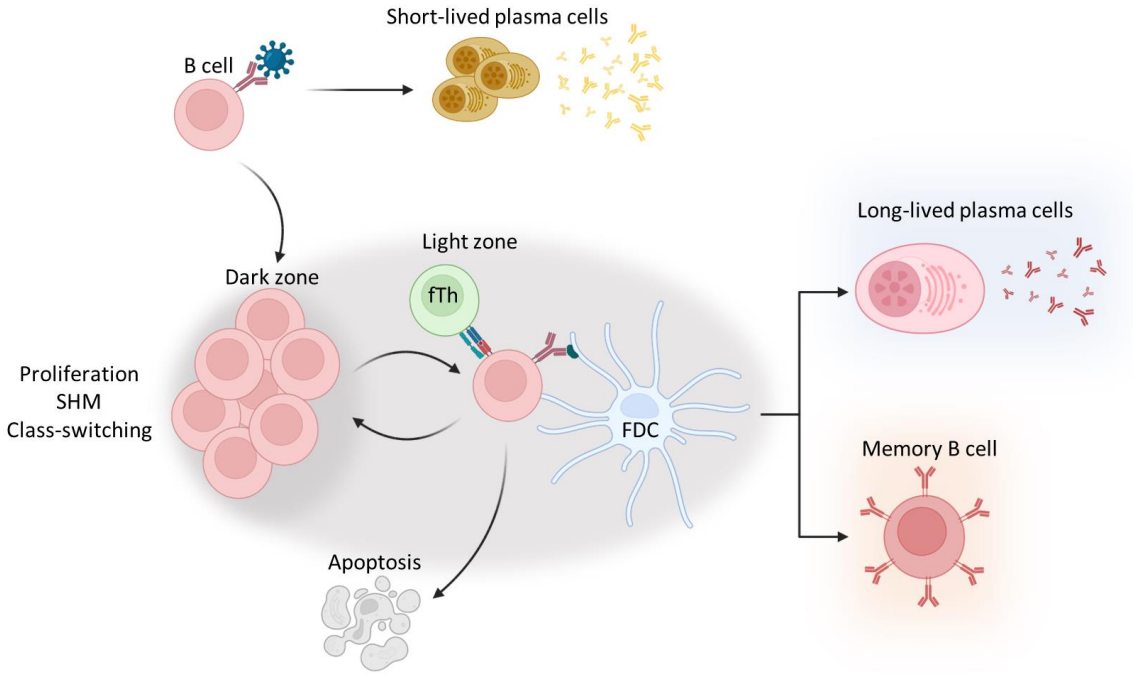


Figure 3. ***GC reaction and B cell differentiation.*** *Interaction with cognate antigens triggers B cell activation and further differentiation. Some B cells differentiate into short-lived plasma cells and readily secrete antibodies. Others form temporary structures called germinal centers. Germinal centers are composed of proliferating and mutating B cells zones, called dark zone, and selection zones, called light zones. B cells that develop affinity matured BCRs receive positive signals from fTh cells and differentiation into either memory B cells or long-lived plasma cells. Memory B cells circulate and patrol for a second encounter with the same antigen. Long-lived plasma cells migrate to the BM and continually secrete antibodies.*

Memory B cells re-enter the circulation to patrol the body and prepare to quickly respond to future encounters with the same antigen and converting into antibody-secreting cells. Memory B cells can also re-enter GCs and undergo further rounds of SHM. However, how often this happens is debated in the field[41,42]. Long-lived plasma cells migrate to BM niches and exclusively work as antibody-secreting cells, continuously replenishing circulating antibodies. Immunological memory durability varies between infection and immunizations

settings. Understanding the factors regulating long-lasting immunity are crucial to improve vaccine elicited immune responses.

1.1.3 Adaptive immune receptors

1.1.3.1 T cell receptor (TCR)

Each TCR is a heterodimer of either alpha (α) and beta (β) chains or gamma (δ) and delta (γ) chains (Figure 5A) [43]. These chains belong to the superfamily of immunoglobulins and are expressed exclusively as membrane-bound molecules. Overall, TCRs are designed to be somewhat polyreactive and can engage with peptide-MHC complexes with different levels of affinities[28,44]. Due to this restriction, most TCR antigen recognition is limited to linearized peptides.

Beta and delta chains are formed by somatic rearrangements of variable (V), diversity (D) and joining (J) genes. Alpha and gamma chains undergo similar gene rearrangement with the difference that their locus lack D genes. In humans, the beta and gamma loci are found on chromosome 7, while the alpha and delta locus is on chromosome 14 (Figure 4). Notably, the delta gene are located between the V and J genes of the alpha locus, with which it shares a limited number of Vs[43].

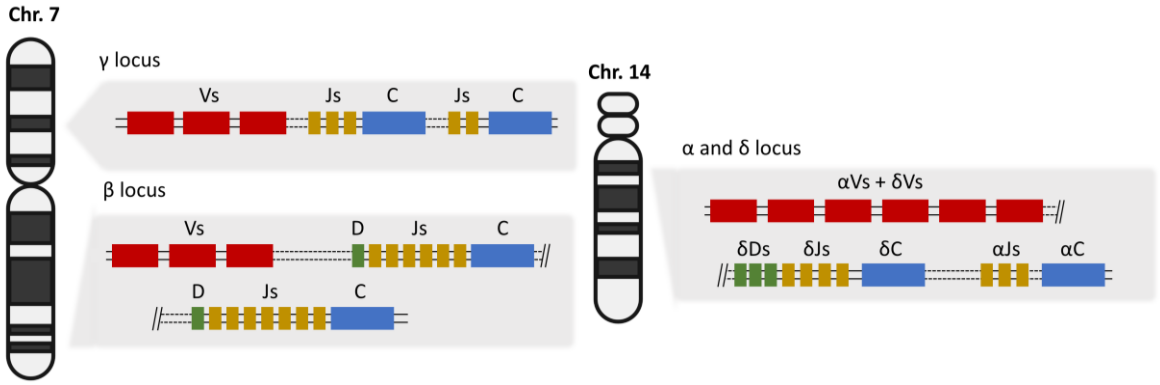


Figure 4. ***Representation of human TCR loci.*** *In human TCR loci are located in chromosomes 7 and 14. The gamma locus is located on the p arm of chromosome 7 while the beta locus is located close to the telomeric end of the q arm of chromosome 7. The delta locus is located between V gene and J genes of the alpha locus near the centromeric end of the q arm of chromosome 14.*

Each individual genome contains numerous V, D and J genes arranged in cluster over the TCR loci. Each rearrangement event joins only one V, one D (in beta and delta chains) and one J gene. The combinatorial assembly of a large number of V, D and J genes results in the generation of a tremendously diverse TCR repertoire, which is further enhanced by stochastic insertion and deletion events at the VD and DJ junctions during rearrangements and differs over time as new naïve T cells are produced. Both the TCR and BCR loci have undergone multiple duplication events in the genome throughout evolution,

resulting in non-functional copies of V, D and J genes as pseudogenes, or open reading frames (ORFs)[3,45].

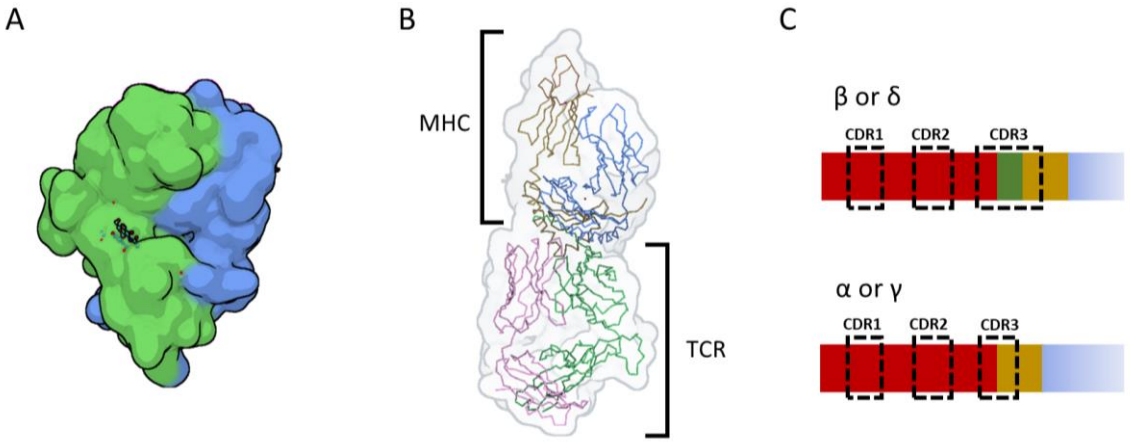


Figure 5. ***TCR overview.*** *A) TCR structure from PDB entry 1TCR. B) Interaction between MHC and TCR. Structure from PDB entry 4GRL. C) Representation of linear V(D)J arrangement. CDRs are highlighted in dotted boxes.*

Structurally, each TCR chain is composed of three regions of variability, called complementary determining regions (CDR1, CDR2 and CDR3) and four more conserved framework regions (FR1, FR2, FR3 and FR4) which provide structural stability (Figure 5B). The CDR1 and CDR2 are germline encoded by the V gene while the CDR3 is determined by the recombination of V(D)J genes. An extra region of variability, called hypervariable region 4 (HV4), is identified in the germline-encoded FR3[46–48].

Canonically, CDRs are responsible for antigen binding by direct interaction with peptide-MHC complexes with the CDR1 and CDR2 mainly involved in MHC recognition and the CDR3 mainly involved in recognizing the presented peptide (the antigen)[28]. Both chains contribute equally to this interaction with the TCR sitting on the top of the peptide cleft at different angles. However, different non-canonical interactions have been observed for semi-invariant T cell population and γδ T cells, where non-peptide antigens trigger T cell activation.[49–51].

Investigating TCR repertoires is important for our understanding of antigen interactions and the subsequent T cell response[52–58]. For instance, some peptides may be selectively recognized by TCRs that use specific germline genes, while others depend more on the CDR3. For instance, germline variation in the CDR2 loop of TRBV9 genes were showed to affect specific epitopes recognition[56]. Moreover, response against M1 influenza epitope generate public clonotypes using TRBV19, TRAV38, and TRAJ52[59].

1.1.3.2 B cell receptor (BCR)

BCRs are heterodimers of a heavy and light chain[60,61]. While there is only one heavy chain, the light chains can comprise either a kappa (κ) or a lambda (λ) chain. These chains belong to the superfamily of immunoglobulins and, unlike TCRs, can be expressed as both membrane-bound and secreted molecules, known as antibodies or immunoglobulins (Ig). BCRs main feature is the identification of antigens in their tertiary or quaternary structures.

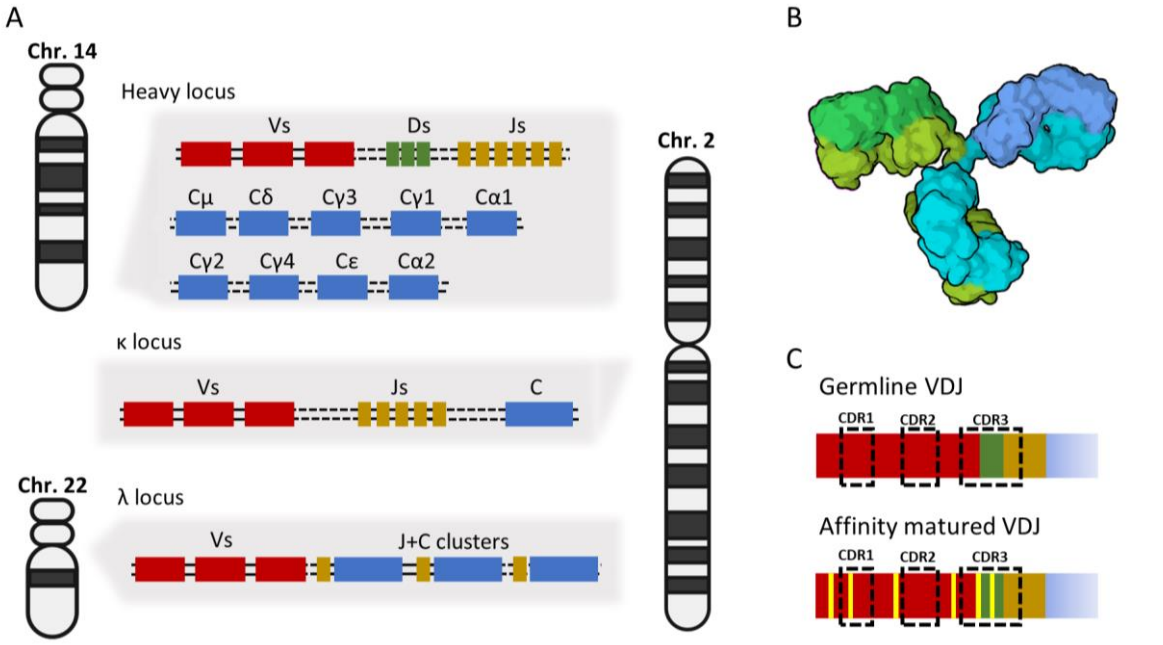


Figure 6. ***Representation of human BCR loci and antibody structure.*** *A) In human BCR loci are located in chromosomes 2, 14 and 12. The heavy chain locus is located at the telomeric end of chromosome 14 q arm. The kappa and lambda loci are located near the centromeric end of chromosomes 2 and 22 q arms, respectively. B) Antibody structure from PDB entry 1IGT. C) Representation of germline and mutated VDJ. SHM is represented in yellow.*

Heavy and light chains formation follow the same somatic rearrangement of V, D and J genes described for TCRs chains. The heavy chain is formed by V, D and J genes rearrangement and the light chain by V and J genes. Furthermore, like the TCR chains, they also have three CDRs and four FR regions with the CDR1 and CDR2 being germline-encoded by the V gene and the CDR3 resulting from rearrangement of V, D and J genes. The heavy chain locus features different heavy chain constant regions, which determines the BCR isotype and its effector function when secreted. Briefly, in humans, IgM provides a first line of defense against pathogens and, lastly, IgD still represent an enigma to immunologists but may have a regulatory role on naïve B cells[62–64], IgA (subclasses IgA_1 and IgA_2) is mainly involved in mucosal tissue protection and can be found in many secretions, IgG (subclasses IgG_1, IgG_2, IgG_3 and IgG_4) makes up the core humoral immune response against viruses and bacteria, IgE is involved in parasite infections but it is more

infamous for being involved in allergies[60]. In humans, the heavy chain locus is on chromosome 14, the kappa locus on chromosome 2 and the lambda locus on chromosome 22 (Figure 6A).

Structurally, BCRs are formed by two identical heavy chains and two identical light chains to make up a characteristic Y-shape molecule with the variable part at the tips and the constant part at the base. In its membrane-bound form, BCRs bind antigens to provide activation, proliferation, and differentiation signals to the cell (Figure 6B). when B cells differentiate into antibody-secreting cells (plasmablasts and plasma cells) the surface BCRs is downregulated[65–67]. In their secreted form, antibodies use different strategies to fight pathogens. Neutralizing antibodies block crucial interactions with host receptors for viral entry and are the correlates of protection for most vaccines[68]. In addition, non-neutralizing antibodies can tag pathogens for phagocytosis in a process call opsonization or can activate the complement system to promote inflammation and pathogen destruction[69].

During GCs reactions, BCRs undergo further modifications by SHM (Figure 6C). Mutations introduced by SHM are random but the selection process in the light zone select BCRs without deleterious mutations that have improved affinity and for the selected antigens[70]. Often, SHM provides additional interactions between the antibody-antigen contact surface, decreasing the antibody off-rate resulting in higher affinity[71].

BCR repertoire studies are important to dissect the role of B cell clones in health and diseases. Analysis of expressed repertoires can identify signature features of the response in the context of tumors, infections, allergy, and autoimmune diseases[72–76]. For instance, an abnormal clonal expansion can be indicative of certain types of malignancies such as leukemia and lymphomas[77,78]. Another example is provided by the identification of features such as biased gene family usage, and different levels of SHM observed in autoimmune patients[79–82]. Different studies report how specific response against pathogens, or their antigens, is dependent on certain germline immunoglobulin alleles[83–86]. Remarkable examples are provided by the need of IGHV1-2*02 for the generation of a specific class of broadly neutralizing antibodies (bnAbs) against HIV-1[87,88] or by more recent observations of biased gene usage in response against SARS-CoV-2 S protein[89–93].

1.2 AIRR-seq

With adaptive immune receptor repertoire sequencing (AIRR-Seq) we obtain thousands of cells in one sequencing run, allowing analysis of TCR or BCR gene usage, clonal expansion, affinity maturation (for BCRs) and selection. Prior to the development of deep sequencing techniques, gene usage in BCR and TCR repertoire studies utilized low-

throughput technologies such as Sanger sequencing of cloned transcripts or FACS analysis with unpractical massive sets of V-specific antibodies[94].

Despite the availability of full genomic assemblies[95,96], the adaptive immune receptor loci remain poorly defined due to their high complexity[97–99]. The extremely polymorphic nature and presence of several repeated sequences and frequent gene copy number variation makes them prone to mis-assemblies and misassignments using conventional short-read sequencing approaches. This has hampered the development of comprehensive databases of TCR and BCR germline V, D and J alleles, a critical requirement for accurate AIRR studies[99,100].

At present, germline gene databases for BCR and TCR V, D and J genes from different species are provided by the international ImMunoGeneTics information system (IMGT), which used different genome reference assemblies to identify these genes. Recent studies have shown how these sources do not encompass the full diversity and complexity of adaptive immune receptor loci as they are based only on a limited number of European individuals[101].

Recent studies performed by us and other groups utilize different approaches and novel technologies for repertoire analysis, briefly described in the following paragraphs [99,100,102–104].

1.2.1 Next generation sequencing

Next generation sequencing (NGS) technologies (also known as high-throughput sequencing) allow for deep parallel sequencing of different targets compared to standard sequencing methods (e.g., Sanger sequencing)[105]. There are different NGS technologies applied to AIRR-Seq protocols, each of them having advantages and disadvantages[106]. Here, I mainly focus on the technology provided by Illumina®, which is one utilized in the work performed for this thesis to produce libraries.

Illumina® utilizes a "sequencing by synthesis" approach, which can cover up to 600 bp with their MiSeq platform. Briefly, during each MiSeq run, the sequencing target is immobilized on a flow cell and amplified to form clusters. Each sequence then undergoes different sequencing cycles where, in each cycle, a single fluorophore-labelled dNTP is added and the signal visualized. This provides efficient single base pair reading for roughly 300 bp and is performed at both sequence termini for a total of 600 bp. This coverage is sufficient to retrieve full V(D)J sequences from BCRs and TCRs transcripts. Library preparation is a crucial step for NGS runs, as it's one of the main determinants of the quality of the output. AIRR-Seq library preparations utilize either 5'RACE or 5'MTPX protocols[107]. 5'RACE library preparation doesn't require extensive primer design but produce long amplicons that might exceed the coverage limit. Conversely, 5'MTPX protocols produce shorter amplicons but need careful primer design to incorporate all

target genes. Generally, 5′RACE are more apt to be used for repertoire analysis of species with little available adaptive immune receptor loci information while 5′MTPX can harness available data for precise primer design. Past and current members of my lab have worked extensively to improve library preparation protocols[100,107]. For human and NHPs libraries, the lab has established 5′MTPX protocols with constant-specific primers targeting proximal sequences to the V(D)J segment and a series of leader-specific primers to avoid primer bias against V gene variants.

AIRR-Seq studies need generation of computational pipelines to handle and curate big datasets[99,106,108–112]. The work of this thesis is extensively based on the use of IgDiscover, a BCR and TCR germline gene inference tool developed by the Karlsson Hedestam lab[99]. Germline inference focuses on detecting germline sequences obtained from expressed rearranged V(D)J transcripts. The group has continued to develop IgDiscover, and the program now includes the module "plotallele", which allows visualization of haplotypes, the module "clonotypes", which clusters related V(D)J transcripts for the identification of clonal lineages and "corecount" which provides a means to validate genotypes and identify polymorphisms at the 3′ end of the transcript, otherwise difficult to capture using inference approaches. Our group constantly refines the computational pipelines for AIRR and germline gene inference, while other groups work on genomic approaches using long-read sequencing[113,114].

1.2.2 Single cell sequencing

The capability of isolating and sequencing single cells provides a powerful tool in many areas of biology. The main advantage provided to AIRR-Seq is the ability to identify paired (heavy and light for BCRs and alpha/beta for TCRs) chains sequences from single cells. The sequences can be used for the production of recombinant monoclonal antibodies (mAbs) or TCRs. In contrast, standard protocols involve sorting cells of interest into microtiter plates, amplifying the target of interest via PCR, and Sanger sequence paired chain transcripts[103,115]. This approach relies on standard molecular biology techniques and requires close to zero computational knowledge. However, these protocols are time consuming and rely heavily on the available workforce.

