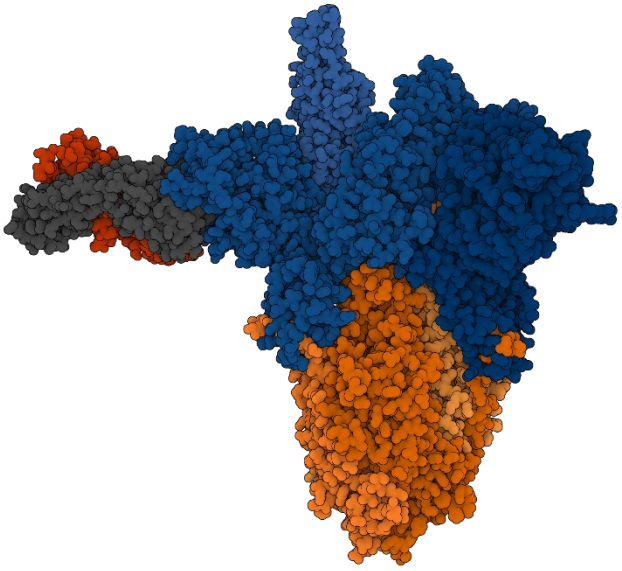
# Neutralising antibody responses against HIV-1 and SARS-CoV-2

Graham, Carl

*Awarding institution:*
King's College London

# Neutralising antibody responses against HIV-1 and SARS-CoV-2.

Carl Graham



A thesis submitted to the University of London for the degree of Doctor of Philosophy

King's College London
Department of Infectious Diseases
School of Immunology & Microbial Sciences

# Declaration

I, Carl Graham, confirm that the work presented in my thesis is my own. Where information has been derived from other sources, I confirm this has been indicated in the thesis. London, 10/07/2023.

# Abstract

Understanding how viruses and antibodies evolve in response to one another can help inform future vaccines and therapies. This project investigated longitudinal samples from individuals infected with HIV-1 and SARS-CoV-2, the causative viruses of AIDS and COVID-19 respectively.

Multiple lineages of antibodies that can neutralise a wide range of HIV-1 variants have been previously isolated. They target conserved epitopes on the HIV-1 env glycoprotein, the only virally encoded antigen presented on the viral surface. Understanding how bnAbs develop during natural infection could inform future vaccination strategies. Initially, this project sought to isolate bnAbs from a HIV-1 infected individual across multiple timepoints, which required the optimization of numerous methods. These included biotinylated antigenic bait production, PBMC staining, FACS sorting and a function monoclonal antibody library generation pipeline. During this project, antibodies capable of cross-binding the gp120s of two HIV-1 global panel viruses were isolated. Through site-directed mutagenesis and other molecular cloning techniques, insights were gained into the role of the V3 and V1 loops in viral escape in donor SJU.

With the arrival of the COVID-19 pandemic, the focus of this project was shifted to the antibody response to SARS-CoV-2. As such, many of the aforementioned methods were updated for work with the novel virus. Next, the polyclonal neutralising antibody responses of subjects infected with SARS-CoV-2 was investigated longitudinally. This revealed that neutralising titres sharply declined throughout the first 3-months following infection and that disease severity had an impact on the magnitude of the response. To understand these responses at a monoclonal level, B-cells were sorted from three SARS-CoV-2 infected donors and a library of 108 monoclonal antibodies was generated. Characterization of this library revealed that Spike specific antibodies had low levels of somatic hypermutation and that neutralisers could be separated into 7 competition groups. Further, it was shown that particular competition groups were unable to neutralise the B.1.1.7 variant of concern as a result of specific mutations within the receptor binding domain and NTD. To ascertain whether vaccination with ChAdOx1-nCoV resulted in a similar antibody response to natural infection, 44 SARS-CoV-2 specific monoclonal antibodies were isolated from a single donor. All 7 competition groups were found to be present and an additional NTD-specific group was discovered. Despite low levels of plasma neutralisation, the monoclonal antibodies isolated form this subject showed higher levels of SHM, increased potency and better breadth (against a panel of 5 variants of concern) compared to natural infection monoclonal antibodies. IgG spike specific B-cells were still measurable in this donor 9-months post-boost, at which time plasma neutralisation was undetectable. The conclusions of this work have highlighted the importance monitoring antibody and B-cell decline following SARS-CoV-2 infection and vaccination. Furthermore, it has highlighted sites of vulnerability on the SARS-CoV-2 spike and how these are impacted by viral escape mutations.

# Preface

This thesis is comprised of 6 chapters. Chapter 1 is a comprehensive literary review of the scientific topics fundamental to the thesis and outlines the objectives of the research presented. Chapters 4 and 5 are revised versions of manuscripts that have been published in scientific journals, of which the candidate is the co-first author. The contributions of all authors on these manuscripts are listed in the "Contribution of authors" section. These chapters have their own introduction, results and discussion sections, whilst details on the materials and methods can be found in Chapter 2. The results presented in this thesis are summarised and discussed in Chapter 6. The references for Chapters 1-6 are listed at the end of the thesis.

# Contribution of authors

## Appearing in Chapter 4

"Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans."

Jeffrey Seow*, **Carl Graham***, Blair Merrick*, Sam Acors*, Suzanne Pickering, Kathryn J. A. Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Rui Pedro Galao, Gilberto Betancor, Harry D. Wilson, Adrian W. Signell, Helena Winstone, Claire Kerridge, Isabella Huettner, Jose M. Jimenez-Guardeño, Maria Jose Lista, Nigel Temperton, Luke B. Snell, Karen Bisnauthsing, Amelia Moore, Adrian Green, Lauren Martinez, Brielle Stokes, Johanna Honey, Alba Izquierdo-Barras, Gill Arbane, Amita Patel, Mark Kia Ik Tan, Lorcan O'Connell, Geraldine O'Hara, Eithne MacMahon, Sam Douthwaite, Gaia Nebbia, Rahul Batra, Rocio Martinez-Nunez, Manu Shankar-Hari, Jonathan D. Edgeworth, Stuart J. D. Neil, Michael H. Malim and Katie J. Doores.

Published in *Nature microbiology* (October 2021).

K.J.D., B.M., S.J.D.N., M.H.M. and J.D.E. designed the study. J.S., **C.G**., S.A., K.J.A.S., O.H., A.O'B., N.K., I.H. and K.J.D. performed ELISAs. J.S., **C.G**. and S.P. performed neutralization assays. H.W., C.K., R.M.N., M.J.L. and J.M.J.-G. prepared pseudovirus or wild-type virus. J.S., **C.G**., B.M., S.A., K.J.D., M.H.M. and S.J.D.N. analysed and interpreted the data. S.P., R.P.G., G.B., H.D.W. and A.W.S. curated hospital serum samples. B.M., K.B., A.P., M.K.I.T., L.O.C., G.O'H., E.M., S.D., G.N. and R.B. assisted in collection of samples from hospitalized patients. B.M., L.S., K.B., A.M., A.G., L.M., B.S., J.H., A.I.-B., G.A., A.P. and R.B. assisted in collection of samples from HCWs. G.O'H., E.M., S.D. and G.N. assisted in project administration. N.T. provided new reagents. K.J.D., M.H.M., S.J.D.N., B.M., J.D.E., J.S., **C.G**., S.P., R.P.G. and M.S.H. drafted the manuscript or substantially revised it.

## Appearing in Chapter 5

"Neutralization potency of monoclonal antibodies recognizing dominant and subdominant epitopes on SARS-CoV-2 Spike is impacted by the B.1.1.7 variant."

**Carl Graham***, Jeffrey Seow*, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose M. Jimenez-Guardeño, Adam G. Laing, Yin Wu, Magdalene Joseph, Luke Muir, Marit J. van Gils, Weng M. Ng, Helen M.E. Duyvesteyn, Yuguang Zhao, Thomas A. Bowden, Manu Shankar-Hari, Annachiara Rosa, Peter Cherepanov, Laura E. McCoy, Adrian C. Hayday, Stuart J.D. Neil, Michael H. Malim, and Katie J. Doores.

Published in *Immunity* (June 2021).

Conceptualization, K.J.D., M.H.M., **C.G**., J.S.; Methodology, **C.G**., J.S., I.H., H.K., N.K., S.A., H.W., S.P., R.P.G, L.D., M.J.L., J.M.J.-G., L.M., W.M.N., H.M.E.D., Y.Z., T.A.B.; Formal analysis, K.J.D., **C.G**., J.S., I.H., H.K., W.N.G., T.A.B.; Resources, A.G.L., Y.W., M.J., M.S.-H., A.R., P.C., L.E.M., A.C.H., M.J.v.G.; Supervision, K.J.D., M.H.M., S.J.D., A.C.H., L.E.M., P.C.; Writing – original draft, K.J.D., **C.G**., J.S., I.H., H.K., M.H.M.; All authors reviewed and edited the manuscript.

"ChAdOx1 nCoV-19 vaccine elicits monoclonal antibodies with cross-neutralizing activity against SARS-CoV-2 viral variants."

Jeffrey Seow*, **Carl Graham***, Sadie R. Hallett, Thomas Lechmere, Thomas J.A. Maguire, Isabella Huettner, Daniel Cox, Hataf Khan, Suzanne Pickering, Rebekah Roberts, Anele Waters, Christopher C. Ward, Christine Mant, Michael J. Pitcher, Jo Spencer, Julie Fox, Michael H. Malim, and Katie J. Doores.

Published in *Cell reports* (May 2022).

Conceptualization, K.J.D., J. Seow, and **C.G**.; Methodology, J. S, **C.G**., S.R.H., T.L., T.J.A.M., I.H., D.C., H.K., and S.P.; Formal analysis, K.J.D., J. Seow, **C.G**., S.R.H., H.K., I.H., and M.J.P.; Resources, R.R., A.W., C.C.W., C.M., M.J.P., and J. Spencer; Supervision, K.J.D., J.F., M.H.M., and J. Spencer; Writing – original draft, K.J.D., J. Seow, and **C.G**. All authors re-viewed and edited the manuscript.

# Table of contents

# Index of figures

# Index of tables

# Abbreviation

## *A*

Ammonium–chloride–potassium
(ACK)
Acquired immune deficiency syndrome
(AIDS)
Activation-induced cytidine deaminase
(AID)
Adenovirus early genes
(E1 and E3)
Alphacoronavirus
(Alpha-CoV)
Alveolar type 1
(AT1)
Alveolar type 2
(AT2)
Amino-terminal domain
(NTD)
Angiotensin-converting enzyme 2
(ACE2)
Antibody-dependent cellular cytotoxicity
(ADCC)
Antibody-dependent cellular phagocytosis
(ADCP)
Antiretroviral therapy
(ART)
Apurinic/apyrimidinic endonuclease
(APE)
American Type Culture Collection
(ATCC)

## *B*

Base excision repair
(BER)
B-cell receptor
(BCR)
Betacoronavirus
(Beta-CoV)
Broadly neutralising antibodies
(bnAbs)
Bruton's tyrosine kinase
(Btk)

## *C*

Capsid
(CA)
C-C chemokine receptor
(CCR)
CD4 binding site
(CD4bs)
Circulating recombinant form
(CRF)
Class-switch recombination
(CSR)
Cluster of differentiation
(CD)
Column volume
(CV)
Common lymphoid progenitor
(CLP)
Common myeloid progenitor
(CMP)
Complement component 1q
(C1q)
Complementarity-determining region
(CDR)
Complement-dependent cytotoxicity
(CDC)
Cryogenic electron microscopy
(cryo-EM)
Coronavirus disease 2019
(COVID-19)
C-X-C motif chemokine ligand
(CXCL)
CXC-chemokine receptor
(CXCR)
Cytoplasmic tail
(CT)

## D

Deltacoronavirus
(Delta-CoV)
Deoxyribonucleic acid
(DNA)
Dimethylsulfoxide
(DMSO)
Dithiothreitol
(DTT)
Diversity gene
(D)
DNA-dependent protein kinase
(DNA-PK)
DNA-dependent protein kinase catalytic subunit
(DNA-PKcs)
Double distilled water
(ddH2O)
Double-membrane vesicle
(DMV)
Double stranded break
(DSB)
Dulbecco's modified eagle medium
(DMEM)
Dulbecco's phosphate-buffered saline
(DPBS)

## E

Ethylenediaminetetraacetic acid
(EDTA)
Endosomal sorting complex required for transport
(ESCRT)
Envelope
(E) - SARS-CoV-2
(*env*/Env; gene/protein) - HIV-1
Enzyme-linked immunosorbent assay
(ELISA)
Escherichia coli
(E. coli)
Exonuclease-1
(Exo1)
Extracorporeal membrane oxygenation
(ECMO)

## F

Fc-alpha receptor
(FcαR)
Fc-epsilon receptor
(FcεR)
Fc-gamma receptor
(FcγR)
Fluorescence-activated cell sorting
(FACS)
fms-like tyrosine kinase 3
(FLT3)
Foetal bovine serum
(FBS)
Follicular dendritic cell
(FDC)
Fraction of inspired oxygen
($FiO_2$)
Fragment antigen binding
(Fab)
Fragment crystallisable
(Fc)
Frame-work region
(FR)
Fusion peptide
(FP)

## G

Gammacoronavirus
(Gamma-CoV)
Group specific antigen
(*gag*/Gag; gene/protein)
Geometric mean titre
(GMT)
Germinal centre
(GC)
Glycoprotein 120
(gp120)
Glycoprotein 41
(gp41)
Granulocyte-macrophage progenitor

(GMP)
Guy's and St Thomas' NHS Foundation Trust
(GSTFT)

## *H*

Healthcare worker
(HCW)
Heat inactivated
(HI)
Hematopoietic stem cell
(HSC)
Heparin-induced thrombocytopenia
(HIT)
Heptad repeat
(HR)
High mobility group box 1
(HMGB1)
High performance liquid chromatography
(HPLC)
HIV-1 env constant regions
(C1-C5)
HIV-1 variable env regions
(V1-V5)
Horseradish peroxidase
(HRP)
Human coronavirus NL63
(HCoV-NL63)
Human coronavirus-229E
(HCoV-229E)
Human coronavirus-HKU1
(HCoV-HKU1)
Human coronavirus-OC43
(HCoV-OC43)
Human immunodeficiency virus
(HIV)

## *I*

Immunoglobulin G-degrading enzyme of *Streptococcus pyogenes*
(IdeS)
Immunoglobulin surrogate light chain proteins
(λ5 and VpreB)
ImmunoGenetics
(IMGT)
Immunoglobulin
(Ig)
Immunoglobulin constant domains
($CH_{1\text{-}3/4}$ and $C_L$)
Immunoglobulin heavy chain
($Ig_H$)
Immunoglobulin light chain
($Ig_L$)
Immunoglobulin light chain kappa
($Ig_L\kappa$)
Immunoglobulin light chain lambda
($Ig_L\lambda$)
Immunoglobulin signal transducing proteins
(Igα and Igβ)
Immunoglobulin variable domains
($V_H$ and $V_L$)
Inhibitory concentration for 50% neutralisation
($IC_{50}$)
Inhibitory dilution for 50% neutralisation
($ID_{50}$)
Insertion/deletion
(indel)
Integrase
(IN)
Interferon-gamma
(IFN-γ)
Interleukin
(IL)
Intervening exon
(I-exon)

## *J*

Janus kinase
(JAK)
Joining gene
(J)
Joining-chain
(J-chain)

## *L*

Lipid nanoparticle
(LNP)
Lipopolysaccharide
(LPS)
Long terminal repeat
(LTR)
Long-term hematopoietic stem cell
(LT-HSC)
Luciferase
(Luc)
Lysogeny broth
(LB)

## *M*

Major histocompatibility complex
(MHC)
Matrix
(MA)
Megakaryocyte-erythrocyte progenitor
(MEP)
Membrane
(M)
Membrane proximal external region
(MPER)
Men who have sex with men
(MSM)
Middle East respiratory syndrome coronavirus
(MERS-CoV)
Mismatch repair
(MMR)
Milli-Q water
(MQ)
MLH1/PMS2 complex
(MutLα)
Monoclonal antibodies
(mAbs)
Messenger ribonucleic acid
(mRNA)
MSH2-MSH6 heterodimer
(MutSα)
Multipotent progenitor
(MPP)

## *N*

Native flexibly linked
(NFL)
Natural killer
(NK)
Negative factor
(*nef*/Nef; gene/protein)
Non-homologous end joining
(NHEJ)
Non-nucleoside reverse transcriptase inhibitor
(NNRTI)
Non-template-nucleotides
(N-nucleotides)
Non-structural protein
(nsp)
Nuclear pore complexe
(NPC)
Nucleocapsid
(N) - SARS-CoV-2
(NC) - HIV-1
Nucleoside reverse transcriptase inhibitor
(NRTI)

## *O*

Open reading frame
(ORF)
Optical density
(OD)

## *P*

Palindromic-nucleotide
(P-nucleotide)
Peripheral blood mononuclear cell
(PBMC)
Phosphate-buffered saline
(PBS)
Piperazine-N,N'-bis(2-ethanesulfonic acid
(PIPES)
Poly(ADP-ribose) polymerase 1
(PARP1)
Polyethylene glycol

(PEG)
Polyethyleniminely
(PEI)
Polymerase
(*pol*/Pol; gene/protein)
Polymerase chain reaction
(PCR)
Polynucleotide kinase
(PNK)
Polypurine tract
(ppt)
Post-onset of symptoms
(POS)
Potential N-glycosylation sites
(PNGS)
Pre-B-cell receptor
(pre-BCR)
Pre-integration complex
(PIC)
Proliferating cell nuclear antigen
(PCNA)
Protease
(PR)

## *R*

Real-time quantitative PCR
(RT-qPCR)
Receptor binding domain
(RBD)
Receptor binding motif
(RBM)
Recombination signal sequence
(RSS)
Recombination signal sequence with 12 base pair linker
(12RSS)
Recombination signal sequence with 23 base pair linker
(23RSS)
Recombination-activating gene
(RAG)
Regulator of expression of virion proteins
(*rev*/Rev; gene/protein)
Regulatory T-cell
($T_{reg}$)
Replication and transcription complex
(RTC)
Reverse transcriptase
(RT)
Reverse transcription complex
(RTC)
Reverse transcription polymerase chain reaction
(RT-PCR)
Revolutions Per Minute
(rpm)
Ribonucleoprotein
(RNP)
RNA-dependent RNA polymerase
(RdRp)
Roswell park memorial institute
(RPMI)

## *S*

SARS-CoV-2 Spike subdomains
(SD1 and SD2)
SARS-related coronavirus
(SARSr-CoV)
Secondary lymphoid organs
(SLO)
Severe acute respiratory syndrome coronavirus
(SARS-CoV)
Severe acute respiratory syndrome coronavirus 2
(SARS-CoV-2)
Short pulse antiretroviral therapy at seroconversion
(SPARTAC)
Short-term hematopoietic stem cell
(ST-HSC)
Shrimp alkaline phosphatase
(rSAP)
Simian immunodeficiency virus
(SIV)
Simian virus 40
(SV40)
Single genome amplification

(SGA)
Single-membrane vesicle
(SMV)
Sodium dodecyl-sulfate polyacrylamide gel electrophoresis
(SDS-PAGE)
SJU027003
(SJU)
Somatic hyper mutation
(SHM)
Spike
(S)
Stem-cell factor
(SCF)
Subgenomic RNA
(sgRNA)

## *T*

Tangential flow filtration
(TFF)
T-cell independent
(TI)
T-cell receptor
(TCR)
Terminal deoxynucleotidyl-transferase
(TdT)
T-follicular helper cell
($T_{FH}$ cell)
Tetramethylethylenediamine
(TEMED)
Timepoint
(TP)
Toll-like receptor
(TLR)
Trans-activator of transcription
(*tat*/Tat; gene/protein)
Transcription regulatory sequence
(TRS)
Transcription regulatory sequence leader sequence
(TRS-L)
Translesion synthesis
(TLS)
Transmembrane domain
(TMD)
Transmembrane protease serine 2
(TMPRSS2)
Type-I interferons
(IFN-I)
Type 2 respiratory failure
(T2RF)

## *U*

Uncleaved pre-fusion-optimised
(UFO)
Unmutated common ancestor
(UCA)
Untranslated region
(UTR)
Uracil-DNA-glycosylase
(UNG)
Ultraviolet
(UV)

## *V*

V3-loop fused to non-human IgG1 Fc region
(V3-Fc)
Variable gene
(V)
Variant of interest
(VOI)
Variant of concern
(VOC)
Vascular cell adhesion molecule 1
(VCAM-1)
Very late antigen-4
(VLA-4)
Viral protein R
(*vpr*/Vpr; gene/protein)
Viral protein U
(*vpu*/Vpu; gene/protein)
Virion infectivity factor
(*vif*/Vif; gene/protein)

## *W*

Wild type
(WT)
World Health Organisation
(WHO)

## *X*

X-ray repair cross complementing
(XRCC)

# Chapter 1: introduction

## 1.1 Components of the immune system

### 1.1.1 Haematopoiesis

In the human body, blood cells are made and consistently replenished through a process called haematopoiesis. The haematopoietic system is generally considered to be hierarchical in nature, ultimately producing distinct specialised cells which perform a wide range of important functions (Figure 1.1) [1].

All blood cellular components are produced from a common precursor known as the hematopoietic stem cell (HSC). In adults, HSCs are found in the bone marrow and are capable of self-renewal and

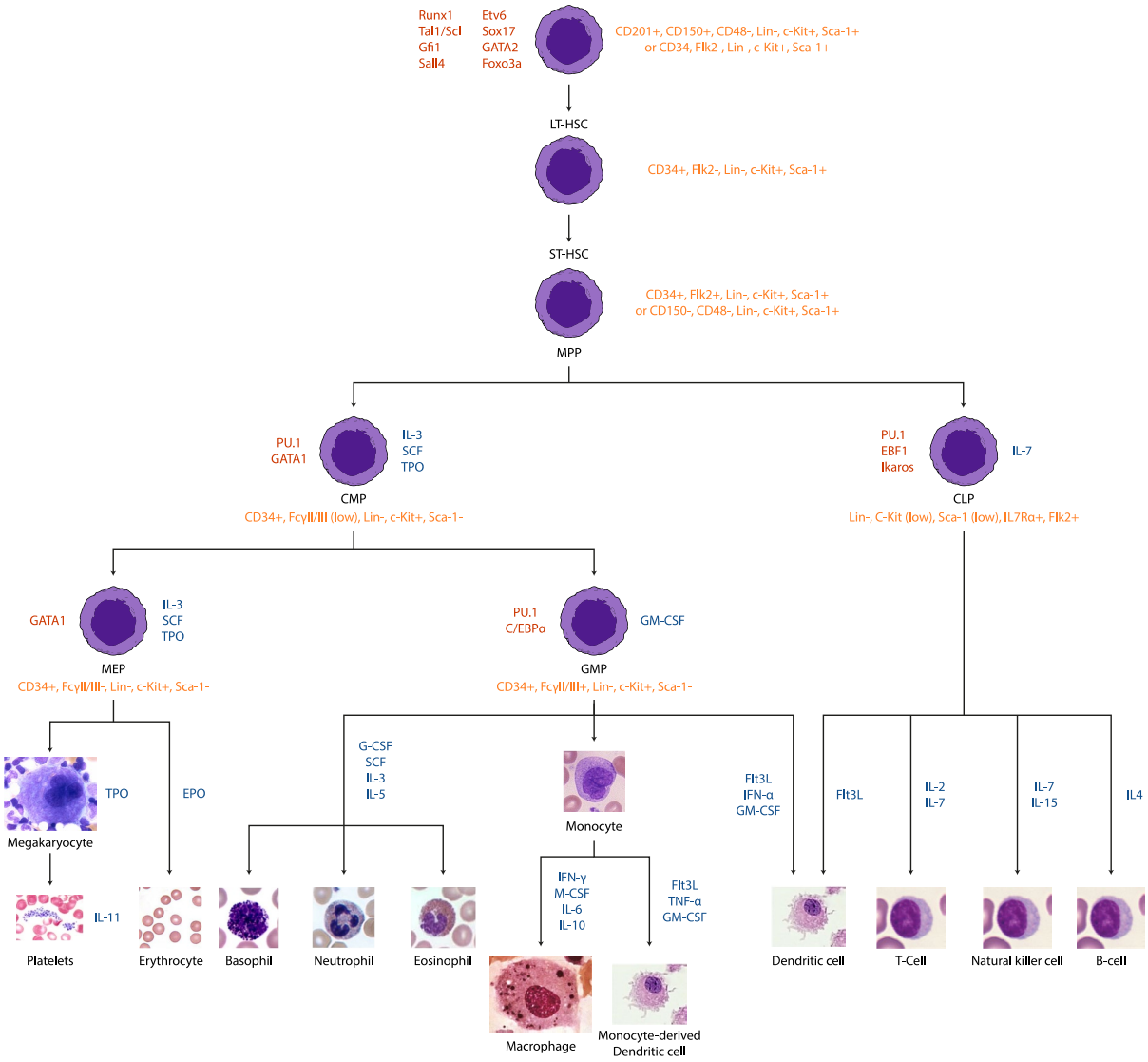



*Figure 1.1: Classical model of Haematopoiesis.*
*Shown here is the classical model of the haematopoiesis hierarchy. LT-HSCs differentiate into ST-HSCs, and then into MPPs with reduced self-renewal ability. MPPs differentiate into either CMPs or CLPs, in the first lineage commitment. CMPs can give rise to MEPs or GMPs, whilst CLPs generate lympthocytes and dendritic cells. MEPs differentiate into megakaryocytes/platelets and erythrocytes. GMPs become granulocytes, macrophages and dendritic cells. Haemopoietic differentiation is controlled by extrinsic cytokines (dark blue) and intrinsic transcription factors (red), the different cell populations are distinguishable by surface markers (orange). Adapted from [7].*

multipotent differentiation [2]. HSCs can be separated into two sub-populations based on their expression of surface marker cluster of differentiation (CD) 34. Long-term HSCs (LT-HSC) are $CD34^+$, whilst short-term HSCs (ST-HSC) are $CD34^-$ [3]. LT-HSCs maintain self-renewal capability for longer (>3-4 months) compared to ST-HSCs (<1 month), but are rarer and quiescent.

LT-HSCs give rise to ST-HSCs, which further differentiate into cells that lack reconstitution ability called multipotent progenitors (MPP) [4]. MMPs can differentiate into either the common myeloid progenitors (CMP) or common lymphoid progenitors (CLP). CMPs can differentiate into granulocyte-macrophage progenitors (GMP) or megakaryocyte-erythrocyte progenitors (MEP), whilst CLPs can give rise to B-cells, T-cells, natural killer (NK) cells and dendritic cells. GMPs differentiate into basophils, neutrophils, eosinophils, monocytes (which mature into macrophages or monocyte-derived dendritic cells) and dendritic cells. Lastly, MEPs can give rise to megakaryocytes (responsible for producing platelets) or erythrocytes.

This process of sequential differentiation from LT-HSCs to mature blood cells is controlled by a complex orchestration of intrinsic transcription factors and extrinsic cytokines [5,6]. Advancements in single-cell technologies have enabled amendments to be made to this classic model of haematopoiesis, however they will not be discussed here [7].

### 1.1.2 Innate vs adaptive immunity

The immune system can be broadly considered to consist of two lines of defence, innate and adaptive immunity (Table 1.1). Innate immunity is the first line of defence against pathogens, acting immediately or within hours of exposure via non-specific mechanisms. The innate immune response does not generate immunological memory, and would not recognise the same antigen in the event of a subsequent exposure. Adaptive immunity acts as the second line of defence against pathogens, reacting to antigens in a specific manner but requiring a lapse in time from exposure to maximal response. Adaptive immunity is able to generate immunological memory, allowing a fast and effective response to a second antigen exposure. The mechanisms of innate and adaptive immunity are not mutually exclusive and instead work in tandem to efficiently tackle infection [8].