Advancement in fluidics engineering resulted in the generation of microfluidic systems where minute quantity of fluid can be manipulated at will. Microfluidic systems can encapsulate single cells in droplets with all the necessary reagents needed for an experiment. For AIRR-Seq protocols, single cells are encapsulated with cDNA synthesis reagents to amplify BCR and TCR transcripts which can be then sequenced using NGS technologies. Microfluidics platforms employ two strategies to achieve paired chain sequencing: physically linking the receptor transcripts for cDNA synthesis or capturing cells with beads containing cell-specific barcodes to be used for downstream data

processing[116–119]. Microfluidic systems provide higher throughput than standard techniques but require post-sequencing processing of data.

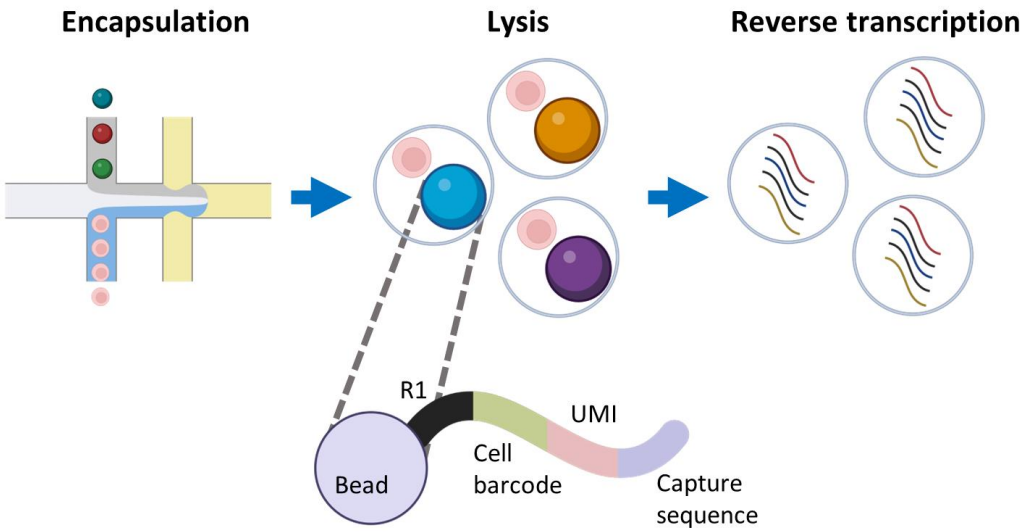


Figure 7. ***10X single cell sequencing schematic.*** *Cells are encapsulated in droplets generated via water in oil emulsion. Each droplet contains a single cell and a gel bead covered with oligonucleotides stretches. The cell is lysed, and the mRNA fragmented for the bead to capture. Reverse transcription produces cDNA with UMIs, cell barcodes and adaptors for library preparation.*

Eventually, commercialization of microfluidics systems granted access to researchers without extensive engineering background. As of now, 10X genomic provides an extremely user-friendly service with their Chromium platforms. The chromium platform encapsulates each cell with a solution of reagents and a bead containing a cellular barcode. After cell lysis, the transcript is fragmented and captured on the bead. The bead incorporates the cellular barcode in each transcript during the cDNA synthesis step. After sequencing, the short reads can be reconstructed for each individual cell. Availability of V(D)J kits allows for high throughput paired chain data generation.

Future advancement in single cell sequencing is likely to improve resolution and throughput while improving accessibility to more labs.

1.3 Anti-viral vaccine research

The principle of vaccination is to elicit a pathogen-specific protective response that provides immunological memory without stimulating the adverse effect triggered by a potential infection with the same pathogen. The first anti-viral vaccines consisted of live attenuated pathogens that had limited replication capability or inactivated pathogens with no replication capability[120].

Modern approaches aim at focusing the immune response on specific target antigens by generating sub-unit vaccines[121,122]. While initial sub-unit vaccines relied on protein extracts obtained from processed pathogens, modern products rely on multi-disciplinary expertise to identify and design target antigens, usually the surface antigen of viruses as these are targets for neutralizing antibodies (nAbs).

This section briefly introduces concepts related to recent vaccine research findings related to viruses. To focus the discussion, I will discuss this section in the context of SARS-CoV-2 and usage of NHP models.

1.3.1 SARS-CoV-2

The 30th of January 2020, the World Health Organization (WHO) declared the status of Public Health Emergency of International Concern after the outbreak of a novel coronavirus in December 2019 in the city of Wuhan (Hubei, China). The virus, called SARS-CoV-2 (severe acute respiratory syndrome coronavirus 2), was identified as the cause of a wide range of symptoms, mostly affecting the respiratory system, characterizing the disease called coronavirus disease 2019 (COVID-19). On the 11th of March 2020, the status of pandemic was declared by the WHO, causing not only massive concerns to public health but generating enormous socio-economical disorders all over the world.

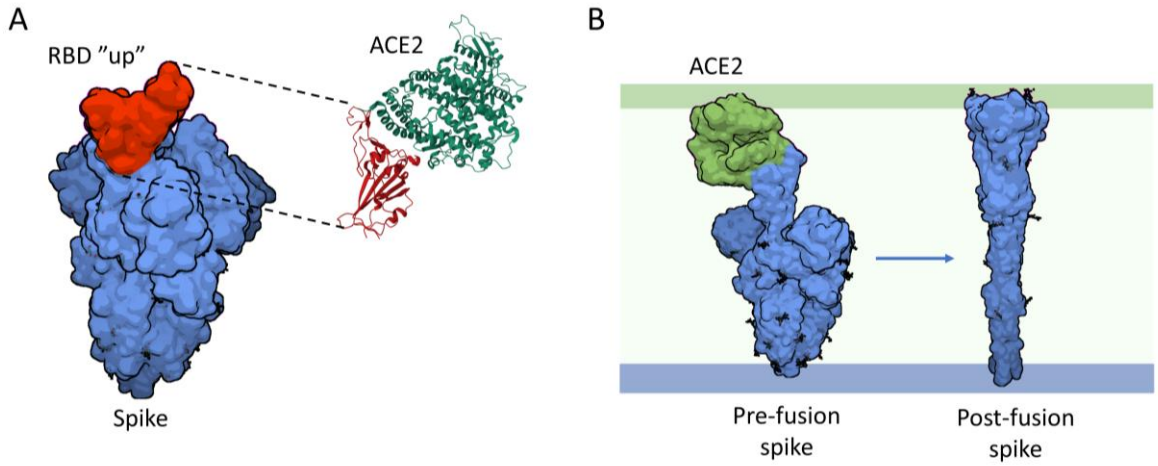


Figure 8. ***SARS-CoV-2* S *structure* overview.** *A) Pre-fusion S trimer with one RBD (in red) in the "up" conformation.* Zoomed in interaction of RBD with ACE2. S s*tructure from PDB entry 7T67.* RBD-ACE2 structure from PDB entry 7E3J. *B) RBD interaction with ACE2 (in green) triggers conformational changes which lead to viral-host membrane fusion. S-ACE2 structure from PDB entry 6CS2. Post-fusion S structure from PDB entry 7E9T.*

SARS-CoV-2 is a positive single-stranded RNA virus of the genus of betacoronavirus that initiates infections by interacting with host angiotensin-converting enzyme 2 (ACE2)[123,124]. The virus interaction with ACE2 is mediated by the viral S protein, a homotrimer composed of two subunits, S1 and S2[125,126]. S1 contains the receptor binding domain (RBD),

a mobile domain that directly interface with ACE2, triggering viral membrane fusion and release of viral material. Once inside the cell, the virus highjacks host cellular pathways to promote viral proliferation while avoiding detection[127]. The S protein, and especially the RBD, is a major target for neutralizing antibodies and of central interest in vaccine studies[128].

The response of the scientific community to the pandemic was extraordinary, with vaccines being developed and rolled out at record speed. National and international efforts resulted in the production of vaccines with various platforms with most of the global population receiving multiple doses of the following vaccines: adenovirus vectored vaccines (AstraZeneca, Johnson & Johnson/Janssen pharmaceuticals, Sputnik V, Bharat Biotech)[129–132], mRNA vaccines (Moderna, Pfizer-BioNTech)[133,134], recombinant protein vaccines (Novavax)[135] and inactivated vaccines (Sinopharm, Sinovac)[136,137].

The viral landscape has change from the initial outbreak, with different SARS-CoV-2 variants emerging at regular waves[138]. These variants are characterized by mutations granting advantages features, such as enhance virulence, enhanced infectivity or escape of neutralizing immunity. At the time this thesis is written, the subvariants of the Omicron clade make up most of the circulating viral landscape[139,140]. Moderna and Pfizer-BioNTech have updated their vaccine to include Omicron variants in an attempt to repristinate global immunity[141].

1.3.2 Non-humane primate (NHP) model

The usage of animal models in research is crucial to study the complex networks of reactions that characterize the immune system. Despite the development of new technologies and the ever-expanding power of computational tools, immunology research still cannot rely solely on in vitro or in silico experiments. Despite the fact that mouse models can be adapted for a variety of studies, NHPs are often required for final proof of concepts in infection and vaccination pre-clinical studies.

The term NHP refers to many species of monkey from the old and the new world. In pre-clinical studies the primary species of choice are Asian origin macaques such as *Macaca mulatta* (Rhesus macaques) or *Macaca fascicularis* (Cynomolgus macaques). Macaques display high homology to humans in both genetic, molecular, and cellular features and they are susceptible to many human infecting pathogens[142–147].

The usage of NHPs was crucial at the beginning of the pandemic to reveal the effect of infections, vaccination, and pharmacological therapies, opening the door to clinical trials. For example, NHPs have been used to study the virus's replication and distribution in different tissues[148–154]. How this correlates with disease severity and development of certain symptoms. Antiviral drugs were used on NHPs to test the capability of controlling infection and hampering replication[155,156]. The immune response against the virus was

analyzed through challenge and immunizations studies, with all the vaccine platform undergoing NHPs test before clinical trials[128,138,148,152,157–172].

2 Research aims

The goal of this thesis was to improve our understanding of adaptive immunity by combining sequencing and immunoassays for germline gene identification, AIRR analysis and functionality.

In each individual paper the aims were:

- **Paper I**: perform a comprehensive analysis of human germline TCR loci by including individuals from different populations and comparing TCR genes in modern human genomes with ancient human and NHPs genomes.
- **Paper II**: characterize the humoral immune response in mouse and NHPs immunized with pre-fusion subunit SARS-CoV-2 S proteins aministred in adjuvant.
- **Paper III**: determine the effect of an heterotypic protein boosts on the humral immune response in the context of a SARS-CoV-2 landscape with emerging variants.
- **Paper IV**: track SARS-CoV-2 S immunization-elicited B cell lineages in NHPs, isolate mAbs and describe B cell lineages features, especially those with interesting neutralization profiles

3 Materials and methods

3.1 Ethical considerations

3.1.1 Human subjects

The study in **paper I** involved samples obtained from human donors. In this setting is crucial to protect individual privacy. The study was performed by following the regulations of national and international bodies. To summarize the most important aspects of these regulations:

1) The purpose of the research study must be explained to the participants and only those who give their consent to provide samples can be included in the study.
2) Their identity is kept secret and confidential by using anonymizing codes.
3) The raw data is kept on secure servers and can only be accesed via data sharing agreement.

Since our study involved handling of sequencing information, possible dilemmas can arise when disease associated genetic variants are identified. This type of communication is not straightforward as the individual set of moral and ethical principles might be too complex to predict the best course of action. However, such associations are currently not known for TCR genes as they have not been studied in detail. Thus, I did not foresee any ethical dilemmas of this sort.

The studies have been conducted after approval of the following ethical permits and amendments: #2006/893-31/4, #2011/222-31/1, #2013/550-32, #2013/549-32 and #2018/2354-32.

3.1.2 Non-human primates

The NHP model was used in **papers II, III and IV**. Despite technological advancements in biomedical research, there are no comprehensive *in-vitro* or *in-silico* replacements that mimic complex system interactions. Of all animal models, NHPs provide the highest level of homology with humans and this model is therefore important for immunology and infection studies.

During the course of the studies present in this thesis, I adhered to the three Rs of animal research

- Replacement: we use our macaques after appropriate results have been observed in vitro or in small mammals.
- Reduction: our macaques were part of previous vaccine studies, and the smallest sample size was employed for our SARS-CoV-2 projects.

- Refinement: our macaques were housed at the Astrid Fagraeus Laboratory facility where they udnerwent optimal housing conditions, enriched caring routines and regular checks.

The germline alleles identified in these projects will be added to the macaque immunoglobulin germline V, D, J database housed by our group (http://kimdb.gkhlab.se/). In the long run, the information collected in these studies might benefit the collective understanding we have of these animals in and outside pre-clinical contexts. Furthermore, I hope that the information gathered in my study will contribute to more efficient future NHP studies.

The work performed aimed to add useful knowledge about vaccine research. Vaccines represent the most efficient health intervention strategy against pathogens and diseases currently available.

The studies have been conducted after approval of the following ethical permits and amendments: #18427-2019 and #10895-2020.

3.2 Sampling

In **paper I** PBMCs were isolated from whole blood samples collected with informed consent from healthy volunteer donors that enabled the donors to self-identify their population group and from a cohort of malaria-infected individuals from central Africa.

In paper **II, III and IV**, blood, LNs, BM and spleen samples obtained throughout the vaccination regimens were collected after were performed under sedation with 10-15 mg/kg ketamine (Ketaminol 100 mg/ml, Intervet, Sweden) administered intramuscularly.

Mononuclear cells were isolated from blood and BM samples via density-gradient centrifugation using Ficoll-Paque. Mononuclear cells were isolated from LNs and spleen single cell suspensions using 70 μm cell strainers.

3.3 Immunizations

In **paper II and IV**, macaques were immunized with stabilized spike trimer (100 μg) mixed in 75 μg of Matrix-M (Novavax AB). Macaques were immunized intramuscularly with half of the dose administered in each quadricep.

In **paper III**, macaques were immunized with stabilized spike trimer (2 μg) mixed in 50 μg of Matrix-M (Novavax AB) and subsequently boosted with either 2 μg (H05), 10 μg (H06), or 50 μg (H07) of soluble RBD in 50 μg of Matrix-M. Macaques were immunized intramuscularly with half of the dose administered in each quadricep.

3.4 Immunoassays

3.4.1 ELISA

Enzyme-linked immunosorbent assays (ELISA) were used to characterize binding specificity of plasma (**paper II**) and mAbs (**paper IV**).

Overall, ELISAs were performed by coating ELISA plates overnight at 4°C with 100 µL of prefusion-stabilized S protein or RBD at a concentration of 1 µg/ml in 1x PBS. Plasma samples or mAbs, serially diluted in blocking solution, were added, and incubated for 2 hours at room temperature. Antibody-antigen interaction was detected using 100 µL HRP-conjugated anti-monkey IgG Fc (Nordic MUbio) diluted 1:20,000 in PBS-T for plasma. Detection of mAbs was performed by using 100 µL of HRP-conjugated anti-human Fcγ Ab (Jackson ImmunoResearch) diluted to 1:10 000 in PBS-T. EC_{50} titers were calculated from the posterior median value midway between the plate minimum and maximum.

3.4.2 Pseudovirus neutralization assay

Pseudovirus neutralization assays were used to detect neutralization capability of plasma (**paper II, III**) and mAbs (**paper IV**).

Plasma samples were heat inactivated at 56°C for 60 min. Pseudotyped lentiviruses displaying S from the SARS-CoV-2 pandemic founder variant (Wu-Hu-1) variants of concerns and packaging a firefly luciferase reporter gene were generated by the co-transfection of HEK293T cells using Lipofectamine 3000 (Invitrogen) per the manufacturer's protocols. On **paper II** pseudotyped viruses sufficient to generate ~100,000 relative light units (RLUs) were incubated with serial dilutions of plasma for 60 min at 37°C in a 96-well plate, and then ~15,000 HEK293T-ACE2 cells were added to each well. On **paper III** pseudotyped viruses sufficient to generate ~50,000 RLUs were incubated with serial dilutions of plasma for 60 min at 37°C in a 96-well plate, and then ~15,000 HEK293T-ACE2 cells were added to each well. On **paper IV** pseudotyped viruses sufficient to generate ~30,000 RLUs were incubated with serial three-fold dilutions of mAbs for 60 min at 37°C in a 96-well plate, and then ~10,000 HEK293T-ACE2 cells were added to each well. IC_{50} titers were interpolated as the reciprocal serum dilution at which RLUs were reduced by 50% relative to control wells in the absence of serum.

3.4.3 B cell ELISpot

In **paper II**, ELISpot immunoassays were used to assess the presence of S-specific memory B cells.

ELISpot plates were coated overnight at 4°C with 10 µg/ml anti-human Fcγ (Jackson ImmunoResearch). Serially diluted PBMCs were plated and incubated overnight at 37°C, 5% CO2 after previous cultured for 72 h in supplemented complete media supplemented. The following biotinylated probes were used: 2.5 ng/ml goat anti-human Fcg (Jackson

ImmunoResearch) to detect total IgG, 1 mg/ml prefusion-stabilized spike protein or 3 mg/ml ovalbumin to detect, respectively, total IgG and antigen-specific IgG. Spots were counted using an Immunospot analyzer.

3.4.4 Analysis of T cell response

In **paper II**, T cells reactivity was determined by intracellular staining of target cytokines IFNγ, IL-2 and IL-13.

Cryopreserved PBMCs were thawed and rested for 3 hours at 37°C in a 5% CO2 incubator. Afterwards, PBMCs were incubated with SARS-CoV-2 S overlapping peptide pool (OLP, JPT Peptide Technologies) at 2 µg/mL or recombinant Spike trimer 57 at 10 µg/mL. DMSO was used as a negative control, and Staphylococcal Enterotoxin B (Sigma-Aldrich) was used as a positive control. Brefeldin A (Sigma-Aldrich) was added to block Golgi apparatus transport and inhibit cytokines secretion. Surface marker staining preceded intracellular staining. Stained cells were fixed using 1% formaldehyde and acquired with LSRFortessa flow cytometer (BD Biosciences).

3.4.5 Analysis of γδTCR reactivity to BTNLs

In **paper I**, TRGV4 allelic variation effect on BTNLs reactivity was tested in vitro using Jurkat reporter cell lines JE6-1 cell lines. TRGV4*02 and TRGV4*02_SO072 were analyzed in this assay.

TCR-expressing JE6-1 reporter cells were added on top of BTNL3/8- or empty vector-transfected HEK293T cells and cocultured for 18 hours before staining. As a positive control, the reporter cells were cultured in 96- well plate pre-coated with 10 µg/ml anti-CD3ε antibody. Samples activities were acquired on BD LSR Fortessa flow cytometer (BD Biosciences).

3.5 Monoclonal antibody production

In **paper IV**, V(D)J sequences from single cell analysis were used to produce mAbs.

Sequences were re-adapted for Gibson assembly cloning by inserting complementary overlapping sequences to the expression vectors human Igγ1 H, Igκ1 or Igλ2 leader and constant regions. Gibson cloning was conducted by mixing the digested vector and the V(D)J insert with 2X Gibson Assembly Master Mix (New England BioLabs) as per manufacturer instructions. The reaction product was transformed into XL10-Gold Ultracompetent cells (Agilent) and cultures were scaled up to obtain a suitable quantity of plasmid for expression in HEK293F cells. Transfection was carried out by adding a transfection mix to HEK293F cells composed of heavy chain vector, light chain vector, FreeStyle™ MAX Reagent and Opti-MEM. The supernatant was harvested after 7 days and mAbs were purified using gravity driven column purification with Protein G Sepharose

(Cytivia). After elution, mAbs were diluted in PBS and concentrated using Pierce™ Protein Concentrators with a cutoff of 30 kDa.

3.6 Sequencing

3.6.1 Sanger genomic validation

In **paper I**, a set of novel alleles were validated by Sanger sequencing following target PCR amplification using primers located upstream of exon 1 and downstream of exon 2 of each V gene or targeting the genomic sequence flanking a J gene. Novel alleles were identified from DNA extracted from our donors and from DNA samples of 1000 Genome project samples.

3.6.2 NGS library preparation

In **paper I**, human TCR cDNA libraries were created using chain specific cDNA synthesis primers. Individual primers specific for the constant regions of TRA, TRB, TRG and TRD to produce, using the Sensiscript cDNA synthesis kit (Qiagen). This cDNA synthesis primers also includes a 21 nucleotide UMI that is present in the amplified product and is used in downstream computational analysis. TCR library amplification was performed separately for TRA, TRB, TRG and TRD using chain specific 5'MTPX primer sets, along with a universal reverse primer that recognizes the amplification target site present in the 5' tail of the chain specific cDNA synthesis primer. The TCR library PCR products were subsequently gel purified and indexed for sequencing using the index primer sets and conditions described in Vazquez Bernat et al, 2019[107].

In **paper IV**, macaque BCR cDNA libraries were created using heavy chain IgM and IgG specific cDNA synthesis primers[100]. This cDNA synthesis primers also includes a 21 nucleotide UMI that is present in the amplified product and is used in downstream computational analysis. Libraries were prepared using a 5'MTPX protocol described in Vazquez Bernat et al, 2019[107].