The cellular mediators of the innate immune response are phagocytes, dendritic cells, mast cells, basophils, eosinophils and NK cells (Table 1.2). Phagocytes include neutrophils and macrophages, both of which engulf their targets and eliminate them via bactericidal pathways. Neutrophils are short-lived cells, whilst macrophages are long-lived cells. Neutrophils also have granules and enzymatic pathways which help to protect the host from pathogenic microbes. In addition to phagocytosis, macrophages can present antigens to other members of the immune system. Similarly, dendritic cells can also phagocytose pathogens and present them to other cells. Mast cells and basophils have many features in common, both functioning to initiate acute inflammatory responses, for example allergic and

*Table 1.1: Comparing features of innate and adaptive immunity.*
*Comparison of characteristics displayed by the innate and adaptive arms of the immune system. Adapted from [627].*

| | **Innate** | **Adaptive** |
|---|---|---|
| *Features* | Primitive and broad | Highly specific |
| *Speed of onset* | Immediate | ~3 day lag |
| *Potency* | Lower | Higher |
| *Kinetics* | Fast (hours-days) | Slow (days-weeks) |
| *Amplification* | No (insignificant) | Yes |
| *Duration* | Short (days) | Long (months/years) |
| *Memory* | No | Yes |
| *Activity* | Always present | Normally silent |
| *Specificity* | Unspecific | Highly specific |

asthmatic reactions. Additionally, mast cells are known to be early generators of cytokines as a result of injury or infection. Mast cells are usually located within the connective tissue around blood vessels, whilst basophils are found in the circulation. Eosinophils are granulocytes capable of phagocytosis and are important in the defence against parasites that are generally too large to be engulfed. Further, eosinophils are also involved in acute inflammatory responses. NK cells are important for the

*Table 1.2: List of immune cell characteristics.*
*Shown here are the key functions, lifetimes and main targets of the immune cells. B-cells are not included, as they are the topic of 1.2. Adapted from [8].*

| Cell type | % in adults | Functions | Lifetime | Main target |
|---|---|---|---|---|
| **Macrophage** | Varies | • Phagocytosis<br>• Antigen presentation<br>• Degranulation | Month-years | Various |
| **Neutrophil** | 40-75% | • Phagocytosis<br>• Degranulation | 6 hours-few days | • Bacteria<br>• Fungi |
| **Eosinophil** | 1-6% | • Degranulation<br>• Release of enzymes, growth factors and cytokines | 8-12 days (circulate for 4-5 hours) | • Parasites<br>• Various allergic tissues |
| **Basophil** | <1% | • Degranulation<br>• Release of histamines, enzymes and cytokines | Likely a few hours-few days | Various allergic tissues |
| **Mast cell** | Common in tissues | • Degranulation<br>• Release of histamines, enzymes and cytokines | Months-years | • Parasites<br>• Various allergic tissues |
| **T-cell** | 20-40% | • T-cell (CD4+): immune response mediators<br>• T-cell (CD8+): cell destruction | Weeks-years | • T-cell (CD4+): intracellular bacteria<br>• T-cell (CD8+): virus infected cells and tumour cells |
| **Monocyte** | 2-6% | Differentiate into macrophages and dendritic cells to elicit immune responses | Hours-days | Various |
| **NK cell** | 15% (varies) | • Tumour rejection<br>• Destruction of infected cells<br>• Release of perforin and granzymes which induce apoptosis | 7-10 days | • Virus infected cells<br>• Tumour cells |

elimination of cells infected by viruses and the rejection of tumours. To destroy infected cells, NK cells release perforins and granzymes to induce apoptosis. NK cells also produce the cytokine interferon-gamma (IFN-γ), which mobilises antigen-presenting cells and activates other arms of the immune response [8]. The complement system is a non-cellular biochemical pathway of innate immunity, which can identify and opsonise particular pathogens. The complement cascade makes pathogens more vulnerable to phagocytosis and is even capable of eliminating some pathogens or infected cells directly. Further, it can lead to the induction of adaptive immunity via mobilisation and activation of antigen-presenting cells [8].

The cellular mediators of the adaptive immune response are T-cells. Each T-cell expresses a unique antigen-specific membrane bound receptor, known as a T-cell receptor (TCR). T-cells are activated by interaction with an antigen-presenting cell which has digested and displayed the appropriate antigen bound to a major histocompatibility complex (MHC). This stimulates the T-cell to differentiate into either a cytotoxic T-cell ($CD8^+$) or a helper T-cell ($CD4^+$). Cytotoxic T-cells are activated by MHC class I molecules, and are mainly involved in the elimination of infected cells and tumours. Helper T-cells are activated by MHC class II molecules, and release cytokines which influence the activity of other immune cells. Several types of helper T-cell can be generated and they are characterised by the cytokines they release and the downstream cells that they activate. Upon activation both types of T-cell rapidly proliferate, but once the infection has been resolved the majority of effector cells die and are removed by phagocytes. However, some of these cells remain as memory cells that can be quickly activated if the same antigen is encountered [8].

Similarly to T-cells, B-cells also express a unique antigen-specific membrane bound receptor called a B-cell receptor (BCR). Simplistically speaking, upon activation B-cells can differentiate into short-lived plasma cells and long-lived memory B-cells. The plasma cells are responsible for the secretion of soluble versions of their BCRs called antibodies, which enter the circulation targeting foreign antigens. Although the majority of plasma cells undergo apoptosis when the infection has been cleared, memory B-cells persist and can been activated during a secondary response. The primary function of B-cells is the production of antibodies, hence B-cells are considered the mediators of the humoral response in adaptive immunity [8].

## 1.2 B-Cell biology

### 1.2.1 Antibody structure and function

The first description of the humoral immune response came in 1890, where it was observed that cell-free blood could provide immunity to tetanus and diphtheria [9]. Antibodies or immunoglobulins (Ig), representing the key component of humoral immunity, are Y-shaped glycoproteins which are expressed on the surface of B-cells and secreted by plasma cells (Figure 1.2) [10,11]. Antibodies bind to harmful entities such as pathogens, toxins and dead cells, neutralising or tagging them for removal.

When an antibody molecule is digested by the proteolytic enzyme papain, two fragments of identical size (referred to as fragment antigen binding or Fab) are produced alongside another slightly larger fragment (fragment crystallisable or Fc) [12]. As their namesake suggest, Fab regions contain the portion of the antibody responsible for antigen binding. On the other hand, Fc regions were named on the basis that they crystalize readily. It was later discovered that the highly conserved Fc regions, and their associated N-linked carbohydrates, play an important role in mediating antibody effector functions [13]. The Fab and Fc are connected by a flexible region known as the hinge. Indeed, electron microscopy studies of antibodies binding to divalent hapten complexes have observed that various ranges of motion can be achieved by flexion at the hinge region [14].

Antibody molecules are composed of two identical heavy chains ($Ig_H$) and two identical light chains ($Ig_L$) held together by a series of disulfide bonds [15]. There are two types of $Ig_L$, kappa ($Ig_L\kappa$) located on chromosome 2 and lambda ($Ig_L\lambda$) located on chromosome 22, presented in humans at a ratio of ~2:1 in humans (κ:λ) [16]. As a result of isotypic exclusion B-cells express only $Ig_L\kappa$ or $Ig_L\lambda$, however rare exceptions have been documented [17].

Early examination of antibody amino acid sequences revealed two remarkable features. Firstly, each chain was composed of linearly arranged homologous (but non-identical) sequences, ~110 amino acids in length. An individual homologous region represented a discrete and compactly folded protein unit, which would come to be known as a domain. Interestingly, this indicated that antibody chains had evolved through successive duplication of an ancestral gene corresponding to a single domain. The second feature was that the first amino-terminal domain sequence varied greatly between antibodies, while the rest of the domains remained constant across antibodies with matching isotypes [18]. Hence, the variable portions of the $Ig_H$ and $Ig_L$ were considered the variable domains ($V_H$ and $V_L$) and the constant portions were referred to as the constant domains ($C_{H1\text{-}3/4}$ and $C_L$). The $V_H$ and $V_L$ combine to form the variable region at the apex of an antibody Fab, enabling specific binding to a vast array of antigens. The $C_H$ and $C_L$ form the constant region of an antibody molecule, which spans both the Fab and the Fc. On Fab molecules, it has been shown that the junction between the variable and constant regions is capable of bending and rotating [19]. This freedom, combined with the flexibility achieved at the hinge region, permits either arm of an antibody to bind antigens exposed in various spatial presentations.

### 1.2.2 V(D)J recombination

In a single individual, the total number of available antibody specificities is referred to as the antibody repertoire. The variable regions of antibodies must be extremely diverse to ensure specific responses are mounted against any harmful agents encountered. It has been estimated that an antigen-naïve repertoire can contain >$10^{18}$ possible variable region sequences [20]. As the amount of possible variable region sequences exceeds the number of B-cells present in the body, it would appear that such diversity cannot be encoded as germlines. Instead, variable region genes are created via random recombination of a pool of specific gene segments found in the genome. These gene segments can be categorized as variable (V), diversity (D) or joining (J). A functional $V_H$ gene is made from a combination of single V, D and J genes. Alternatively, a functional $V_L$ gene is constructed by recombining single V and J genes. There are different numbers of functional gene segments available in each category for

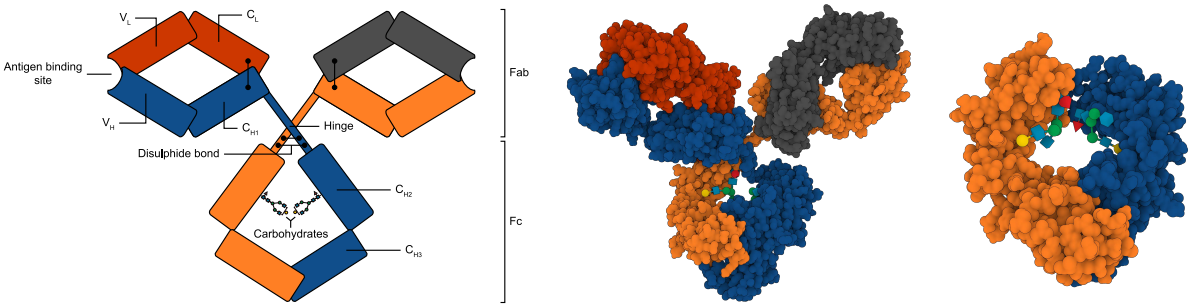



*Figure 1.2: Annotated antibody structure.*

*Antibodies are Y-shaped glycoproteins, that can be broken down into two fragments (Fab and Fc). These molecules are composed of $Ig_H$ (dark blue and orange) and $Ig_L$ chains (red and grey) held together by disulfide bonds. The antibody chains can be divided into discrete domains, which are variable ($V_H$ and $V_L$) or constant ($C_{H1\text{-}4}$), as well as a flexible hinge region. The Fc portion of antibodies harbour N-linked glycans that are important for mediating an array of effector functions. Adapted from [628,629]. PDB code = 1IGY.*

both the $V_H$ and $V_L$ regions (Table 1.3) [21–23]. As a result of genetic polymorphism, not all individuals possess the same number of functional gene segments [24]. Recombination of the antibody variable region follows a well-defined order during lymphoid development. To begin, the D and J genes of the $V_H$ are recombined and subsequently joined to an upstream V gene (Figure 1.3a). The $C_H$ gene is added to the newly formed VDJ gene through splicing of the primary RNA. Following this, the $V_L$ is formed by recombination of the V and J genes (Figure 1.3b). The arrangement of J and $C_L$ genes differs between $Ig_L\kappa$ and $Ig_L\lambda$ using antibodies. For the $Ig_L\kappa$ loci, a single $C_L$ gene exists downstream of multiple J segments. In the $Ig_L\lambda$ loci, each J gene has its own $C_L$ gene located downstream. Hence, the $C_L$ used in $Ig_L\lambda$ antibodies is dependent on the J gene. The $C_L$ genes are added to the assembled VJ genes via RNA splicing [25].

The recombination process is site-specific, meaning it occurs precisely at the coding end of each connectable element, because each gene segment is flanked by a recombination signal sequence (RSS). The nucleotide arrangement of RSSs is well understood, conserved heptamer and nonamer motifs are separated by linker sequences which are 12 (12RSS) or 23 (23RSS) base pairs in length (Figure 1.3c) [26]. In vivo recombination preferably occurs between a 12RSS and a 23RSS, commonly referred to as the 12-23 rule [27]. The 12-23 rule ensures that only corresponding V, D and J gene segments are assembled. Indeed, the $V_L$ of $Ig_L\lambda$ antibodies have V genes with 23RSSs downstream and J genes with 12RSSs upstream. In the $V_L$ of $Ig_L\kappa$ antibodies this pattern is reversed, with V genes having 12RSSs downstream and J genes having 23RSSs upstream. The D genes of the $V_H$ are bordered by 12RSSs on both sides, whilst V genes have 23RSSs downstream and J genes have 23RSSs upstream [28].

V(D)J recombination is initiated by the recombination-activating gene (RAG) enzyme complex, composed of RAG1, RAG2 and high mobility group box 1 (HMGB1) proteins (Figure 1.4a). The RAG complex binds to an RSS flanking a gene segment (e.g D gene 12RSS of $V_H$) and then recruits a second RSS on a separate gene segment (e.g J gene 23RSS of $V_H$), following the 12-23 rule. Next, the RAG complex makes an endonucleolytical nick to the captured elements precisely between the gene segments and the heptamer motifs. The cutting points at the 12RSS and 23RSS heptamers generate two separate 3' -OH groups, one for each of the captured genes. The 3' -OH groups act as nucleophiles in transesterification attacks on phosphodiester bonds located on the antiparallel deoxyribonucleic acid (DNA) strands. This closes the coding V, D or J gene segments with hairp in loops (coding ends), whilst producing a blunt double-stranded break at the end of the RSSs (signal ends) [29].

The coding ends are then bound by Ku, a heterodimer of Ku70 and Ku80, which forms a complex called the DNA-dependent protein kinase (DNA-PK) by recruiting its catalytic subunit (DNA-PKcs) [30]. Artemis, a protein with nuclease activity, associates with DNA-PK creating the DNA-PK:Artemis complex. DNA-PK phosphorylates Artemis resulting in a nick at a random position in each of the coding sequences, close to where the hairpins were initially generated [31]. This forms single-stranded nucleotide tails at the end of the nicked gene segments. The tails may contain short palindrome sequences, because the original coding and added nucleotides used to be complementary in the double-stranded DNA. Nucleotides added via this mechanism are called P-nucleotides, due to their tendency to be

*Table 1.3: V(D)J and C gene segments.*
*Shown here are the numbers of functional V, D, J and C gene segments found at the IgH, IgLκ and IgLλ loci. Adapted from [628].*

| Segment | Light chains | | Heavy chains |
|---|---|---|---|
| | $Ig_L\kappa$ | $Ig_L\lambda$ | $Ig_H$ |
| **V** | 34-38 | 29-33 | 38-46 |
| **D** | 0 | 0 | 23 |
| **J** | 5 | 4-5 | 6 |
| **C** | 1 | 4-5 | 9 |

palindromic in nature. Next, the tails can be modified by terminal deoxynucleotidyl-transferase (TdT) which adds nucleotides at random (with a GC bias), increasing the length of the overhangs. Nucleotides added through this process are call non-template or N-nucleotides. Throughout this, nucleotides can also be removed from the coding ends through the action of a currently unidentified exonuclease (possibly Artemis). Nucleotides are added or removed at the opposing overhangs until they eventually form complementary sequences and pair. An exonuclease (also possibly Artemis) removes any mismatched nucleotides and the remaining gaps are filled by template-dependant DNA polymerases. The coding ends are linked by a complex of DNA ligase IV and X-ray repair cross-complementing protein 4 (XRCC4), forming a coding joint [32]. The joining of double stranded breaks (DSBs) via this mechanism is referred to as the classical non-homologous end joining (NHEJ) pathway.

At the blunt signal ends, Ku70:80 binds to the RSSs and the DNA ligase IV:XRCC4 complex precisely joins them to generate a signal joint. Generally, the coding genes being recombined are arranged in

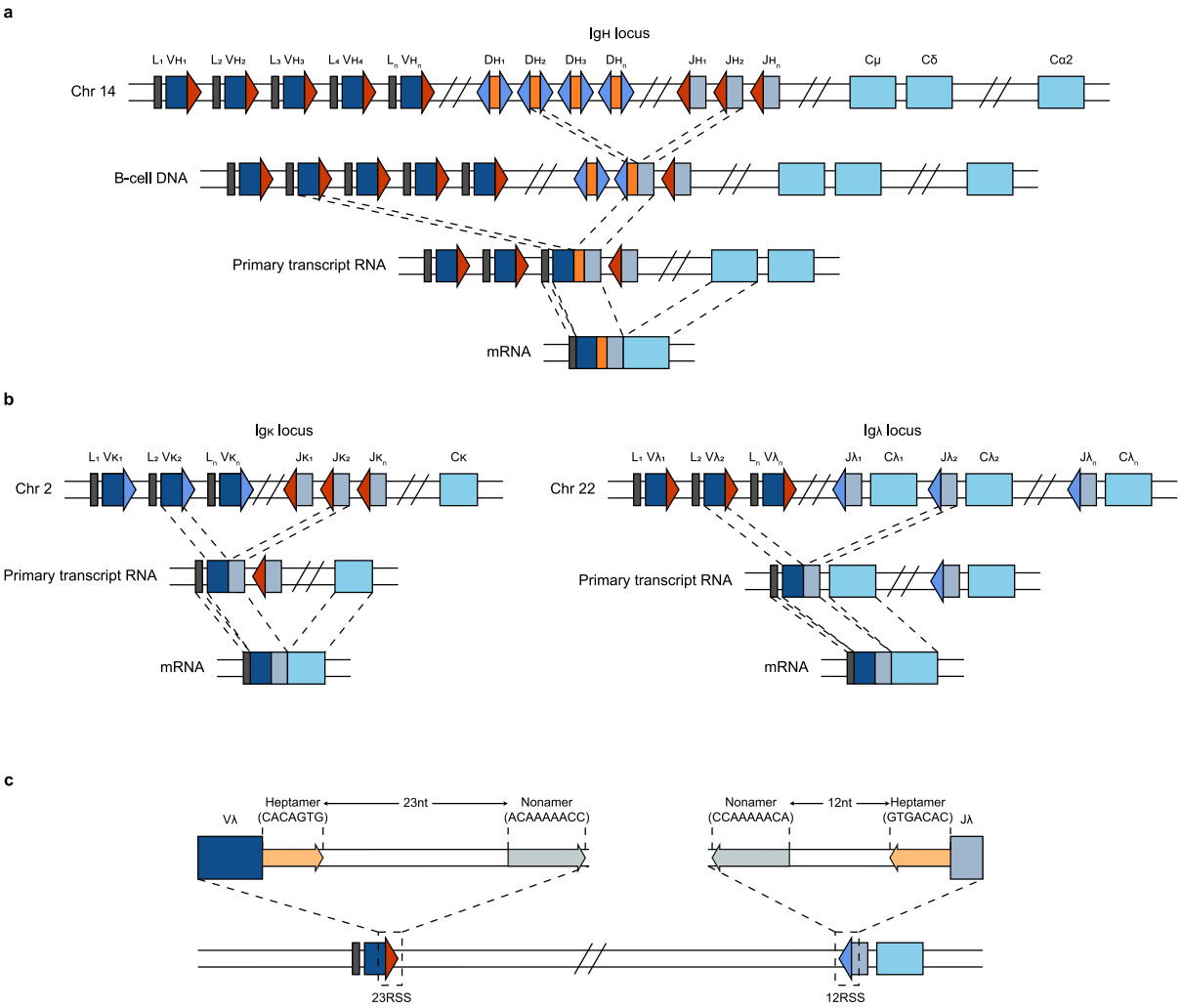



*Figure 1.3: Structure of V(D)J genes and recombination process.*
*V(D)J gene segments are arranged in a non-functional state in the germline, but they are brought together through a series of recombination events to form functional variable region sequences. At the $Ig_H$ locus (**a**), the D and J genes are rearranged first and subsequently joined to an upstream V gene. The L and $C_H$ are joined with the newly formed VDJ gene via RNA splicing. At the $Ig_L$ loci (**b**), the V and J genes are recombined and then spliced between the L and $C_L$. The arrangement of the $Ig_L\kappa$ and $Ig_L\lambda$ genes differ, with all $Ig_L\kappa$ VJ genes using the same Cκ and the $Ig_L\lambda$ VJ genes each using a designated Cλ. The gene segments are rearranged in the correct orientation due to the presence of complementary 12RSS and 23RSS sequences (**c**), which can be found adjacent to the gene segments. This type of diversity is known as combinatorial diversity. Adapted from [630].*

the same transcriptional orientation (Figure 1.5). In such instances, the formation of the signal joint results in the intervening DNA being released as an excision circle. However, coding genes can also be oriented in opposing transcriptional directions. For example, half of $Ig_{L}\kappa$ V gene segments are found in the opposite orientation to their corresponding J gene elements. Joining such RSSs requires looping of the DNA and results in a functional $V_L$ and integration/inversion of the normally excised intervening DNA [33].

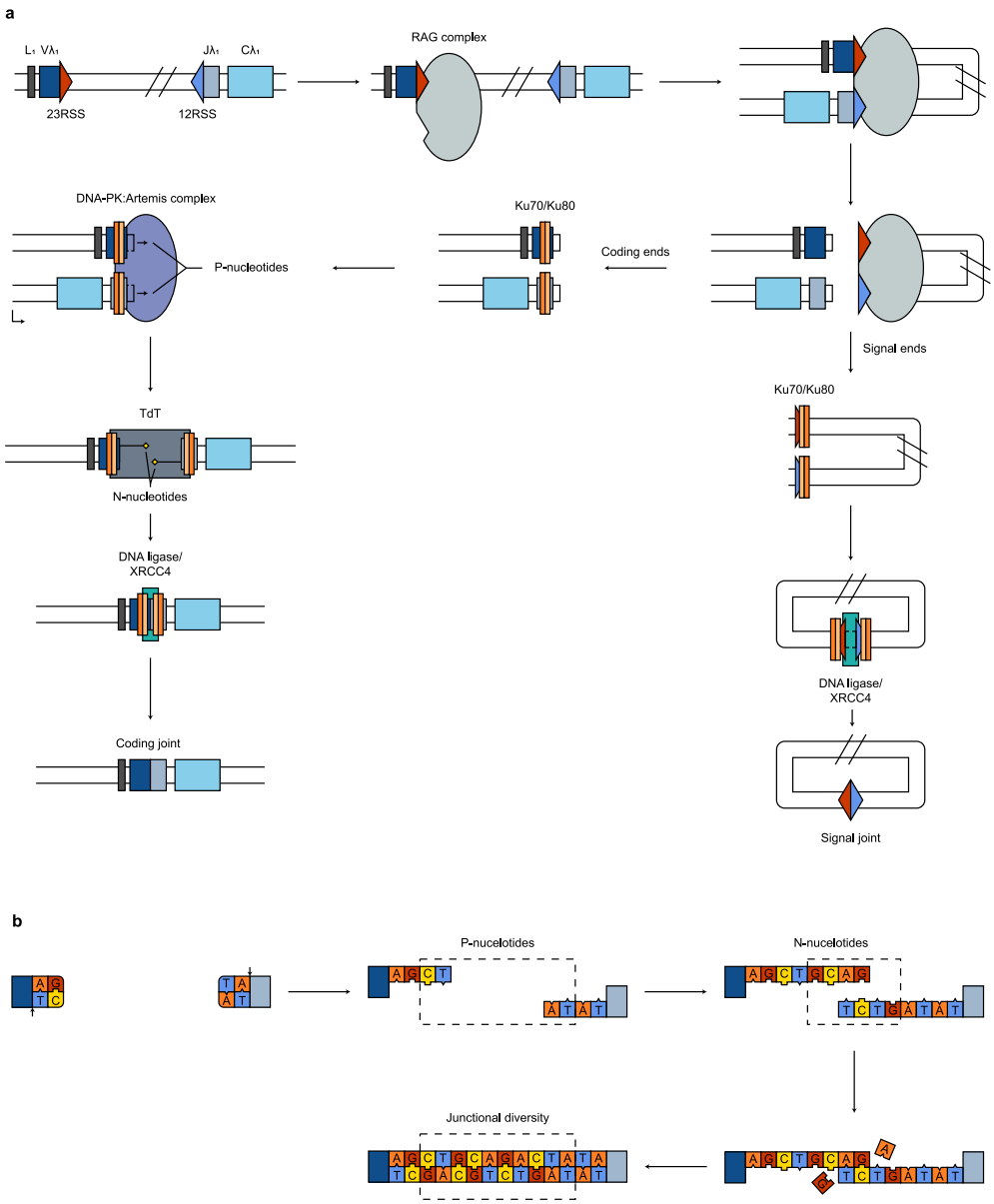



*Figure 1.4: V(D)J recombination pathway.*
*V(D)J recombination is initiated by the RAG complex (**a**), which binds to RSSs flanking the gene segments. RAG complex endonuclease activity results in single strand cuts being made in the DNA backbone between the gene segments and their RSSs. This leads to the formation of hairpin loops on the coding ends and blunt double stranded breaks on the signal ends. The two types of DNA ends are resolved in different ways. At the coding ends (left), Ku70:80 bind to the hairpins, and the DNA-PK:Artemis complex is recruited. The DNA-PK:Artemis complex nicks the hairpins at a random site, opening the loops and allowing TdT to add or remove nucleotides. The coding ends are ligated by DNA ligase IV:XRCC4. The signal ends (right) are resolved by the actions of Ku70:80 and DNA ligase IV:XRCC4, resulting in inversion or excision of the intervening DNA. Recombination in this manner (**b**) is imprecise and results in the introduction of P-nucleotides and N-nucleotides. This type of diversity is known as junctional diversity. Adapted from [628].*

The variation introduced via random pairing of antibody chains and V(D)J recombination is known as combinatorial diversity (Figure 1.4b). Combinatorial diversity may be slightly more limited in practice compared to theory. There is considerable variation in the usage of different V gene segments, which could be the result of polymorphisms in the RSSs [34]. Further, the number of V(D)J segments available for recombination can differ between individuals, and some $Ig_H$-$Ig_L$ pairs are unable to produce functional antibodies [24]. This is an important feature to consider, for example HIV-1 research has shown that only certain germline V(D)J arrangements possess the inherent characteristics required for broad neutralisation [35]. However, such germlines can be rare and be difficult to stimulate with immunogens, making this a substantial roadblock to HIV-1 vaccine development [36]. The alteration that is introduced between gene segments during V(D)J recombination is known as junctional diversity. The contributions that P/N-nucleotide addition or exonuclease removal make to the coding joint vary [37]. For instance, P-nucleotides are often absent in antibody sequences and a lower number of N-nucleotides are found in $Ig_L$ chains [31,38].

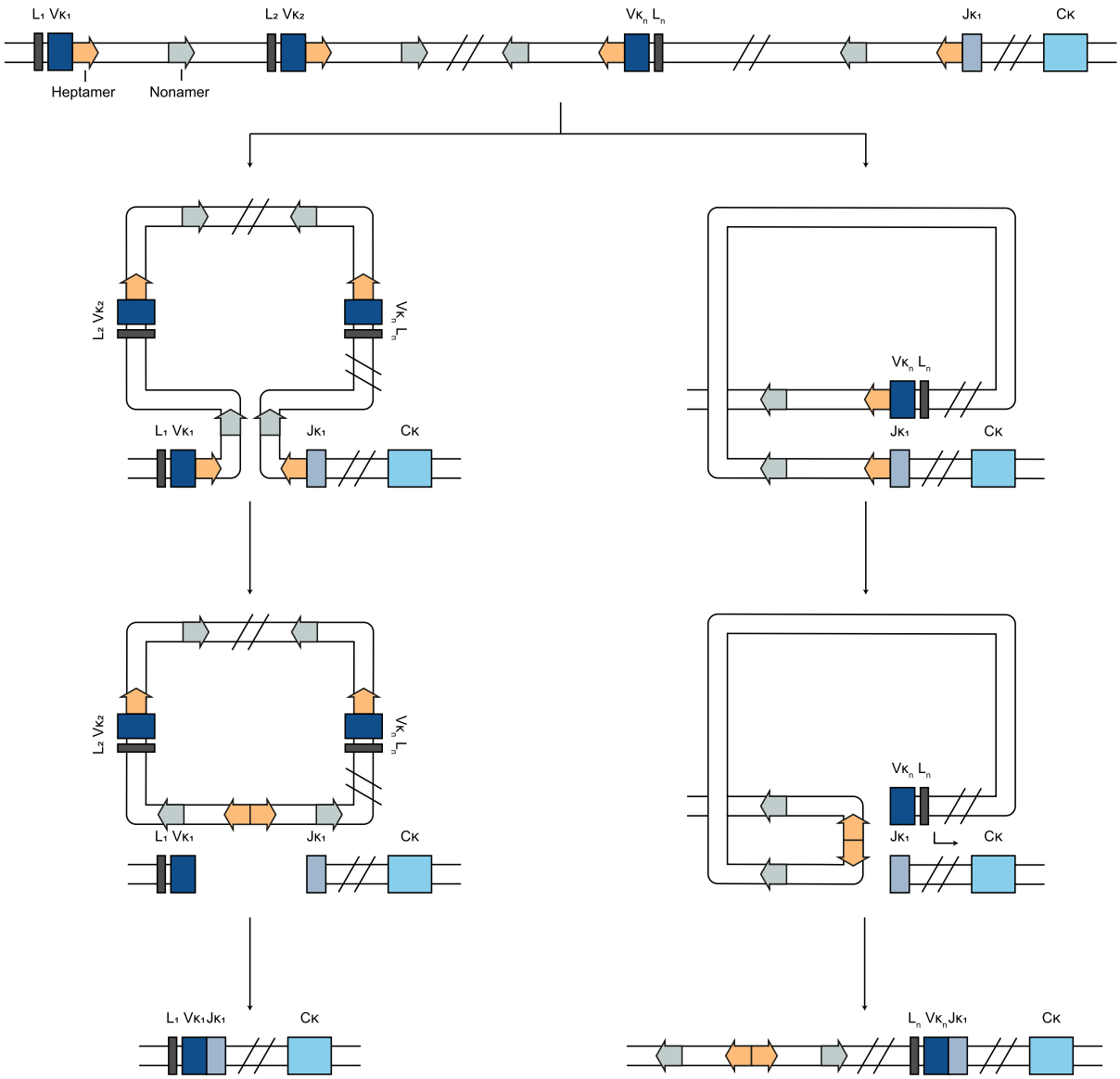



*Figure 1.5: V(D)J recombination dependency on gene transcriptional orientation.*
*In most cases, gene segments undergoing recombination are arranged in the same transcriptional orientation (left). Juxtaposition of the RSSs leads to looping out of the intervening DNA and the release of an excision circle. In other instances, the gene segments are found in an opposite transcriptional orientation and must adopt a coiled topology to rearrange. This results in inversion and integration of the intervening DNA. Adapted from [628].*

### 1.2.3 Complementarity-determining regions and framework regions

Antibody domains are comprised of two compact antiparallel β-sheets, joined via a disulfide bond, forming a cylindrical structure known as an immunoglobulin fold (Figure 1.6a). The β-sheet motifs are created from laterally connected sections of polypeptide arranged in conformations called β-strands.

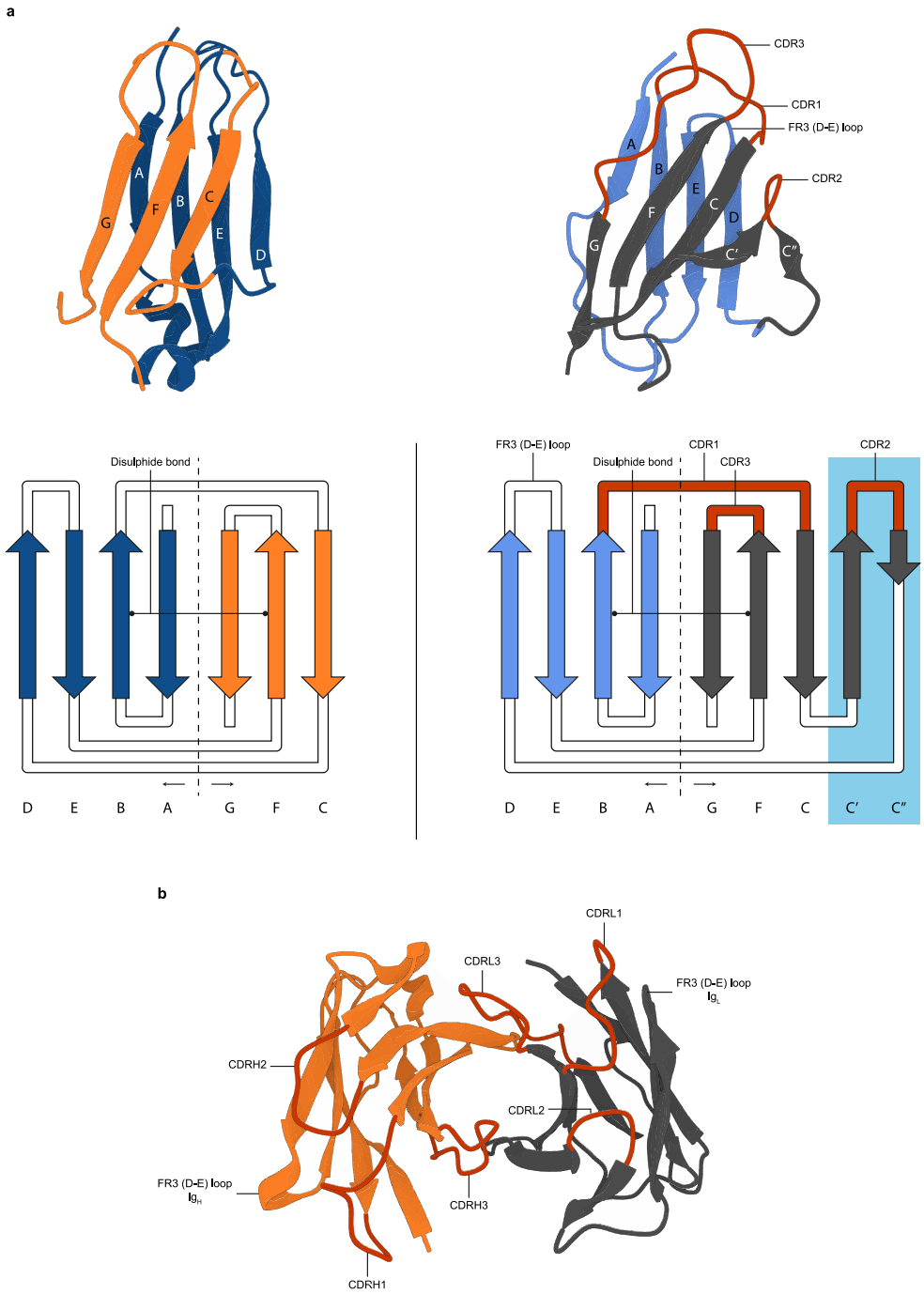



*Figure 1.6: Structure of antibody variable and constant domains.*
*Antibody variable and constant domains are barrel-shaped structures, composed of antiparallel β sheets (made from β strands) held together by a disulfide bond (**a**). The β strands are lettered sequentially in the order they appear in the amino acid sequence. Variable domains (light blue and grey) have two additional β strands (C' and C''), which are not found in constant domains (dark blue and orange). There are three areas within the variable domains that display a high level of sequence variation (CDRs). The CDRs form three loops at the apex of variable regions (**b**), which formulate the majority of the contacts with the antigen. Occasionally a loop from the FR (FR3) also participates in antigen binding. Adapted from [628,629]. PDB code = 1IGY.*

The β-strands are displayed as arrows on structural diagrams and labelled in alphabetical order as they appear in the amino acid sequence from N-terminus to C-terminus [39].

In constant domains, the first β-sheet has four β-strands (↓A ↑B ↓E ↑D) and the second has three (↓C ↑F ↓G). The two β-sheets are joined by a disulfide bond located between the β-strands ↑B and ↑F. The constant domains are generally tightly packed and have short loops linking the β-strands. In the variable domain, the arrangement of the first β-sheet and the location of the disulfide bond is the same as the constant domain. However, the second β-sheet of the variable domain has five β-strands (↓C'' ↑C' ↓C ↑F ↓B), two more than what is found in the constant domain. Further, variable domains are usually less compact, with longer loops than constant domains [39].

Early variability studies showed that the sequence variation found in antibody variable regions is not distributed equally. Three areas display a high level of variability, these are called complementarity-determining regions (CDR; Figure 1.6b). Additionally, there are four areas not encompassed by the CDRs that show less variability, these are known as frame-work regions (FR) [40]. The FRs generate the β-sheets, ensuring the structure of the antibody is maintained. The CDRs form three of the loops connecting the β-strands, these are located on the external side of the immunoglobulin fold. The $V_H$ and $V_L$ both contribute three CDRs, when joined these loops formulate the majority of the contacts made with the antigen [39]. Although not generally considered a CDR, there is a fourth loop located in the FR3 of variable domains. In some instances, the FR3 loop has been shown to participate in antigen binding and lead to an enhancement in the neutralisation of viruses such as HIV-1 [41,42].

The first two CDRs of the $V_H$ and $V_L$ domains are encoded solely by their V gene segment. In contrast, the CDR3 of the $V_H$ is partially encoded by its corresponding V, D and J gene segments. Likewise, the CDR3 of the $V_L$ is partially encoded by its corresponding V and J gene segments. The CDR3 loops of both the $V_H$ and $V_L$ can therefore incorporate junctional diversity into their sequence. The CDRH3 is the most diverse loop because it forms two junctions when undergoing recombination. Consequently, there are two opportunities for the acquisition or loss of P/N-nucleotides, enabling a wide array of CDRH3 lengths to be generated in the human repertoire [43]. In turn, the CDRL3 is considered to be the second most diverse loop of the CDRs. Both the CDRH3 and the CDRL3 have been shown to play crucial roles in the binding of antibodies to antigens [44,45]. Long CDRH3s have been shown to be a crucial feature of HIV-1 bnAbs, enabling these antibodies to reach through the glycan shield of their target and contact vulnerable conserved protein regions underneath [46].

### 1.2.4 Somatic hypermutation

In addition to combinatorial and junctional diversity, variation can also be introduced into the DNA sequences of antibodies through a process called somatic hyper mutation (SHM). SHM introduces point mutations and insertion/deletions (indel) into the $V_H$ and $V_L$ genes. As a result of SHM, the variable regions of antibodies have a mutation rate that is approximately a million times higher than what is observed spontaneously in somatic cells. Point mutations introduced by SHM can be considered as transitional or transversional. A transitional mutation is when a pyrimidine (C or T) is substituted by another pyrimidine or a purine (A or G) is substituted by another purine. A transversional mutation is when a pyrimidine is substituted with a purine or vice versa [47].

SHM is mediated by an enzyme called activation-induced cytidine deaminase (AID), which acts on single-stranded DNA (Figure 1.7) [48]. AID is only expressed in B-cells that have been activated, deaminating C to U via nucleophilic attack of the pyrimidine ring found in C. AID has a preference for hotspots, denoted by the consensus motif DGYW (or WRCH on the antiparallel strand), in which D = A/G/T, Y = C/T, W = A/T, H = T/C/A, R = A/G [49]. The U generated can be considered a dual lesion because it is not normally found in DNA and is mismatched with the G on the opposite strand. If a U produced

by AID persists until S-phase, DNA synthesis via any known DNA polymerase will result in a C/G > T/A transitional mutation. This occurs because U has a close similarity to T and will instruct DNA polymerases to insert an A opposite it when used as a template. However, the lesion can be repaired with high fidelity via the action of conventional base excision repair (BER) or mismatch repair (MMR) pathways, prior to this [50].

In conventional BER, the uracil is excised by uracil-DNA-glycosylase (UNG) creating an abasic site. At the abasic site the DNA backbone is cleaved by an apurinic/apyrimidinic endonuclease (APE), generating a nick. At the nick, poly(ADP-ribose) polymerase 1 (PARP1) is recruited and signals the recruitment of X-ray repair cross complementing 1 (XRCC1), polymerase β, and DNA ligase (I or III). Polymerase β processes the nick to a single nucleotide gap and fills it with high fidelity. Finally, the nick is sealed by the action of a DNA ligase (I or III) [51]. In conventional MMR, the lesion is recognized by the MSH2-MSH6 heterodimer (MutSα) and a complex of MLH1 and PMS2 (MutLα). MutLα nicks the DNA adjacent to the lesion allowing proliferating cell nuclear antigen (PCNA) associated exonuclease-1 (Exo1) to be recruited. Exo1 generates a patch of single-stranded DNA, starting from

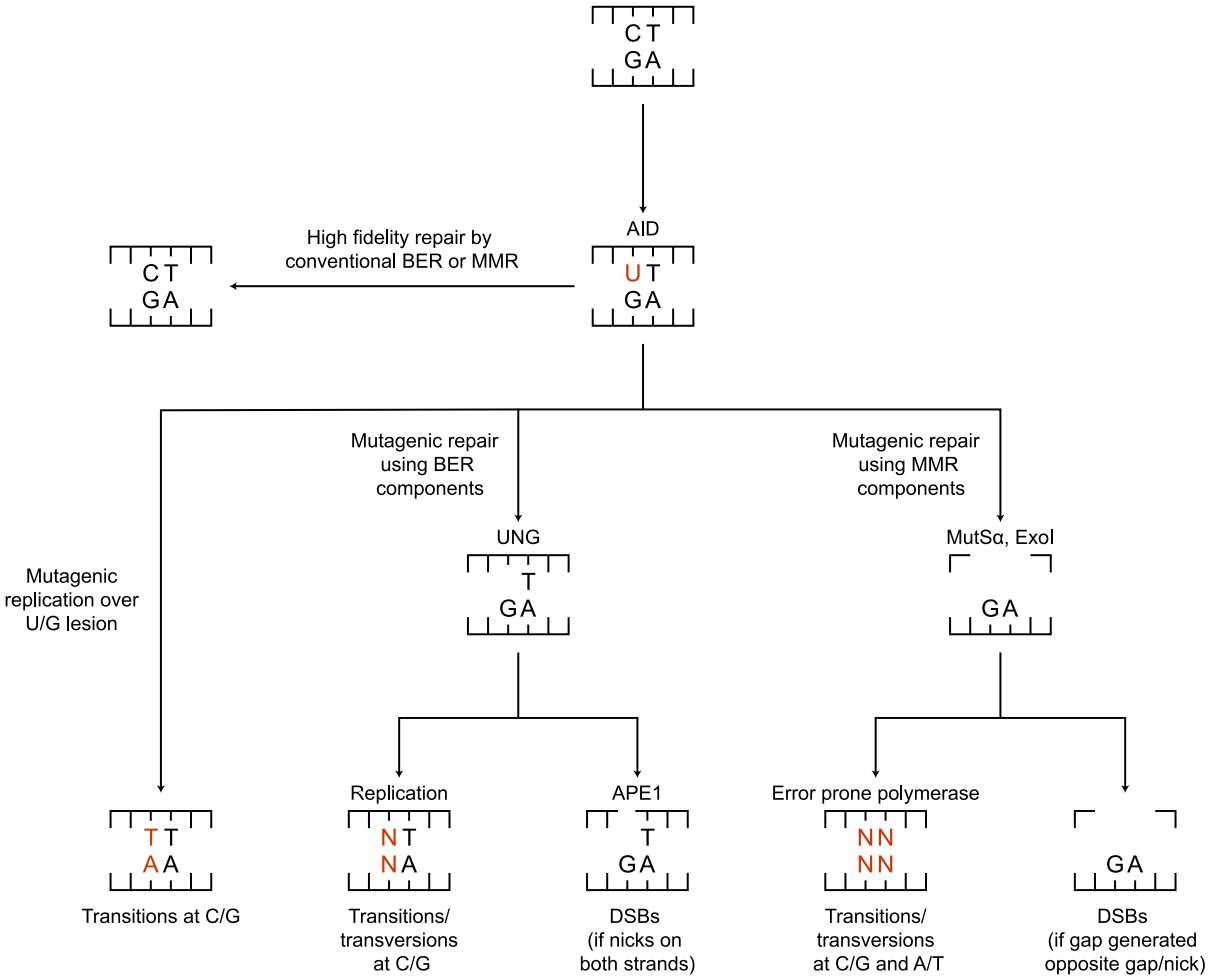



*Figure 1.7: Mechanisms of SHM.*
*AID deaminates C to U. U/G lesions are typically repaired with high fidelity by the BER or MMR pathways, but in B-cells they generate variation causing SHM. Mutagenic replication over the U/G lesion results in a transition mutation at C/G base pairs. UNG of the BER pathway can excise the U creating an abasic site. Replication over the abasic site can yield transition and transversion mutations at C/G base pairs. APE may create a nick at the abasic site, and overlapping nicks can result in double stranded DNA breaks. MutSα of the MMR pathway can nick the DNA adjacent to the U/G lesion, allowing Exo1 to excise the patch of DNA containing the mismatch. Error-prone polymerases resynthesise the patch leading to spreading of mutations to A/T base pairs. Overlapping gaps or nicks may lead to double stranded DNA breaks. Adapted from [54].*

the nick and extending past the lesion. Finally, PCNA recruits another high-fidelity polymerase (polymerase δ) which fills the gap, with DNA ligase I sealing the nick [52].

In B-cells the BER and MMR pathways produce mutations and double-stranded breaks in response to AID induced lesions, instead of performing their usual tasks of maintaining genetic fidelity. SHM is limited without the action of these pathways and only the aforementioned C/G > T/A transitional mutations can occur. Currently, the understanding of these processes are less fully understood compared to their conventional counterparts. Following the BER pathway, an AID mediated uracil can be excised by UNG resulting in an abasic site being formed. Due to the non-instructive nature of abasic sites, replication following the action of translesion synthesis (TLS) polymerases results in transitional and transversional mutations at C/G positions [50]. TLS polymerase Rev1 can only insert C at an abasic site and could therefore be responsible for the production of C/G > G/C transversional mutations [53]. It is unknown which TLSs are responsible for the other types of mutations produced via the BER pathway [50]. To generate mutations at A/T residues, the MMR must be employed. MutSα recognises the lesion and Exo1 cuts a patch of nucleotides encompassing the mismatch. A low fidelity polymerase such as polymerase η is recruited to resynthesize the gap. Hence, mutations can be spread from the original U/G lesion to surrounding A/T nucleotides [54].

Double-stranded DNA breaks may also occur as a consequence of the SHM process. After a uracil is excised by UNG, an APE could generate a nick at the abasic site. The presence of nicks on opposing DNA strands, at an AGCT motif for instance, could generate a double-stranded break. Gaps produced by the MMR pathway could overlap with each other or a nick produced by the BER pathway, both resulting in a double-stranded DNA break. Repair of such breaks may result in the introduction of indels to the variable region of antibody genes [54]. It is likely that antibody variable regions that contain indels are selected against because of their tendency to disrupt the reading frame of the gene. However, the variation added to antibody genes via indels has been shown to be beneficial in some instances, such as the six amino acid insertion in the CDRH2 of the HIV-1 targeting antibody PGT-128 [55,56].

In $V_H$ and $V_L$ regions SHM tends to be concentrated at the CDRs, although it is not fully understood why [57]. As the CDRs interact with antigen, mutations which improve antibody binding could be selected for. Mutations in the FRs could be selected against because they interfere with the antibody structure [58]. AID may preferentially target the CDRs because they contain a higher number of hotspots, in contrast to the rest of the variable exon [54]. Further, AID induced lesions in the CDRs could also be repaired using different mechanisms compared to those used in the rest of the variable region [59].

### 1.2.5 Class-switch recombination

An antibody can be categorized as one of five main classes, these are IgM, IgD, IgG, IgE and IgA. There are four subclasses of IgG (IgG1, IgG2, IgG3 and IgG4) and two subclasses of IgA (IgA1 and IgA2). The different types of $C_H$, which define the class of an antibody, are referred to as isotypes and are denoted with the equivalent Greek alphabet μ, δ, γ, ε and α [60]. The sequences encoding the $C_H$ domains of separate isotypes bear differences, yielding antibodies with distinct features. Indeed, each isotype has a different number/location of disulfide bonds, distribution of N-linked carbohydrates, number of $C_H$ domains and hinge arrangement (Figure 1.8). Consequently, isotypes serve separate purposes via interacting with different receptors, each fulfilling distinct functions as part of the immune response [61].

The different $C_H$ isotype genes are encoded in the $Ig_H$ loci downstream of the J genes, in the order of Cμ, Cδ, Cγ3, Cγ1, ψCε, Cα1, Cγ2, Cγ4, Cε and Cα2. One of the Cε genes (annotated with ψ) is no longer active and is therefore referred to as a pseudogene [62]. Each of the $C_H$ genes can be separated into

numerous exons, each encoding an antibody domain or hinge region. The Cμ is the first gene available downstream of a newly assembled VDJ region, and is therefore the first isotype expressed by developing B-cells. The Cμ gene is spliced adjacent to the VDJ gene, removing any leftover obstructive J gene segments, forming a complete $Ig_H$ transcript. The next $C_H$ gene in line is Cδ, and IgD is expressed alongside IgM on the surface of mature B-cells. In such cells, an extended primary transcript is produced and then spliced to make messenger ribonucleic acid (mRNA) containing either Cμ or Cδ. When a B-cell becomes activated, it stops producing IgD/IgM as a result of an irreversible DNA recombination event. During this event the Cμ/Cδ genes are removed and a different $C_H$ isotype is positioned downstream of the VDJ, this process is called class-switch recombination (CSR; Figure 1.9). As CSR only acts on the $C_H$ domains, antibodies of different classes are produced with identical variable regions and antigen specificities [60].

With the exception of Cδ, each $C_H$ gene is preceded by a cytokine/activation inducible promoter, an intervening exon (I-exon) and a switch region. Switch regions are long stretches of repetitive DNA located in introns upstream of their corresponding $C_H$ gene. Although there is some sequence variation between the switch regions of different isotypes, they are all G-rich and contain a high density of AID

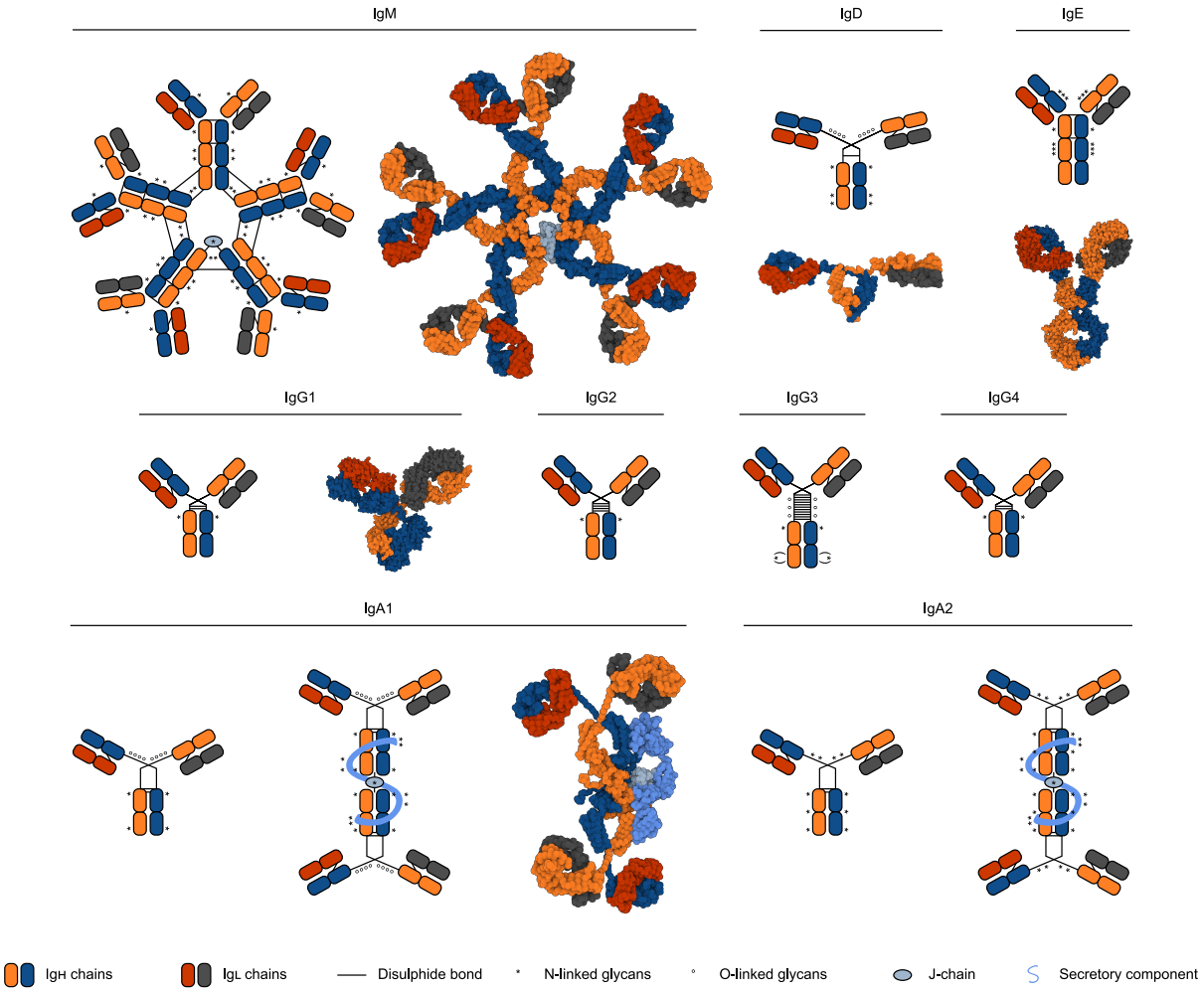



*Figure 1.8: Annotated structures of antibody isotypes and subclasses.*
*In humans, there are five antibody isotypes (IgM, IgD, IgG, IgA and IgE) and six subclasses (IgG1-4, IgA1 and IgA2). Isotypes and subclasses exist as monomers and/or multimers, joined by disulfide bonds, polypeptide J chains or a secretory component. They display a variety of hinge lengths/flexibilities and have differing numbers/locations of disulfide bridges and O/N-linked glycans. Changes in isotype or subclass enable different interactions with other immunes cells, therefore deploying distinct immune effector functions. Adapted from [67,629,631,632,633, 634,635]. PDB codes = 1IGY, 2RCJ, 1ZVO, 6EYO, 2R56 and 3CHN.*

hotspots. CSR is guided by a process called germline transcription, in which transcription occurs through the switch regions of the isotypes that are going to recombine [63]. Germline transcription commences at the promoter and continues through the I-exon, switch region and $C_H$, then the primary transcript undergoes polyadenylation. In a mature germline transcript, the I-exon is spliced adjacent to the $C_H$ gene. Germline transcription does not produce proteins of significant length, and this is due to the many stop codons contained in the I-exons [64]. However, it does create long stretches of single-stranded DNA formed by the non-template strands (also known as R-loops) [65].

Similarly to SHM, CSR is mediated by AID. Hence, the formation of R-loops in the switch regions will render them suitable AID-substrates. AID deaminates C to U in the switch regions, allowing members of the BER and MMR pathways to generate double-stranded DNA breaks. The two switch regions are brought together via the NHEJ pathway, and the intervening DNA is released as an excision circle. Transcription of the $Ig_H$ will now go on to yield an antibody with a different $C_H$ isotype. As switch regions are contained in introns, a productive antibody will always be produced because the merged switch regions will be removed during RNA splicing. Further CSR events can occur but only with downstream elements because the intermediate DNA was permanently removed [60].