All libraries were sequenced with 15% PhiX174 DNA and the Illumina Version 3 (2 × 300 bp) sequencing kit using a MiSeq 3x300 platform (Illumina).

3.6.3 Single cell sequencing

In **paper IV**, single B cells were sorted with a FACSAria cell sorter (BD Bioscience) using a biotinylated S probe. Prior to sorting, TotalSeq-C hashing barcodes for downstream sample identification were added to terminal samples from H03 and PBMCs from blood samples obtained from I10.

The sorted cells were processed with the 10X Chromium, sequencing '5' V(D)J' enriched libraries, and TotalSeq-C feature barcode libraries. As Rhesus macaques are not a species

directly supported by 10X, we adapted the "Human" 10X kit by spiking-in a mix of Ig constant-region primers from Brochu et al., 2020[173].

3.7 Bioinformatic tools

3.7.1 IgDiscover

IgDiscover was used in **paper I and IV** to infer the germline adaptive immune receptor repertoire of each studied individual. Different IgDiscover modules have been used to accommodate specific analysis.

In **paper I**, TCR libraries were processed using IgDiscover (v.1.0.1). In brief, the R1 and R2 sequence reads were merged, the UMIs were extracted, and the sequences were assigned to a starting database, consisting of the IMGT TCR reference database for each TCR chain. The IgDiscover program assigns each TCR sequence to the closest database reference sequence and calculates specific nucleotides that differ from this assignment. The identification of a consensus sequence that differs from a starting database sequence results in a potential candidate novel germline sequence. The program then uses a series of filters designed to identify features of germline sequences and to remove non-germline sequences from the set of candidate sequences. These filters, primarily based on the identification of multiple CDR3s or Js associated with candidate Vs or multiple CDR3s and Vs associated with candidate Js, function to identify unique V and J sequences that are utilized in multiple independent rearrangements. In addition, the program filters sequences for clonal expansion to handle expanded TCR clones can interfere with germline inference processes by obscuring heterozygous alleles of the same gene.

The *corecount* module was used to analyze the chain-specific genotype of each individual. The *corecount* module inputs the set of filtered sequences produced by IgDiscover. For each gene, unique UMI containing sequences identical to alleles in the database are counted. In contrast to the germline inference process the *corecount* genotypic analysis module enables identification of low expressed genes. *Corecount* genotypic analysis was performed on V and J sequences from all libraries and D sequences from TRB and TRD libraries.

The module *plotalleles* from IgDiscover was used to infer individual's haplotype. The module is based on the principle that heterozygous J or D alleles can be used to map associated V alleles to the appropriate chromosomal region. V(D)J recombination occurs locally on a single chromosome and so a heterozygous J or D allele will only recombine with V alleles present on that chromosomal region and not on the corresponding locus of the other heterozygous chromosome. Hence maternally derived V alleles will associate with the maternally derived J or D allele and paternally derived V alleles will likewise only associate with the paternally derived J or D alleles on that chromosome. The *plotalleles* module creates a phased map of V and J alleles present based on the heterozygous J or

D content of that case. In addition, heterozygous V alleles enable the identification of J or D gene deletion within the library. Haplotype analysis is utilized for two purposes. First it is a useful validation process. The clear separation of a heterozygous novel germline allele serves as a means of validating the germline inference process. Secondly, the process enables the identification of structural or expression-based variation in each genotype, enabling the identification of duplications, deletions or loss of expression that may be due to recombination associated genomic variation. IgDiscover based inferred haplotype analysis was performed as described previously in Vázquez Bernat et al, 2021[100], using the appropriate heterozygous TCR anchors.

In **paper IV**, IgM libraries synthesized from blood samples obtained before the vaccination regimen were used to generate individual germline repertoire via inference analysis using IgDiscover version v1.0.0 with default settings. The reference database for the analysis was obtained from KIMDB (http://kimdb.gkhlab.se/) by pooling both Rhesus and Cynomolgus databases and removing duplicates.

Bulk IgG libraries were processed with IgDiscover for assignment of individual VDJ germline repertoire. The module *clonotype* was used to determine B cell clonal lineages using post-processed 10X single cell data as reference together with bulk IgG libraries. Lineages were defined by identical V and J allele assignments, identical CDR3 lengths, and permitting a maximum of 20% mismatches to nucleotide CDR3 sequences. If single cell sequences with different light chain assignments were assigned to the same lineage, the lineage was redefined by also considering same V and J allele assignment of the light chain and same CDR3 length.

3.7.2 10X data processing

The single cell VDJ read assembly identified in **paper IV** was run in CellRanger's "de novo" mode, to avoid reference bias, since CellRanger lacks a Rhesus macaque reference database. The filtered contigs files were then assigned to individualized H03 or I10 IGH database and IGK + IGL IMGT database. Non-productive sequences or sequences without CDR3 identification were removed. Subsequently, we filtered cells with multiple IGH, IGK or IGL assignments and cells with presence of both IGK and IGL. Finally, we set a strict threshold for hashing barcode assignment. The sequences with maximum value of counts per hash lower than 20 were classified as "Low_counts". Hashes accounting for >= 60% of all counts were selected for assignment. Otherwise, the sequences were classified as "Unassigned".

3.7.3 Custom scripts

Different custom scripts were written using python, R or Julia for specific analysis or plotting throughout all the papers.

In **paper II, III and IV**, ELISAs curves were fitted using a four parametric logistic regression after minimum-maximum normalization for each plate.

In **paper IV**, bulk IgG libraries were polished using a denoising processing tool called the Fast Amplicon Denoising (FAD) described in Kumar et al. (2019) and removed chimeric sequences (likely due to PCR recombination between unrelated antibodies) using a hidden Markov model designed for this purpose.

During B cell lineage identification, when light chain allele assignment was ambiguous, we manually assigned the light chain by computing a maximum-likelihood analysis using FastTree with the full germline database. Sequences clustered together were assigned to the closest germline light chain. In case of divergent light chains, lineage assignment was determined by minimum levenshtein distance using each single cell heavy chain sequence as reference. Sequences from each lineage were then aligned using MAFFT v7.490. The alignment was used as input to FastTree (compiled with the double-precision flag) to compute phylogenetic trees. The function phylo::reroot from R package "phytools v1.2-0" was used to root the tree to an inferred germline sequence. Germline CDR3 inference was performed using the Julia package "MolecularEvolution" using a HKY85 substitution model.

4 Results & discussion

4.1 Paper I – Individualized TCR genotyping reveal extensive diversity and regions of introgression from Neanderthal ancestors

The current knowledge on the TCR germline loci is based on a limited number of individuals of European origin, collected in the IMGT database[61,101]. While genetic inter-individual diversity is increasingly appreciated for other immune related loci[174–177], AIRR-Seq faced unprecedent challenges due to the genetic features of the TCR and BCR loci[101]. In **paper I**, we used IgDiscover to determine V, D and J germline composition of alpha, beta, gamma, and delta loci from 45 individuals from Africa, Europe, East Asian and South Asia (Figure 9). Additionally, we included publicly available TCR libraries from five monozygotic twin pairs in the analysis.

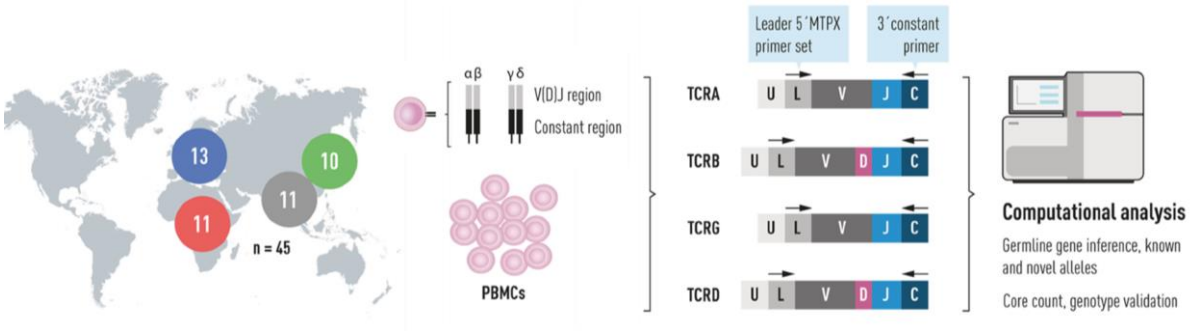


Figure 9. ***Study design overview.*** *We obtained blood samples from participants from different populations. mRNA from PBMCs was used to produce TCR chains-specific libraries. The libraries were analyzed with IgDiscover for personalized germline gene inference.*

Generation of personalized databases allowed the identification of a total of 175 novel V and J allelic variants: 76 TRAV (of which 8 shared with the delta locus), 65 TRBV, 15 TRGV, one TRDV, 17 TRAJ and 4 TRBJ. All the novel J alleles and 63.4% of novel V alleles consisted in variants with at least one codon change.

Comparison of individuals genotypes at both genetic and allelic levels revealed an extremely diverse repertoire landscape. We observed higher level of heterozygosity in the V genes of the alpha locus (Figure 10A) compared to the beta locus (Figure 10B) and instances of structural variations in form of homozygous or hemizygous loss of expression of several genes. Loss of expression can be explained by either complete absence of the genomic sequence or by presence of mutations that ablate expression. In our dataset, we confirmed genomic absence of TRGV4 and TRGV5 in one of the donors as well as presence of variations resulting in stop codons or splicing site removal in others.

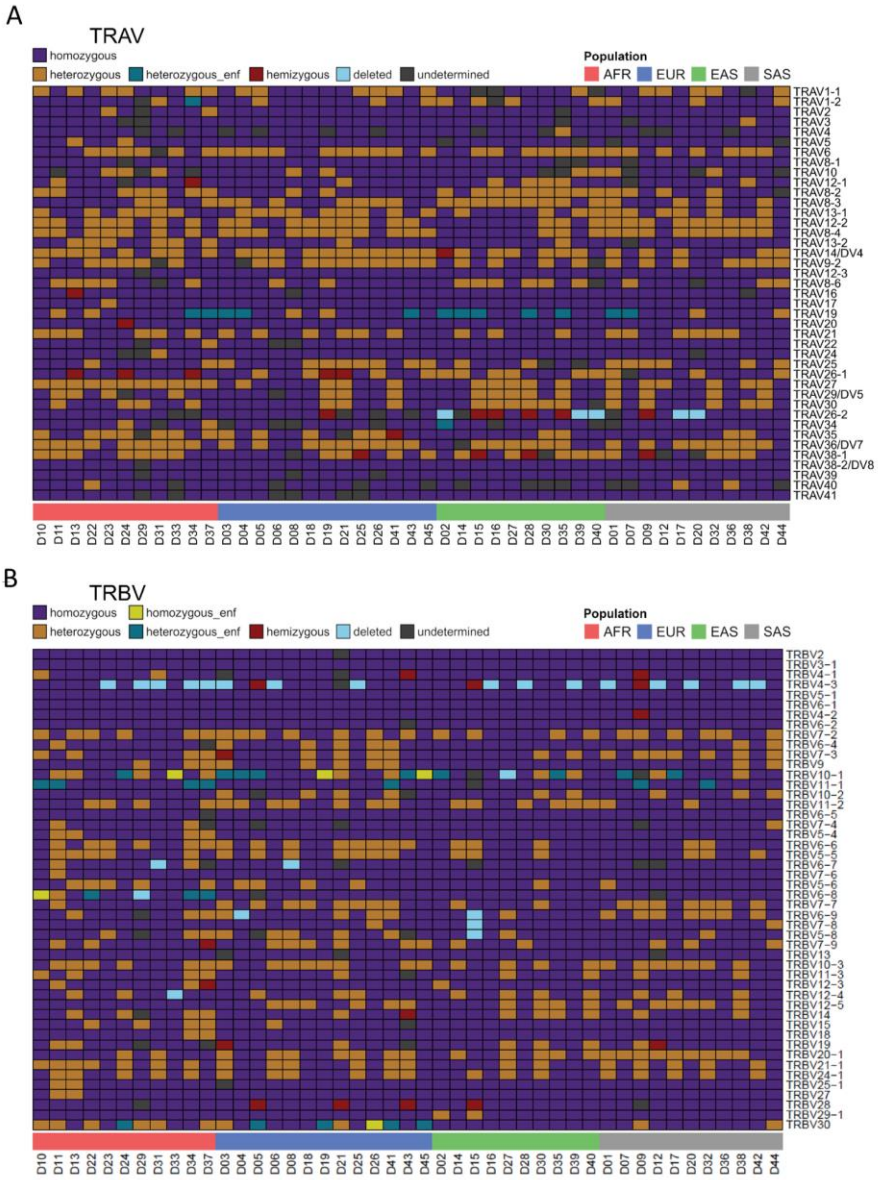


Figure 10. ***Individuals TRAV and TRBV gene content.*** *A) Individuals TRAV and B) TRBV genes variation. Donors ID arranged on the x axis are grouped by population. V genes on the y axis are arranged by chromosomal order.*

Loss of expression of some genes might influence T cell subsets with semi-invariant TCRs. For instance, we reported the example of TRAJ12, where loss of expression was homozygous for two individuals and hemizygous for other two individuals. TRAJ12 usage determines one of the three main subsets MAIT cells, the TRAV1-2/TRAJ12 subset, identifiable by CDR3 sequence CAXXDSSYKLIF[178]. The two individuals with homozygous loss of TRAJ12 did not show any presence of TRAV1-2/TRAJ12 CDR3s while the other MAIT CDR3s subsets were identifiable. Interestingly, one of the individuals with hemizygous loss of TRAJ12 showed absence of TRAV1-2/TRAJ12 CDR3s. Haplotype analysis revealed that its functional TRAJ12 was placed on the same haplotype of TRAV1-2*03_S6094, a non-functional allele (Figure 10A, donor D34). Hence, the TRAV1-

2/TRAJ12 composition on each chromosome lacked the presence of functional genes to generate TRAV1-2/TRAJ12 CDR3s.

Despite the finding that most novel alleles were found in all the population included in the study, each locus presented cases with remarkable different allele usage frequencies, where certain population, particularly apparent in the African donors, carried some exclusive alleles. These observations are consistent with previous reports on sub-Saharan Africans genomic diversity[179]. Our observation was further validated by comparison of SNPs data from the 1000 Genome Project (1KGP). Overall, several alleles were present in multiple individuals from the African population group but absent in others including one TRAJ, 11 TRAV, 17 TRBV and two TRGV alleles. Furthermore, one TRAV allele and three TRBV alleles were exclusive to Europeans, two TRAV and two TRBV alleles to East Asians and one TRAV and one TRBV allele exclusive to South Asians.

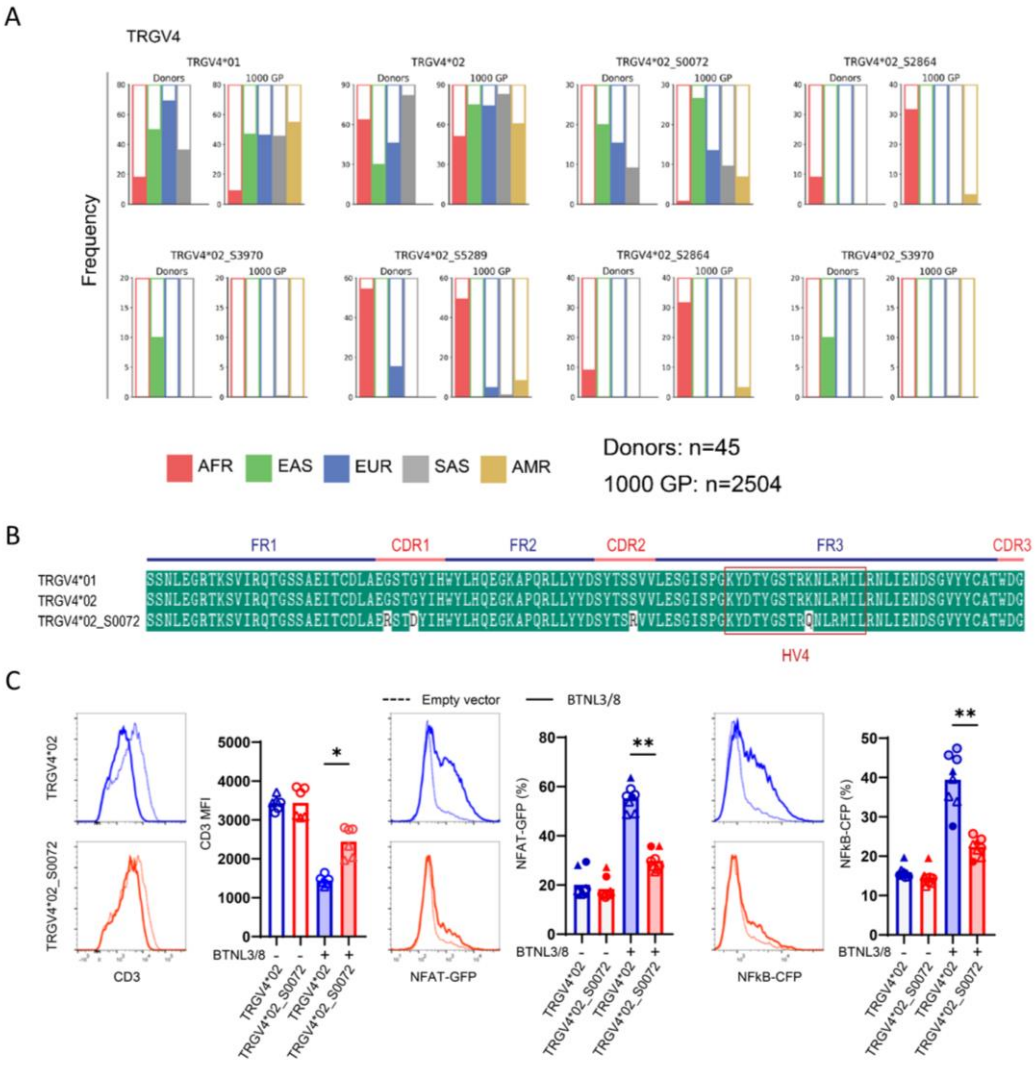


Figure 11. ***TRGV4 alleles genomic and functional analysis.*** *A) Population-based allele frequencies. For each allele, donor usage frequencies (left) are compared to frequencies from 1KGP (right). B) Amino acid sequence alignment of TRGV4*01, TRGV4*02, and TRGV4*02_S0072. C) BTNL3 reactivity analysis of TRGV4*02 vs TRGV*02_S0072. From left to right, regulation of surface CDR3, NFAT, and NFκB. Two independent experiments were performed with each reporter cell line, one with a single replicate (filled symbols) and one with 3 technical replicates (empty symbols). *<0.05 and ** < 0.005 by Wilcoxon-signed rank test.*

We also reported the opposite trend, several alleles were common but absent in specific populations. The most prominent example is represented by TRGV4 alleles. TRGV4 showed a surprising level of allelic variation with TRGV4*02_S2864, TRGV4*02_S5289 and TRGV4*02_S9926 with African specificity while TRGV4*02_S0072 was frequent in the European, South Asian, and East Asian groups (Figure 11A).

We then compared our findings with TCR gene sequences from four high coverage archaic reference assemblies of three Neanderthals, namely Vindija, Chagarskya and Altai, and the single Altai Denisovan. Despite we observed that most of the TCR alleles are shared between modern and archaic humans, we identified three regions of introgression. These regions encompass four alleles shared between non-Africans and the Vindija Neanderthal: TRAV12-2*02_S6060, TRAJ24*02_S1049, TRAJ26*01_S5236, and TRGV4*02_S0072.

The introgressed TRGV4*02_S0072 allele showed similarity to TRGV4 alleles of great apes and contained 8 nucleotide differences compared to the closest known human TRGV4 allele, TRGV4*02, resulting in four amino acid changes in the CDR1, CDR2 and HV4 regions of the V gene (Figure 11B). We tested if this allele results in functional changes in γδTCRs. Previous reports described the interaction between γδTCRs and BTNL3/BTNL8 via HV4 and CDR2 loops of TRGV4 binding to BTNL3[47,180,181]. In co-culture experiments with BTNL3/BTNL8-expressing target cells we showed that TRGV4*02_S0072-using TCR exhibited markedly reduced TCR downregulation as well as reduced NFAT-GFP or NFκB-CFP reporter activity (Figure 11C).

In conclusion, the data reported in **paper I** revealed previously unreported novel alleles and high levels of diversity in the TCR loci. Novel alleles with coding changes variation may impact on function. Comparative analysis with available primate assemblies reveals similar degrees of diversity, especially in the CDRs sequences, and presence of variants with stop codons.

AIRR analysis is important to study the immune systems dynamics in health and diseases. The TCR repertoire may determine how the response against certain antigens is modulated[182–186]. It was reported that the T cell response to the M1 epitope of influenza was limited and utilized a public TCR that involved the use of TRBV19 and TRAV38/TRAJ52[59]. TRAV26-2 was shown to be the most commonly used TRAV gene in response to an HLA-A2 restricted cytomegalovirus epitope[187]. TRAJ12, TRAJ20 and TRAJ33 genes are frequently found rearranged to TRAV1-2 MAIT cells semi-invariant TCRs[178,188,189]. MAIT cells have been shown to differentially respond to bacterial and viral challenge[190,191]. Collectively, our data reports several instances of allelic variants with loss of function on these genes.

Many novel alleles show variation in the CDR3 region, particularly polymorphisms in the 5' end of the J gene. Reports indicate that germline-encoded variations affecting CDRs may

shape the T cell repertoire[185,192,193]. Other studies indicated differential binding to peptide-MHC modulated by single allelic variations[56].

Concrete evidence of germline function modulation is represented by TRGV4*02_S0072, identified in this study, which alters cell activation by BTNL3/BTNL8 interaction. BTNL3/BTNL8 is an important regulator of intestinal intraepithelial Vg4^{+} T cells, which have immunomodulatory effects during inflammation and cancer[181,194,195].

The results of **paper I** indicate extensive genetic diversity in expressed TCR genes at the population level. The origin of this diversity likely involves demographic evolutionary events such as mutation, natural selection, population bottlenecks and gene flow between modern populations and, historically, between modern and archaic human populations[196]. This is demonstrated by identification of Neanderthal introgressed regions and comparison of repertoires from NHP assemblies. The database of human TCR germline alleles produced here is intended to assist future research in identifying important aspects of TCR biology that may have previously been overlooked.

4.2 Paper II - Adjuvanted spike protein immunization elicits potent and durable neutralizing humoral immune response

Shortly after the outbreak of SARS-CoV-2 and pandemic was declared, it was clear that rapid and effective prophylactic measures were needed, including protective vaccines.

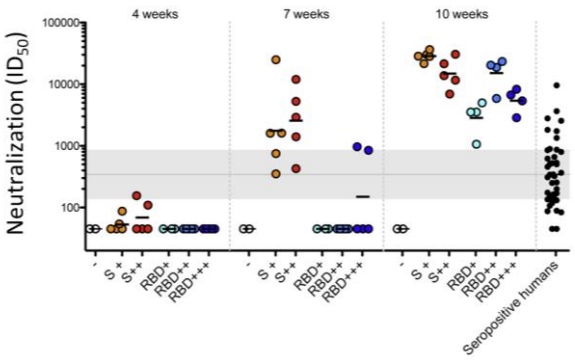


Figure 12. ***Immunized mice serum neutralization.*** *Serum ID_{50} titers from S immunized, RBD immunized, and unimmunized mice (empty dots). S+ (5 mg S, orange dot); S++ (25 mg S, red dots); RBD+ (5 mg RBD, cyan dots); RBD++ (25 mg RBD, blue dots); RBD+++ (50 mg RBD, navy dots). ID_{50} titers from seropositive donors (black dots) obtained from Castro Dopico X. et al.*[197]*. Shaded grey band corresponds to the median and interquartile range of the neutralization potency of aforementioned donors.*

In **paper II**, we sought to investigate the humoral immune response elicited by immunization regimens consisting of S subunits in mice and NHPs. We first tested in our mouse model the effect of immunization regimens composed of either adjuvated pre-fusion stabilized S or RBD. A total of 24 mice were injected with three immunizations at week 0, 4 and 7 of either 5 µg S, 25 µg S, 5 µg RBD, 25 µg RBD or 50 µg RBD. In S immunized mice, neutralizing antibody responses were already detectable after priming, with escalating potency after each boost. In contrast, RBD immunized mice showed detectable neutralization only after two immunizations with 50 µg RBD, while a third immunization (second boost) elicited potent neutralizing antibodies regardless of the RBD dose (Figure 12).

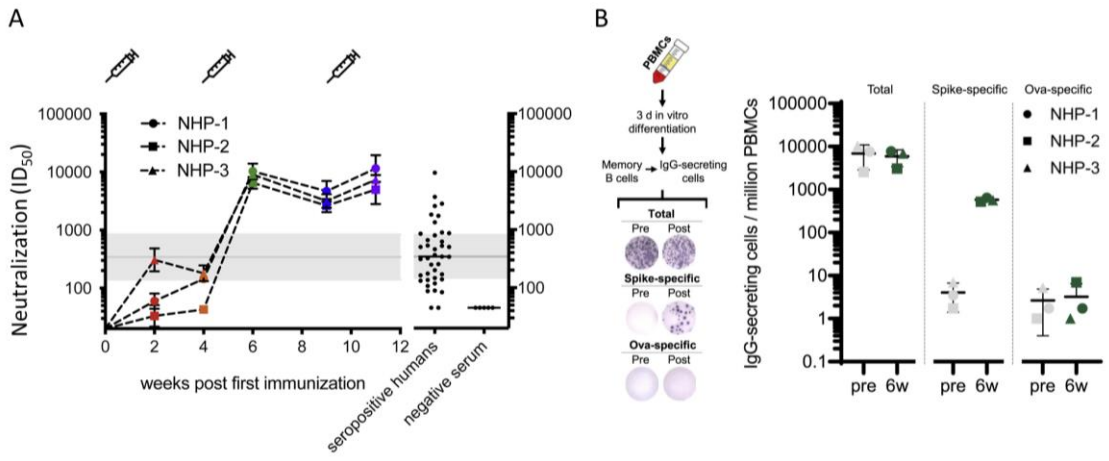


Figure 13. ***NHP plasma neutralization and S-specific memory B cell formation.*** *A) Plasma ID_{50} titers from S immunized NHPs. Seropositive donors depicted as in figure 12. B) B cell ELISPOT analysis of in vitro differentiated memory B cells 2 weeks after the second immunization.*

Consequently, we sought to investigate the response in three Rhesus macaques immunized with 100 µg pre-fusion stabilized S mixed in 75 µg of Matrix-M. The immunization regimen consisted of three immunizations at weeks 0, 4 and 9. Detectable antigen-specific IgG and neutralizing responses were measured already 2 weeks after priming (Figure 13A). Remarkably, a second dose boosted the response to extremely high neutralization titers. Additional boosting did not further raise peak neutralizing titers. Two immunizations of adjuvated S were sufficient to establish S-specific memory B cells (Figure 13B).

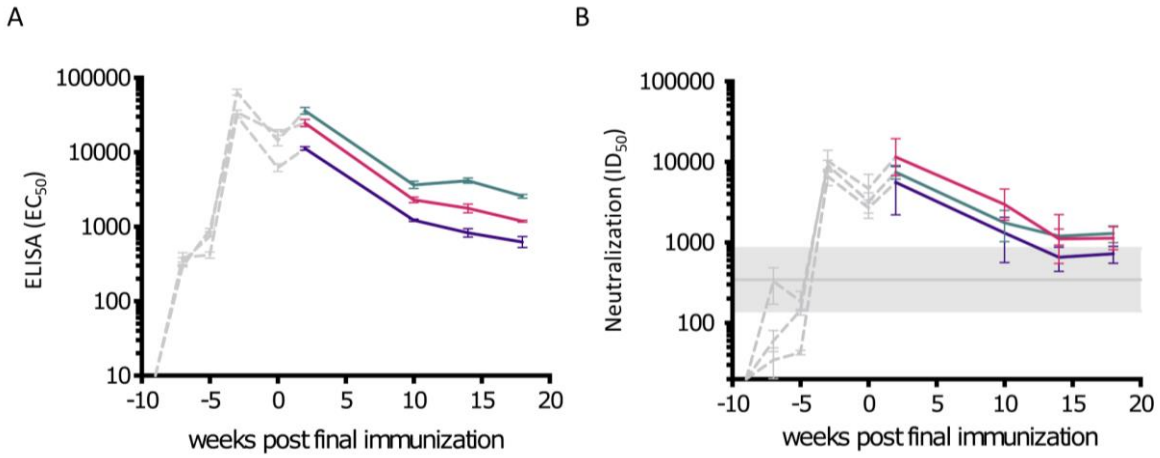


Figure 14 Antibody response durability. A) Binding antibody titers and B) neutralization antibody titers in NHPs over a following period of 4 months following the last immunization.

At the time we had these results (Jun 2020 – Sep 2020), studies of natural infection and pre-clinical vaccine studies reported that neutralizing antibodies were elicited in both settings[167,198–203]. However, data on long lasting durability of elicited response were lacking. To address this, we regularly sampled the macaques over a 4-month period after the last immunization. Binding titers (Figure 14A) and neutralization titers (Figure 14B) decreased significantly in the first 8 weeks, with an estimate of 10-fold and 4-fold decrease from peak response, respectively. Over the following weeks, waning response stabilized, slowly decaying to a slightly less 2-fold decrease, indicative of a lost-lasting response. However, additional time points taken considerably later would be needed to draw firm conclusion abouts the durability of the response following vaccination. Furthermore, we only followed antibody titers without performing challenge experiments. Consequently, we did not have any readout on antibody related protective effects. From studies performed by others, we know that adjuvated protein sub-unit immunizations generate detectable antibody neutralizing titers in upper and lower airways[204–206]. Intranasal and intratracheal challenge revealed reduced viral replication in this compartment with little signs of inflammation. Furthermore, passive transfer experiments show grant protection in mouse models. This is reflected in other studies performed with other vaccine platforms and passive transfer experiments[167,171,199,201,207–210].

Collectively, the data reported in **paper II** provided promising perspective for protein-based vaccines. Elicitation of potent long-lasting neutralization titers and detection of S-specific memory B cells are optimal outcomes of immunization regimens.

4.3 Paper III - Heterotypic RBD boost elicits cross-neutralizing protective antibody responses

The race for vaccine developments against SARS-CoV-2 registered record speed, with several candidates reaching phase III clinical trials. However, the global pandemic landscape was changing as several variants appeared with increased frequency in different parts of the world. Variants with immune escape mutations were of particular concern as they could undermine acquired immunity[211–213].

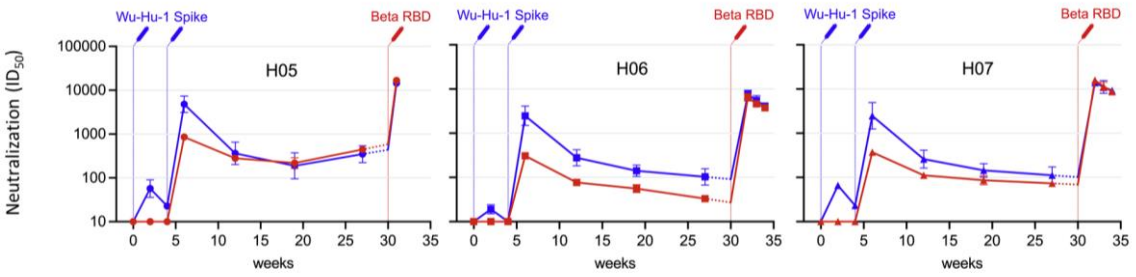
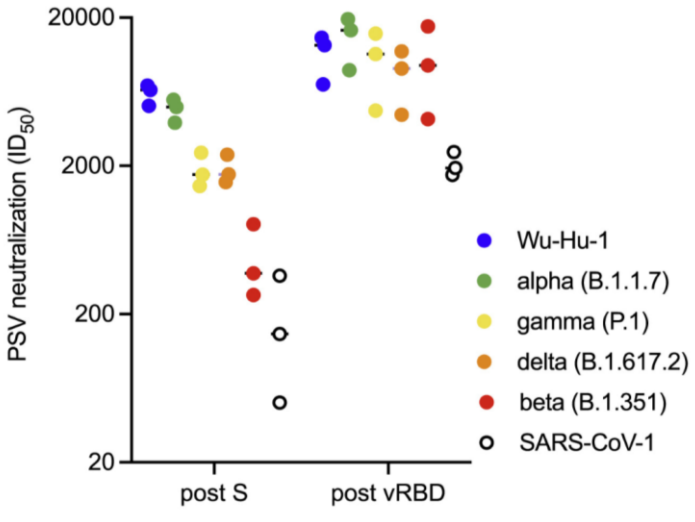


Figure 15. ***Heterotypic immunization regimen.*** *NHP individual plasma ID_{50} throughout the whole immunization regimen. Ancestral strain neutralization in blue, beta strain neutralization in red. Immunization and antigens indicated with syringes.*

In **Paper III**, we investigated the effect of heterotypic immunization regimens using a boosting with the RBD from the beta variant. Three macaques were immunized at weeks 0 and 4 with 2 µg pre-fusion stabilized S protein derived from ancestral strain. Neutralizing titers were already detectable after a single dose against ancestral strain, but not the beta variant (Figure 15). Response after boosting was considerably more potent and detectable also against beta. However, beta neutralization was on average 9-fold less potent than neutralization of the ancestral stain. Consistent with the findings of **paper II**, neutralizing titers showed a significant decrease in the first weeks following a second immunization with subsequent waning stabilization.

Figure 16. ***SARS-CoV-2 variants cross-neutralization.*** *Post-2 vs post-vRBD cross-neutralization of SARS-CoV-2 variants (color coded dots: Wu-Hu-1 (blue), beta (red), alpha (green), gamma (yellow), and delta (orange). SARS-CoV-1 was included as well (empty dots). Each animal is represented with a dot.*



Six months after first immunization, macaques were boosted with a beta RBD at different doses: 2 µg for H05, 10 µg for H06 and 50 µg for H07. We then followed up the response in the next 4 weeks. H05 had to be terminated earlier due to an unrelated illness. Potent cross-neutralizing response was readily detected in all macaques, suggesting recruitment of cross-neutralizing memory B cells elicited by former immunizations (Figure 15). We then investigated the effect of the heterotypic boost in eliciting cross-neutralization to unrelated variants (Figure 16). Lateral comparison of plasma from post second S boost (post-S) and post beta-RBD boost (post-vRBD) revealed superior neutralization of alpha, gamma, and delta variants, as well as SARS-CoV-1. Plasma cross-neutralization was investigated also for the macaques involved in **paper II**, whose regimen costing of three ancestral S protein immunization. Reduced variants immunization persisted throughout the whole immunization regimen, indicating cross-reactive B cells were recalled only by the heterotopic boost.

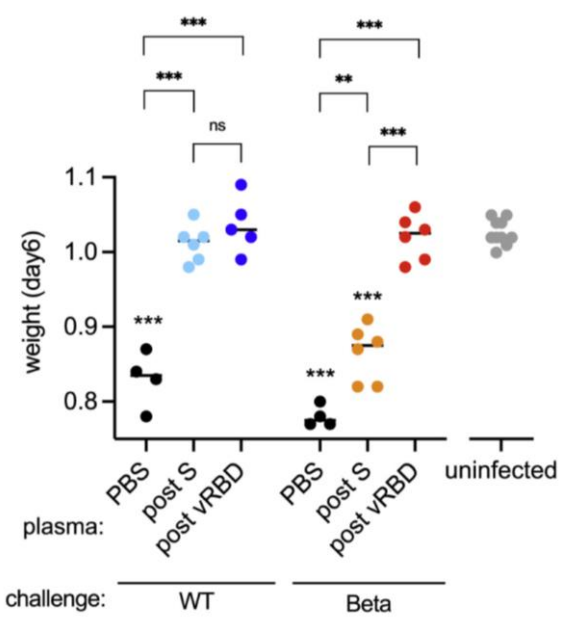


Figure 17. ***Passive infusion of plasma obtained after heterotopic boost confers protection in mouse model.*** *Mice weight loss at day 6 after challenge. Unchallenged mice are shown in gray. Groups with significant weight loss compared to uninfected mice are annotated directly above the dots of that group. Pairwise group comparison displayed above square brackets. **p < 0.01 and ***p < 0.001; ns, not significant.*

We investigated the prophylactic properties of cross-reactive plasma in our K18-hACE2 mouse model (Figure 17). Briefly, mice were passively infused with post-S plasma or post-vRBD plasma before being challenged with lethal dose of either ancestral or beta virus. The mice were monitored for weight loss, a proxy of disease severity. Control mice were injected only with PBS. While both mice infused with post-S plasma and post-vRBD were completely protected against weight loss after ancestral challenge, only post-vRBD infused mice were protected against beta challenge.

Reduced efficacy of acquired immunity against beta and other variants was reported in infected and vaccinated individuals[214–221]. While there was concern that original antigenic sin, as observed for influenza A virus[222–224], could compromise immunity against infections with other variants, our data demonstrated that heterotypic boosting elicited potent cross-neutralizing responses in our setting. This effect was observed even at low

immunogen dosage and is consistent with similar observation from other groups[225–228]. Despite the finding from our mouse experiments in **paper II**, suggesting that RBD is a poor priming immunogen, boosting with RBD resulted in potent neutralization titers making it a promising candidate for vaccine recall formats.

4.4 Paper IV – tracing of immunization elicited B cell linages elucidate dissemination to different compartments and identify broadly neutralizing antibodies

In **paper IV**, we sought to investigate more in detail the elicited response in two of the macaques reported in **paper II**. Here, we identified B cell lineages elicited by ancestral S protein immunizations and followed their evolution throughout the immunization regimen by investigating different immune related compartments.

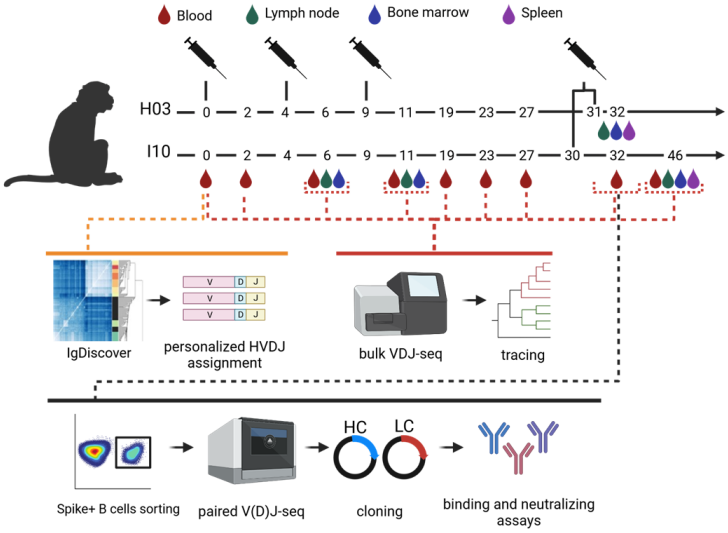


Figure 18. ***Study design.*** *Two NHPs, H03 and I10, were inoculated four times with a pre-fusion stabilized S adjuvated with Matrix-M, indicated by syringes. Samplings of blood, LNs, BM and spleen are indicated with a red, green, blue, and purple drop, respectively. The samples obtained were used to produce bulk IgG heavy chain libraries. Samples from week 0 were used to determine the individual NHP germline heavy chain database with IgDiscover. Samples from week 32 were used to sort B cells to undergo 10X single cell V(D)J sequencing.*

Macaques H03 and I10 underwent regular immunization and sampling phases, summarized in figure 14A, of blood, BM and LNs. An extra boost was performed at week 30 and 31 to I10 and H03., followed by termination of H03 and blood sampling of I10 at week 32 (Figure 18). We then used a C-terminally biotinylated prefusion-stabilized S trimer to sort B cells from I10 blood and four different compartments collected at week 32: two distinct iLN-R (iLN-R1 and iLN-R2), spleen, and pooled non-draining LNs (mesLN, medLN, axLN, and perLN, all together as LNOther). Sorted cells single cell 5′ V(D)J sequencing using Chromium (10X Genomics) modifying the 10X V(D)J enrichment steps by spiking in NHP-specific constant-region primers. The resulting paired heavy and light chain sequences were used as basis for identifying S-specific lineages on bulk IgG libraries prepared from samples obtained through the whole vaccination regimen. To improve lineage identification, we generated individualized germline databases of heavy chain V and J alleles for each animal and combined it with light chain alleles available from IMGT.