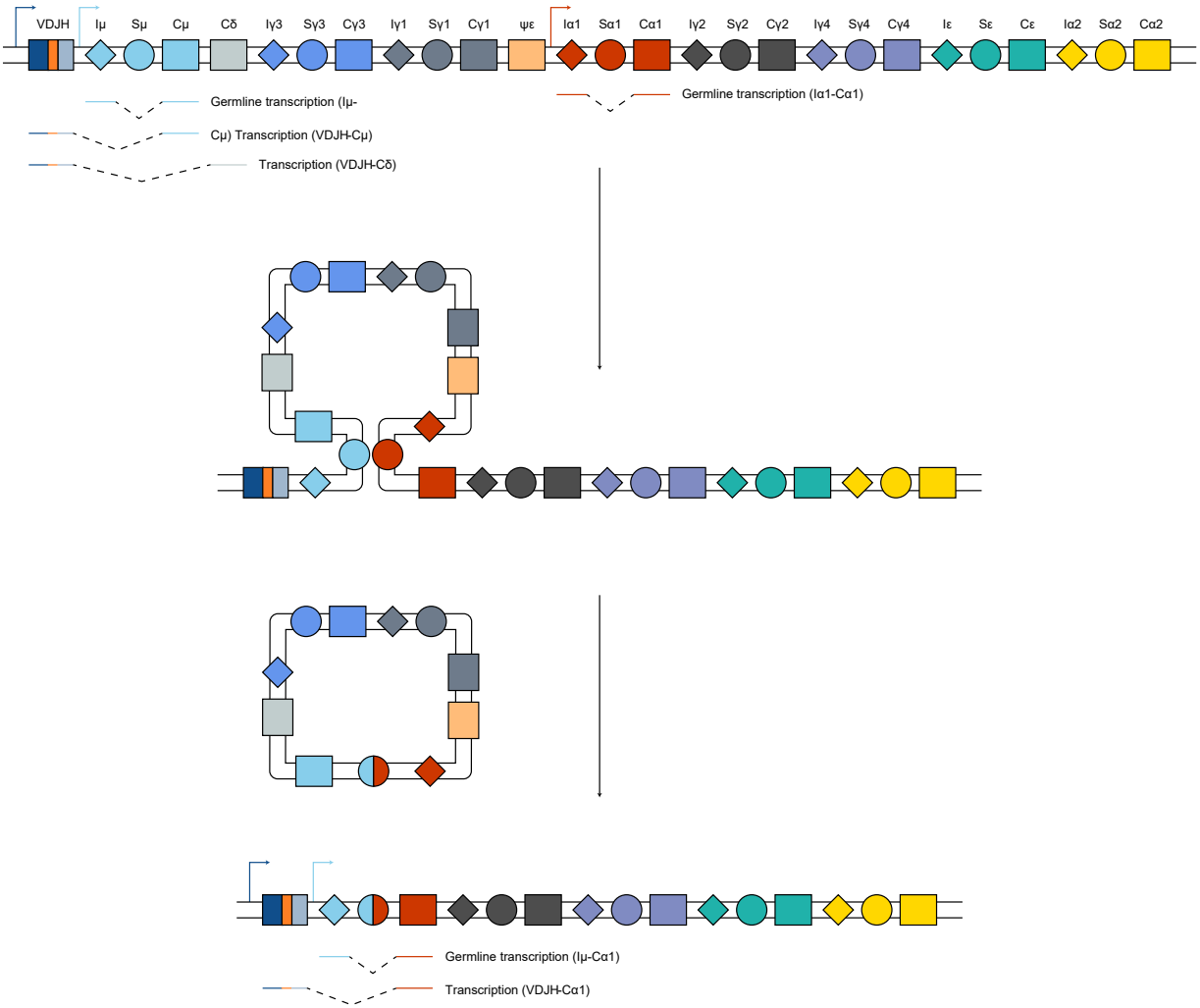



*Figure 1.9: Mechanism of CSR.*
*CSR is the process in which the $C_H$ genes of an antibody are exchanged with a downstream set of genes. This requires the action of AID, which is guided by germline transcription, and the formation of double stranded DNA breaks in switch regions. Repair of these DSBs by the NHEJ pathway results in juxtaposition of the VDJ gene with a downstream $C_H$ exon cluster, and the release of the intervening DNA as an excision circle. Adapted from [62].*

Some CSR events are more common than others, with IgM/IgD B-cells being most likely to switch to either IgG1 or IgA1. Further, subclass switching between members is more common than switching to another class [66]. CSR patterns can be skewed by altering germline transcription via activation of the appropriate promoters using immune stimulants or cytokines [63].

### 1.2.6 Effector functions of antibody classes

The various antibody isotypes serve different effector functions and it is important the most suitable is selected by CSR during an immune response. Antibodies perform their effector functions via interactions between their Fc portions and Fc receptors available on other immune cells. Such interactions can lead to a variety of effector functions such as antibody-dependent cellular cytotoxicity (ADCC), antibody-dependent cellular phagocytosis (ADCP) and initiating complement [67].

As discussed previously, monomeric IgM is expressed on naïve B-cells as BCRs and throughout particular stages of B-cell development. Additionally, two types of secreted IgM can be produced, natural IgM and adaptive IgM. The production of natural IgM does not require antigenic stimulation and is thought to provide an early innate-like protection against pathogens. Natural IgM antibodies have low levels of SHM, weak affinity and are often polyreactive [68]. Conversely, adaptive IgM antibodies undergo affinity maturation and can be mutated, bind with higher affinity and generate a memory response [69].

Secreted IgM molecules can join together to make a pentamer structure with ten antigen binding sites, allowing high avidity. The IgM monomers are connected to each other via disulfide bonds within the $C_{H4}$ regions, a joining chain (J-chain) is also incorporated into the structure to enable secretion at mucosal surfaces [61]. IgM pentamers which have bound to an antigen can initiate the compliment cascade [67].

IgD is co-expressed with IgM on the surface of naïve B-cells. Less is known about the role that IgD antibodies play in the immune response. When a naïve B-cell encounters a suitable antigen IgD is downregulated, although B-cells do sometimes class-switch to IgD. Plasma cells secreting IgD antibodies are found within the upper respiratory tract and the tonsils [70].

The most abundant antibody class found in human serum is IgG, of which there are four subclasses. The constant region genes of the four IgG subclasses have >90% homology, but display significant differences in the hinge and $C_{H2}$ regions [71]. These differences relate to structural changes affecting interactions with Fc-gamma receptors (FcγR) and complement component 1q (C1q), which are important for ADCC/ADCP and complement-dependent cytotoxicity (CDC) respectively. This means that each subclass of IgG can enable a different effector function after binding antigen [72].

The subclasses IgG1 and IgG3 share similar roles, with both members mainly targeting soluble and membrane-bound proteins. The Fc regions of IgG1 and IgG3 antibodies can bind to C1q and all three of the FcγRs, allowing the initiation of the complement cascade and ADCC [73]. Despite having similar effector functions, IgG1 dominant responses are found much more frequently. This may be due to IgG3 antibodies having very extensive hinge regions, which are more susceptible to proteolysis, and a short half-life [74]. IgG2 antibodies mainly bind to polysaccharides and are important for IgG responses to bacterial capsular polysaccharide [75]. IgG2 antibodies do not potently activate complement and can only bind to particular FcγRIIa allelic variants [76]. IgG4 mainly binds to protein antigens and is often generated in response to allergens. IgG4 does not activate complement but can interact with FcγRII and FcγRIII, albeit at a significantly lower affinity compared to IgG1 [61]. IgG4 antibodies are generated as a result of long-term repeated exposure to an antigen and are thought to regulate inflammatory IgE responses [77].

IgA is the main class found at most mucosal surfaces, and the second most prevalent in human serum. IgA antibodies can form dimers by associating with a J-chain and another polypeptide chain called the secretory component [67]. Monomeric IgA is found in the blood and produced in the lymph nodes and bone marrow, whereas dimeric IgA is found in the mucosal surfaces and produced by the nearby lymphoid tissue. There are two subclasses of IgA, and the main difference between them is that IgA1 has a longer hinge region compared to IgA2. The longer hinge may increase IgA1 avidity by enabling bivalent interactions with more spatially separated targets, but it could also increase vulnerability to proteolysis [78]. IgA antibodies can induce the alternative and lectin complement pathways, alongside the Fc-alpha receptor (FcαRI) found on cells such as neutrophils. Serum IgA triggers inflammatory responses such as phagocytosis and ADCC, whilst secretory IgA is non-inflammatory and mainly blocks pathogenic interactions with the mucosal surface [79,80].

IgE class antibodies have the shortest half-life and are found at the lowest concentration in human serum [81]. IgE antibodies are generated in response to parasitic worm infections and play a role in allergic reactions. The Fc regions of IgE antibodies can bind to two Fc-epsilon receptors (FcεRI and FcεRII). IgE antibodies have a very high affinity for FcεRI, which is expressed by mast cells, basophils, antigen-presenting cells, monocytes and eosinophils. This interaction with FcεRI results in mast cell degranulation and strong activation of an inflammatory response [61,82].

### 1.2.7 B-cell development and central tolerance

In the bone marrow of adults, B-cells develop from their progenitors via a series of definable stages. These stages are referred to as early pro-B-cell, late pro-B-cell, large pre-B-cell, small pre-B-cell and immature B-cell (Figure 1.10) [83]. The bone marrow provides a specific microenvironment that delivers signals which govern differentiation from HSCs to mature B-cells. Some of these signals are produced by stromal cells which create a network of non-lymphoid connective tissue within the bone marrow. Stromal cells form specific cellular contacts with developing B-cells through cell-adhesion molecules and their associated ligands. Further, they can produce soluble or membrane-bound cytokines/chemokines which direct B-cell proliferation and differentiation [84].

It has been shown that stromal cells play an essential role in committing developing lymphocytes to pro-B-cell lineages (Figure 1.11). Stromal cells express the ligand for fms-like tyrosine kinase 3 (FLT3) which binds to FLT3 on MMPs. They also produce the chemokine C-X-C motif chemokine ligand 12 (CXCL12) which helps to retain MMPs to appropriate locations within the bone marrow. FLT3 signalling results in the differentiation of MMPS to CLPs, which express the receptor for interleukin-7 (IL-7). Stromal cells release IL-7, and it acts as a vital signal for the growth and survival of murine developing B-cells (role in human B-cell development is debated). Stromal cells produce vascular cell adhesion molecule 1 (VCAM-1), alongside other CAMs, which binds to the integrin very late antigen-4 (VLA-4) on MMPs. This promotes binding of the receptor tyrosine kinase Kit on the surface of pro-B-cells and the membrane-bound cytokine stem-cell factor (SCF) on stromal cells, resulting in further proliferation [84].

Developmental steps from early pro-B-cell to immature B-cell are categorized by the rearrangement and expression of antibody genes. The first recombination occurs in the $Ig_H$ between D and J gene segments, typically happening at the early pro-B-cell stage on both alleles [85]. The majority of DJ recombination events are productive because most D genes do not contain stop codons when translated in any of the three reading frames [86].

At the late pro-B-cell stage, additional $Ig_H$ recombination occurs between a V gene segment and a DJ gene. Initially this only occurs on one chromosome and if successful the second $Ig_H$ chromosome does not undertake any further recombination. This is called allelic exclusion and it stops productive VDJ

recombination happening at both $Ig_H$ alleles, ensuring that each B-cell only produces a single specific receptor. However, if recombination on the first chromosome is unsuccessful, a rearrangement of the second chromosome does ensue. In the event that neither of the chromosomes yield a productive $Ig_H$, the pro-B-cell is eliminated [87].

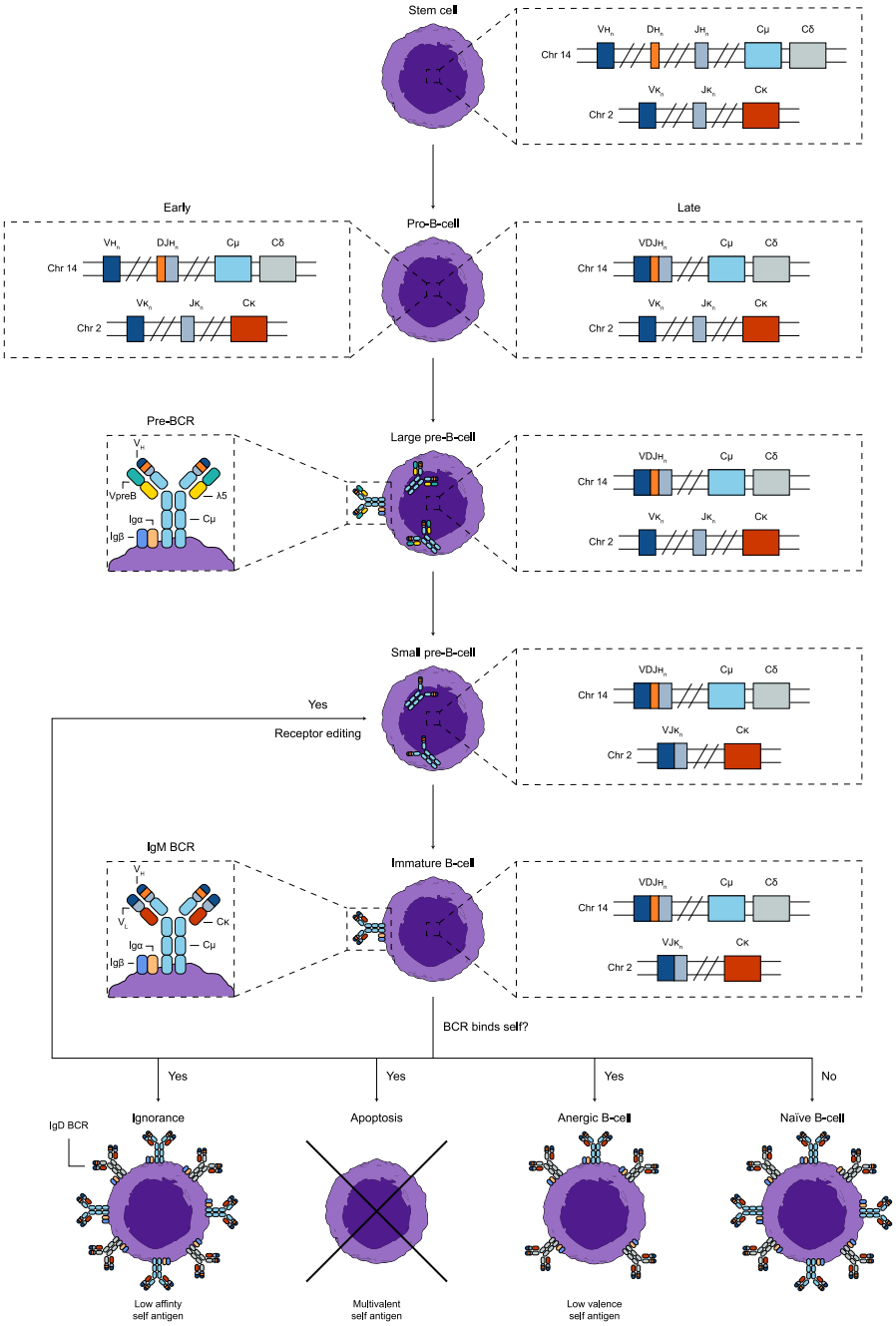



*Figure 1.10: Timeline of B-cell development and central tolerance.*
*The development of a B-cell lineage proceeds through a series of definable stages. In stem cells, the V(D)J gene segments are found in their germline configuration. Rearrangement of the $Ig_H$ DJ gene can begin in CLPs but mostly occurs in early pro-B-cells, which gives rise to late pro-B-cells that generate a complete $V_H$ gene. The $V_H$ genes are then expressed with a Cμ, two surrogate $Ig_L$ proteins (λ5 and VpreB) and two signal transducing proteins (Igα and Igβ), as a pre-BCR in the cytoplasm and to some extent, on the surface of large-pre-B-cells. This yields small pre-B-cells, which stop producing surrogate $Ig_L$ and express the $Ig_H$ only in the cytoplasm. Instead, the $Ig_L$ VJ genes are recombined and IgM BCRs are eventually expressed on the surface of immature B-cells. At the immature B-cell stage, it is determined whether the newly generated BCR is autoreactive. If the BCR is autoreactive there are four possible fates, including deletion, anergy, receptor editing and ignorance. Adapted from [628].*

The $V_H$ genes are expressed in conjunction with a Cμ, two surrogate proteins that resemble an $Ig_L$ (λ5 and VpreB) and two signal transducing proteins (Igα and Igβ), forming a pre-B-cell receptor (pre-BCR). Successfully assembled pre-BCRs form dimers or oligomers at the cell surface, signalling that a productive $Ig_H$ has been generated and marking the transition from the late pro-B-cell to large pre-B-cell. Large pre-B-cells go through a period of proliferation, giving rise to many small pre-B-cells [87,88].

In small pre-B-cells $Ig_L$ VJ recombination occurs and similarly to $Ig_H$, this process exhibits allelic exclusion. Although, there is added redundancy at this stage due to the organization of the $Ig_L$ genes. If a particular VJ recombination is not productive, further rearrangements may occur preventing the use of the second chromosome. Additionally, the $Ig_L$ generally only produce either an $Ig_L$κ or $Ig_L$λ, with $Ig_L$κ recombination tending to precede that of $Ig_L$λ. If either of the $Ig_L$κ alleles produce a suitable $Ig_L$, rearrangement at the $Ig_L$λ loci will cease and this is known as isotypic exclusion. However, if all of the recombination events at both of the $Ig_L$κ alleles are unsuccessful, then an $Ig_L$λ chromosome will begin to undergo rearrangement. If all of the recombination events at the $Ig_L$κ and $Ig_L$λ loci are failures, the pre-B-cell is lost. After a productive $Ig_H$ and $Ig_L$ have successfully paired, they are expressed as an IgM BCR on the cell surface and the pre-B-cell becomes an immature B-cell [87,88].

At the immature B-cell stage, it is determined whether the newly expressed BCR has reactivity to self-antigens. The goal of this process, known as central tolerance, is to remove or inactivate autoreactive B-cells. If immature B-cells are not autoreactive, they are released into the periphery where they circulate and mature to naïve B-cells expressing IgD as well as IgM. However, B-cells that are found to bind self-antigens during central tolerance have four possible fates, these include deletion, anergy, receptor editing and ignorance. Deletion results in the removal of autoreactive B-cells via apoptosis and this appears to occur mainly when the self-antigen is multivalent. Anergy is when B-cells become unresponsive and are not stimulated by antigen binding, this usually occurs when the self-antigen is of low valence. The migration of anergic B-cells within the peripheral lymphoid organs is impeded and they are unable to compete with non-autoreactive B-cells for survival signals. However, there is a chance that an autoreactive BCR can be rescued via undergoing further rearrangements of the $Ig_L$ genes (and perhaps the $Ig_H$ genes), this process is called receptor editing. If the newly recombined $Ig_L$ yields a BCR that no longer binds self-antigens, then its development will proceed as normal.

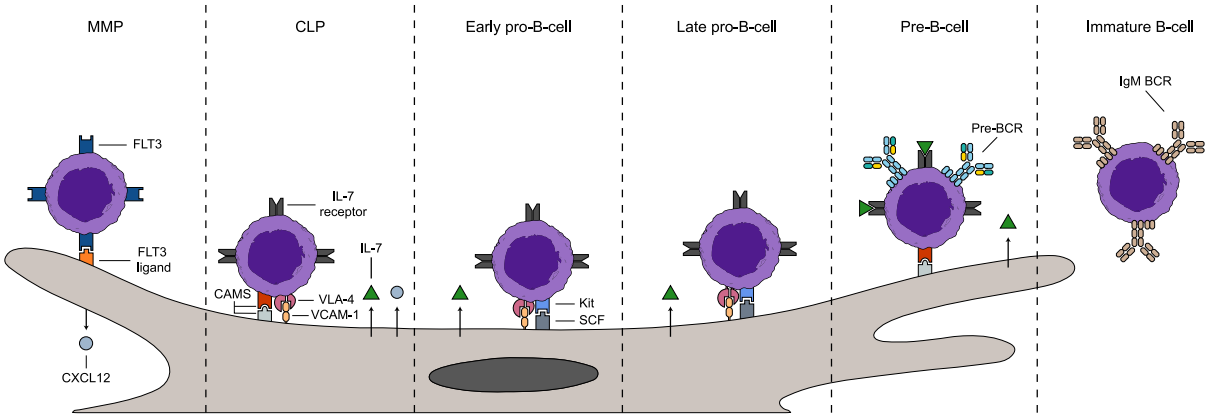



*Figure 1.11: Stromal cells aid B-cell development.*
*Bone marrow stromal cell interactions with B-cell progenitors are required for successful development to the immature B-cell stage. MMPs express tyrosine kinase FLT3, which binds to its ligand on stromal cells. MMPs are retained due to their attraction to the chemokine CXCL12, continued FLT3 signalling enables differentiation to the next stage. CLPs express the receptor for IL-7, and IL-7 is produced by stromal cells allowing further development. The CLPs bind to adhesion molecules VCAM-1 through their integrin VLA-4. This interaction promotes the binding of the receptor tyrosine kinase Kit on the surface of pro-B-cells with membrane-bound cytokine SCF on stromal cells, inducing proliferation of the subsequent B-cell progenitors. Adapted from [628].*

Numerous rounds of rearrangement can occur but if the BCR remains autoreactive, it is eliminated. Ignorance is when the contact between the BCR and its self-antigen (if made at all) is not strong enough to produce an appropriate signal during development. As a result of ignorance, self-targeting B-cells may be released into the periphery [89].

### 1.2.8 B-cell activation and peripheral tolerance

Naïve B-cells first come into contact with antigens in the periphery at secondary lymphoid organs (SLO) such as lymph nodes and the spleen (Figure 1.12) [90]. Within SLOs, T-cells and B-Cells inhabit two distinct zones and these are known as the T-cell areas and primary follicles respectively. When naïve B-cells enter a SLO they migrate to the primary follicles due to the attraction of CXC-chemokine receptor 5 (CXCR5) to its chemokine ligand CXCL13, which is released by the residing follicular dendritic cells (FDC). When specialized macrophages or FDCs present an appropriate antigen to a naïve B-cell within the primary follicle, the B-cell starts to express the chemokine receptor C-C chemokine receptor (CCR7). The naïve B-cell is attracted to the border of the T-cell area, where chemokine ligands for the CCR7 are produced by stromal cells and dendritic cells. Conversely, naïve T-cells express CCR7 but not CXCR5, hence their localization within the T-cell area. When a naïve T-cell is primed with a suitable peptide by a dendritic cell, CXCR5 expression begins to occur. Whilst some T-cells will mature into effector cells and leave the SLO, others will differentiate into T-follicular helper cells ($T_{FH}$ cells). The $T_{FH}$ cells move to the border of the primary lymphoid follicle, where they can now encounter B-cells which have been activated [91,92].

When a suitable antigen is presented to a B-cell, it becomes internalised and processed into peptides, which are expressed bound to MHC class II molecules (Figure 1.13). Importantly, the peptide presented on the B-cell surface is not necessarily the antigen that was originally bound by the BCR. For instance, a BCR that recognises a viral spike epitope could cause internalisation of an entire virion, and any of the peptides from this pathogen could be displayed as a peptide:MHC complex. Through a process called linked recognition, B-cells are only activated by helper T-cells that have been primed to recognise the peptide presented on the MHC class II molecule. Alongside such an interaction, additional contacts are made between the CD40 and its ligand, expressed by B-cells and helper T-cells respectively. The CD40 based engagement triggers proliferation of both cell types, and cytokine release by the helper T-cell (IL-4, IL-5 and IL-6) [93].

Following successful interactions with $T_{FH}$ cells, naïve B-cells appear to have three possible fates during the first phase of a primary response. Some naïve B-cells can differentiate into short-lived plasma cells, secreting low affinity antibodies of switched and non-switched isotype. Short-lived plasma cells tend to have lifespans restricted to the course of the infection, accumulating within the medullary cords of lymph nodes and the red pulp of the spleen [94]. Taking another differentiation route, some naïve B-cells migrate with their associated T-cells back to a lymphoid follicle (now referred to as a secondary lymphoid follicle) where they continue to proliferate, eventually forming a structure called a germinal centre (GC). Such B-cells, known as GC B-cells, differentiate further during the second phase of a primary response [92]. Other naïve B-cells can differentiate directly into memory B-cells, without forming GCs first [95].

The second phase of a primary response takes place in GCs and involves antigen affinity selection of GC B-cells. GCs mostly consist of proliferating B-cells but resting B-cells, helper T-cells, stromal cells and FDCs are also present. Proliferating GC B-cells shift the resting B-cells towards the follicle edge, forming a mantle zone around two histologically distinguishable regions called the light and dark zones. Dark zone GC B-cells are called centroblasts, initially they proliferate rapidly whilst expressing the chemokine receptors CXCR4 and CXCR5, with lower levels of surface antibody (particularly IgD).

The stromal cells within the dark zone produce CXCL12, the ligand of CXCR4, as a means to retain the centroblasts [96]. During dark zone clonal expansion, SHM is initiated and it alters the genetic code of the IgM centroblasts. The mutations introduced have the potential to make the BCRs bear a higher or

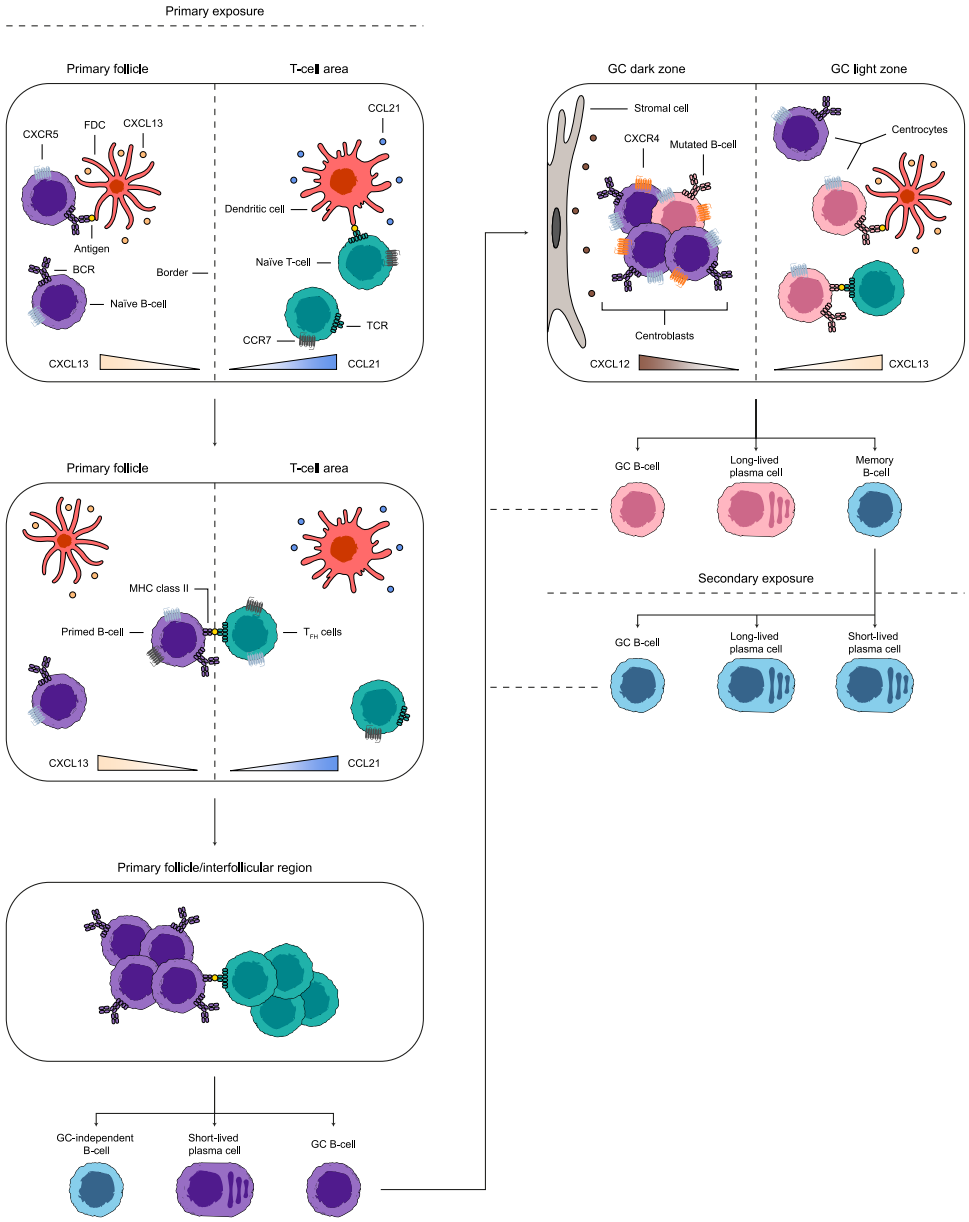



*Figure 1.12: Phases of B-cell activation and affinity maturation.*
*Naïve B-cells first come into contact with antigens in the periphery at SLOs, where T-cells and B-cells occupy distinct zones (T-cell areas and primary follicles). They migrate to these areas due to the attraction of chemokine receptors to their ligands. When B-cells and T-cells are primed with an appropriate ligand by antigen presenting cells, they begin to express the chemokine receptor specific for the ligand found in the opposing region. Both cell types migrate to the boarder of the T-cell area and primary follicle, and complementary cells encounter each other. Both cell types proliferate and B-cells differentiate into GC-independent memory B-cells, short-lived plasma cells or GC B-cells. GC B-cells migrate with their associated T-cells to a secondary lymphoid follicle where they continue to proliferate, eventually forming a GC. B-cells occupy two regions in the GC (dark and light zones) due to the attraction of chemokine receptors to their ligands. Dark zone GC B-cells (centroblasts) undergo proliferation and SHM. The centroblasts eventually lower their cell division rate and stop expressing CXCR4, enabling migration to the light zone, becoming centrocytes. Here, the centrocytes compete for a limited antigen supply. GC B-cells that survive have three possible fates, including long-lived plasma cells, memory B-cells or GC B-cells. Following a secondary exposure, memory B-cells become re-activated, repeating these processes. They can differentiate into long or short lived plasma cells or re-enter the GC to undergo further rounds of affinity maturation. Adapted from [92,628].*

lower affinity for their associated antigens, and could also make them autoreactive [97]. Eventually, centroblasts lower their cell division rate, stop expressing CXCR4 and produce greater levels of surface antibody. Such cells, now referred to as centrocytes, are able to migrate to the light zone where FDCs are producing CXCL13. Centrocytes do undergo proliferation within the light zone, but at a markedly reduced rate compared to the dark zone. Within the light zone, centrocytes compete for limited levels of antigen presented by FDCs. Higher affinity centrocytes are able to take up antigens and present them to nearby helper T-cells, thus receiving signals important for survival and CSR. Conversely, low affinity or autoreactive centrocytes are unable to receive such signals and are eliminated [96,97].