A total of 199 and 41 clonal lineages were identified for H03 and I10, respectively (Figure 19). Lineages were traced in blood, LNs, and BM samples collected at different time points, as well as from spleen samples obtained at termination. In general, lineages that could be

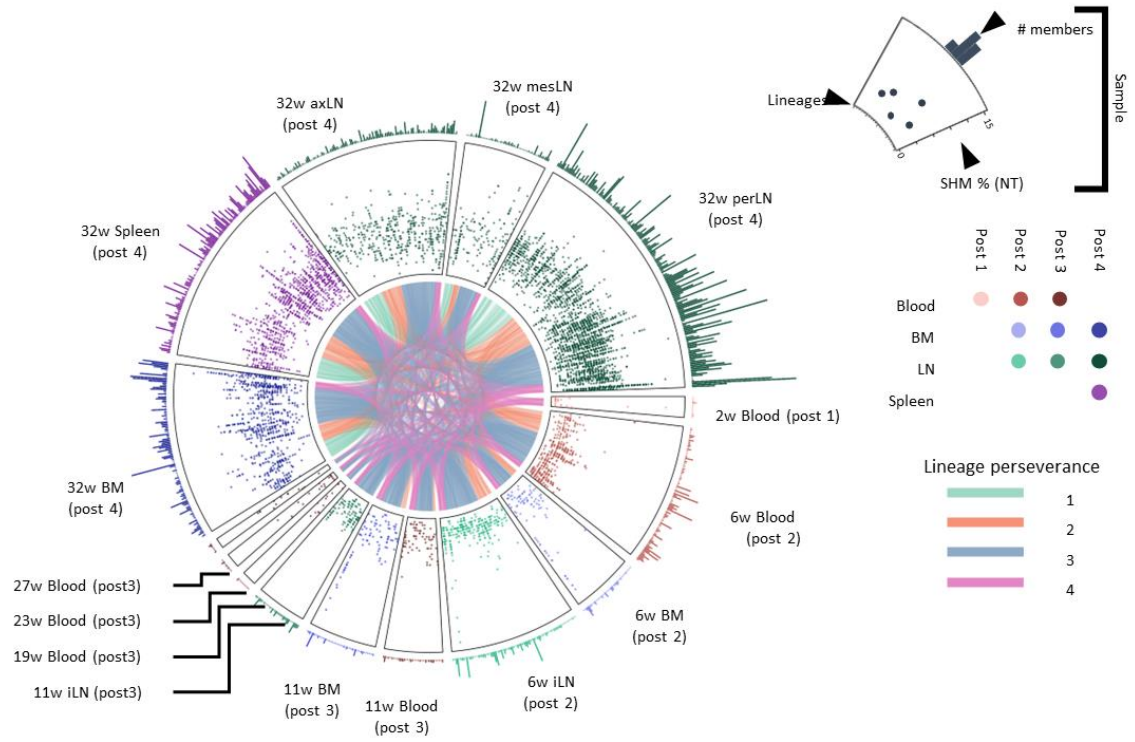


Figure 19. ***H03 lineage tracing overview.*** *Each sector represents a sample. Starting from degree 0 the samples are ordered, clockwise, by sample time. The base of each sector represents a curved x axis with each lineage as a point. Each sector is divided into three levels. The outer level shows the number of members of that lineage, the mid-level shows each member's SHM, and the inner common level connects lineages that were present in multiple samples. The inner level is color-coded based on lineage persistence. Lineage persistence indicates the number of appearances after each immunization (post 1, post 2, post 3 or post 4) on the traced data. The other sectors are color-coded with red, blue, green, and purple for blood, BM, LN, and spleen, respectively, with darker shades for increasing weeks.*

traced at early time points corresponded to larger clonal disseminations and showed greater persistence throughout the immunization regimen. For instance, most of the lineages detected in blood after priming (week 2) were identified in all 96 samples obtained thereafter. H03 lineage dataset proved to be richer than I10, which prompted us to pursue the study of its lineages more in details. Interestingly, perLN and spleen were characterized by elevated count of lineage members, especially if compared to axLN and mesLN obtained at the same time, denoting preferential recruitment to those compartments. Lineages were regularly detected in BM as early as week 6 and showed moderately elevated counts of members at week 32. Only a limited number of lineages could be traced in blood samples obtained at weeks 19, 23, and 27, indicating contraction of detectable circulating lineages. Overall, the lineages steadily accumulated SHM during the course of the experiment, indicating persisting GCs, or possibly re-entry of memory B cells to GCs. However, lineages with few unmutated members can still be detected at late time points, suggesting recruitment of cells from the naïve pool of B cells with advantageous germline BCRs following the later immunizations.

IGHV gene usage was broad, as seen in most polyclonal responses with frequent usage of genes that are highly used also in the total IgG repertoire[103,229–231]. However, some S-specific lineages used genes representing the mid/low total IgG usage frequency range. For instance, H03 showed enrichment of IGHV1-NL_2*01_S2809 while I10 showed preferential usage of IGHV1-138*01_S6980. The human orthologues of these alleles are IGHV1-2*02 (91.9% homology) and IGHV1-69*10 (95.3% homology), respectively. These genes are often used in human SARS-CoV-2 S responses[232–237], suggesting the existence of conserved binding modes between some NHPs and human S-specific antibodies.

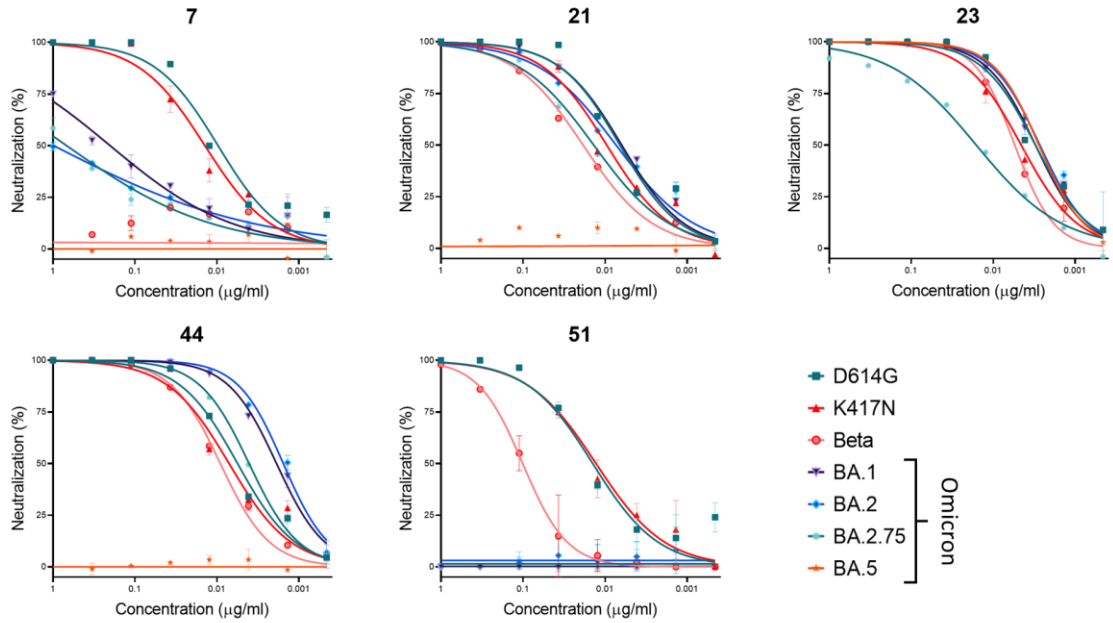


Figure 20. ***Isolated bnAbs from H03****. Neutralization curves from pseudovirus assay of top 5 bnAbs.*

We next selected a set of heavy and light chain pairs from H03 for synthesis of mAbs. The selection aimed to cover different IGHV genes, SHM levels, lineage sizes and compartments of origin. Of 41 S-binding mAbs, 13 targeted the FRBD. Ten of the RBD-binders displayed neutralizing activity, measured by pseudovirus neutralization assay. Potency varied between the neutralizing mAbs, with IC_{50} ranging from 0.004 µg/ml to 2.502 µg/ml against the ancestral variant. Mabs showed different degrees of cross-reactivity against a panel of omicron variants, beta, and K 417N strain (Figure 20). Mab 23 exhibited the most potent and broadest neutralizing activity against all the variants of the panel and was selected for further characterization. We further evaluated the neutralization breadth of mAb 23 against circulating Omicron subvariants BA.4.6 and BA.2.75.2, known to resist neutralization from different reported nAbs. Mab 23 showed detectable neutralization albeit with reduced potency against both variants.

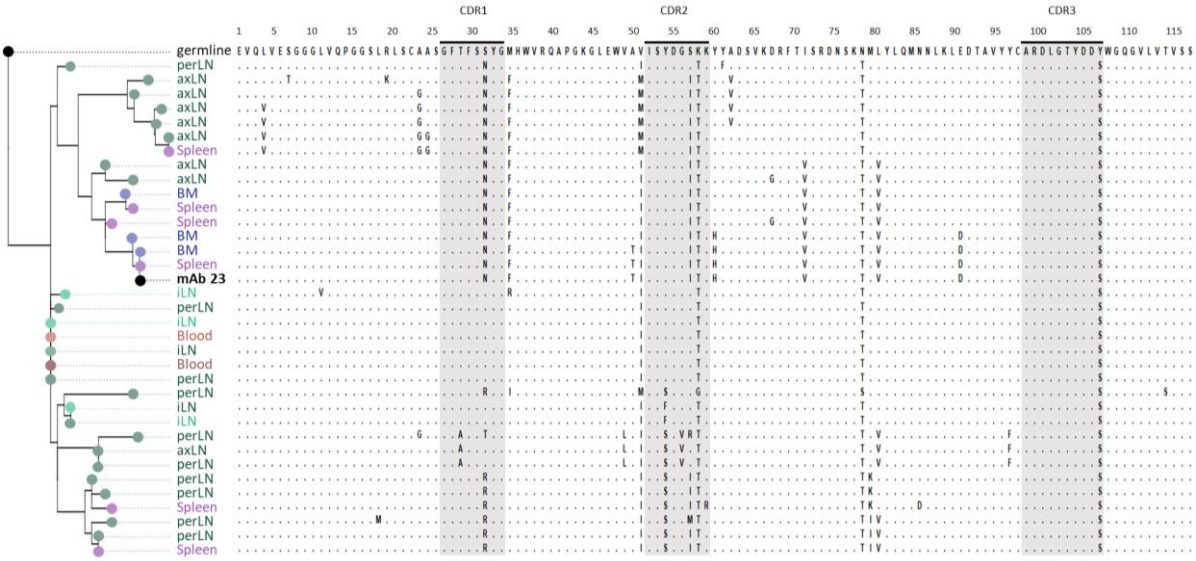


Figure 21. ***MAb 23 lineage.*** *Phylogenetic analysis of mAb 23 lineage. The phylogenetic tree is computed with a maximum likelihood approach. The tree is rooted to an inferred germline. Amino acid alignment of each lineage member is displayed on the right with highlighted CDRs.*

Phylogenetic analysis of mAb 23 revealed dissemination to several compartments with equal heavy chain VDJ sequences found in BM and spleen (Figure 21). MAb 23 lineage did not show any variance in the HCDR3 amino acid sequences of any traced sequence. This alone suggested that HCDR3 is likely crucial for RBD recognition and selection of the lineage members. This was further corroborated by a cryo-EM structure of mAb 23 Fab bound to the ancestral S trimer (Figure 22). The structure revealed that binding was primarily mediated by HCDR3 residues D99, Y103 and D104 with additional contribution from HCDR2 residue Y53 and adjacent light chain residues Y49 and A50. Mab 23 epitope recognition was similar to other broadly neutralizers, such as Bebtelovimab (LY-CoV1404) and 2-7. All these mAbs interact with residues overlapping with the receptor binding motif (RBM) and can bind with the RBD in both "up" and "down" conformations but do so with uncanonical modalities compared to described class II binders. Most of the residues recognized by mAb 23 are reported to be conserved with two notable exceptions: N440K variations features all the omicron subvariants but doesn't seem to impact cross-neutralization; R346T instead is mutated in BA.4.6 and BA.2.75.2 strains and justify the reduced potency against these strains.

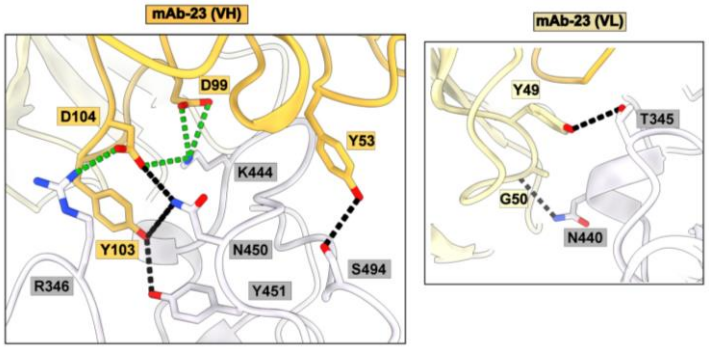


Figure 22. ***Structural analysis of mAb23–ancestral S interaction.*** *Cryo-EM analysis of mAb23 Fab interaction with S. Zoomed representation of heavy chain (left) and light chain (right) interaction against RBD. mAb 23 residues are depicted in yellow, RBD residues in white.*

In summary, the objective of **paper IV** was to detect and characterize immunization elicited B cell lineages, including those mediating neutralizing activity. Lineages were consistently traced in spleen and non-draining LNs such as axLNs, mesLNs and perLNs, indicating widespread dissemination through the circulatory and lymphatic network. Remarkably, the vast majority of the lineages were traceable in the perLNs. Our observation corroborates other studies conducted in NHPs showing an interconnection between inguinal and periaortic lymph nodes[238,239]. Lineages were consistently traced in BM samples, a compartment dedicated to long-term immunological memory. Despite the fact that the study did not identify B cell phenotypes, tracing of elicited lineages in this compartment provides promising data that might prompt future dedicated studies.

Identification of macaque orthologous matching human genes reported to be used with high frequencies in the SARS-CoV-2 response support the use of NHP as a valid pre-clinical model.

Characterization of the interaction between mAb 23 and S and comparison with other available structures highlighted the presence of conserved epitopes and interesting neutralization profiles, which can be used to guide future vaccine efforts.

5 Conclusions

The work presented in this thesis focused on characterizing the two arms of the adaptive immune system: T cells and B cells. My work revolved around the study of adaptive immunity repertoires and focused on both their genomic and functional features. Collectively, the findings presented in this thesis provide a comprehensive characterization of adaptive immune receptors, updating our current knowledge on population repertoires and vaccine elicited response.

The key findings from the studies included in this thesis can be summarized with the following statements.

From **paper I**:

- The collective TCR germline repertoire is more diverse than we knew previously appreciated.
- Several TCR germline alleles are present at various frequencies in different opulations.
- TCR germline encoded variants can modulate the presence of defined T cell subsets.
- The TCR loci present regions of introgressions inherithed from Neanderthal ancestors.

From **paper II and III**:

- SARS-CoV-2 sub-unit protein immunization elicit potent and durable humoral immune response.
- SARS-CoV-2 sub-unit protein immunization stimulate production of memory B cells.
- Heterotypic boosting stimulate a cross-neutralizing humoral immune response.

From **paper IV**:

- SARS-CoV-2 S immunization-elicited B cell lineages distribute in multiple immune compartments, including non-draining LNs, BM and spleen.
- B cell lineages elicited in NHPs share IGHV genes usage with orthologues of human IGHV genes.
- SARS-CoV-2 ancestral S immunization elicits broadly neutralizing antibodies.
- mAb 23 neutralization breadth is determined by interaction with convserved residues on the RBM.

Overall, this thesis emphasizes the power of AIRR-Seq analysis and its capacity to provide detailed information that helps us understand adaptive immunity.

6 Points of perspective

In this thesis, I focused on the study of AIRR of both TCRs and BCRs in perspective of genomic characterization and study immunization elicited responses.

The findings of **paper I** produced an extensive human germline TCR database that can be used as reference for future studies. These studies highlight the extent of variation in human TCR germline genes, which can only be appreciated when performing personalized genotyping as shown here. The information obtained will aid in future studies of immunological mechanisms and will pave the way for disease association studies, which up until now have not included the TCR genes. It is important to highlight that the generation of this type of data was possible only by including donors from different population groups and additional diverse populations can be included in future studies. We show how germline encoded variants can impact certain T cells populations, such as MAIT cells and γδ T cells. Finally, we reported three regions of genomic introgression from Homo neanderthalensis. The genes we encode today are the result of millions of years and adaptation to different environments, not the least the loci involved in immunity. Studies of gene flow between modern man and previous hominids is of paramount importance to understand human evolution.

The work performed for **paper II, III and IV** revolved around the study of B cell responses after immunizations with SARS-CoV-2 S protein subunits. At the time this thesis work was performed, the pandemic underwent different waves of SARS-CoV-2 variants, with Omicron subvariants currently dominating. Vaccines have been rolled out globally, with individuals receiving up to 5 doses. Despite the fact that the findings of **paper II and III** can be considered outdated given the pace of research in the SARS-CoV-2 field, they were new at the time they were published, and data provided can guide future protein-based vaccine regimen design. The global landscape of different viruses is unstable and subjected to changes over time. Heterotypic vaccine regimen design might be a future standard for some diseases. In **paper IV**, we followed up on selected macaques from **paper II** to characterize their B cell lineages at great depth. I believe this type of approach is still in its infancy but will become more conventional with further technological development. We found that each lineage tells a different story underscoring the complexity of our response to vaccination. All together, these stories assemble into a picture with extremely high resolution that we probably cannot completely appreciate yet. B cell lineage evolution studies can become extremely powerful in different contexts. For instance, coupling lineage evolution and mapping of interesting epitopes can guide immunogen design. Inclusion of different anatomical compartment can elucidate lineage trafficking and clonal relationship. Or more, inclusion of B cell phenotypes can improve cell population characterization.

In conclusion, the data reported on this thesis laid the foundation for several other potential studies. The possibilities of detecting detailed differences in repertoires provide for exciting and important new frontiers of investigation. I hope that myself and others will exploit these opportunities to make new discoveries and gain an improved understanding about the adaptive immune system.

7 Acknowledgements

I am grateful for the opportunity to have spent my time at Karolinska Institute working on this thesis and the people I have met along the way. Without their presence, the experience would not have been as fun and enjoyable as it was.

The first person I need to thank is my main supervisor, **Gunilla Karlsson Hedestam**. Your passion for science was obvious already from the first meeting we had and never witnessed declined since I joined the lab. If anything, the opposite. Your interest and dedication are ever expanding, and you always manage to be pragmatic and efficient. If I was able to dedicate so much to this field is also thanks to the way you live science. Joining your lab is a badge of honor which I might have bragged about too often.

I'm extremely grateful to have **Ben Murrell** as my co-supervisor. I'm fascinated by how you manage to find interesting messages in every little speckle of data. Your enthusiasm for coding is contagious and I hope to reach half of your proficiency with that. I might have boasted sometimes to have the most engaged co-supervisor of the whole institute.

A special thanks also to **Martin Corcoran**. You are indeed a secret co-supervisor and always provide pertinent knowledge every time you're asked. I'm amazed by the number of ideas you devise. Your humor is exhilarating, and I cannot imagine what the lab would feel without you. My only regret is not having a boat to invite you for a fishing trip.

I have to acknowledge and thank past and present members of the **GKH lab**. It was a pleasure to share the journey with all of you. I like to think that I've learned a little bit from each one of you. **Néstor**, I miss your objectivity in discussing science. Thanks for your enthusiasm on training and for introducing me to obstacle running and python. It was fun sharing moments in and outside the lab with you. I'll try to adopt your passion for scuba dividing in my future trips. **Sharesta**, thanks for always being present and being one the best workout buddy I've ever had. I'll try to match the perseverance you showed me when you ran a marathon. It was fun partying around with you. **Ganesh**, my former desk-mate, thanks for paving the way to lineage tracing analysis. I was truly inspired by your work. **Uta**, thanks for always being up to a scientific discussion, your background always results to new points of view. **Mateusz**, thanks for the patience in explaining bioinformatics and computer science, your arrival was a blessing for the group. **Masha**, thanks for always bringing sweets and teaching Russian anecdotes, your help with the 10X data was truly appreciated. **Mark**, thanks for your immense help with python, R and Julia. I hope I'll be able to code like you. Thanks for also sharing the late hours and the ping pong matches. **Xaquin**, my desk-mate, thanks for always sharing your veterinary knowledge and showing interest in every scientific topic we end up discussing. It's always fun hearing about the adventure you had around the world. Thanks for the invitation to your place and sharing Rudy with us. **Sanjana**, thanks for being the nice person you are. You are unmatched in

your kindness and not reluctant to share it. Thanks for taking care of my plants and sharing your Indian cuisine with me. Thanks a lot for the help in the lab. Best wishes with your PhD, You're one of the next. **Sungyong**, thanks for your help with the Fab and always sharing jokes in the lab. Thanks for introducing me to Korean cuisine, I'll try my best to experience it all. **Monika**, thank you for sharing your expertise on mice and B cells and helping in many experiments. I'll never reach the love you have for FACS and for that I wish I'd always have you during my scientific career. You make life in the lab a very pleasant and enjoyable experience. **Pradee**, thanks for your everlasting laughter, for introducing me to the NHPs and for your help in many experiments. Thanks for expanding my knowledge on Indian cuisine and Indian music together with Sanjana. Best of luck with your upcoming defense. One last sprint till the end! **Andrea**, thanks for providing the Italian influence this lab needs. It's always nice and challenging speaking science in my mother tongue with you. You're clearly an enthusiast, you're going to be fine in any group you end up with. **Martina**, thanks for your immense help with administration and bureaucratic stuff. You work behind the curtains but you're extremely helpful in the group.

A big thanks to the students that helped me during my PhD. A big thanks especially to **Liset**, your contribution to the NHP tracing paper was extremely needed. You were quick to grasp the protocol and eager to learn new things. I truly wish you're going to have an enjoyable PhD.

Thanks to the members of **Ben's group**. The collaboration with you was crucial for my PhD journey. **Dan**, for running massive amounts of neutralization assays and always providing accurate feedback. **Kim**, for all the other neutralization assays and always providing with materials in a timely fashion. Thanks for your company on the tennis field as well. **Julian**, for your help with the project as well and for always inviting me to play futsal, despite my numerous postponements.

Thanks to past and present members of **Jonathan's group**. **Jonathan** first of all for the expertise in T cells and involvement in several projects. **Leona**, **Chris,** and **Julian**, for cheering up the lab and massive help in film making. Sharing the lab space with you was always a pleasure. **Egon** and **Junjie**, thanks for your contribution to different projects and being part of the "late evening" company.

Thanks to past and present members of **Gerry's group**. **Gerry,** for your expertise about virology and expanding the Irish humor with Martin. **Leo,** for sharing your protein expertise and always providing us with loads of proteins to play with. I'll manage to do an Ironman at some point. **Marc**, for sharing countless conversations in the cell culture room. You're very considerate and idealistic. I wish you all the best for your future. **Laura,** for always being available to share material despite unanticipated requests. **Ainhoa**, thanks for always being helpful, especially when I had to submit the defense application. **Siwen**, thanks for being part of the "late evening" company. It makes the late work more bearable.

David, thanks for always inviting me to social events. Your never-ending energy is truly appreciated. **Ben** and **Lifeng**, for cheering up the KI gym, working out with you was always enjoyable.

Thanks to all the collaborators in and out of KI that helped with the projects. First of all, the staff of **AFL**, especially **Bengt** and **Jenny**, for your immense support in managing the NHPs. It was always enjoyable to come over and see the macaques. The work environment always felt professional and relaxed. The **Lorè group**, especially Klara for your help with the T cell cytokines detection assay. The **Jakobsson group** in Uppsala, especially **Carolina** and **Maximilian**, for their expertise and help in ancient genome analysis. The **Hällberg** group, especially Hrishikesh, for quickly solving mAb 23 structure and providing structural biology expertise. **Taras group**, especially **Kewei**, for helping with the gamma/delta TCR assay.

Big thanks to the administration of MTC. Especially **Gesan, Eva, Åsa, ad Marta** for caring about the department and the immense service to students, researchers and the rest of the stuff.

Thanks to all the people I met through MTC, MSA and DSA. Past members of other groups, **Silke** and **Mitch** for always being cheerful and sociable. **Vanessa**, for your Portuguese energy. **Lidia**, for having the patience of organizing the MSA journal clubs. **Mariana**, for being very friendly and being one of the last symbols of the old MTC. **Jojo**, for holding hard for MTC students at the MSA. **Sebastian**, for your incredible commitment to science and doctoral education. **Leonie**, for truly caring about students' wellbeing and taking over the MSA. **Okan**, for following up after Leonie and keeping the MSA alive. **Roberta**, for always caring a smile in the quarter and the occasional chats in the corridor. **Giuseppe** and **Alexandra,** for the occasional climbing, fishing, and hiking outings. **Valeria**, **Jan, Yildiz** and **Yujiao**, for the massive involvement at the DSA. **Natalie**, for fun moments around Stockholm and the archipelago. **Henna**, for always being cheerful and enthusiastic. **Benedek,** for always providing fun company and your laid-back personality. **Inga** and **Loic**, for the amazing adventure organizing the KiiM retreat.

Thanks to all the people I met outside of work, despite most of them are affiliated to KI. **Eveline** and **Alexander**, my current flat mates, for being extremely nice at home and providing occasional scientific chats. **Despoina**, for your Greek lessons, the long walks around Stockholm, your love for pizzas and daily conversations. **Luca** and **Sara**, for your cheerful company, night-long conversations at your place, Bud spencer movies and online company. **Marine**, for always providing a caring company, your French lessons, the lazy evenings and for always being compassionate. **Emanuela**, for refreshing my memories of my region with your accent from Marche. **Nicole**, **Nathan** and little **Noah**, for being fun and wonderful people and always the ideal flat mates. **Harpa**, for teaching me how to roll snow to make a snowman and for bringing Icelandic fermented shark. **Thibaud**, for your

suggestions on workouts and new great metal groups. **Natalia**, for always being welcoming and easygoing and animating the midsommars. **Jonas**, for sharing with us your homemade snaps and bits of Swedish culture.

Thanks to the second family I acquired in Stockholm. There's so much I should write for each one of you but I would need to write another thesis just for that. **Joep**, for always being up to do something and being present all through my time in Stockholm. For the unforgivable rice and chicken dinner and climbing outings. Definitely, for sharing part of the Stockholm housing adventure together. You're next, it's going to be great!!! **Jemina**, for always being so cheerful and fun, for always inviting us all to do stuff and trying to include everyone, and for your love for the sun, despite your Finnish origin we consider you a Mediterranean. Best wishes for your next adventure in the Netherlands!!! **Mirco**, for your enthusiasm about science, literature, philosophy and also bureaucracy that ends with interesting conversations. For your immense care, reflected by your cuisine and advices, and for always sharing your place with all of us. You never give up, I'm confident your time is about to come!!! **Gabriele**, for always being available to explain bioinformatics and updating us on the latest tech, feeding us delicious food and the gaming experiences at your place. Your drawings were always great and you're further improving. Best wishes for your career!!! **Nuria**, for having the best humor and the best combinations of gins and plants. For your carefree mindset and for organizing quizzes for everybody and being always welcoming for games and drinks. All the best for your PhD as well, you're almost done!!! **Carlos.** Thanks to you as well for being fun, the padel matches and for sharing your expertise on travels and flights. It's always refreshing to hear life science outsiders. Keep it up in the flying industry!!! **Shane**, for being welcoming to everyone and your willingness to always party. I hope we're going to share many shots in the future. I hope I'll manage to see the rise of Professor S. Wright!!! **Maria**, for your energy and your honesty and always having positive words. For always being ready to dance and have fun. All the best navigating the corporate world!!! **Alek**, for your enthusiasm about science, music, and the Soul series. Your randomness is always fun to experience. Keep it up, I'm looking forward your next artistic endeavor!!! **Björn**, for your midsommars and your historical anecdotes. For organizing Karaoke parties and sharing bits of Swedish culture with us. Let's have a great time in Colombia!!! **Stefania**, for your energy, reflected in your voice, and your spontaneity. Your presence definitely animates the environment every time you discover something new. Your Italian is already very good. Never give up, you have the energy to achieve everything!!! **Kostas**, for being easy-going and sharing with Stefania the Greek culture. For being daring and always up to new experiences. It's always a pleasure spending time with you, looking forward more moments together!!! A big thanks to the ones living far away. **Anna**, for your spirited personality and obsession with Catalonia. **Uri**, for always being chill and relaxed. Thanks to both of you for welcoming us to your place. **Alice** for your pleasant company and your mother's amazing grappa. For always being in touch, we miss you up here but we hope you enjoy more sun than us. **Lorenzo**, for the

extremely needed DJing service at all the CMM parties. For your friendly, serene, and cordial personality.

A huge thanks to all the scattered around Italy or around the world. **Leonardo, Flaminia, Francesca, Melo e Bury**. We struggle to keep in touch, but we always manage. All my friends back home that I can't wait to meet again every holidays. The ones that keep me company online and update me on the most recent events. **Manuè, Monte, Kobra, Federica la fagianella, Valentina la capa fagiana, Vergà, Riccani, Ricky, Paoletto, Peppe, Samuè, Federica, Paolo, Andrea, Giada, Leonardo, Robertino** and many many others**.**

Un caloroso grazie a tutta la mia famiglia, che mi accoglie sempre calorosamente ogni volta che torno a casa e che si tiene in contatto ogni settimana. A **Lauretta** che mi manda sempre cuoricini e animaletti nei messagi e che mi dà regolarmente il buon giorno e la buona notte. A **Simone**, che mi racconta vari anneddoti familiari e nonostante tutto si preoccupa di tutti. I miei migliori auguri per la tua carriera. A **Zia**, che mi accoglie sempre con lacrime e sorrisi ed è sempre pronta a prepararmi i miei piatti preferiti. A **Mamma**, che mi aggiorna constatemente di tutte le uscite e le attività di casa ed è sempre pronta a darmi consigli. A **Papà**, che è sempre al tempo che fà a Stoccolma e mi ricorda sempre che sono in gamba. Mi mancate molto ma sapete che siete sempre nel mio cuore, nonostante la distanza e le litigate occasionali.

8 References

1. Flajnik MF, Kasahara M. Origin and evolution of the adaptive immune system: Genetic events and selective pressures. *Nat Rev Genet*. 2010;11(1):47-59. doi:10.1038/nrg2703

2. Buchmann K. Evolution of innate immunity: Clues from invertebrates via fish to mammals. *Front Immunol*. 2014;5(SEP):1-8. doi:10.3389/fimmu.2014.00459

3. Cooper MD, Alder MN. The evolution of adaptive immune systems. *Cell*. 2006;124(4):815-822. doi:10.1016/j.cell.2006.02.001

4. Chaplin DD. Overview of the immune response. *Journal of Allergy and Clinical Immunology*. 2010;125(2 SUPPL. 2):S3-S23. doi:10.1016/j.jaci.2009.12.980

5. Moser M, Leo O. Key concepts in immunology. *Vaccine*. 2010;28(SUPPL. 3):C2. doi:10.1016/j.vaccine.2010.07.022

6. Kaufman J. Unfinished Business: Evolution of the MHC and the Adaptive Immune System of Jawed Vertebrates. *Annu Rev Immunol*. 2018;36:383-409. doi:10.1146/annurev-immunol-051116-052450

7. Ahmed R, Gray D. Immunological Memory and Protective Immunity: Understanding Their Relation. *Science (1979)*. 1996;272(April):54-61.

8. Zinkernagel RM. On immunological memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2000;355(1395):369-371. doi:10.1098/rstb.2000.0576

9. Koch U, Radtke F. Mechanisms of T cell development and transformation. *Annu Rev Cell Dev Biol*. 2011;27. doi:10.1146/annurev-cellbio-092910-154008

10. Hardy RR, Hayakawa K. B cell development pathways. *Annu Rev Immunol*. 2001;19. doi:10.1146/annurev.immunol.19.1.595

11. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. *Cell*. 2002;109(2 SUPPL. 1):45-55. doi:10.1016/S0092-8674(02)00675-X

12. Alt FW, Oltz EM, Young F, Gorman J, Taccioli G, Chen J. VDJ recombination. *Immunol Today*. 1992;13(8):306-314. doi:10.1016/0167-5699(92)90043-7

13. Schatz DG, Oettinger MA, Schlissel MS. V(D)J recombination: molecular biology and regulation. *Annu Rev Immunol*. 1992;10:359-383. doi:10.1146/annurev.iy.10.040192.002043

14. Galson JD, Trück J, Fowler A, et al. In-depth assessment of within-individual and inter-individual variation in the B cell receptor repertoire. *Front Immunol*. 2015;6(OCT):1-13. doi:10.3389/fimmu.2015.00531

15. Woodsworth DJ, Castellarin M, Holt RA. Sequence analysis of T-cell repertoires in health and disease. *Genome Med*. 2013;5(10). doi:10.1186/gm502

16. Nemazee D. Mechanisms of central tolerance for B cells. *Nat Rev Immunol*. 2017;17(5):281-294. doi:10.1038/nri.2017.19

17. Hogquist KA, Baldwin TA, Jameson SC. Central tolerance: Learning self-control in the thymus. *Nat Rev Immunol*. 2005;5(10):772-782. doi:10.1038/nri1707

18. Chan TD, Wood K, Hermes JR, et al. Elimination of Germinal-Center-Derived Self-Reactive B Cells Is Governed by the Location and Concentration of Self-Antigen. *Immunity*. 2012;37(5):893-904. doi:10.1016/j.immuni.2012.07.017

19. Brink R, Phan TG. Self-Reactive B Cells in the Germinal Center Reaction. *Annu Rev Immunol*. 2018;36(1):339-357. doi:10.1146/annurev-immunol-051116-052510

20. Bouillet P, Metcalf D, Huang DCS. Proapoptotic Bcl-2 relative Bim required for certain apoptotic responses, leukocyte homeostasis, and to preclude autoimmunity. *Science (1979)*. 1999;286(5445):1735-1738. doi:10.1126/science.286.5445.1735

21. Enders A, Bouillet P, Puthalakath H, Xu Y, Tarlinton DM, Strasser A. Loss of the pro-apoptotic BH3-only Bcl-2 family member Bim inhibits BCR stimulation-induced apoptosis and deletion of autoreactive B cells. *Journal of Experimental Medicine*. 2003;198(7):1119-1126. doi:10.1084/jem.20030411

22. Sakaguchi S, Wing K, Yamaguchi T. Dynamics of peripheral tolerance and immune regulation mediated by Treg. *Eur J Immunol*. 2009;39(9):2331-2336. doi:10.1002/eji.200939688

23. Mueller DL. Mechanisms maintaining peripheral tolerance. *Nat Immunol*. 2010;11(1):21-27. doi:10.1038/ni.1817

24. Redmond WL, Sherman LA. Peripheral tolerance of CD8 T lymphocytes. *Immunity*. 2005;22(3):275-284. doi:10.1016/j.immuni.2005.01.010

25. Brink R. The imperfect control of self-reactive germinal center B cells. *Curr Opin Immunol*. 2014;28(1):97-101. doi:10.1016/j.coi.2014.03.001

26. Akkaya M, Kwak K, Pierce SK. B cell memory: building two walls of protection against pathogens. *Nat Rev Immunol*. 2020;20(4). doi:10.1038/s41577-019-0244-2

27. Kreslavsky T, Gleimer M, Garbe AI, von Boehmer H. αβ versus γδ fate choice: Counting the T-cell lineages at the branch point. *Immunol Rev*. 2010;238(1):169-181. doi:10.1111/j.1600-065X.2010.00947.x

28. Rossjohn J, Gras S, Miles JJ, Turner SJ, Godfrey DI, McCluskey J. T Cell Antigen Receptor Recognition of Antigen-Presenting Molecules. *Annu Rev Immunol*. 2015;33(1):169-200. doi:10.1146/annurev-immunol-032414-112334

29. Morath A, Schamel WW. αβ and γδ T cell receptors: Similar but different. *J Leukoc Biol*. 2020;107(6):1045-1055. doi:10.1002/JLB.2MR1219-233R

30. Singer A, Adoro S, Park JH. Lineage fate and intense debate: Myths, models and mechanisms of CD4- versus CD8-lineage choice. *Nat Rev Immunol*. 2008;8(10):788-801. doi:10.1038/nri2416

31. Savage PA, Klawon DEJ, Miller CH. Regulatory T Cell Development. *Annu Rev Immunol*. 2020;38(1):421-453. doi:10.1146/annurev-immunol-100219-020937

32. Khairallah C, Chu TH, Sheridan BS. Tissue Adaptations of Memory and Tissue-Resident Gamma Delta T Cells. *Front Immunol*. 2018;9(NOV):1-22. doi:10.3389/fimmu.2018.02636

33. Ribot JC, Lopes N, Silva-Santos B. γδ T cells in tissue physiology and surveillance. *Nat Rev Immunol*. 2021;21(4):221-232. doi:10.1038/s41577-020-00452-4

34. Hayday AC. γδ T Cell Update: Adaptate Orchestrators of Immune Surveillance. *The Journal of Immunology*. 2019;203(2):311-320. doi:10.4049/jimmunol.1800934

35. Bendelac A, Savage PB, Teyton L. The Biology of NKT Cells. *Annu Rev Immunol*. 2007;25(1):297-336. doi:10.1146/annurev.immunol.25.022106.141711

36. Provine NM, Klenerman P. MAIT Cells in Health and Disease. *Annu Rev Immunol*. 2020;38(1):203-228. doi:10.1146/annurev-immunol-080719-015428

37. Allman D, Wilmore JR, Gaudette BT. The continuing story of T-cell independent antibodies. *Immunol Rev*. 2019;288(1):128-135. doi:10.1111/imr.12754

38. Mesin L, Ersching J, Victora GD. Germinal Center B Cell Dynamics. *Immunity*. 2016;45(3):471-482. doi:10.1016/j.immuni.2016.09.001

39. Stavnezer J, Guikema JEJ, Schrader CE. Mechanism and Regulation of Class Switch Recombination. *Annu Rev Immunol*. 2008;26(1):261-292. doi:10.1146/annurev.immunol.26.021607.090248

40. Radbruch A, Muehlinghaus G, Luger EO, et al. Competence and competition: The challenge of becoming a long-lived plasma cell. *Nat Rev Immunol*. 2006;6(10):741-750. doi:10.1038/nri1886

41. Mesin L, Schiepers A, Ersching J, et al. Restricted Clonality and Limited Germinal Center Reentry Characterize Memory B Cell Reactivation by Boosting. *Cell*. 2020;180(1):92-106.e11. doi:10.1016/j.cell.2019.11.032

42. de Carvalho RVH, Ersching J, Barbulescu A, et al. Clonal replacement sustains long-lived germinal centers primed by respiratory viruses. *Cell*. 2022;186(1):131-146.e13. doi:10.1016/j.cell.2022.11.031

43. Lefranc, M.-P. and Lefranc G. *The T Cell Receptor FactsBook*.; 2001. doi:10.1016/b978-0-12-441352-8.x5000-0

44. Szeto C, Lobos CA, Nguyen AT, Gras S. TCR Recognition of Peptide–MHC-I: Rule Makers and Breakers. *Int J Mol Sci*. 2021;22(1). doi:10.3390/ijms22010068

45. Dean J, Emerson RO, Vignali M, et al. Annotation of pseudogenic gene segments by massively parallel sequencing of rearranged lymphocyte receptor loci. *Genome Med*. 2015;7(1):1-8. doi:10.1186/s13073-015-0238-z

46. Paletta D, Fichtner AS, Hahn AM, et al. The hypervariable region 4 (HV4) and position 93 of the α chain modulate CD1d-glycolipid binding of iNKT TCRs. *Eur J Immunol*. 2015;45(7):2122-2133. doi:10.1002/eji.201545534

47. Willcox CR, Vantourout P, Salim M, et al. Butyrophilin-like 3 Directly Binds a Human Vγ4+ T Cell Receptor Using a Modality Distinct from Clonally-Restricted Antigen. *Immunity*. 2019;51(5):813-825.e4. doi:10.1016/j.immuni.2019.09.006

48. Hongmin L, Llera A, Tsuchiya D, et al. Three-dimensional structure of the complex between a T cell receptor β chain and the superantigen staphylococcal enterotoxin B. *Immunity*. 1998;9(6):807-816. doi:10.1016/S1074-7613(00)80646-9

49. López-Sagaseta J, Dulberger CL, Crooks JE, et al. The molecular basis for Mucosal-Associated Invariant T cell recognition of MR1 proteins. *Proc Natl Acad Sci U S A*. 2013;110(19). doi:10.1073/pnas.1222678110

50. Macho-Fernandez E, Brigl M. The extended family of CD1d-restricted NKT cells: Sifting through a mixed bag of TCRs, antigens, and functions. *Front Immunol*. 2015;6(JUL). doi:10.3389/fimmu.2015.00362

51. Deseke M, Prinz I. Ligand recognition by the γδ TCR and discrimination between homeostasis and stress conditions. *Cell Mol Immunol*. 2020;17(9):914-924. doi:10.1038/s41423-020-0503-y

52. Gras S, Kjer-Nielsen L, Burrows SR, McCluskey J, Rossjohn J. T-cell receptor bias and immunity. *Curr Opin Immunol*. 2008;20(1). doi:10.1016/j.coi.2007.12.001

53. Hallmayer J, Faraco J, Lin L, et al. Narcolepsy is strongly associated with the T-cell receptor alpha locus. *Nat Genet*. 2009;41(6):708-711. doi:10.1038/ng.372

54. Brennan RM, Petersen J, Neller MA, et al. The Impact of a Large and Frequent Deletion in the Human TCR β Locus on Antiviral Immunity. *The Journal of Immunology*. 2012;188(6). doi:10.4049/jimmunol.1102675

55. Brennan RM, Burrows JM, Elliott T, et al. Missense single nucleotide polymorphisms in the human T cell receptor loci control variable gene usage in the T cell repertoire. *Br J Haematol*. 2014;166(1). doi:10.1111/bjh.12830

56. Gras S, Chen Z, Miles JJ, et al. Allelic polymorphism in the T cell receptor and its impact on immune responses. *Journal of Experimental Medicine*. 2010;207(7). doi:10.1084/jem.20100603

57. Shomuradova AS, Vagida MS, Sheetikov SA, et al. SARS-CoV-2 Epitopes Are Recognized by a Public and Diverse Repertoire of Human T Cell Receptors. *Immunity*. 2020;53(6). doi:10.1016/j.immuni.2020.11.004