GC B-cells that survive affinity maturation during the second phase of a primary immune response have three possible fates. Firstly, some GC B-cells become long-lived plasma cells and they are found within the bone marrow. Long-lived plasma cells actively release large amounts of antibody and persist for years, sometimes a lifetime, despite lack of antigen or secondary exposure [98]. Some GC B-cells give rise to quiescent long-lived memory B-cells that take up residence in niches within SLOs, promoting their subsequent exposure to antigen. Other GC B-cells re-enter the GC dark zone, undergoing further rounds of SHM and selection, with survivors differentiating to memory B-cells or long lived-plasma cells. In the event of a secondary exposure, memory B-cells become activated and give rise to either short-lived plasma cells, long-lived plasma cells or GC B-cells, thus repeating the cycle of affinity maturation [92,99].

Memory B-cells as well as long-lived and short-lived plasma cells can also be generated by activating B-cells with T-cell independent (TI) mechanisms. There are two types of TI antigen that can activate B-cells, TI-1 antigen and TI-2 antigen. TI-1 antigens consist of lipopolysaccharide (LPS), bacterial DNA and particular viral elements [100,101]. TI-1 antigens interact with toll-like receptors (TLR) and BCRs expressed by B-cells, leading to activation. TI-1 activated B-cells only generate antibodies of the IgM class and do not go through SHM or CSR. TI-2 antigens consist of large, highly repetitive carbohydrate or protein epitopes. Large amounts of TI-2 antigen are displayed on the surface of pathogens, hence multiple BCRs can be stimulated simultaneously. This triggers a Bruton's tyrosine kinase (Btk)-dependent

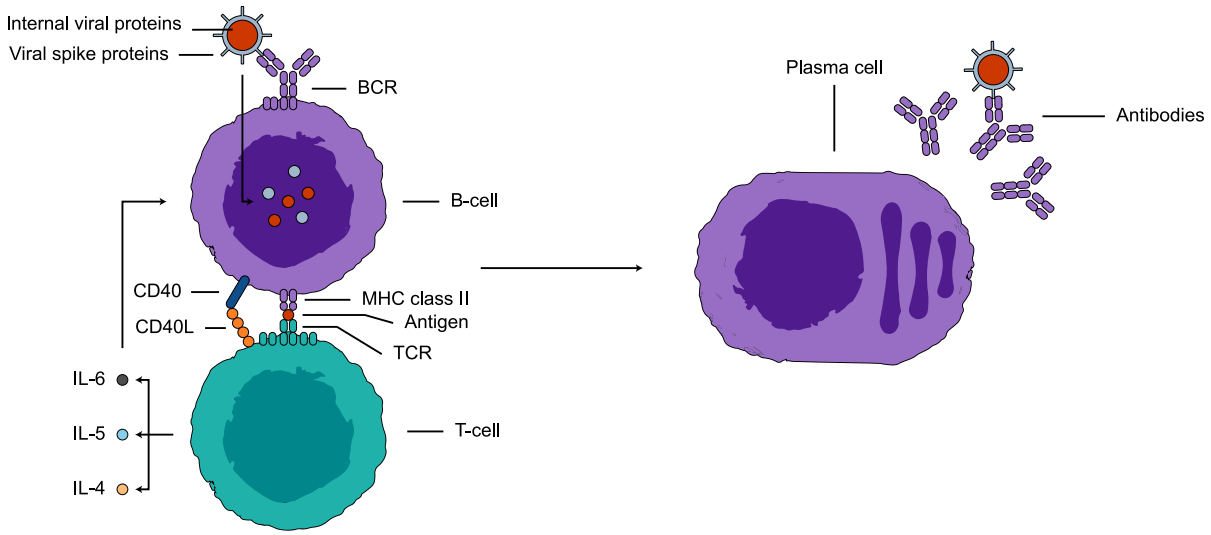



*Figure 1.13: B-cell stimulation and linked recognition.*
*When a B-cell encounters the appropriate antigen, it internalises and processes it into peptides that can be expressed on an MHC class II molecule for helper T-cells. Additional contacts are made between B-cell CD40 and its ligand found on T-cells. This triggers proliferation of both cell types, and cytokine release by the helper T-cell. The peptide presented on the MHC class II molecule does not have to be the same as the one originally bound by the BCR. As seen here, the B-cell can bind a viral coat protein, internalise the entire virion and then present peptides from the internal viral proteins. Adapted from [628].*

signalling cascade, activating the B-cell in the absence of T-cell or TLR interaction [102]. TI-2 antigens have been shown to be capable of generating memory B-cell responses. However, such B-cells are phenotypically distinct from memory B-cells developed using T-cell help, in addition to having low levels of SHM and CSR [103,104].

It is possible for some self-reactive B-cells to escape into the periphery despite the actions of central tolerance, through anergy or ignorance. Moreover, the SHM that occurs during affinity maturation could result in the generation of autoreactive B-cells within the SLOs [105]. The likelihood of such B-cells becoming activated is reduced through a process called peripheral tolerance, which uses mechanism such as anergy and deletion.

Anergic B-cells show signs of prior stimulation, such as downregulation of surface IgM, but do not become fully activated. This is because they are unable to recruit the necessary T-cell help, thus chronic exposure leads to a state of unresponsiveness [106]. Similar to B-cells, immature T-cells undergo negative selection following exposure to a variety of self-antigens by the autoimmune regulatory protein [107]. Additionally, the development of potentially autoimmune helper T-cells is suppressed by the action of a population of cells called regulatory T-cells ($T_{reg}$) [108]. The chances that an autoreactive T-cell is released into the periphery, that is specific for the same self-antigen as an autoreactive B-cell is therefore low.

When immature B-cells are first released into the periphery, they are classed as transitional B-cells. Transitional B-cells can be distinguished from mature B-cells by their surface marker expression, short half-life and anti-IgM inducible apoptosis. This deletion of transitional B-cells following early activation of the BCR is considered to be a means of removing autoreactive cells [109]. Additionally, it has been shown that $T_{regs}$ can act directly on autoreactive B-cells and induce apoptosis [110].

The fate of B-cells during immune responses is tightly regulated by signals from the innate immune arm, in addition to those produced by the adaptive immune arm. Sustaining a balance between activating and inhibitory signals ensures prevention of unwanted autoreactive responses, whilst maintaining protection against invading pathogens [111].

It has been shown that macrophages and dendritic cells express a range of TLRs, enabling them to recognise microbes [112]. Activation of TLRs on these innate immune cells results in the release of inflammatory cytokines, which ultimately influence the fate of B-cells. For instance plasmacytoid dendritic cells secrete high levels of type-I interferons (IFN-I) following viral infection, which can lead to an increase in the expression of TLR7 in naïve B cells [113,114]. Further, plasmacytoid dendritic cell derived IFN-I increases memory B-cell sensitivity to TLR7 ligand, supporting differentiation to plasma cells [115]. Activation of TLRs on innate immune cells can also have a suppressive impact on B-cell populations. For example, studies have shown that LPS stimulation of macrophages and myeloid dendritic cells inhibits the differentiation of autoreactive B-cells into plasma cells. This suppressive effect was dependent on the secretion of soluble factors such as CD40L and IL-6 [116,117]. Overall, the innate arm of the immune system plays an important role in modulating B-cell responses.

### 1.2.9 IMGT numbering scheme

Antibody based research requires a numbering system that allows investigators to accurately identify and publicise the location of functionally important variable region amino acid residues. Although others exist, the ImmunoGenetics (IMGT) variable region numbering scheme is used most commonly and is the same for $Ig_H$, $Ig_L$ and TCRs. Hence, the IMGT system will be employed throughout this thesis.

Variable regions, known as V-DOMAINs in the IMGT system, are numbered from amino acid positions 1-129 and conserved residues are always assigned the same numbers (Figure 1.14a). The FR1, FR2,

FR3 and FR4 span the amino acid positions 1-26, 39-55, 66-104 and 118-129 respectively. Of note, the FR4 starts with a tryptophan for $Ig_H$ and a phenylalanine for $Ig_L$. The CDR1, CDR2 and CDR3 span the amino acid positions 27-38, 56-65 and 105-117 respectively. The CDR1 can be between 5-12 amino acids in length, the CDR2 can be between 0-10 amino acids in length and the CDR3 can be between 5-13 amino acids in length. If any of the CDRs are shorter than the maximum default lengths, gaps are introduced at the top of the loops and numbers are removed symmetrically as the loops shorten

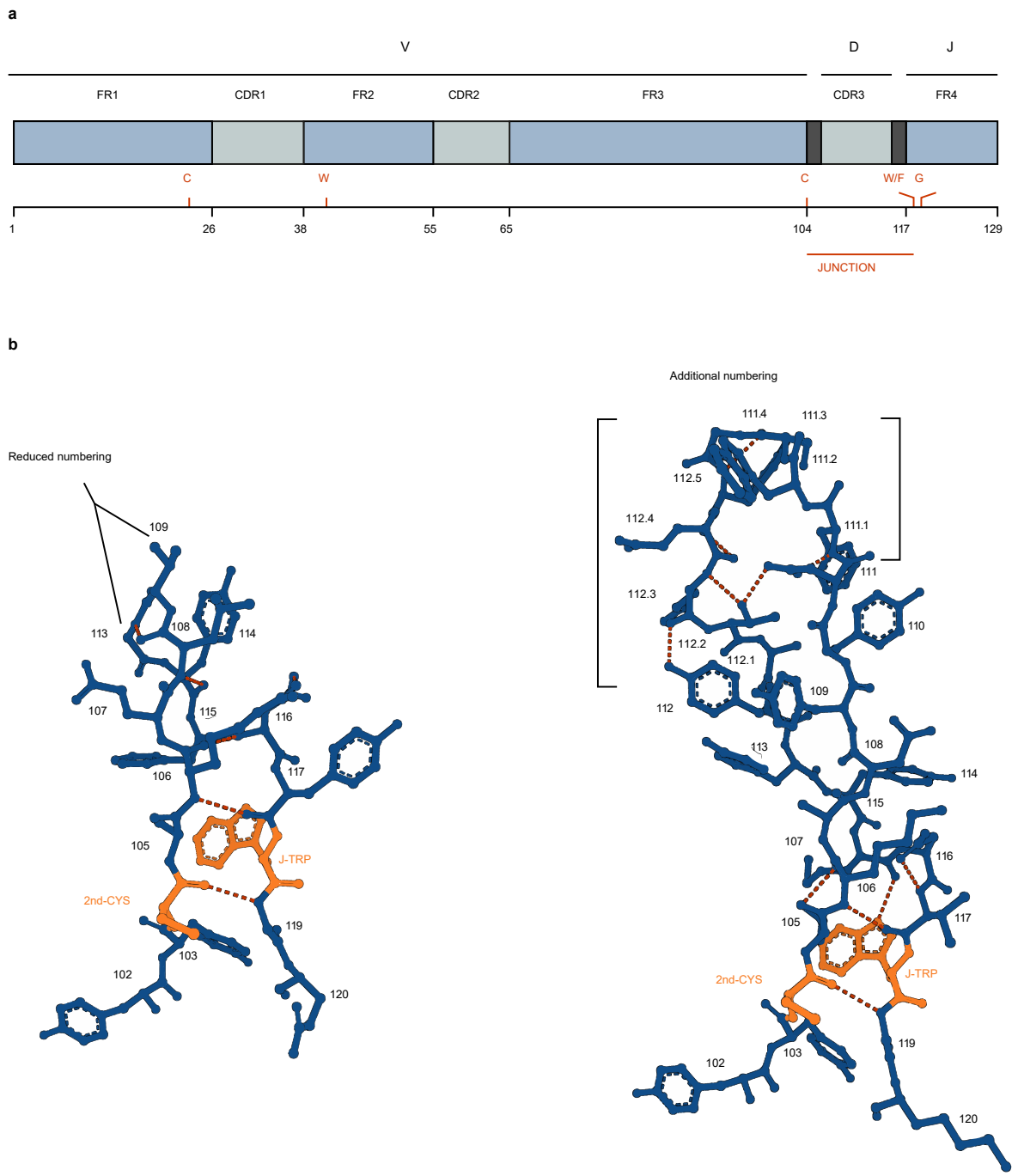



*Figure 1.14: IMGT numbering scheme and the impact of CDR length.*
*The IMGT region numbering system is the most commonly used and is the same for $Ig_H$, $Ig_L$ and TCRs. Variable regions are labelled from 1-129 (**a**) and conserved residues (red) are always assigned the same numbers. If the CDRs are shorter than the maximum default lengths (**b**), gaps are introduced at the top of the loops and numbers are removed symmetrically as the loops shorten. As CDR3s tend to have more variation, a system is in place that adds extra entities symmetrically between positions 111 and 112 as the loops get longer. Models were made using ABodyBuilder with sequences from P008_047 and P008_004* [647].

(Figure 1.14b). As CDR3s tend to have more variation, a system is in place that adds extra entities symmetrically between positions 111 and 112 as the loops get longer. Thus, the CDRs can be shortened or, for the case of the CDR3s, lengthened without altering the numbering of the structurally important amino acids. There are conserved cysteine residues at positions 23 (1st-CYS) and 104 (2nd-CYS). Further, there is a conserved glycine that forms a hydrogen bond with 2nd-CYS at position 119 (Glycine 119) and a conserved tryptophan at position 41 (CONSERVED-TRP). The region starting at 2nd-CYS and finishing with a J region tryptophan or phenylalanine (J-TRP/J-PHE) is known as the JUNCTION [118].

## 1.3 HIV-1

### 1.3.1 The AIDS pandemic and the origins of HIV

The first documented cases of the disease that would come to be known as acquired immune deficiency syndrome (AIDS) were recorded in 1981 [119]. AIDS case studies reported previously healthy subjects succumbing to atypical opportunistic infections and rare malignancies, that were normally limited to the severely immunosuppressed [120]. Initially, it was observed that AIDS most commonly affected men who have sex with men (MSM) and injection drug users. Cases in MSM could be linked sexually, and additional cases were emerging in patients with haemophilia and blood transfusion recipients. In later years, cases were also discovered in heterosexual individuals and infants [121]. These epidemiological findings began to implicate that AIDS was arising from the sexual, percutaneous or perinatal transmission of a pathogen. By 1984, at least three research groups had independently isolated and characterised a novel retrovirus from patients that had, or were considered to be at risk of AIDS [122–124]. Although a variety of names had already been proposed, in 1986 it was announced that the causative agent of AIDS would be officially known as human immunodeficiency virus (HIV) [125]. With the discovery of HIV, it was believed that a marketable vaccine would be produced within the following 2-3 years [126].

Since the start of the AIDS pandemic, there have been ~80 million HIV infections and over 36 million HIV-associated deaths worldwide. In 2020, it was estimated that ~37 million people were living with HIV and ~680,000 individuals had died from HIV-related illnesses globally [127]. The introduction of highly active antiretroviral therapy (ART) significantly slowed disease progression to AIDS in HIV patients and reduced the number of new infections. However, access to ART is not universal and AIDS is still a leading cause of death in endemic regions such as sub-Saharan Africa [120].

HIV originated via cross-species zoonotic transfers of simian immunodeficiency viruses (SIV), likely as a consequence of bush meat hunting (Figure 1.15a) [120,128]. SIV has been found in 40 species of non-human primate, which represent the natural reservoir of these viruses [129]. Infected primates are mostly considered to be asymptomatic (but not always), each hosting different strains of SIV (denoted as $SIV_{host\text{-}species}$) [130]. There are two genetically distinct versions of HIV, referred to as HIV-1 and HIV-2, and both are capable of causing AIDS. HIV-1 can be separated into four lineages (M, N, O and P), with each group arising from independent zoonotic transmission events. HIV-1 group M, otherwise known as the major group, caused the global pandemic and was transferred to humans from Chimpanzees ($SIV_{cpz}$). HIV-1 group N also resulted from a spill-over event from Chimpanzees, whilst O and P were derived from Western Gorillas ($SIV_{gor}$). HIV-2 was transferred to humans from Sooty Mangabes ($SIV_{smm}$) and is considered to be less virulent, have a lengthier asymptomatic stage and be less transmissible [120].

HIV-1 group M evolved into 9 subtypes/clades (A, B, C, D, F, G, H, J and K) and over 40 distinct circulating recombinant forms (CRF). Most clades were derived in Africa and spread undetected worldwide for decades, with different subtypes establishing themselves in distinct geographical regions (Figure 1.15b) [131]. HIV-1 group O comprises ~1% of global HIV-1 infections and is primarily found in Western Africa. There have only been 15 confirmed cases of HIV-1 group N, all of which were in or around Cameroon [132]. Only 2 confirmed cases of HIV-1 group P have been identified, both of which were in Cameroon [133]. Estimations of HIV-2 infection rates are impeded because health systems in endemic areas lack a recognised surveillance system and are unable to differentiate between HIV-1 and HIV-2. Nonetheless, it is believed that HIV-2 is mainly circulating in Western Africa, with a total of ~1-2 million individuals infected globally [134].

## 1.3.2 HIV-1 genome and structure

HIV-1 is an enveloped retrovirus (*Retroviridae*), and is part of the lentivirus genus. The HIV-1 positive single-stranded RNA genome is ~9kb in length and encodes 9 genes, which can be translated to yield

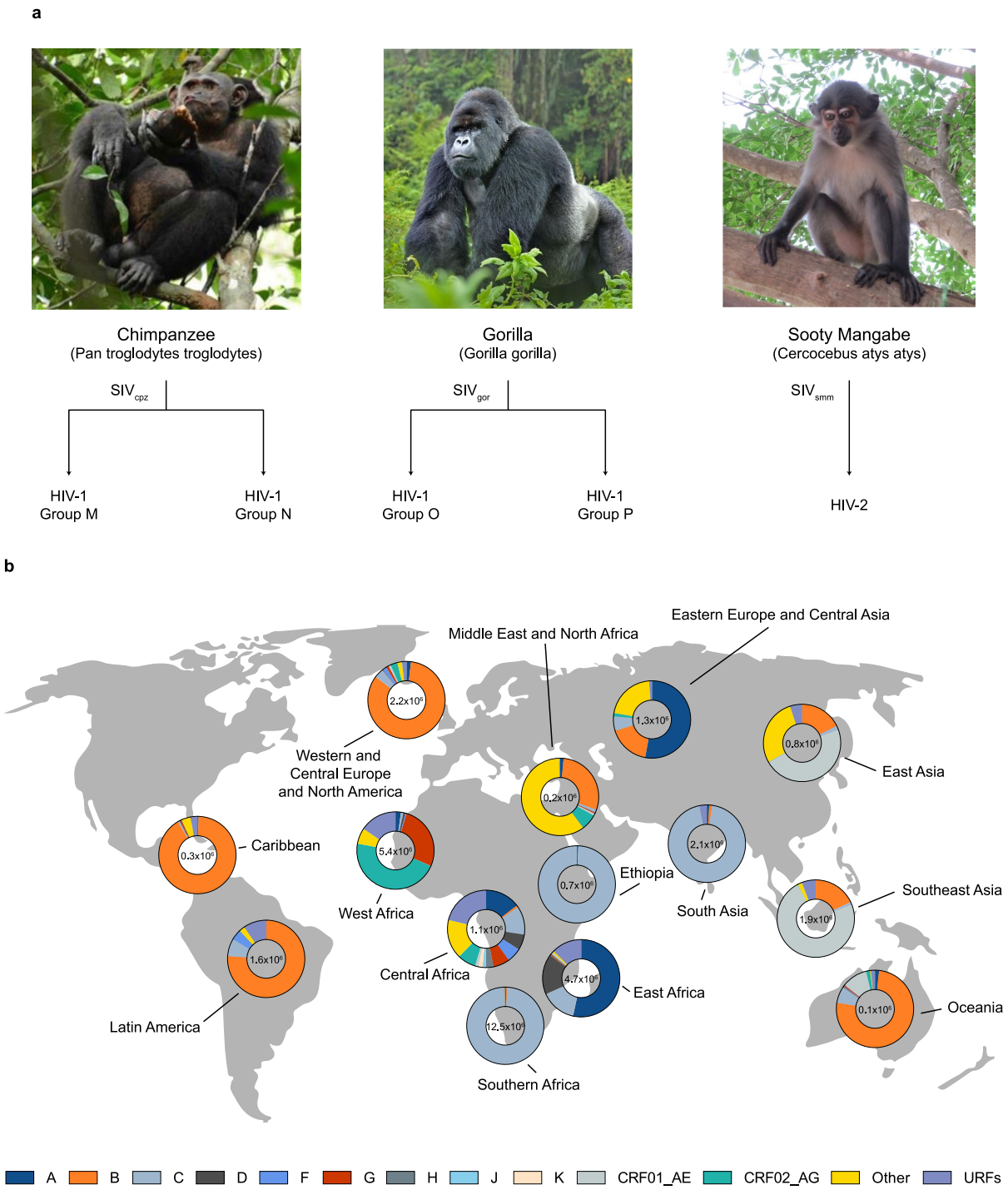



*Figure 1.15: Animal origins of HIV and regional distribution of HIV-1 clades.*
*HIV originated via cross-species zoonotic transfers of SIV (**a**). HIV-1 group M and N were transferred to humans by Chimpanzees, whilst HIV-1 group O and P were derived from Gorillas. HIV-2 was transferred to humans from Sooty Mangabes, but is considered to be less virulent, have a lengthier asymptomatic stage and be less transmissible. HIV-1 group M, which is responsible for the global pandemic, evolved into 9 clades (A, B, C, D, F, G, H, J and K) and over 40 distinct CRFs. Different subtypes established themselves in distinct geographical locations (**b**). Shown here are the global HIV-1 clade distributions from 2010-2015, and the number of people living with HIV-1 in each region is indicated in the centre of the pie charts. Adapted from [131].*

15 different proteins (Figure 1.16a). These include the structural genes group specific antigen (*gag*), polymerase (*pol*) and envelope (*env*), the regulatory genes trans-activator of transcription (*tat*) and regulator of expression of virion proteins (*rev*), and the accessory genes virion infectivity factor (*vif*), viral protein R (*vpr*), viral protein U (*vpu*) and negative factor (*nef*). The structural genes are translated as polyproteins, with *gag* incorporating matrix (MA), capsid (CA), nucleocapsid (NC) and p6, *pol* incorporating protease (PR), reverse transcriptase (RT) and integrase (IN), and *env* incorporating glycoprotein 120 (gp120) and glycoprotein 41 (gp41) [135]. To generate all 15 proteins the virus utilises leaky scanning, frameshifts, alternative splicing and post-translational cleavage of polyproteins [136]. The viral genome is flanked by stretches of untranslated regions (UTR), which consist of long terminal repeats (LTR). As it is single-stranded, complementary nucleotides within the viral genome can interact and form secondary structures [137].

A typical mature HIV-1 particle is ~100nm in diameter and spherical in appearance (Figure 1.16b). The outer shell of the virus is called the envelope, a phospholipid bilayer derived from the host cell. The

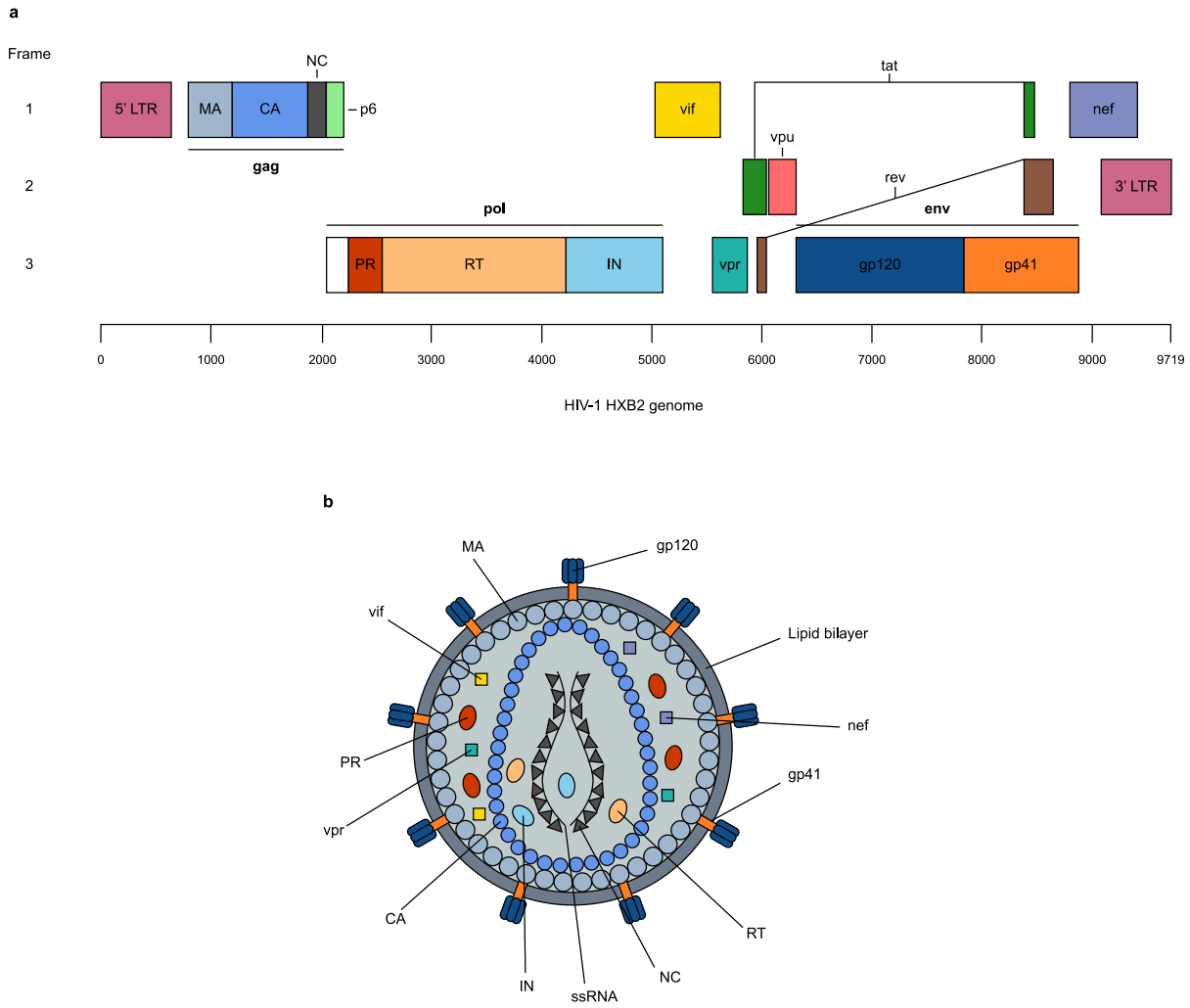



*Figure 1.16: Annotated HIV-1 genome and virion.*
*HIV-1 has a positive single-stranded RNA genome (~9kb) that encompasses 9 genes, and is flanked by UTRs that consist of LTRs (**a**). The genes encode three structural (gag, pol and env), two regulatory (tat and rev) and four accessory (vif, vpr, vpu and nef) proteins. The structural proteins are expressed as polyproteins, which are later processed into MA/CA/NC/p6 (gag), PR/RT/IN (pol) and gp120/gp41 (Env). HIV-1 virions are spherical in shape (**b**), and are broadly composed of an envelope (derived from the host cell lipid bilayer), an outer capsid membrane (made from MA proteins) and a conical shaped capsid (consisting of CA proteins). There are Env glycoproteins embedded in the viral envelope. Carried within the virions are two copies of the viral genome, alongside other essential viral proteins. Adapted from [636].*

envelope is laid over the symmetrical outer capsid membrane, which is made from MA proteins, and gives the virion its round morphology [135]. There are ~14 Env glycoproteins embedded in the envelope of each virus, alongside other cell surface proteins expressed by the host cell [138]. The Env molecules form transmembrane interactions between their gp41 subunits and the MA proteins of the outer capsid membrane. The conically shaped capsid occupies the centre of virions, it is composed of CA proteins and attaches to the outer capsid membrane via its tapered pole [135]. Two identical molecules of the viral genome are stored in the capsid, forming strong interactions with NC and p6 proteins. Viral proteins PR, RT, IN, Vif, Vpr and Nef are also incorporated in mature particles, whereas regulatory proteins Tat and Rev are absent.