58. Han F, Faraco J, Dong XS, et al. Genome Wide Analysis of Narcolepsy in China Implicates Novel Immune Loci and Reveals Changes in Association Prior to Versus After the 2009 H1N1 Influenza Pandemic. *PLoS Genet*. 2013;9(10). doi:10.1371/journal.pgen.1003880

59. Song IY, Gil A, Mishra R, Ghersi D, Selin LK, Stern LJ. Broad TCR repertoire and diverse structural solutions for recognition of an immunodominant CD8+ T cell epitope. *Nat Struct Mol Biol*. 2017;24(4):395-406. doi:10.1038/nsmb.3383

60. Schroeder HW, Cavacini L. Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology*. 2010;125(2 SUPPL. 2). doi:10.1016/j.jaci.2009.09.046

61. Lefranc MP, Lefranc G. *The Immunoglobulin Factsbook*.; 2001.

62. Maity PC, Blount A, Jumaa H, Ronneberger O, Lillemeier BF, Reth M. B cell antigen receptors of the IgM and IgD classes are clustered in different protein islands that are altered during B cell activation. *Sci Signal*. 2015;8(394):ra93. doi:10.1126/scisignal.2005887

63. Übelhart R, Hug E, Bach MP, et al. Responsiveness of B cells is regulated by the hinge region of IgD. *Nat Immunol*. 2015;16(5):534-543. doi:10.1038/ni.3141

64. Carsetti R, Köhler G, Lamers MC. A role for immunoglobulin D: interference with tolerance induction. *Eur J Immunol*. 1993;23(1):168-178. doi:10.1002/eji.1830230127

65. Shapiro-Shelef M, Lin KI, McHeyzer-Williams LJ, Liao J, McHeyzer-Williams MG, Calame K. Blimp-1 is required for the formation of immunoglobulin secreting plasma cells and pre-plasma memory B cells. *Immunity*. 2003;19(4):607-620. doi:10.1016/S1074-7613(03)00267-X

66. Horcher M, Souabni A, Busslinger M. Pax5/BSAP maintains the identity of B cells in late B lymphopoiesis. *Immunity*. 2001;14(6):779-790. doi:10.1016/S1074-7613(01)00153-4

67. Shaffer AL, Lin KI, Kuo TC, et al. Blimp-1 orchestrates plasma cell differentiation by extinguishing the mature B cell gene expression program. *Immunity*. 2002;17(1):51-62. doi:10.1016/s1074-7613(02)00335-7

68. Plotkin SA. Updates on immunologic correlates of vaccine-induced protection. *Vaccine*. 2020;38(9):2250-2257. doi:10.1016/j.vaccine.2019.10.046

69. Lu LL, Suscovich TJ, Fortune SM, Alter G. Beyond binding: Antibody effector functions in infectious diseases. *Nat Rev Immunol*. 2018;18(1):46-61. doi:10.1038/nri.2017.106

70. Nakagawa R, Calado DP. Positive Selection in the Light Zone of Germinal Centers. *Front Immunol*. 2021;12(March):1-8. doi:10.3389/fimmu.2021.661678

71. Victora GD, Nussenzweig MC. Germinal centers. *Annu Rev Immunol*. 2012;30:429-457. doi:10.1146/annurev-immunol-020711-075032

72. Trück J, Ramasamy MN, Galson JD, et al. Identification of Antigen-Specific B Cell Receptor Sequences Using Public Repertoire Analysis. *The Journal of Immunology*. 2015;194(1). doi:10.4049/jimmunol.1401405

73. Hoehn KB, vander Heiden JA, Zhou JQ, Lunter G, Pybus OG, Kleinstein SH. Repertoire-wide phylogenetic models of B cell molecular evolution reveal evolutionary signatures of aging and vaccination. *Proc Natl Acad Sci U S A*. 2019;116(45). doi:10.1073/pnas.1906020116

74. de Bourcy CFA, Lopez Angel CJ, Vollmers C, Dekker CL, Davis MM, Quake SR. Phylogenetic analysis of the human antibody repertoire reveals quantitative signatures of immune senescence and aging. *Proc Natl Acad Sci U S A*. 2017;114(5). doi:10.1073/pnas.1617959114

75. Croote D, Darmanis S, Nadeau KC, Quake SR. Antibodies Cloned From Single. *Science (1979)*. 2018;1309(December):1306-1309.

76. Parks T, Mirabel MM, Kado J, et al. Association between a common immunoglobulin heavy chain allele and rheumatic heart disease risk in Oceania. *Nat Commun*. 2017;8(May):1-10. doi:10.1038/ncomms14946

77. Zhang J, Hu X, Wang J, et al. Immune receptor repertoires in pediatric and adult acute myeloid leukemia. *Genome Med*. 2019;11(1):1-11. doi:10.1186/s13073-019-0681-3

78. Zelenetz AD, Chen TT, Levy R. Clonal expansion in follicular lymphoma occurs subsequent to antigenic selection. *J Exp Med*. 1992;176(4):1137-1148. doi:10.1084/jem.176.4.1137

79. Volpe JM, Kepler TB. Genetic correlates of autoreactivity and autoreactive potential in human Ig heavy chains. *Immunome Res*. 2009;5(1):1-8. doi:10.1186/1745-7580-5-1

80. Bashford-Rogers RJM, Smith KGC, Thomas DC. Antibody repertoire analysis in polygenic autoimmune diseases. *Immunology*. 2018;155(1). doi:10.1111/imm.12927

81. Tipton CM, Fucile CF, Darce J, et al. Diversity, cellular origin and autoreactivity of antibody-secreting cell population expansions in acute systemic lupus erythematosus. *Nat Immunol*. 2015;16(7). doi:10.1038/ni.3175

82. Liu S, Hou XL, Sui WG, Lu QJ, Hu YL, Dai Y. Direct measurement of B-cell receptor repertoire's composition and variation in systemic lupus erythematosus. *Genes Immun*. 2017;18(1). doi:10.1038/gene.2016.45

83. Jackson KJL, Liu Y, Roskin KM, et al. Human responses to influenza vaccination show seroconversion signatures and convergent antibody rearrangements. *Cell Host Microbe*. 2014;16(1). doi:10.1016/j.chom.2014.05.013

84. Murugan R, Scally SW, Costa G, et al. Evolution of protective human antibodies against Plasmodium falciparum circumsporozoite protein repeat motifs. *Nat Med*. 2020;26(7):1135-1145. doi:10.1038/s41591-020-0881-9

85. Robbiani DF, Bozzacco L, Keeffe JR, et al. Recurrent Potent Human Neutralizing Antibodies to Zika Virus in Brazil and Mexico. *Cell*. 2017;169(4):597-609.e11. doi:10.1016/j.cell.2017.04.024

86. Davis CW, Jackson KJL, McElroy AK, et al. Longitudinal Analysis of the Human B Cell Response to Ebola Virus Infection. *Cell*. 2019;177(6):1566-1582.e17. doi:10.1016/j.cell.2019.04.036

87. Wu X, Yang ZY, Li Y, et al. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. *Science (1979)*. 2010;329(5993). doi:10.1126/science.1187659

88. Huang J, Kang BH, Ishida E, et al. Identification of a CD4-Binding-Site Antibody to HIV that Evolved Near-Pan Neutralization Breadth. *Immunity*. 2016;45(5). doi:10.1016/j.immuni.2016.10.027

89. Fagiani F, Catanzaro M, Lanni C. Molecular features of IGHV3-53-encoded antibodies elicited by SARS-CoV-2. *Signal Transduct Target Ther*. 2020;5(1). doi:10.1038/s41392-020-00287-4

90. Yuan M, Liu H, Wu NC, et al. Structural basis of a shared antibody response to SARS-CoV-2. *Science (1979)*. 2020;369(6507). doi:10.1126/science.abd2321

91. Barnes CO, West AP, Huey-Tubman KE, et al. Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell*. 2020;182(4). doi:10.1016/j.cell.2020.06.025

92. Yuan M, Wu NC, Zhu X, et al. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science (1979)*. 2020;368(6491). doi:10.1126/science.abb7269

93. Yuan M, Liu H, Wu NC, Wilson IA. Recognition of the SARS-CoV-2 receptor binding domain by neutralizing antibodies. *Biochem Biophys Res Commun*. 2021;538. doi:10.1016/j.bbrc.2020.10.012

94. Liu X, Wu J. History, applications, and challenges of immune repertoire research. *Cell Biol Toxicol*. 2018;34(6). doi:10.1007/s10565-018-9426-0

95. Craig Venter J, Adams MD, Myers EW, et al. The sequence of the human genome. *Science (1979)*. 2001;291(5507). doi:10.1126/science.1058040

96. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822). doi:10.1038/35057062

97. Watson CT, Breden F. The immunoglobulin heavy chain locus: Genetic variation, missing data, and implications for human disease. *Genes Immun*. 2012;13(5). doi:10.1038/gene.2012.12

98. Watson CT, Steinberg KM, Huddleston J, et al. Complete haplotype sequence of the human immunoglobulin heavy-chain variable, diversity, and joining genes and characterization of allelic and copy-number variation. *Am J Hum Genet*. 2013;92(4). doi:10.1016/j.ajhg.2013.03.004

99. Corcoran MM, Phad GE, Bernat NV, et al. Production of individualized v gene databases reveals high levels of immunoglobulin genetic diversity. *Nat Commun*. 2016;7. doi:10.1038/ncomms13642

100. Vázquez Bernat N, Corcoran M, Nowak I, et al. Rhesus and cynomolgus macaque immunoglobulin heavy-chain genotyping yields comprehensive databases of germline VDJ alleles. *Immunity*. 2021;54(2). doi:10.1016/j.immuni.2020.12.018

101. Peng K, Safonova Y, Shugay M, et al. Diversity in immunogenomics: the value and the challenge. *Nat Methods*. 2021;18(6):588-591. doi:10.1038/s41592-021-01169-5

102. Chernyshev M, Kaduk M, Corcoran M, Karlsson Hedestam GB. VDJ Gene Usage in IgM Repertoires of Rhesus and Cynomolgus Macaques. *Front Immunol*. 2022;12. doi:10.3389/fimmu.2021.815680

103. Phad GE, Pushparaj P, Tran K, et al. Extensive dissemination and intraclonal maturation of HIV Env vaccine-induced B cell responses. *Journal of Experimental Medicine*. 2020;217(2). doi:10.1084/jem_20191155

104. Kaduk M, Corcoran M, Karlsson Hedestam GB. Addressing IGHV Gene Structural Diversity Enhances Immunoglobulin Repertoire Analysis: Lessons From Rhesus Macaque. *Front Immunol*. 2022;13(March):1-11. doi:10.3389/fimmu.2022.818440

105. Mardis ER. Next-generation sequencing platforms. *Annual Review of Analytical Chemistry*. 2013;6:287-303. doi:10.1146/annurev-anchem-062012-092628

106. Benichou J, Ben-Hamo R, Louzoun Y, Efroni S. Rep-Seq: Uncovering the immunological repertoire through next-generation sequencing. *Immunology*. 2012;135(3):183-191. doi:10.1111/j.1365-2567.2011.03527.x

107. Bernat NV, Corcoran M, Hardt U, et al. High-quality library preparation for NGS-based immunoglobulin germline gene inference and repertoire expression analysis. *Front Immunol*. 2019;10(APR). doi:10.3389/fimmu.2019.00660

108. Bolotin DA, Poslavsky S, Mitrophanov I, et al. MiXCR: Software for comprehensive adaptive immunity profiling. *Nat Methods*. 2015;12(5). doi:10.1038/nmeth.3364

109. Li S, Lefranc MP, Miles JJ, et al. IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. *Nat Commun*. 2013;4. doi:10.1038/ncomms3333

110. Zhang W, Wang IM, Wang C, et al. IMPre: An accurate and efficient software for prediction of T- and B-cell receptor germline genes and alleles from rearranged repertoire data. *Front Immunol*. 2016;7(NOV). doi:10.3389/fimmu.2016.00457

111. Gupta NT, vander Heiden JA, Uduman M, Gadala-Maria D, Yaari G, Kleinstein SH. Change-O: A toolkit for analyzing large-scale B cell immunoglobulin repertoire sequencing data. *Bioinformatics*. 2015;31(20):3356-3358. doi:10.1093/bioinformatics/btv359

112. vander Heiden JA, Yaari G, Uduman M, et al. PRESTO: A toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. *Bioinformatics*. 2014;30(13):1930-1932. doi:10.1093/bioinformatics/btu138

113. Gibson WS, Rodriguez OL, Shields K, et al. Characterization of the immunoglobulin lambda chain locus from diverse populations reveals extensive genetic variation. *Genes Immun*. 2022;(July):1-11. doi:10.1038/s41435-022-00188-2

114. Omer A, Peres A, Rodriguez OL, et al. T cell receptor beta germline variability is revealed by inference from repertoire data. *Genome Med*. 2022;14(1):1-19. doi:10.1186/s13073-021-01008-4

115. Martinez-Murillo P, Tran K, Guenaga J, et al. Particulate Array of Well-Ordered HIV Clade C Env Trimers Elicits Neutralizing Antibodies that Display a Unique V2 Cap Approach. *Immunity*. 2017;46(5):804-817.e7. doi:10.1016/j.immuni.2017.04.021

116. Dekosky BJ, Ippolito GC, Deschner RP, et al. High-throughput sequencing of the paired human immunoglobulin heavy and light chain repertoire. *Nat Biotechnol*. 2013;31(2). doi:10.1038/nbt.2492

117. DeKosky BJ, Kojima T, Rodin A, et al. In-depth determination and analysis of the human paired heavy- and light-chain antibody repertoire. *Nat Med*. Published online 2015. doi:10.1038/nm.3743

118. Turchaninova MA, Britanova O v., Bolotin DA, et al. Pairing of T-cell receptor chains via emulsion PCR. *Eur J Immunol*. 2013;43(9). doi:10.1002/eji.201343453

119. Wang B, Dekosky BJ, Timm MR, et al. Functional interrogation and mining of natively paired human v H:V L antibody repertoires. *Nat Biotechnol*. 2018;36(2):152-155. doi:10.1038/nbt.4052

120. Josefsberg JO, Buckland B. Vaccine process technology. *Biotechnol Bioeng*. 2012;109(6):1443-1460. doi:10.1002/bit.24493

121. Rappuoli R, Bottomley MJ, D'Oro U, Finco O, de Gregorio E. Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. *Journal of Experimental Medicine*. 2016;213(4):469-481. doi:10.1084/JEM.20151960

122. Rappuoli R. Reverse vaccinology Rino Rappuoli. *Curr Opin Microbiol*. Published online 2000:445-450.

123. Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell*. 2020;181(2). doi:10.1016/j.cell.2020.02.058

124. Li W, Moore MJ, Vasllieva N, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*. 2003;426(6965). doi:10.1038/nature02145

125. Huang Y, Yang C, Xu X feng, Xu W, Liu S wen. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. *Acta Pharmacol Sin*. 2020;41(9). doi:10.1038/s41401-020-0485-4

126. Wrapp D, Wang N, Corbett KS, et al. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science (1979)*. 2020;367(6483). doi:10.1126/science.aax0902

127. Harrison AG, Lin T, Wang P. Mechanisms of SARS-CoV-2 Transmission and Pathogenesis. *Trends Immunol*. 2020;41(12). doi:10.1016/j.it.2020.10.004

128. Mandolesi M, Sheward DJ, Hanke L, et al. SARS-CoV-2 protein subunit vaccination of mice and rhesus macaques elicits potent and durable neutralizing antibody responses. *Cell Rep Med*. 2021;2(4). doi:10.1016/j.xcrm.2021.100252

129. Logunov DY, Dolzhikova I v., Shcheblyakov D v., et al. Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia. *The Lancet*. 2021;397(10275). doi:10.1016/S0140-6736(21)00234-8

130. Voysey M, Clemens SAC, Madhi SA, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *The Lancet*. 2021;397(10269). doi:10.1016/S0140-6736(20)32661-1

131. Sadoff J, Gray G, Vandebosch A, et al. Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. *N Engl J Med*. 2021;384(23):2187-2201. doi:10.1056/NEJMoa2101544

132. Ella R, Reddy S, Jogdand H, et al. Safety and immunogenicity of an inactivated SARS-CoV-2 vaccine, BBV152: interim results from a double-blind, randomised, multicentre, phase 2 trial, and 3-month follow-up of a double-blind, randomised phase 1 trial. *Lancet Infect Dis*. 2021;21(7):950-961. doi:10.1016/S1473-3099(21)00070-0

133. Polack FP, Thomas SJ, Kitchin N, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *New England Journal of Medicine*. 2020;383(27). doi:10.1056/nejmoa2034577

134. Baden LR, el Sahly HM, Essink B, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *New England Journal of Medicine*. 2021;384(5). doi:10.1056/nejmoa2035389

135. Keech C, Albert G, Cho I, et al. Phase 1–2 Trial of a SARS-CoV-2 Recombinant Spike Protein Nanoparticle Vaccine. *New England Journal of Medicine*. 2020;383(24). doi:10.1056/nejmoa2026920

136. Xia S, Duan K, Zhang Y, et al. Effect of an Inactivated Vaccine Against SARS-CoV-2 on Safety and Immunogenicity Outcomes: Interim Analysis of 2 Randomized Clinical Trials. *JAMA*. 2020;324(10):951-960. doi:10.1001/jama.2020.15543

137. Zeng G, Wu Q, Pan H, et al. Immunogenicity and safety of a third dose of CoronaVac, and immune persistence of a two-dose schedule, in healthy adults: interim results from two single-centre, double-blind, randomised, placebo-controlled phase 2 clinical trials. *Lancet Infect Dis*. 2022;22(4):483-495. doi:10.1016/S1473-3099(21)00681-2

138. Sheward DJ, Mandolesi M, Urgard E, et al. Beta RBD boost broadens antibody-mediated protection against SARS-CoV-2 variants in animal models. *Cell Rep Med*. 2021;2(11). doi:10.1016/j.xcrm.2021.100450

139. Fan Y, Li X, Zhang L, Wan S, Zhang L, Zhou F. SARS-CoV-2 Omicron variant: recent progress and future perspectives. *Signal Transduct Target Ther*. 2022;7(1). doi:10.1038/s41392-022-00997-x

140. Tegally H, Moir M, Everatt J, et al. Emergence of SARS-CoV-2 Omicron lineages BA.4 and BA.5 in South Africa. *Nat Med*. 2022;28(9):1785-1790. doi:10.1038/s41591-022-01911-2

141. Burki T. COVID vaccine booster doses for omicron variants. *Lancet Respir Med*. 2022;10(10):936. doi:10.1016/S2213-2600(22)00361-7

142. Loré K. Isolation and immunophenotyping of human and rhesus macaque dendritic cells. *Methods Cell Biol*. 2004;2004(75). doi:10.1016/s0091-679x(04)75026-8

143. Vigdorovich V, Oliver BG, Carbonetti S, et al. Repertoire comparison of the B-cell receptor-encoding loci in humans and rhesus macaques by next-generation sequencing. *Clin Transl Immunology*. 2016;5(7). doi:10.1038/cti.2016.42

144. Rotger M, Dalmau J, Rauch A, et al. Comparative transcriptomics of extreme phenotypes of human HIV-1 infection and SIV infection in sooty mangabey and rhesus macaque. *Journal of Clinical Investigation*. 2011;121(6). doi:10.1172/JCI45235

145. Gibbs RA, Rogers J, Katze MG, et al. Evolutionary and biomedical insights from the rhesus macaque genome. *Science (1979)*. 2007;316(5822). doi:10.1126/science.1139247

146. Baroncelli S, Negri DRM, Michelini Z, Cara A. Macaca mulatta, fascicularis and nemestrina in AIDS vaccine development. *Expert Rev Vaccines*. 2008;7(9). doi:10.1586/14760584.7.9.1419

147. Kwissa M, Nakaya HI, Oluoch H, Pulendran B. Distinct TLR adjuvants differentially stimulate systemic and local innate immune responses in nonhuman primates. *Blood*. 2012;119(9). doi:10.1182/blood-2011-10-388579

148. Klasse PJ, Nixon DF, Moore JP. Immunogenicity of clinically relevant SARS-CoV-2 vaccines in nonhuman primates and humans. *Sci Adv*. 2021;7(12):1-23. doi:10.1126/SCIADV.ABE8065

149. Zheng H, Li H, Guo L, et al. Virulence and pathogenesis of SARS-CoV-2 infection in rhesus macaques: A nonhuman primate model of COVID-19 progression. *PLoS Pathog*. 2020;16(11):1-23. doi:10.1371/JOURNAL.PPAT.1008949

150. Munster VJ, Feldmann F, Williamson BN, et al. Respiratory disease in rhesus macaques inoculated with SARS-CoV-2. *Nature*. 2020;585(7824). doi:10.1038/s41586-020-2324-7

151. Chandrashekar A, Liu J, Martino AJ, et al. SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science (1979)*. 2020;369(6505):812-817. doi:10.1126/science.abc4776

152. Garrido C, Curtis AD, Dennis M, et al. SARS-CoV-2 vaccines elicit durable immune responses in infant rhesus macaques. *Sci Immunol*. 2021;6(60):1-17. doi:10.1126/SCIIMMUNOL.ABJ3684