### 1.3.3 HIV-1 entry

HIV-1 can spread within the body in two ways, firstly it can bud from infected cells into the extracellular space, encountering new targets from here. Alternatively, it can exploit the natural interactions between immune cells and infect via virological synapses [139]. HIV-1 can only enter cells that express CD4, which is present on $CD4^+$ T-cells and to a lesser extent macrophages, monocytes and dendritic cells [140]. Additionally, entry requires the presence of a chemokine co-receptor, either CXCR4 or CCR5 [141].

To initiate the entry process, viral gp120 contacts the N-terminus of the host cell CD4 receptor, interacting with a conserved binding site (Figure 1.17). This causes substantial conformational changes in gp120, leading to exposure of the co-receptor binding site. Indeed, gp120 then binds with the N-terminus of the host cell co-receptor, these interactions are facilitated by a particular loop within the glycoprotein (V3 loop). Further conformational changes ensue, allowing the gp41 fusion peptide (FP) to insert into the host cell membrane. The viral and host membranes are merged as a result of gp41 folding back on itself. This is mediated via the formation of a 6-helix bundle by two heptad repeats (HRs) located in gp41 [142].

### 1.3.4 HIV-1 reverse transcription

A successful fusion event results in the release of the capsid into the host cell cytoplasm, with the goal of delivering its genetic payload to the nucleus. Prior to this, the capsid must be uncoated and the viral genome reverse transcribed. Uncoating is not fully understood and different models have been proposed regarding the level of capsid uncoating that can occur before nuclear import (Figure 1.18) [143].

Reverse transcription is catalysed by RT which has two functional domains, DNA/RNA polymerase and RNaseH. As its name suggests, DNA/RNA polymerase can generate a DNA copy of either a DNA or RNA template. The RNaseH domain can degrade RNA from DNA-RNA hybrid structures. Although RT is the main mediator of this process, MA, CA, NC, IN and Vpr are also required, and together form the reverse transcription complex (RTC) [144].

Reverse transcription begins with the binding of a host tRNA primer (Lys3) to the primer binding site, located near the 5’ end of the viral genome (Figure 1.19). RT generates a complementary DNA copy of the viral 5’ UTR, as this occurs RNaseH degrades the template RNA. The generated negative sense single-stranded DNA can be transferred to the 3’ UTR. This strand transfer is possible because the ends of the viral RNA are direct repeats, also referred to as R sequences. The transfer can be made to either of the two copies of the viral genome. Reverse transcription of the complete negative sense DNA genome continues until the primer binding site is reached. The template RNA is degraded by RNaseH, except for two regions known as the 5’ and 3’ polypurine tract (ppt). RT uses the 3’ ppt as a primer and synthesises the positive sense 3’ end of the viral genome, including the first 18 nucleotides of Lys3. Hence, the positive sense section generated is complementary for the negative sense primer

binding site and can be transferred to the 5' end. Finally, RT completes extension of the positive and negative sense strands, resulting in double-stranded linear viral DNA with 5' and 3' UTRs at both ends [144].

The DNA/RNA polymerase domain of RT lacks proof-reading competency and can introduce errors at an elevated rate of $1.4 \times 10^{-5}$ mutations per base pair cycle [145]. Further, recombination events can occur during reverse transcription when two genetically distinct genomes are carried in an infecting virion. Such a scenario can arise when a cell has been infected by two different HIV-1 particles [146]. This rapid introduction of genetic diversity enables HIV-1 to evade host immune responses [135].

### 1.3.5 HIV-1 integration

At the late stages of reverse transcription, the RTC becomes the pre-integration complex (PIC) which is channelled into the host cell nucleus through nuclear pore complexes (NPC; Figure 1.18) [144]. Once inside, PIC directs integration of the proviral DNA to sites in close proximity to the NPCs. Integration tends to occur at areas that have high gene density and transcriptional activity, this process is further aided by host cell proteins [147]. Integration begins with IN multimerisation at the LTRs of the proviral DNA (Figure 1.20). In most cases, two nucleotides are removed from the 3' ends of both strands of the proviral DNA. Conserved C-A sequences are generated on the 3' strands, whilst the 5' strands are left with a two base pair overhang. The 3' C-A motifs possess highly reactive OH groups that act as nucleophiles in transesterification attacks on opposing strands of the host DNA. The positions of attack are 5 base pairs apart, around which the 3' ends of the viral DNA are covalently joined to the host DNA. The 5' overhangs are removed and the 5 base pair unit is separated, yielding an integrated

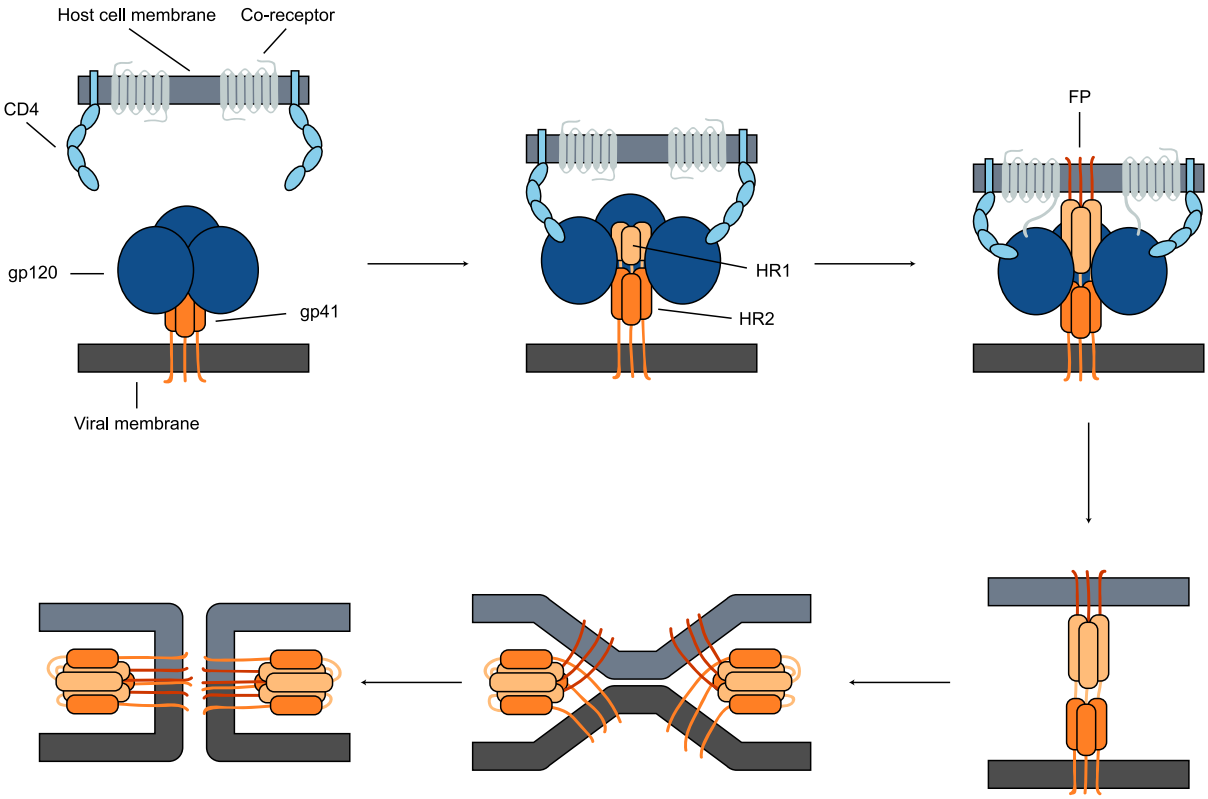



*Figure 1.17: HIV-1 Env mediates fusion of viral and host cell membranes.*
*HIV-1 entry into host cells is facilitated by Env. To begin, gp120 binds to the N-terminus of the host cell CD4 receptor. This results in conformational changes that expose the co-receptor binding site, and the V3 loop of gp120 interacts with the host cell co-receptor. Further conformational changes ensue, and the FP of the gp41 is inserted into the host cell membrane. The viral and host cell membranes are merged as the result of a 6-helix bundle formation, by the two HRs in gp41. Adapted from [142].*

genome flanked by two single-stranded gaps that used to be complementary. Finally, these single-stranded 5 base pair gaps are repaired by host cell damage repair pathways [148].

### 1.3.6 HIV-1 viral protein production

Following integration in the host cell, the proviral DNA can act as a template for the production of viral mRNA transcripts and single-stranded RNA genomes (Figure 1.21). Transcription is mediated by the 5'

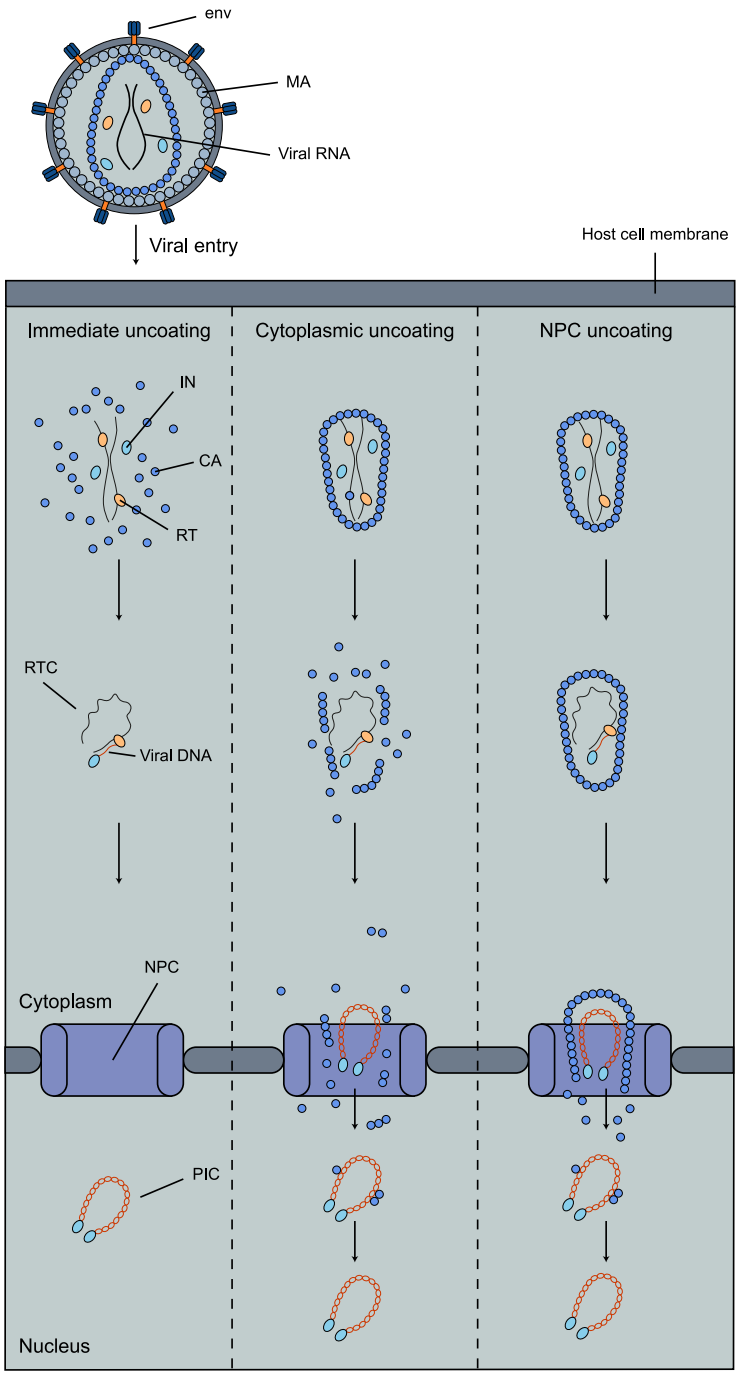



*Figure 1.18: HIV-1 capsid uncoating mechanisms.*

*HIV-1 capsid uncoating is not fully understood but different models have been suggested. The first model suggests that the core disassembles rapidly and relatively completely, soon after membrane fusion. The second suggests that core disassembly occurs in the cytoplasm, but a detectable amount of CA remains associated with the RTC and mediates interactions with host factors that enable nuclear import. The third model suggests that the core remains intact until it reaches the NPC, protecting the replicating genome from cytosolic DNA sensors. Adapted from [143].*

LTR promotor, which is recognised by host cell transcription factors and machinery, subsequently leading to the recruitment of RNA polymerase II [149]. Transcribed units of the HIV-1 genome are capped and polyadenylated, they undergo extensive alternative splicing to generate at least 70 different mRNAs. These viral transcripts can be categorised as fully spliced mRNAs (~2kb), incompletely spliced mRNAs (~4kb) or full-length unspliced mRNAs (~9kb). Fully spliced mRNAs encode the regulatory proteins *tat* and *rev*, as well as the accessory protein nef. Incomplete mRNAs encode the other accessory proteins (*vif*, *vpr* and *vpu*), as well as the structural protein *env*. Full-length unspliced mRNAs encode all of the viral proteins, and can be packaged inside new virions to serve as their genome [150].

Complete removal of introns via splicing is a prerequisite for the export of cellular pre-mRNAs, and until this occurs they are retained in the nucleus by spliceosomal factors [151]. Immediately after infection, only fully spliced proviral mRNAs are exported from the nucleus for translation [152]. Further, only low levels of these short mRNAs are transcribed due to RNA polymerase II stalling at the 5' LTR

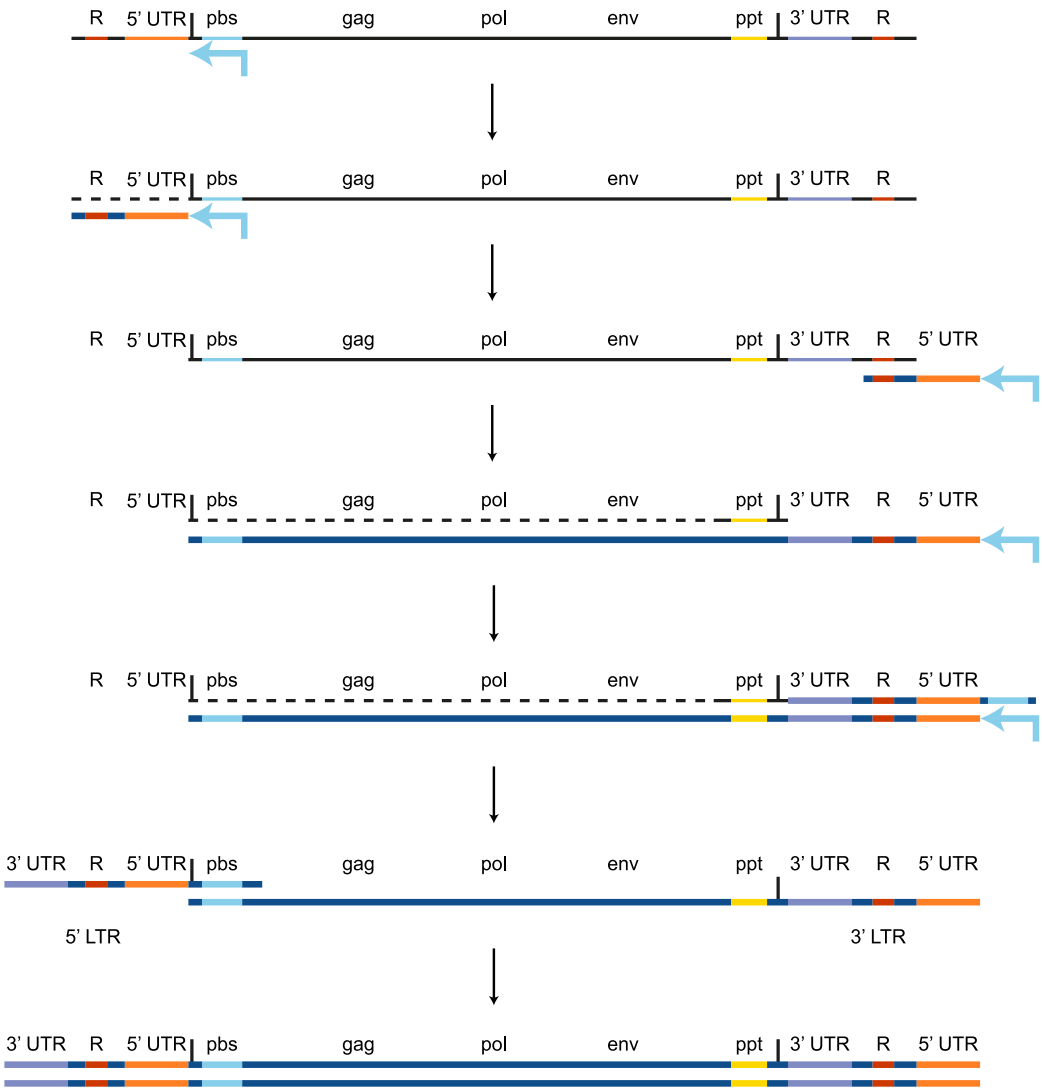



*Figure 1.19: HIV-1 reverse transcription process.*
*Reverse transcription begins with Lys3* binding *to the pbs of the HIV-1 genome. RT generates a* complementary *copy of the 5' UTR and RNaseH degrades the template RNA. The negative sense single stranded DNA is* transferred *to the 3' UTR, due to* the *complementary R sequences found in both regions. RT generates negative sense single stranded DNA* up to *the pbs. The template RNA is degraded by RNaseH* with *the exception of the ppt. RT uses the ppt as a primer to synthesise the positive sense 3' end of the genome, including the first 18 nucleotides of Lys3. This is transferred to the 5' end of the HIV-1 genome,* as it is *complementary* to the *pbs. The complete double stranded DNA genome is then synthesised by RT. Adapted from [144].*

promotor [153]. These limitations are resolved by the action of the regulatory proteins Tat and Rev, which return to the nucleus post-translation. In newly generated mRNAs Tat binds to the transactivation-response element, ultimately preventing RNA polymerase II stalling and enhancing proviral transcription. Through interaction with the Rev response element, Rev enables the trafficking of incomplete and full-length mRNAs to the cytoplasm via NPCs [152].

The open reading frames (ORF) of *gag* and *pol* overlap, with Gag being generated during the conventional translation process. However a -1 programmed frameshift, mediated by a slippery sequence and pseudoknot structure in the RNA, results in the production of Gag-Pol fusion proteins in 2–10% of ribosomes [154]. After translation, Gag and Gag-Pol proteins are trafficked to the plasma membrane by actin-dependent cellular secretory pathways [155]. The MA portion of Gag is myristoylated and accumulates at lipid rich regions of the host cell membrane via protein-lipid interactions [156]. The NC domain of Gag incorporates two copies of the viral RNA genome, which Lys3 tRNA molecules have already associated with [157,158]. The p6 section of Gag recruits Vpr and the endosomal sorting complex required for transport (ESCRT) machinery required for viral budding [159,160].

All *env* transcripts have an overlapping *vpu* ORF located upstream, but each gene is translated in a separate frame. This is achieved via leaky scanning, in which the ribosome scans through and skips the first ORF, instead initiating at a downstream start codon [154]. HIV-1 *env* is translated in the rough endoplasmic reticulum as a precursor called gp160. The Envs generated are heavily glycosylated, with each monomer possessing ~30 potential N-glycosylation sites (PNGS) [161]. Glycosylation begins with the transfer of precursor oligosaccharides from dolichol to N-X-S/T sequons, where X is any amino acid

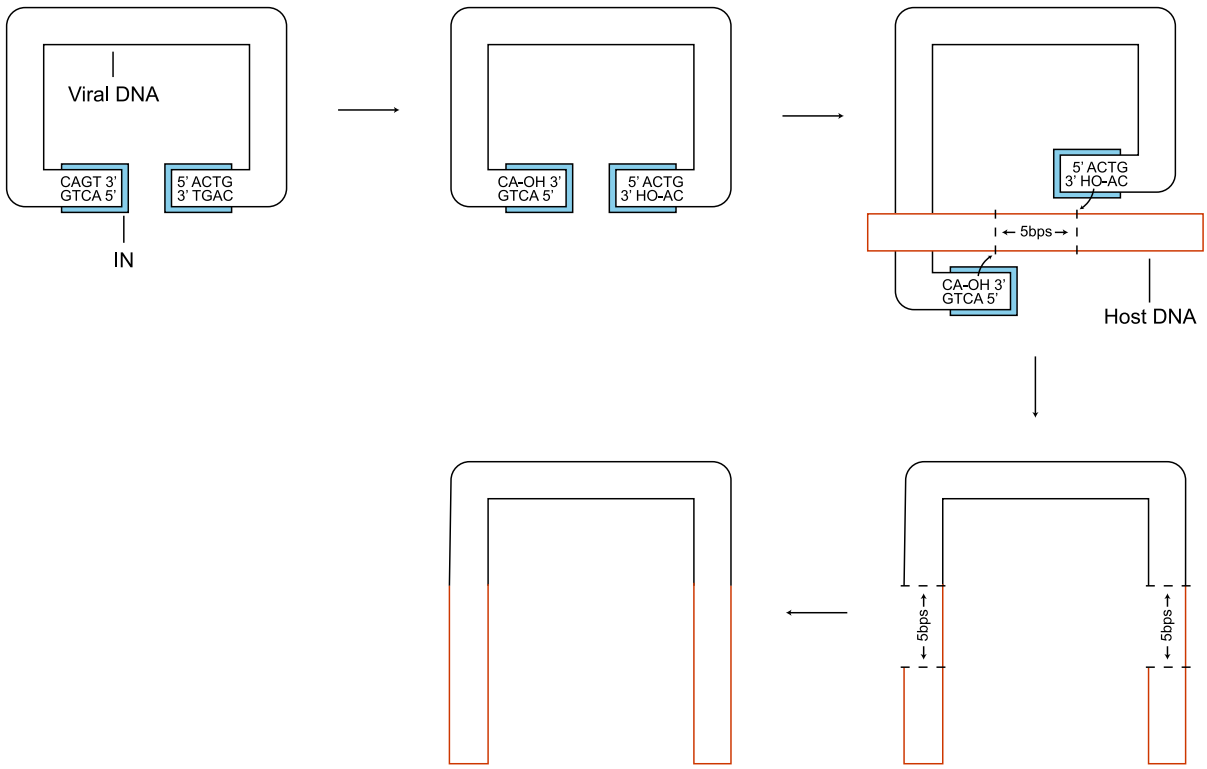



*Figure 1.20: Mechanism of HIV-1 genome integration.*
*Integration of proviral DNA into the host cell genome is mediated by the PIC. Generally, two nucleotides are removed from both 3' strands of the proviral DNA. This produces conserved C-A motifs on both ends of the 3' strands, whilst the 5' strands are left with a two base pair overhang. The 3' C-A motifs have OH groups that perform transesterification attacks on the opposing strands of the host DNA, at sites that are 5 base pairs apart. The overhangs are removed and the 5 base pair unit is separated, flanked by two single stranded DNA gaps that were once complementary. These gaps are repaired by host cell damage pathways. Adapted from [148].*

except proline, by oligosaccharyltransferase [162]. The terminal glucose-monosaccharides are trimmed by glucosidases I and II, yielding monoglucosylated N-linked oligosaccharides that can be recognised by chaperones calnexin and calreticulin [163]. The gp160 monomers undergo chaperone facilitated folding and trimerization [164]. If the proteins are folded correctly they can transfer to the Golgi, whilst incorrectly folded proteins are re-glucosylated by uridine diphosphate glucose glucosyltransferase. From here they can undergo further rounds of chaperone-mediated folding or be degraded [165]. In the Golgi the N-glycans can be trimmed by α-mannosidases, and the addition of N-acetylglucosamine to particular oligosaccharides can induce diversification to hybrid and complex-type glycans [162]. However, the high density of N-glycans means that access by α-mannosidases can be hindered sterically, resulting in under-processed high-mannose type oligosaccharides [166]. In the Golgi, gp160 trimers are

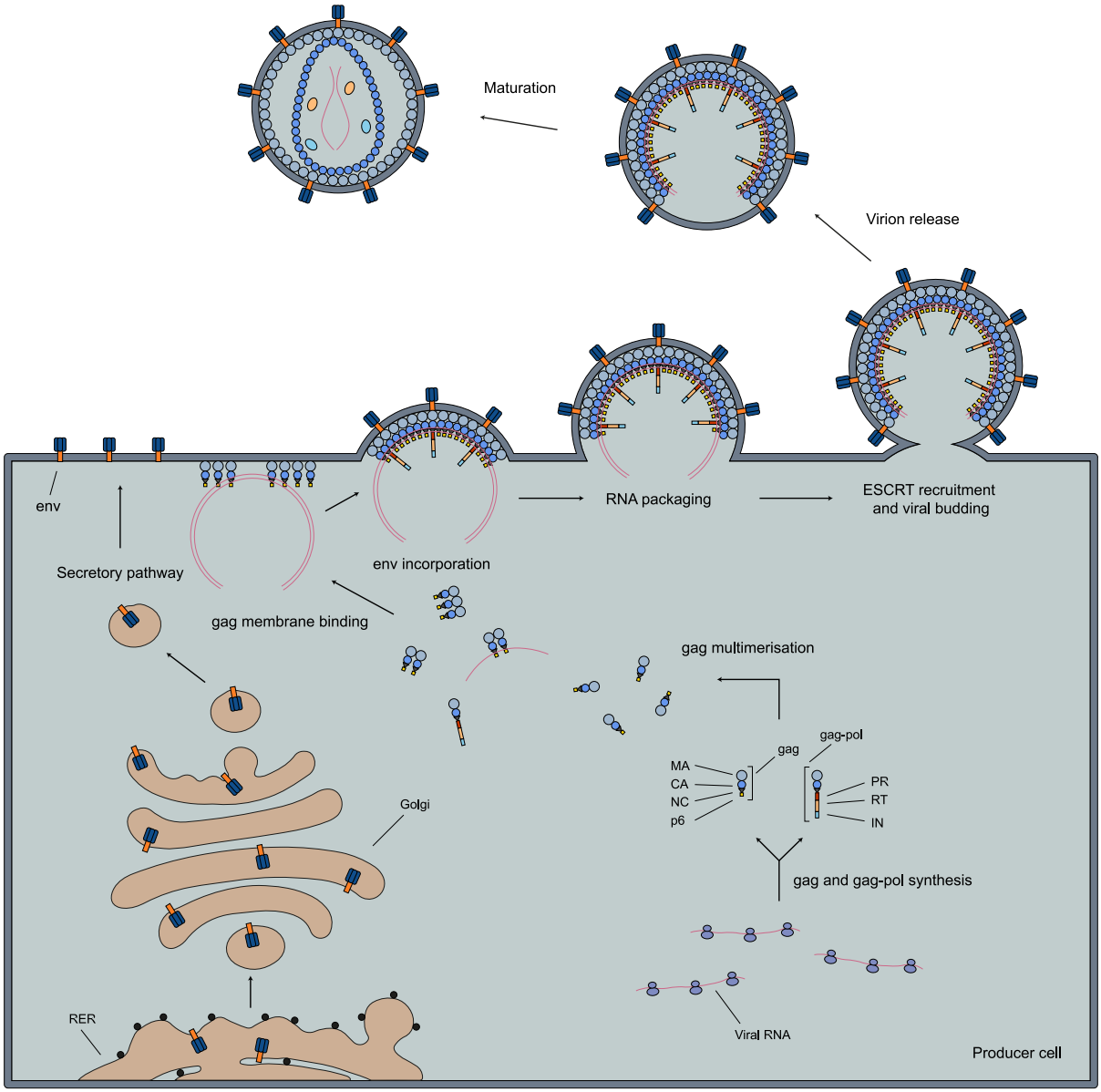



*Figure 1.21: HIV-1 translation and assembly.*
*HIV-1 gag and gag-pol polyproteins are translated in the host cell cytoplasm and trafficked to the plasma membrane. The myristoylated MA domain of gag anchors the polyprotein to phosphatidylinositol bisphosphate in the plasma membrane. The NC domain of gag binds to two HIV-1 genomes via zinc finger motifs. HIV-1 Env is translated in the rough endoplasmic reticulum and transferred through the Golgi apparatus, where it is subjected to post-translational modification. The Env is trafficked to the plasma membrane, where it is incorporated into budding virions through interactions with the MA domain of gag. The p6 domain of gag recruits the host cell ESCRT machinery to facilitate viral budding. Following release, the viral polyproteins are digested by PR, which results in a mature virion. Adapted from [167].*

cleaved by host cell furin at designated cleavage sites R-X-X-R (preferred R-X-K/R-R). This generates mature Env, trimers of non-covalently linked heterodimers of gp120 and gp41, which are trafficked to the plasma membrane [164]. Here Env can incorporate into virions via interactions with the MA domain of Gag [135].