153. Shan C, Yao YF, Yang X lou, et al. Infection with novel coronavirus (SARS-CoV-2) causes pneumonia in Rhesus macaques. *Cell Res*. 2020;30(8):670-677. doi:10.1038/s41422-020-0364-z

154. Salguero FJ, White AD, Slack GS, et al. Comparison of rhesus and cynomolgus macaques as an infection model for COVID-19. *Nat Commun*. 2021;12(1). doi:10.1038/s41467-021-21389-9

155. Williamson BN, Feldmann F, Schwarz B, et al. Clinical benefit of remdesivir in rhesus macaques infected with SARS-CoV-2. *Nature*. 2020;585(7824):273-276. doi:10.1038/s41586-020-2423-5

156. Maisonnasse P, Guedj J, Contreras V, et al. Hydroxychloroquine use against SARS-CoV-2 infection in non-human primates. *Nature*. 2020;585(7826):584-587. doi:10.1038/s41586-020-2558-4

157. Corbett KS, Werner AP, Connell SO, et al. mRNA-1273 protects against SARS-CoV-2 beta infection in nonhuman primates. *Nat Immunol*. 2021;22(10):1306-1315. doi:10.1038/s41590-021-01021-0

158. Waickman AT, Victor K, Newell K, et al. mRNA-1273 vaccination protects against SARS-CoV-2-elicited lung inflammation in nonhuman primates. *JCI Insight*. 2022;7(13). doi:10.1172/jci.insight.160039

159. Li D, Martinez DR, Schäfer A, et al. Breadth of SARS-CoV-2 neutralization and protection induced by a nanoparticle vaccine. *Nat Commun*. 2022;13(1). doi:10.1038/s41467-022-33985-4

160. Francica JR, Flynn BJ, Foulds KE, et al. Vaccination with SARS-CoV-2 Spike Protein and AS03 Adjuvant Induces Rapid Anamnestic Antibodies in the Lung and Protects Against Virus Challenge in Nonhuman Primates. *bioRxiv*. Published online 2021. doi:10.1101/2021.03.02.433390

161. Francica JR, Flynn BJ, Foulds KE, et al. Protective antibodies elicited by SARS-CoV-2 spike protein vaccination are boosted in the lung after challenge in nonhuman primates. *Sci Transl Med*. 2021;13(607):1-18. doi:10.1126/scitranslmed.abi4547

162. Corbett KS, Gagne M, Wagner DA, et al. Protection against SARS-CoV-2 Beta variant in mRNA-1273 vaccine–boosted nonhuman primates. *Science (1979)*. 2021;374(6573):1343-1353. doi:10.1126/science.abl8912

163. Corbett KS, Flynn B, Foulds KE, et al. Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. *New England Journal of Medicine*. 2020;383(16):1544-1555. doi:10.1056/nejmoa2024671

164. Joyce MG, King HAD, Elakhal-Naouar I, et al. A SARS-CoV-2 ferritin nanoparticle vaccine elicits protective immune responses in nonhuman primates. *Sci Transl Med*. 2022;14(632):1-16. doi:10.1126/scitranslmed.abi5735

165. Guebre-Xabier M, Patel N, Tian JH, et al. NVX-CoV2373 vaccine protects cynomolgus macaque upper and lower airways against SARS-CoV-2 challenge. *Vaccine*. 2020;38(50):7892-7896. doi:10.1016/j.vaccine.2020.10.064

166. Yu J, Tostanosk LH, Peter L, et al. DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science (1979)*. 2020;369(6505):806-811. doi:10.1126/science.abc6284

167. van Doremalen N, Lambe T, Spencer A, et al. ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature*. 2020;586(7830):578-582. doi:10.1038/s41586-020-2608-y

168. Liu J, Yu J, McMahan K, et al. CD8 T cells contribute to vaccine protection against SARS-CoV-2 in macaques. *Sci Immunol*. 2022;7(77). doi:10.1126/sciimmunol.abq7647

169. Brouwer PJM, Brinkkemper M, Maisonnasse P, et al. Two-component spike nanoparticle vaccine protects macaques from SARS-CoV-2 infection. *Cell*. 2021;184(5). doi:10.1016/j.cell.2021.01.035

170. He WT, Yuan M, Callaghan S, et al. Broadly neutralizing antibodies to SARS-related viruses can be readily induced in rhesus macaques. *Sci Transl Med*. 2022;14(657). doi:10.1126/scitranslmed.abl9605

171. Yadav PD, Ella R, Kumar S, et al. Immunogenicity and protective efficacy of inactivated SARS-CoV-2 vaccine candidate, BBV152 in rhesus macaques. *Nat Commun*. 2021;12(1):1-11. doi:10.1038/s41467-021-21639-w

172. King HAD, Gordon Joyce M, Lakhal-Naouar I, et al. Efficacy and breadth of adjuvanted SARS-CoV-2 receptor-binding domain nanoparticle vaccine in macaques. *Proc Natl Acad Sci U S A*. 2021;118(38):1-11. doi:10.1073/pnas.2106433118

173. Brochu HN, Tseng E, Smith E, et al. Systematic Profiling of Full-Length Ig and TCR Repertoire Diversity in Rhesus Macaque through Long Read Transcriptome Sequencing. *The Journal of Immunology*. 2020;204(12):3434-3444. doi:10.4049/jimmunol.1901256

174. Liston A, Humblet-Baron S, Duffy D, Goris A. Human immune diversity: from evolution to modernity. *Nat Immunol*. 2021;22(12):1479-1489. doi:10.1038/s41590-021-01058-1

175. Quintana-Murci L. Human Immunology through the Lens of Evolutionary Genetics. *Cell*. 2019;177(1):184-199. doi:10.1016/j.cell.2019.02.033

176. Mangino M, Roederer M, Beddall MH, Nestle FO, Spector TD. Innate and adaptive immune traits are differentially affected by genetic and environmental factors. *Nat Commun*. 2017;8(May 2016):1-7. doi:10.1038/ncomms13850

177. Karlsson EK, Kwiatkowski DP, Sabeti PC. Natural selection and infectious disease in human populations. *Nat Rev Genet*. 2014;15(6):379-393. doi:10.1038/nrg3734

178. Souter MNT, Eckle SBG. Biased MAIT TCR Usage Poised for Limited Antigen Diversity? *Front Immunol*. 2020;11(August):12-23. doi:10.3389/fimmu.2020.01845

179. Jakobsson M, Scholz SW, Scheet P, et al. Genotype, haplotype and copy-number variation in worldwide human populations. *Nature*. 2008;451(7181):998-1003. doi:10.1038/nature06742

180. Willcox CR, Pitard V, Netzer S, et al. Cytomegalovirus and tumor stress surveillance by binding of a human γδ T cell antigen receptor to endothelial protein C receptor. *Nat Immunol*. 2012;13(9):872-879. doi:10.1038/ni.2394

181. di Marco Barros R, Roberts NA, Dart RJ, et al. Epithelia Use Butyrophilin-like Molecules to Shape Organ-Specific γδ T Cell Compartments. *Cell*. 2016;167(1):203-218.e17. doi:10.1016/j.cell.2016.08.030

182. Benati D, Galperin M, Lambotte O, et al. Public T cell receptors confer high-avidity CD4 responses to HIV controllers. *J Clin Invest*. 2016;126(6):2093-2108. doi:10.1172/JCI83792

183. Arndt MF, Koohestani DM, Chobrutskiy BI, et al. TRBV and TRBJ usage, when paired with specific HLA alleles, associates with distinct head and neck cancer survival rates. *Hum Immunol*. 2020;81(12):692-696. doi:10.1016/j.humimm.2020.08.007

184. Broughton SE, Petersen J, Theodossis A, et al. Biased T Cell Receptor Usage Directed against Human Leukocyte Antigen DQ8-Restricted Gliadin Peptides Is Associated with Celiac Disease. *Immunity*. 2012;37(4):611-621. doi:10.1016/j.immuni.2012.07.013

185. Cole DK, Yuan F, Rizkallah PJ, et al. Germ line-governed recognition of a cancer epitope by an immunodominant human T-cell receptor. *J Biol Chem*. 2009;284(40):27281-27289. doi:10.1074/jbc.M109.022509

186. Vuorela A, Freitag TL, Leskinen K, et al. Enhanced influenza A H1N1 T cell epitope recognition and cross-reactivity to protein-O-mannosyltransferase 1 in Pandemrix-associated narcolepsy type 1. *Nat Commun*. 2021;12(1). doi:10.1038/s41467-021-22637-8

187. Chen G, Yang X, Ko A, et al. Sequence and Structural Analyses Reveal Distinct and Highly Diverse Human CD8+ TCR Repertoires to Immunodominant Viral Antigens. *Cell Rep*. 2017;19(3):569-583. doi:10.1016/j.celrep.2017.03.072

188. Gold MC, McLaren JE, Reistetter JA, et al. MR1-restricted MAIT cells display ligand discrimination and pathogen selectivity through distinct T cell receptor usage. *Journal of Experimental Medicine*. 2014;211(8):1601-1610. doi:10.1084/jem.20140507

189. Lepore M, Lewinsohn DA, Lewinsohn DM. T cell receptor diversity, specificity and promiscuity of functionally heterogeneous human MR1-restricted T cells. *Mol Immunol*. 2021;130:64-68. doi:10.1016/j.molimm.2020.12.009

190. Gibbs A, Healy K, Kaldhusdal V, et al. Preserved Mucosal-Associated Invariant T Cells in the Cervical Mucosa of HIV-Infected Women with Dominant Expression of the TRAV1-2-TRAJ20 T Cell Receptor α-Chain. *J Infect Dis*. 2022;226(8):1428-1440. doi:10.1093/infdis/jiac171

191. Shi J, Zhou J, Zhang X, et al. Single-Cell Transcriptomic Profiling of MAIT Cells in Patients With COVID-19. *Front Immunol*. 2021;12:700152. doi:10.3389/fimmu.2021.700152

192. Shugay M, Bagaev D v, Zvyagin I v, et al. VDJdb: a curated database of T-cell receptor sequences with known antigen specificity. *Nucleic Acids Res*. 2018;46(D1):D419-D427. doi:10.1093/nar/gkx760

193. Bovay A, Zoete V, Dolton G, et al. T cell receptor alpha variable 12-2 bias in the immunodominant response to Yellow fever virus. *Eur J Immunol*. 2018;48(2). doi:10.1002/eji.201747082

194. Reis BS, Darcy PW, Khan IZ, et al. TCR-Vgd usage distinguishes protumor from antitumor intestinal gd T cell subsets. *Science (1979)*. 2022;377(6603):276-284. doi:10.1126/science.abj8695

195. Mayassi T, Ladell K, Gudjonson H, et al. Chronic Inflammation Permanently Reshapes Tissue-Resident Immunity in Celiac Disease. *Cell*. 2019;176(5):967-981.e19. doi:10.1016/j.cell.2018.12.039

196. Corcoran M, Chernyshev M, Mandolesi M, et al. Archaic humans have contributed to large-scale variation in modern human T cell receptor genes. *bioRxiv*. Published online 2022. doi:10.1101/2022.08.25.505097

197. Dopico XC, Hanke L, Sheward DJ, et al. Probabilistic approaches for classifying highly variable anti-SARS-CoV-2 antibody responses. *medRxiv*. Published online 2021. doi:10.1101/2020.07.17.20155937

198. Long QX, Liu BZ, Deng HJ, et al. Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat Med*. 2020;26(6):845-848. doi:10.1038/s41591-020-0897-1

199. Yu J, Tostanosk LH, Peter L, et al. DNA vaccine protection against SARS-CoV-2 in rhesus macaques. *Science (1979)*. 2020;369(6505):806-811. doi:10.1126/science.abc6284

200. Suthar MS, Zimmerman MG, Kauffman RC, et al. Rapid Generation of Neutralizing Antibody Responses in COVID-19 Patients. *Cell Rep Med*. 2020;1(3):100040. doi:10.1016/j.xcrm.2020.100040

201. Wang H, Zhang Y, Huang B, et al. Development of an Inactivated Vaccine Candidate, BBIBP-CorV, with Potent Protection against SARS-CoV-2. *Cell*. 2020;182(3):713-721.e9. doi:10.1016/j.cell.2020.06.008

202. Kreer C, Zehner M, Weber T, et al. Longitudinal Isolation of Potent Near-Germline SARS-CoV-2-Neutralizing Antibodies from COVID-19 Patients. *Cell*. 2020;182(4):843-854.e12. doi:10.1016/j.cell.2020.06.044

203. Ju B, Zhang Q, Ge J, et al. Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature*. 2020;584(7819):115-119. doi:10.1038/s41586-020-2380-z

204. Guebre-Xabier M, Patel N, Tian JH, et al. NVX-CoV2373 vaccine protects cynomolgus macaque upper and lower airways against SARS-CoV-2 challenge. *Vaccine*. 2020;38(50):7892-7896. doi:10.1016/j.vaccine.2020.10.064

205. Francica JR, Flynn BJ, Foulds KE, et al. Protective antibodies elicited by SARS-CoV-2 spike protein vaccination are boosted in the lung after challenge in nonhuman primates. *Sci Transl Med*. 2021;13(607):1-18. doi:10.1126/scitranslmed.abi4547

206. Saunders KO, Lee E, Parks R, et al. Neutralizing antibody vaccine for pandemic and pre-emergent coronaviruses. *Nature*. 2021;594(7864):553-559. doi:10.1038/s41586-021-03594-0

207. Vogel AB, Kanevsky I, Che Y, et al. BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature*. 2021;592(7853):283-289. doi:10.1038/s41586-021-03275-y

208. Mercado NB, Zahn R, Wegmann F, et al. Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature*. 2020;586(7830):583-588. doi:10.1038/s41586-020-2607-z

209. Corbett KS, Flynn B, Foulds KE, et al. Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. *New England Journal of Medicine*. 2020;383(16):1544-1555. doi:10.1056/nejmoa2024671

210. Gao Q, Bao L, Mao H, et al. Development of an inactivated vaccine candidate for SARS-CoV-2. *Science (1979)*. 2020;369(6499):77-81. doi:10.1126/science.abc1932

211. Zhou D, Dejnirattisai W, Supasa P, et al. Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell*. 2021;184(9):2348-2361.e6. doi:10.1016/j.cell.2021.02.037

212. Tegally H, Wilkinson E, Giovanetti M, et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature*. 2021;592(7854):438-443. doi:10.1038/s41586-021-03402-9

213. Wang P, Nair MS, Liu L, et al. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature*. 2021;593(7857):130-135. doi:10.1038/s41586-021-03398-2

214. Davis C, Logan N, Tyson G, et al. Reduced neutralisation of the Delta (B.1.617.2) SARS-CoV-2 variant of concern following vaccination. *PLoS Pathog*. 2021;17(12):1-15. doi:10.1371/journal.ppat.1010022

215. Liu C, Ginn HM, Dejnirattisai W, et al. Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. *Cell*. 2021;184(16):4220-4236.e13. doi:10.1016/j.cell.2021.06.020

216. Madhi SA, Baillie V, Cutland CL, et al. Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *New England Journal of Medicine*. 2021;384(20):1885-1898. doi:10.1056/nejmoa2102214

217. Supasa P, Zhou D, Dejnirattisai W, et al. Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. *Cell*. 2021;184(8):2201-2211.e7. doi:10.1016/j.cell.2021.02.033

218. Shinde V, Bhikha S, Hoosain Z, et al. Efficacy of NVX-CoV2373 Covid-19 Vaccine against the B.1.351 Variant. *New England Journal of Medicine*. 2021;384(20):1899-1909. doi:10.1056/nejmoa2103055

219. Hu J, Peng P, Wang K, et al. Emerging SARS-CoV-2 variants reduce neutralization sensitivity to convalescent sera and monoclonal antibodies. *Cell Mol Immunol*. 2021;18(4):1061-1063. doi:10.1038/s41423-021-00648-1

220. Planas D, Veyer D, Baidaliuk A, et al. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature*. 2021;596(7871):276-280. doi:10.1038/s41586-021-03777-9

221. Elliott P, Haw D, Wang H, et al. Exponential growth, high prevalence of SARS-CoV-2, and vaccine effectiveness associated with the Delta variant. *Science (1979)*. 2021;374(6574). doi:10.1126/science.abl9551

222. Lessler J, Riley S, Read JM, et al. Evidence for antigenic seniority in influenza A (H3N2) antibody responses in southern China. *PLoS Pathog*. 2012;8(7):e1002802. doi:10.1371/journal.ppat.1002802

223. Gostic KM, Ambrose M, Worobey M, Lloyd-Smith JO. Potent protection against H5N1 and H7N9 influenza via childhood hemagglutinin imprinting. *Science (1979)*. 2016;354(6313):722-726. doi:10.1126/science.aag1322

224. Monto AS, Malosh RE, Petrie JG, Martin ET. The Doctrine of Original Antigenic Sin: Separating Good From Evil. *J Infect Dis*. 2017;215(12):1782-1788. doi:10.1093/infdis/jix173

225. Wu K, Choi A, Koch M, et al. Variant SARS-CoV-2 mRNA vaccines confer broad neutralization as primary or booster series in mice. *Vaccine*. 2021;39(51):7394-7400. doi:10.1016/j.vaccine.2021.11.001

226. Pavot V, Berry C, Kishko M, et al. Protein-based SARS-CoV-2 spike vaccine booster increases cross-neutralization against SARS-CoV-2 variants of concern in non-human primates. *Nat Commun*. 2022;13(1). doi:10.1038/s41467-022-29219-2

227. Corbett KS, Gagne M, Wagner DA, et al. Protection against SARS-CoV-2 Beta variant in mRNA-1273 vaccine–boosted nonhuman primates. *Science (1979)*. 2021;374(6573):1343-1353. doi:10.1126/science.abl8912

228. Sui Y, Li J, Andersen H, et al. An intranasally administrated SARS-CoV-2 beta variant subunit booster vaccine prevents beta variant replication in rhesus macaques. *PNAS Nexus*. 2022;1(3):1-11. doi:10.1093/pnasnexus/pgac091

229. Galson JD, Trück J, Clutterbuck EA, et al. B-cell repertoire dynamics after sequential hepatitis B vaccination and evidence for cross-reactive B-cell activation. *Genome Med*. 2016;8(1):1-13. doi:10.1186/s13073-016-0322-z

230. Ehrhardt SA, Zehner M, Krähling V, et al. Polyclonal and convergent antibody response to Ebola virus vaccine rVSV-ZEBOV. *Nat Med*. 2019;25(10):1589-1600. doi:10.1038/s41591-019-0602-4

231. Zhang Y, Yan Q, Luo K, et al. Analysis of B Cell Receptor Repertoires Reveals Key Signatures of the Systemic B Cell Response after SARS-CoV-2 Infection. *J Virol*. 2022;96(4). doi:10.1128/jvi.01600-21

232. Robbiani DF, Gaebler C, Muecksch F, et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature*. 2020;584(7821):437-442. doi:10.1038/s41586-020-2456-9

233. Claireaux M, Caniels TG, de Gast M, et al. A public antibody class recognizes an S2 epitope exposed on open conformations of SARS-CoV-2 spike. *Nat Commun*. 2022;13(1):1-15. doi:10.1038/s41467-022-32232-0

234. Sakharkar M, Rappazzo G, Wieland-Alter WF, et al. Prolonged evolution of the human B cell response to SARS-CoV-2 infection. *Sci Immunol*. 2021;6(16):1-15. doi:10.1126/SCIIMMUNOL.ABG6916

235. Yuan M, Huang D, Lee CCD, et al. Structural and functional ramifications of antigenic drift in recent SARS-CoV-2 variants. *Science (1979)*. 2021;373(6556):818-823. doi:10.1126/science.abh1139

236. Wang Y, Yuan M, Lv H, Peng J, Wilson IA, Wu NC. A large-scale systematic survey reveals recurring molecular features of public antibody responses to SARS-CoV-2. *Immunity*. 2022;55(6):1105-1117.e4. doi:10.1016/j.immuni.2022.03.019

237. Pushparaj P, Nicoletto A, Sheward DJ, et al. Immunoglobulin germline gene polymorphisms influence the function of SARS-CoV-2 neutralizing antibodies Immunoglobulin germline gene polymorphisms influence the function of SARS-CoV-2 neutralizing antibodies. Published online 2023:1-14. doi:10.1016/j.immuni.2022.12.005

238. Luo K, Liao HX, Zhang R, et al. Tissue memory B cell repertoire analysis after ALVAC/AIDSVAX B/E gp120 immunization of rhesus macaques. *Journal of Clinical Investigation*. 2016;1(20):1-17. doi:10.1172/jci.insight.88522

239. Lindsay KE, Bhosle SM, Zurla C, et al. Visualization of early events in mRNA vaccine delivery in non-human primates via PET–CT and near-infrared imaging. *Nat Biomed Eng*. 2019;3(5):371-380. doi:10.1038/s41551-019-0378-3