Following successful assembly and budding of an immature particle, Gag and Gag-Pol proteins must be digested to form a mature infective virion. This requires the proteolytic activity of PR, which is packaged in particles as part of Gag-Pol fusion proteins. PR is an aspartyl protease that cleaves multiple sites in Gag and Gag-Pol triggering maturation. There is variation in the cleavage efficiency of PR at each site, making the process highly ordered [167].

### 1.3.7 HIV-1 pathogenesis

Most HIV-1 infections are established from a single variant, referred to as the founder virus, and in almost all cases it has CCR5 tropism [168]. The pathogenesis of HIV-1 can be separated into distinct stages over time based upon patient clinical data (Figure 1.22). The majority of our knowledge about the stages of HIV-1 infection is inferred from SIV studies using animal models [169].

Following vaginal challenge HIV-1 primarily infects $CD4^+$ T-cells, dendritic cells and macrophages [170,171]. Replication initially proceeds in the local mucosa, until dissemination into the local draining lymphoid tissues occurs. This area is heavily populated with $CD4^+$ T-cells and therefore enables efficient viral propagation [172]. This process is called the eclipse phase and can last ~2 weeks, throughout which the infected individual tends to have no symptoms or detectable viral load. During the eclipse phase, HIV-1 can infect cells that are in a resting state and will not actively produce virus. Therefore, latent viral reservoirs that are protected from the host immune response are generated, and at any time such cells can become activated and begin producing virus [173].

Next the infected cells enter the circulatory system, shedding virus into the blood stream and establishing a systemic infection. This signals the transition into the primary infection phase, which is characterised by a burst of plasma viraemia that peaks at ~6 weeks post-infection. The virus is spread

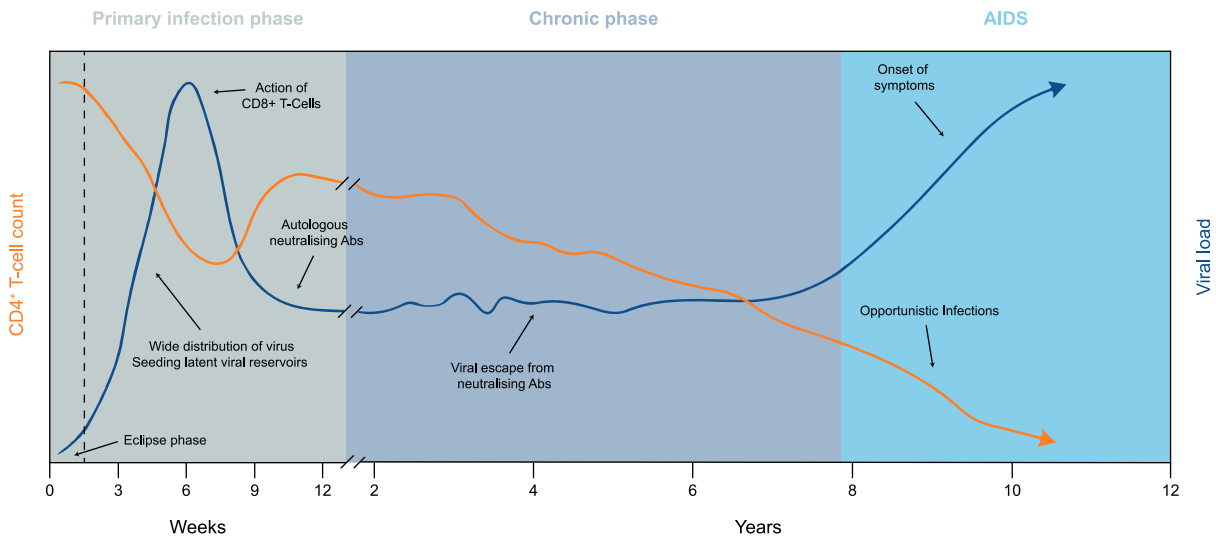



*Figure 1.22: Typical course of HIV-1 infection.*
*Early in HIV-1 infection, viral replication proceeds undetected in the lymphoid tissues (eclipse phase). The virus then sheds into the bloodstream where it can establish a systemic infection, indicated by a drastic increase in viral load and decrease in $CD4^+$ T-cell counts (primary infection phase). At ~12 weeks post-infection, the viral load starts to decline due to the action of $CD8^+$ T-cells (chronic phase). During the chronic phase the viral load remains relatively stable, while the $CD4^+$ T-cell count drops gradually overtime. Transition to AIDS occurs when $CD4^+$ T-cell count drops below 200 cell/mm3 of whole blood. Beneath this threshold, the patient will start to experience symptoms and become susceptible to opportunistic infection and malignancies. Adapted from [637].*

throughout the body, seeding more latent viral reservoirs at peripheral and distal locations [169]. Viral propagation is amplified when $CD4^+$ T-cells become activated and attempt to mount a response against the virus, only to become infected themselves. Accordingly, the circulating $CD4^+$ T-cell count of the patient rapidly decreases as a result of infection and apoptosis [173]. At this stage the patient may display symptoms of seroconversion illness including rash, painful lymphadenopathy, arthropathy and fever, but are more often clinically asymptomatic [169].

The viral load starts to decrease at ~12 weeks post-infection, due to the action of HIV-1 specific cytotoxic $CD8^+$ T-cells, after which the patient is considered to be in the chronic phase of infection. There is variation in the extent of control that each patient can exert against the infection, but typically the viral load remains detectable [173]. The initial level of plasma viraemia detected during this phase is known as the set point viral load, and is a predictor of disease progression. HIV-1 infected individuals that have a lower set point viral load progress more slowly to AIDS and vice versa [174]. Depending on viral and host factors, the chronic phase can last ~1-20 years and is delineated by a slow decline in $CD4^+$ T-cells and a continual increase in viral load [173]. Interestingly, a limited number of individuals infected with HIV-1, referred to as long-term non-progressors, can control their viral load to an unusually low or non-detectable level without clinical intervention [175].

The final phase of HIV-1 pathogenesis is AIDS, in which the patient $CD4^+$ T-cell count falls below ~200 cell/$mm^3$ of whole blood. Beneath this threshold, it is believed that the human body is no longer capable of generating an effective immune response against infections [173].

### 1.3.8 HIV-1 Env structure

The HIV-1 Env mediates entry into host cells, with both the gp120 and gp41 subunits playing key roles throughout this process (Figure 1.23a). The gp120 subunit interacts with CD4 and CXCR4/CCR5 on target cells, whilst the gp41 subunit enables entry by inducing membrane fusion [142]. Being the only viral protein expressed on the surface of the virus, Env represents the sole target for neutralising antibodies [176]. Given the collective importance of Env, an in-depth understanding of its molecular architecture is of significant interest.

The gp120 subunit consists of five variable regions (V1-V5) separated by five constant regions (C1-C5), and is typically held together by 9 disulfide bonds (Figure 1.23b) [177,178]. Each of the variable regions form loop structures on Env, and are commonly referred to as variable loops. The V1, V2, V4 and V5 loops can be highly variable, harbouring point mutations and indels, without compromising infectivity [179]. Conversely, the V3 loop can only tolerate limited indels and remain infectious, this is likely due to its role in co-receptor binding. Further, it has been shown that point mutations that alter V3 charge can influence co-receptor usage [180]. While the variable regions are immunodominant, they tend to escape the host antibody response by evolving as part of a highly heterogenic viral population, often analogised as a moving antigenic target [181]. The constant regions of Env are shielded from the host immune response by glycans that tend to have a low immunogenicity and are often referred to as the glycan shield [182].

Representing one of the most conserved parts of Env, the gp41 subunit is composed of an ecto domain, a transmembrane domain (TMD) and a cytoplasmic tail (CT), and has a single disulfide bond [164,177]. The ecto domain can be further divided into the FP, two HRs (HR1 and HR2) and the membrane proximal external region (MPER). Most sites of vulnerability in the ecto domain are shielded from the immune response by gp120 prior to receptor binding, whilst the TMD and CT are protected by the viral envelope [164].

### 1.3.9 HIV-1 neutralising antibody responses

Neutralising antibodies disrupt viral entry by multiple mechanisms including blocking receptor binding, interfering with membrane fusion and enhancing the decay of Envs [183]. Neutralising antibodies are typically able to aid the complete clearance of pathogens from the body, but this is not the case during HIV-1 infection. This is because HIV-1 integrates into the host cell genome, and establishes latent viral reservoirs which cannot be penetrated by neutralising antibodies [184]. HIV-1 has a high mutation rate that leads to the production of quasispecies, capable of evading neutralising responses [135]. Further, shedding of gp120 from Env can result in non-neutralising, immunodominant

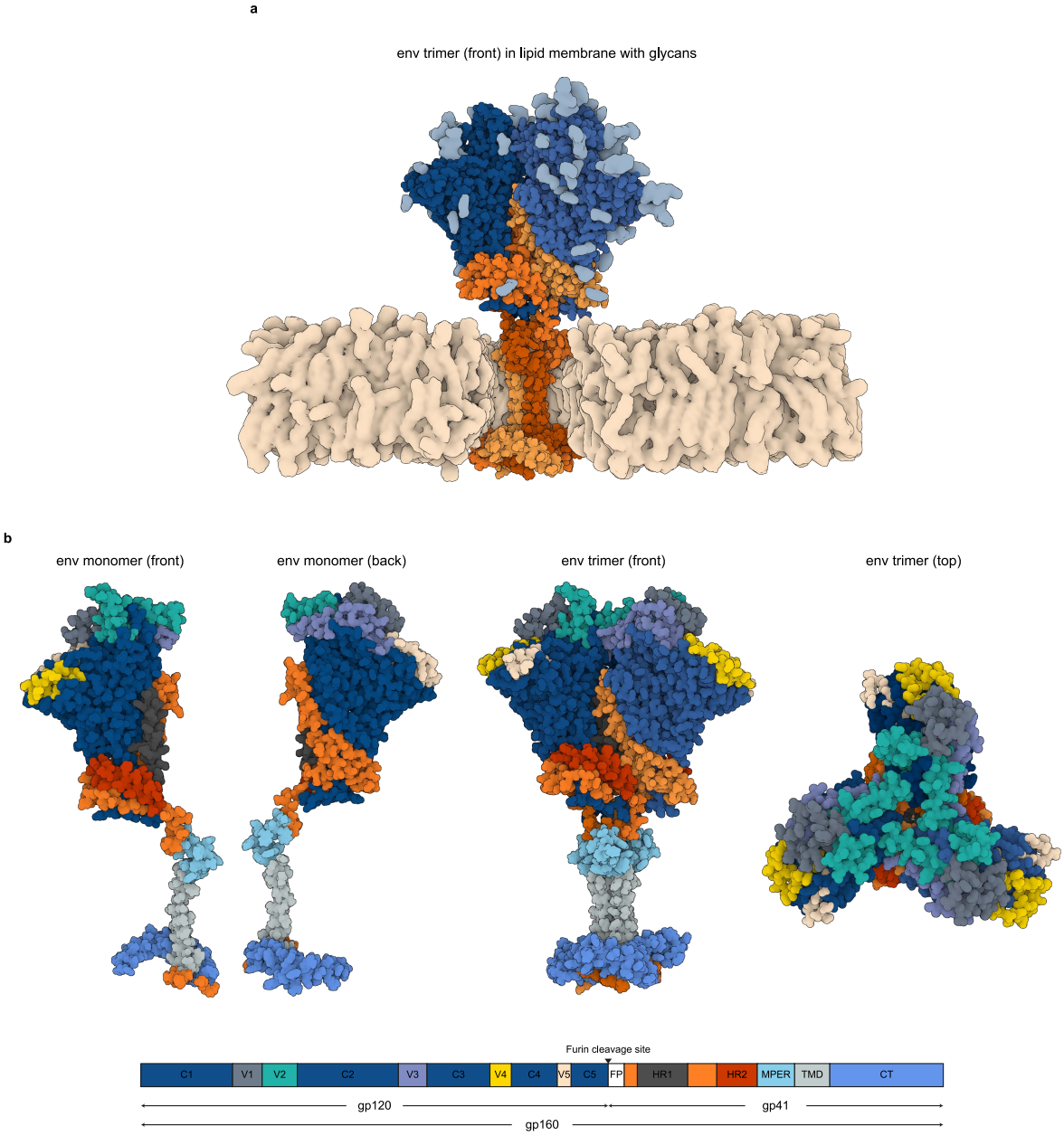



*Figure 1.23: Annotated structure of HIV-1 Env.*
*The HIV-1 Env glycoprotein is a trimer of non-covalently linked gp120 (blue) and gp41 (orange) heterodimers expressed on the surface of virions (**a**). The env molecules are heavily glycosylated, with each monomer possessing ~30 PNGS (partial glycans shown in steel blue). The gp120 subunit is composed of variable (V1-V5) and constant (C1-C5) regions, and is responsible for CD4 and CXCR4/CCR5 receptor binding (**b**). The gp41 subunit contains the FP, HR1, HR2, MPER, TMD and CT, and enables the fusion of the viral and host cell membranes. The env requires cleavage by host cell furin in order to facilitate viral entry, this occurs at the boundary of the gp120 and gp41 subunits. Adapted from [452,638]. PDB codes = 4NCO and 6UJV.*

epitopes being displayed as decoys to the host immune system [185]. However, it is of note that non-neutralising antibodies generated to intact Envs can facilitate ADCC and help to control the virus [186]. There are two types of neutralising antibody responses against HIV-1, these include autologous neutralising antibodies and broadly neutralising antibodies (bnAbs).

Autologous neutralising antibodies only show specificity to particular viral strains present in the infected individual, and are unable to target heterologous strains [187]. This lack of heterologous neutralisation is highlighted in cases of superinfection, a phenomenon in which an individual chronically infected with one strain of HIV-1 subsequently becomes infected with another strain [188]. Autologous neutralising antibodies can be found in circulation by 56 days post-infection, however they are usually observed at 12-20 weeks post-infection [189]. Such antibodies can help to clear their target strains and infected cells, however this has little impact on the overall viral load due to the generation and propagation of viral escape mutants. Escape mutants evade neutralisation by changing their appearance with amino acid substitutions and indels, sometimes altering PNG sites and therefore rearranging the glycan landscape [190,191]. Contemporaneous neutralisation, in which HIV-1 strains isolated from a particular timepoint are susceptible to neutralisation by plasma taken from the same timepoint, is a rare occurrence. In such instances it is believed that the necessary mutations required to evade autologous responses incur a fitness cost to the virus, resulting in incomplete escape [192,193].

Continual viral escape and subsequent adaption by neutralising antibodies can direct antibody responses towards more conserved epitopes on Env, leading to the generation of bnAbs (Figure 1.24). Initial investigations proposed that 10-30% of subjects chronically infected with HIV-1 could develop antibodies which neutralise a wide range of globally circulating HIV-1 variants [194]. Although, one study indicated that such individuals might be of higher prevalence after finding 50% of their cohort had sera capable of broad neutralisation [195]. Unfortunately, bnAbs are slow to develop and it can take several years before HIV-1 infected adults display sera with desirable neutralisation breadth [196]. Typically, bnAbs present with unusual characteristics such as a high level of SHM, indels and a long CDRH3 [55]. The elevated level of SHM found in bnAbs suggests that they have undergone multiple rounds of affinity maturation to acquire breadth [35]. However, it remains uncertain if the extended CDRH3s observed also arise through this gradual process. Indeed, it has been shown that bnAbs can develop from germlines with pre-existing long CDRH3s produced via VDJ recombination. It has been speculated that this could explain how neutralisation breadth develops more rapidly in particular donors [197]. An additional concern is that bnAbs often display auto-reactivity to endogenous factors such as phospholipids [198]. This would imply that elicitation of these antibodies would be disfavoured due to host tolerance mechanisms. However, this is not always the case and promising bnAb lineages with no significant auto-reactivity have been successfully isolated [199].

### 1.3.10 HIV-1 bnAbs

Collectively, the bnAbs discovered so far have been shown to target numerous conserved epitopes located across the HIV-1 Env (Figure 1.25). These include the CD4 binding site (CD4bs), V1/V2 apex, V3-glycan domain, FP/gp120-gp41 interface, silent face and MPER (Table 1.4) [200]. As seen with autologous neutralising antibodies, the circulating virus tends to be resistant to the generated bnAbs implying HIV-1 can escape such responses during chronic infection [201]. Hence, a combination of complementary bnAbs would need to be present to prevent viral escape.

Of the specificities discovered, CD4bs targeting bnAbs tend to display the best combination of breadth and potency. However, they typically take the longest time to develop during natural infection and have the greatest level of SHM on average (30% or more). The CD4bs bnAbs can be separated into two classes, the loop dependent antibodies and the CD4-mimic antibodies. Loop dependant bnAbs

rely on CDRH3 mediated recognition of the CD4bs, some members of this class include b12, HJ16 and the CH103 lineage. CD4-mimic bnAbs show a restricted IgH V gene usage (VH1-2 or VH1-46), mimicking the immunoglobin-like amino-terminal domain of CD4. Potent members of the CD4-mimic class often have a five-residue CDRL3, and are referred to as VRC01-class antibodies [200]. Further, such bnAbs tend to have deletions in the CDRL1, and this is required to accommodate N276 and N462-glycans adjacent to the CD4bs [202]. Examples of VRC01-class antibodies include VRC01, 3BNC117, PCIN63 and VRC27, members of this class possess comparable levels of SHM, potency and breadth as well as a similar binding footprint [203].

Studies of antibody and viral co-evolution have shown how a loop dependant bnAb (CH103) and a CD4-mimic bnAb (CH235) can develop in a single donor. It was shown that CH103 and CH235 cooperated in donor CH505, restricting viral escape through the utilisation of distinct modes of binding. Further, it was observed that the epitope diversification induced by each lineage may have supported the maturation of its counterpart [204–206]. Another co-evolution study followed the generation of a VRC01-class antibody (PCIN63), which had relatively low levels of SHM (13%) but equivalent neutralisation capability to its fellow members. The $Ig_H$ unmutated common ancestor (UCA) and several $Ig_L$ UCA candidates were elucidated. An Env, lacking the N276-glycan and V5-glycans, which was isolated ~6 months prior to the PCIN63 lineage emergence was considered to be the variant responsible for bnAb elicitation [207].

The V1/V2 apex bnAbs are characterised by CDRH3s that are unusually long (24–39 amino acids) compared to average human CDRH3s (~15 amino acids). Despite requiring this unique feature, V1/V2 apex bnAbs are found in ~10% of broad neutralisers and comprise some of the most potent bnAbs discovered. These antibodies bind positively charged residues located on the V2 loop, commonly

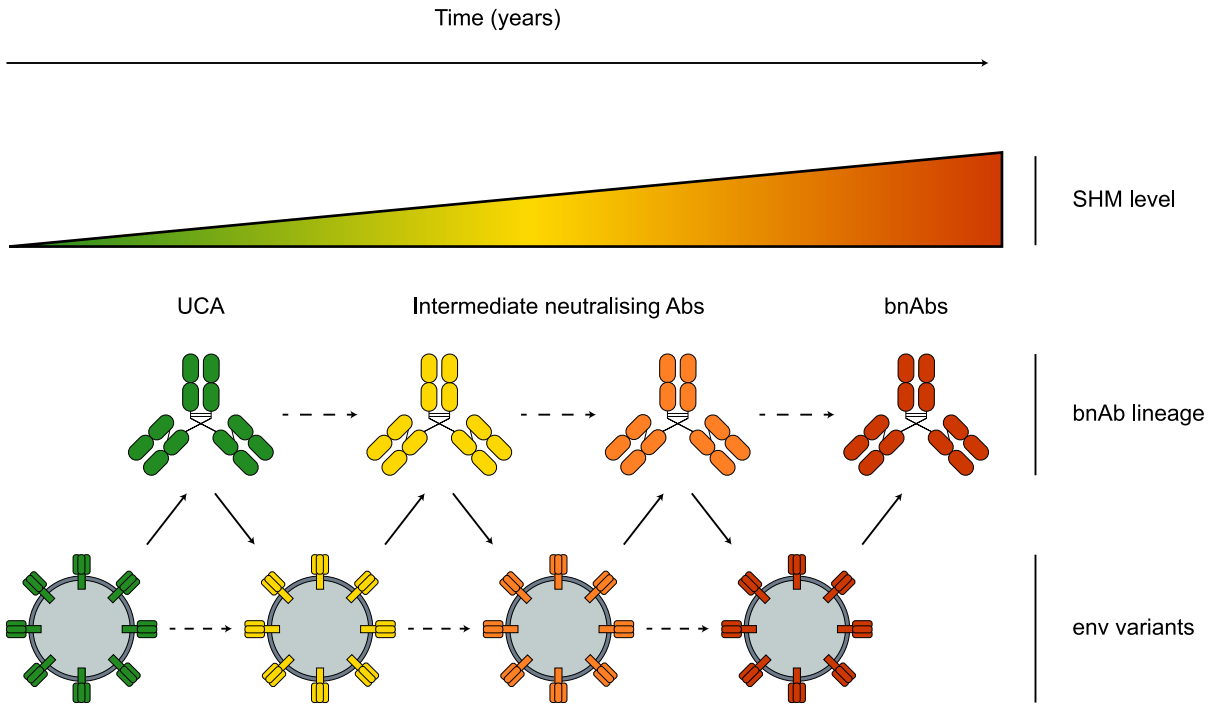



*Figure 1.24: bnAb development is driven by exposure to various Envs.*
*The development of bnAb lineages is driven by sequential exposure to HIV-1 Envs. The bnAb UCA targets the env of a specific variant, leading to affinity maturation and expansion of rare bnAb precursors. Concurrently, a selective pressure is applied on the autologous virus population, leading to the generation of escape mutants. The newly generated Envs can activate bnAbs precursors, repeating this process and directing the lineage to more conserved epitopes. This eventually results in the generation of a bnAbs, which typically display high levels of SHM as a result of multiple rounds of affinity maturation. Adapted from [199].*

relying on interactions with the N160 and N156-glycans, whilst accommodating other V1-glycans. Targeting the Env apex can result in an epitope that crosses multiple protomers, as such these antibodies often prefer (PG9, VRC38 and CH01) or need (PGT-145 and CAP256) an intact trimer for binding [208].

The development of the V1/V2 apex bnAb lineage CAP256-VRC26 was studied in a patient (CAP256) that was infected and superinfected with clade C viruses. It was shown that the Envs that triggered bnAb elicitation in this donor had evolved from the superinfecting virus. Diversification of the virus, collectively as well as at the V1/V2 apex, corresponded with CAP256-VRC26 expansion and the induction of neutralisation breadth. Early escape mutations at important residues within the epitope (166 and 169) were tolerated by members of the family that were broader, implying a mechanism for how breadth was achieved [209,210]. The co-evolution study of another V1/V2 apex bnAb lineage called PCT64 showed that similar mutations at residues 166 and 169 were important for the generation of neutralisation breadth. Further, interaction of the PCT64 UCA with early autologous Envs showed that the viral glycoform heterogeneity was crucial for germline activation [199,211].

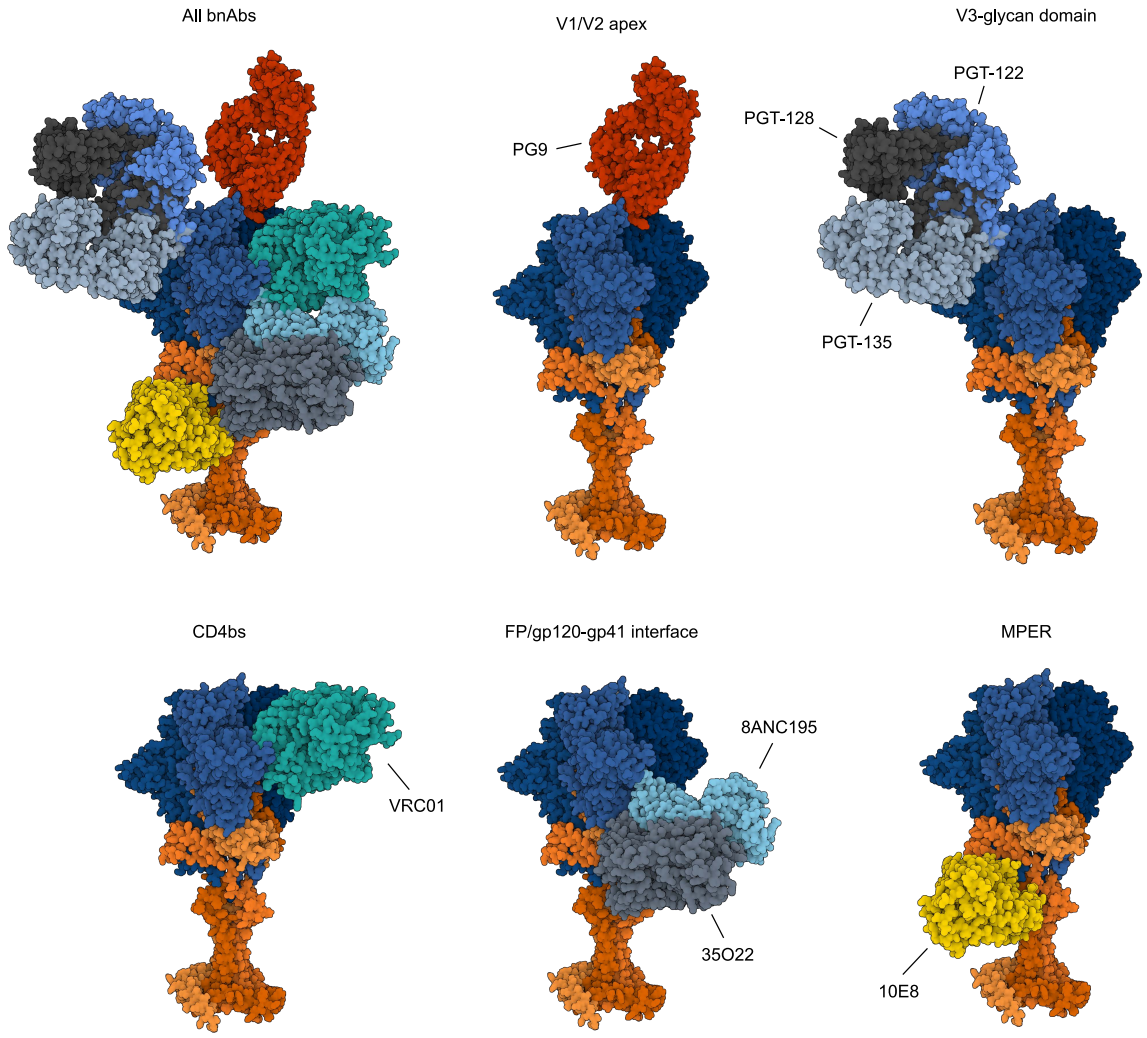



*Figure 1.25: Structures showing bnAb epitopes of HIV-1 Env.*
*There are numerous bnAb epitopes located across the HIV-1 Env. These include the CD4bs, V1/V2 apex, V3-glycan domain, FP/gp120-gp41 interface, silent face and MPER. For some epitopes, such as the V3-glycan domain and FP/gp120-gp41 interface, the bnAbs are less structurally constrained and can use various angles of approach to achieve neutralisation. Adapted from [451,452,453,638,639,640,641,642,648]. PDB codes = 4NCO, 6UJV, 4JM2, 3TYG, 3U2S, 3NGB, 5CJX, 4TVP and 5TFW.*

The V3-glycan domain is the epitope most commonly targeted by the plasma of broad neutralisers [212]. The high density of sugars present in this region sterically hinders the accessibility of glycoprocessing enzymes, resulting in high-mannose glycans which can be recognised as non-self (Figure 1.26) [166,213]. This epitope centres on the glycan at amino acid position N332, and generally requires the presence of the GDIR motif [214]. The V3-glycan domain is less structurally constrained, and the bnAbs against it can use different angles of approach, are less restricted by gene usage and show subtle distinctions in

*Table 1.4: Features of HIV-1 bnAbs.*

*Shown here are the $V_H$ gene usage, CDR3 length, $V_H$ SHM, neutralisation breadth and neutralisation potency of representative bnAbs, as well as the size of the viral panel used to investigate them. Adapted from [183].*

| Target | bnAb | VH gene and allele | CDR3 length (amino acids) | VH SHM (%nt) | Breadth | $IC_{50}$ (µg/mL; median) | Virus panel size |
|---|---|---|---|---|---|---|---|
| **CD4bs** | b12 | VH1-3*01 | 20 | 13.8 | 34% | 2.82 | 162 |
| **CD4bs** | VRC01 | VH1-2*02 | 14 | 32 | 91% | 0.33 | 190 |
| **CD4bs** | N6 | VH1-2*02 | 15 | 31 | 98% | 0.038 | 354 |
| **V1/V2 apex** | PG9 | VH3-33*05 | 30 | 12 | 79% | 0.23 | 162 |
| **V1/V2 apex** | PG16 | VH3-33*05 | 30 | 13.1 | 73% | 0.15 | 162 |
| **V1/V2 apex** | PGT-145 | VH1-8*01 | 33 | 16.2 | 78% | 0.29 | 162 |
| **V1/V2 apex** | PGDM1400 | VH1-8*01 | 34 | 24.3 | 83% | 0.003 | 77 |
| **V3-glycan domain** | 2G12 | VH3-21*01 | 16 | 20.3 | 38% | 2.38 | 162 |
| **V3-glycan domain** | PGT-121 | VH4-59*07 | 26 | 17.8 | 70% | 0.02 | 162 |
| **V3-glycan domain** | PGT-128 | VH4-59*07 | 21 | 19 | 72% | 0.03 | 162 |
| **V3-glycan domain** | PC76-33A | VH4-34*01 | 22 | 11.6 | 47% | 0.53 | 110 |
| **V3-glycan domain** | PGT-135 | VH4-39*07 | 20 | 17.1 | 33% | 0.17 | 162 |
| **FP/gp120-gp41 interface** | PGT-151 | VH3-30*04 | 28 | 18.5 | 66% | 0.29 | 117 |
| **FP/gp120-gp41 interface** | VRC34.01 | VH1-2*02 | 13 | 15 | 49% | 0.218 | 208 |
| **FP/gp120-gp41 interface** | 8ANC195 | V1-69D*01 | 20 | 27.2 | 66% | 0.39 | 118 |
| **Silent face** | SF12 | VH4-59*01 | 23 | 17 | 62% | 0.2 | 119 |
| **Silent face** | VRC-PG05 | VH3-7*01 | 17 | 6 | 27% | 0.8 | 208 |
| **MPER** | 4E10 | VH1-69*06 | 20 | 6.8 | 98% | 1.93 | 181 |
| **MPER** | 2F5 | VH2-5*02 | 9 | 12.1 | 57% | 14.6 | 177 |
| **MPER** | 10E8 | VH3-15*05 | 22 | 19.4 | 98% | 0.35 | 180 |

their amino acid and glycan requirements [208]. Although the N332-glycan is highly conserved throughout HIV-1 subtypes, most isolates can escape V3-glycan bnAbs by acquiring mutations that alter their glycosylation landscape [215]. However, some V3 bnAbs show glycan promiscuity and use other sugars in the absence of the N332-glycan [56]. Although it is not always the case, many of these bnAbs have been shown to have indels in their CDR regions that interact with the glycan shield and enable access to the co-receptor binding site [208]. Some examples of bnAbs from this specificity include 2G12, PGT-121, PGT-128, PGT-135 and 10-1074 [216–218].

When the development of V3-glycan domain targeting bnAbs was characterised in donors PC039 (PCIN39 lineage) and PC076 (PCDN lineage), maturation to breadth was observed in parallel across numerous phylogenetic branches. This suggested that breadth could be achieved via multiple Env structural arrangements, even within the same lineage. Whilst the PCIN39 bnAbs developed insertions in their CDRL1s at several branches, the PCDN lineage acquired breadth without any indels. PCIN39 and PCDN bnAbs employed the VH4-34 germline, which has been linked to autoimmune disorders, but both lineages mutated away from autoreactive features [219,220]. In another donor (CH848), it was

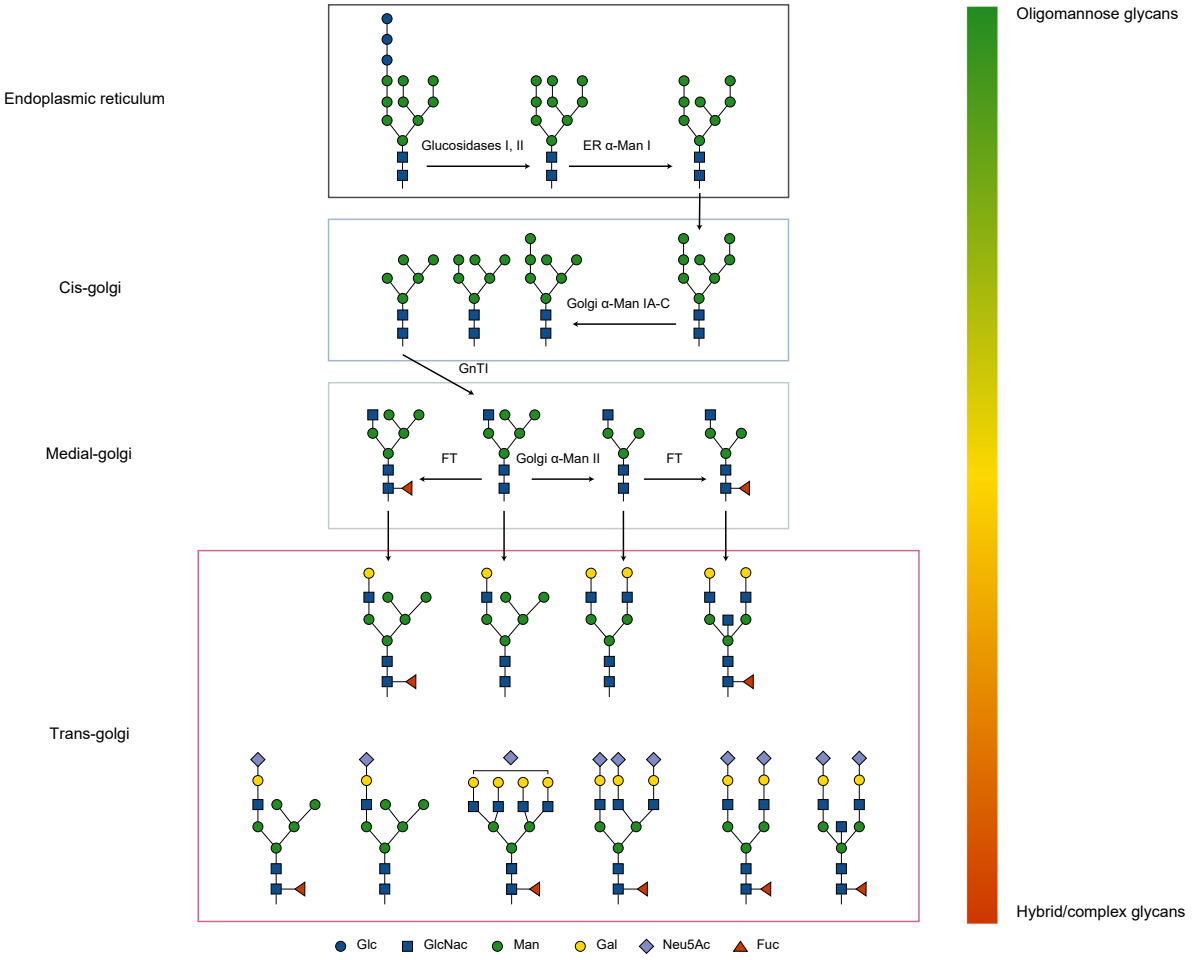



*Figure 1.26: Mammalian N-glycosylation pathway.*
*N-glycosylation begins in the rough endoplasmic reticulum. Following transfer of $Glc_3Man_9GlcNAc_2$ from dolichol-phosphate to the polypeptide, the three terminal glucose residues are trimmed down by glucosidases I and II. The terminal mannose is removed from the D2 arm by ER α-Mannosidase I before being moved to the cis-Golgi. The glycan is further trimmed by Golgi α-Mannosidases IA, IB and IC (Golgi α-Man IA-C) to $Man_5GlcNAc_2$. GlcNAc is added to the D1 arm of $Man_5GlcNAc_2$ by N-acetylglucosaminyltransferase I (GnTI), before the glycan is diversified in the medial-Golgi by α1,6-fucosyltransferase (FT), which adds core-fucosylation or further mannose trimming by Golgi α-Mannosidases II (Golgi α-Man II), followed by core-fucosylation. At the trans-Golgi, glycans are diversified to hybrid and complex-type glycans by further processing through cell-specific glycosyltransferases, sialyltransferases and fucosyltransferases. Adapted form [643].*

shown that non-neutralising antibodies against the autologous Env provided a selective pressure that shortened the V1 loop, enabling development of V3-glycan bnAbs (DH270 lineage) [221].

The FP/gp120-gp41 interface specific bnAbs interact with both the gp120 and gp41 subunits, alongside the glycans that flank this region. Examples of these bnAbs include PGT-151, VRC34.01, 35022, ACS202 and 8ANC195 [222]. Despite targeting the same region, the epitopes of these antibodies do not coincide and they do not compete with each other [223]. PGT-151 only binds to fully cleaved pre-fusion Env trimers, a feature that has been used to aid recombinant protein production and structural studies [200].

Silent face bnAbs target a heavily glycosylated region on the outer domain of gp120 centred on the N448-glycan, this area was previously considered to be immunologically silent and was named accordingly [200]. Currently only two bnAbs specific for this epitope have been isolated, which are VRC-PG05 and SF12. Both of these bnAbs interact with glycans located at positions N448, N262 and N295, and make minimal protein-contact. The prototypical bnAb VRC-PG05 was not particularly broad, however SF12 has since shown that a more promising neutralisation breadth can be achieved via this epitope [224].

The highly conserved MPER region is located on the gp41 subunit and was one of the first neutralising epitopes discovered [225]. The bnAbs of this specificity (2F5, 4E10 and 10E8) are some of the broadest neutralisers around, however they have a weaker potency compared to other classes [200]. MPER bnAbs possess hydrophobic surfaces, long CDRH3 loops and lipid binding qualities which enable them to access their epitope adjacent to the viral membrane [226]. The MPER domain forms a secondary alpha-helix arrangement that is not exposed on Envs in the pre-fusion conformation. Following binding to the CD4 receptor, the Env trimer reveals the MPER epitope and this results in a window of opportunity for bnAb binding [227]. It has been shown that after successful interaction, bnAbs against the MPER region can induce shedding of the gp120 subunit and this could be the mechanism of neutralisation utilised by this class [228].

Examination of the co-evolution of MPER bnAbs led to the identification of three lineages from a single superinfected donor (RV217-40512). The bnAbs VRC42.01, VRC43.01, and VRC46.01 displayed modest levels of SHM (~10%), typical CDR loop lengths and employed different modes of recognition. B-cells presenting the VRC42 UCA could be activated by immunogens bearing the founder virus MPER domain. Further, intermediates of VRC42 could achieve 50% breadth with only 3% SHM, showing that small levels of mutation could result in moderately broad antibodies. Autoreactivity was observed in each of these bnAb lineages, although they were not associated with an autoimmune disorder in this patient. In addition, the broadest lineage (VRC42) had a similar germline usage and mode of binding to 4E10, but showed lower levels of autoreactivity [229].

Understanding how bnAbs and autologous Envs develop in response to one another in HIV-1 infected patients is of great interest to the scientific community. It has been proposed that such studies could provide a vaccination roadmap to the maturation of bnAbs from their associated germlines. However, it is important to note that co-evolution studies are impeded by numerous factors. There is only a small number of cohorts that include untreated HIV-1 infected individuals, and arranging new untreated cohorts would be unethical. Of such cohorts, there are only a limited number of donors from whom adequate longitudinal samples have been taken, allowing a complete co-evolution study to be performed. Out of the subjects that have been suitably sampled, only a fraction will develop neutralisation breadth [208].

### 1.3.11 HIV-1 treatment and prevention

Prior to 1996, there were few options for the treatment of HIV-1 available. Infected individuals were managed clinically with prophylaxis against prevalent opportunistic pathogens, and treated for AIDS related morbidities. As time progressed, novel therapeutics capable of disrupting distinct parts of the HIV-1 life cycle were established and approved for patients (Table 1.5) [230].

There are two types of inhibitors of the reverse transcription process, nucleoside reverse transcriptase inhibitors (NRTI) and non-nucleoside reverse transcriptase inhibitors (NNRTI). NRTIs are analogues of deoxynucleotides that are naturally found in the body, which are preferentially incorporated into the HIV-1 genome during reverse transcription. NRTIs lack a 3'-hydroxyl group, which prevents a 3'-5'-phosphodiester backbone from forming and the subsequent addition of incoming 5'-nucleoside triphosphates. NNRTIs also target the reverse transcription stage of the HIV-1 life cycle, however they interact directly with RT near its binding pocket and inhibit its ability to undergo the transformations needed for the synthesis of new DNA chains. Therapeutics have been developed to prevent the integration of the viral genome into the host cell. These drugs comprise two essential elements, a metal-binding pharmacophore that seizes magnesium ions found at the IN active site and a hydrophobic group that binds to the viral DNA/IN complex [230]. The enzyme PR is required to cleave HIV-1 polyproteins and therefore enables maturation into infectious particles. PR inhibitors compete for the active site of this protein preventing its functionality, and are often co-administered with booster drugs due to their rapid metabolised by the liver [231]. Another type of ART drug is the entry-inhibitors, which can be separated into fusion, CCR5 and post-attachment inhibitors. Enfuvirtide is the only FDA approved fusion inhibitor, it interacts with gp41 and interferes with viral-host membrane fusion [232]. CCR5-inhibtors are antagonists that specifically bind and block this receptor from enabling entry of viruses with the appropriate tropism [233]. Post-attachment inhibitors bind to CD4 but do not hinder its interaction with gp120, instead they decrease CD4 flexibility and prevent subsequent binding to chemokine co-receptors [234]. Studies have shown that bnAbs can be used to control viral load, and in some cases delay viral rebound in infected subjects that have not received or are undergoing an interruption of their ART [235,236]. The viral escape that is usually observed in natural infection can be counteracted by the use of multiple non-competing bnAbs that target distinct non-overlapping epitopes [237].

Due to its rapid mutation rate, HIV-1 is capable of escaping all types of ART drug and the escape mutants generated are transmissible. Hence, ART was developed in which a cocktail of at least three different ART drugs, against at least two HIV-1 life cycle stages were administered. ART can reduce viral load and increase $CD4^+$ T-cell counts, which leads to a lower risk of morbidity and mortality. The reduction in viral load results in a lower chance of subsequent transmission. ART can be used as a post-exposure prophylaxis within 72 hours of exposure, and can also be employed as a pre-exposure prophylaxis in high-risk groups [238].

The first person to be cured of HIV-1, often referred to as the Berlin patient, received two allogeneic hematopoietic stem-cell transplants as treatment for their myeloid leukemia. The stem-cell donor had a natural homozygous mutation in their CCR5 gene (CCR5Δ32), which rendered them as well as the recipient of the transplant resistant to infection by CCR5 tropic HIV-1 [239]. The Berlin patient stopped ART and their viral load has remained undetectable for over 10 years [240]. Likewise, the London patient underwent an allogeneic hematopoietic stem-cell transplantation from a donor carrying the homozygous CCR5Δ32-mutation as treatment for their Hodgkin's lymphoma. The London patient has maintained an undetectable viral load following cessation of ART for over 1.5 years [241].

The discovery of HIV-1 bnAbs and their target epitopes was a tremendous step forward in AIDS research, however vaccination with HIV-1 immunogens in animal studies has still only shown limited potential. Promising candidates identified through this process have been tested in humans, but a vaccination schedule that can elicit a broad response is yet to be achieved. The Thai phase III clinical trial RV144, with over 16,000 participants, was the largest HIV-1 vaccination study ever performed.

*Table 1.5: List of HIV-1 ARTs.*
*Shown here are examples of HIV-1 ARTs, with information regarding the drug class, manufacturer and approval date. Adapted from [142].*

| Brand name | Generic name | Class | Manufacturer | Approval date |
|---|---|---|---|---|
| **Emtriva** | Emtricitabine, FTC | NRTI | Gilead Sciences | 07/02/03 |
| **Epivir** | Lamivudine, 3TC | NRTI | GlaxoSmithKline | 11/17/95 |
| **Hivid** | Zalcitabine, ddC | NRTI | Hoffmann-LaRoche | 06/19/92 |
| **Retrovir** | Zidovudine, AZT | NRTI | GlaxoSmithKline | 03/19/87 |
| **Videx** | Didanosine, ddI | NRTI | Bristol-Myers Squibb | 10/09/91 |
| **Viread** | Tenofovir, TDF | NRTI | Gilead Sciences | 10/26/01 |
| **Zerit** | Stavudine, d4T | NRTI | Bristol-Myers Squibb | 06/24/94 |
| **Ziagen** | Abacavir, ABC | NRTI | GlaxoSmithKline | 12/17/98 |
| **Intelence** | Etravirine, ETV | NNRTI | Tibotec Therapeutics | 01/18/08 |
| **Rescriptor** | Delavirdine, DLV | NNRTI | Pfizer | 04/04/97 |
| **Sustiva** | Efavirenz, EVF | NNRTI | Bristol-Myers Squibb | 09/17/98 |
| **Viramune** | Nevirapine, NVP | NNRTI | Boehringer Ingelheim | 06/21/96 |
| **Agenerase** | Amprenavir, APV | Protease inhibitor | GlaxoSmithKline | 04/15/99 |
| **Aptivus** | Tipranavir, TPV | Protease inhibitor | Boehringer Ingelheim | 06/22/05 |
| **Crixivan** | Indinavir, IDV | Protease inhibitor | Merck | 03/13/96 |
| **Invirase** | Saquinavir, SQV | Protease inhibitor | Hoffmann-LaRoche | 12/06/95 |
| **Lexiva** | Fosamprenavir, APV | Protease inhibitor | GlaxoSmithKline | 10/20/03 |
| **Norvir** | Ritonavir, RTV | Protease inhibitor | Abbott Laboratories | 03/01/96 |
| **Prezista** | Darunavir | Protease inhibitor | Tibotec, Inc. | 06/23/06 |
| **Reyataz** | Atazanavir, ATV | Protease inhibitor | Bristol-Myers Squibb | 06/20/03 |
| **Viracept** | Nelfinavir, NFV | Protease inhibitor | Agouron | 03/14/97 |
| **Fuzeon** | Enfuvirtide, T-20 | Fusion inhibitor | Hoffman-LaRoche and Trimeris | 03/13/03 |
| **Selzentry** | Maraviroc, MVC | CCR5 inhibitor | Pfizer | 08/06/07 |
| **Trogarzo** | Ibalizumab, IBA | Post-attachment inhibitors | TaiMed Biologics | 26/09/2019 |

The vaccine regimen entailed an ALVAC-HIV prime at 0 and 1 months, followed by co-administration of ALVAC-HIV and AIDSVAX B/E as a boost at 3 and 6 months. ALVAC-HIV is a canarypox vector vaccine that results in the expression of subtype B Gag, protease and gp41, as well as CRF01_AE gp120. AIDSVAX B/E is a bivalent recombinant subunit vaccine comprised of two different CRF01_AE gp120s. The trial yielded a modest vaccine efficacy of ~30% at 42 months, but failed to generate bnAb responses [242]. However, non-neutralising IgG responses that targeted the V2-loop were associated with protection. Conversely, elevated levels of IgA in the plasma of subjects were linked to reduced vaccine efficacy. It has been suggested that the non-neutralising IgG antibodies were recognising Env expressed on the surface of infected $CD4^+$ T-cells and facilitating ADCC, a process that the IgA responses could have been disrupting [243]. The RV144 trial was replicated in South Africa (HVTN100/HVTN 702 trial) with 5,407 participants and an alternative bivalent subtype C boost, but was halted early due to the observed lack of efficacy [244].

### 1.3.12 Advances in HIV-1 immunogen development

Efforts to generate bnAb responses via vaccination have so far been unsuccessful, and this was likely due to flaws associated with the selected immunogens. The Env subunit antigens explored in prior clinical trials, such as monomeric gp120 or gp140, are subject to a number of limitations. For instance, monomeric immunogens display non-neutralising immunodominant regions, lack quaternary neutralising epitopes and have different glycosylation profiles compared to the trimeric Env found in nature. Hence, a considerable amount of research has been devoted to improving Env immunogens and significant advancements have been made in this field [245].

The generation of the soluble, native like BG505 SOSIP.664 trimer marked a large step forward in HIV-1 antigen development. The structure was based on the BG505 isolate, a clade A founder virus bearing a T332N mutation that restores the V3-glycan domain bnAb epitope [246]. To ensure solubility, the TMD and CT were eliminated by truncation of the gp41 subunit at amino acid position 664. A covalent bond was introduced between the gp120 and gp41 subunits (A501C and T605C, commonly known as SOS mutation), to prevent dissociation post-cleavage. A point mutation was made at residue I559P (generally referred to as IP), to stabilise the interfaces between the gp41 subunits. To maximise cleavage, the furin cleavage site was enhanced (REKR to RRRRRR) and the BG505 SOSIP.664 glycoproteins were co-expressed with furin. Collectively, these modifications yielded soluble truncated trimers of gp120-gp41 heterodimers, which were fully cleaved but remained intact, and presented their neutralising epitopes correctly [247]. To eliminate the requirement for furin over-expression, native flexibly linked (NFL) trimers altered the SOSIP design by replacing the SOS mutations with a flexible peptide linker [248]. The uncleaved pre-fusion-optimised (UFO) trimers built upon this by substituting the IP mutation with a redesigned HR1 in gp41 [249]. The BG505 SOSIP.664, and the later developed B41 SOSIP.664, were the first Env immunogens to elicit consistent autologous tier 2 and heterologous tier 1 neutralising responses in animal models [250]. It has been shown that monoclonal antibodies (mAbs) capable of potent autologous tier 2 neutralisation, isolated from rabbits immunised with the BG505 SOSIP.664, were targeting an immunodominant region on gp120. This region was exposed due to the presence of a hole in the glycan shield, which was only found in <3% of global HIV-1 isolates. This emphasised that glycan holes are capable of diverting antibody responses away from more desirable epitopes, and should be considered in future studies [251].

These soluble native like trimer platforms have been applied to many Env sequences in the hopes to induce heterologous tier 2 neutralising responses. Numerous consensus structures have been developed that encompass the most common amino acid for each residue from a variety of HIV-1 variant subsets [252]. Mosaic structures, Env sequences that include as many B-cell and T-cell epitopes

as possible, have also been produced [253]. Structures have been made that follow the natural evolution of the virus, attempting to apply the knowledge gained from co-evolution studies [254].

The artificial truncation of the soluble native like trimers exposes a highly immunogenic base, limiting the effectiveness of these glycoproteins [255]. To address this, Env trimer nanoparticle systems have been developed such as virus-like particles, liposomes and ferritin-based structures [256]. Aside from restricting access to the trimer base, nanoparticles also multimerise the immunogen. Multimerisation improves B-cell activation, avidity, GC generation and helps B-cells overcome tolerance mechanisms [257,258].

### 1.3.13 The SPARTAC trial

The short pulse antiretroviral therapy at seroconversion (SPARTAC) trial was performed between 2003 and 2007 in locations including Australia, Brazil, Ireland, Italy, South Africa, Spain, Uganda and the UK. The study cohort consisted of 366 patients with primary HIV-1 infection, which were assigned to one of three groups. The groups were either given ART for 12 weeks, 48 weeks or not at all, with no ART being the standard of care at the time of the investigation. The trial aimed to determine whether administering ART for a short period during primary infection could delay the time it took for patients to reach a $CD4^+$ count below a particular threshold or require ART. It was concluded that a 48-week course of ART delayed disease progression but this was not significantly longer than the length of the treatment [259].

### 1.3.14 Animal models in HIV-1 research

Animal models are crucial for bridging the gap between in vitro and in vivo HIV-1 research. There is a tight host range in immunodeficiency viruses from humans and non-human primates, making it necessary to alter either the host or virus to generate useful animal models [260].

Humanised mice models were developed using immune deficient mice (severe combined immune deficient) engrafted with human PBMCs, HSCs/human bone marrow, liver and thymus (BLT mice), making them susceptible to HIV-1 infection [261]. Despite recapitulating some of the features of HIV-1 establishment and disease pathogenesis, humanised mice are mainly utilised in protections studies. This is because they have relatively short lifespans and do not capture all of the aspects of adaptive immunity or disease progression that are seen in humans [260]. Humanised mice have been used to test the efficacy of passive transfer with bnAbs for preventing HIV-1 infection [262,263]. Further, this animal model has demonstrated the effectiveness of vectored immunoprophylaxis with bnAbs for preventing the intravenous and mucosal transmission of HIV-1 [264].

Experimental inoculation of a number of Asian macaque species (including rhesus, pig-tailed and cynomolgus monkeys), was shown to yield a variety of pathological responses similar to AIDS in humans [265]. However, the SIV Env being used was antigenically distinct to that of HIV-1 and therefore not recognised by bnAbs, limiting the use of this model for antibody-based studies [260]. To rectify this problem, chimeric SIV/HIVs were developed via incorporating the HIV-1 Env into a SIV (typically $SIV_{mac239}$) backbone [266]. This generated a chimeric virus that retained most of the SIV proteins, but was able to display a HIV-1 Env of interest. Addition mutations were sub sequentially added to enhance the binding of the Env to the macaque CD4 receptor [267]. Non-human primate models using SIV/HIV infection have proved indispensable for studying HIV-1 pathogenesis, vaccines, and therapies [268].

### 1.3.15 Parallels between HIV-1 and other chronic viral infections

HIV-1 can be categorised as a chronic viral infection that undergoes continuous productive replication. Other examples of viruses from this category in humans are hepatitis B virus (HBV) and hepatitis C virus (HCV) [269].

HBV and HCV are members of the *Hepadnaviridae* and *Flaviviridae* viral families respectively. There are more than 240 million and 70 million chronically infected individuals globally with HBV and HCV respectively [270]. Both viruses persist in infected liver hepatocytes, eventually leading to the development of liver fibrosis, cirrhosis and hepatocellular carcinoma [271]. Both HBV and HCV have lower mutation rates (1.4-3.2×$10^{-5}$ and 3.5×$10^{-5}$ base substitutions/site/year respectively) compared to HIV-1 (1.3×10-3 base substitutions/site/year) [272–274]. Whilst antibody breadth and viral escape mutations are still a concern for chronic hepatitis research, there is less antigenic shift to overcome compared to HIV-1. However, HBV is known to evade host humoral responses by producing a large excess of decoy non-infectious sub-viral particles [275]. Nonetheless, antibody responses to the surface glycoprotein of this virus have been shown to be associated with recovery from acute infection with HBV, with the potential for protection from re-infection in acute cases [276,277]. Further, inability to generate such antibodies has been linked to chronicity in HBV [278]. Conversely, the development of bnAbs in HCV are typically not protective or able to control infection [279]. Despite there being some parallels between HIV-1 and hepatitis research, antibody responses to these chronic viral infections are more complex and would each require distinct vaccination and treatment plans.

## 1.4 SARS-CoV-2

### 1.4.1 The COVID-19 pandemic and the origins of SARS-CoV-2

Coronaviruses (*Coronaviridae*), named due to their crown-like appearance when examined via electron microscopy, are a large and diverse family of RNA viruses. The subfamily *Coronavirinae* encompasses four genera including *Alphacoronavirus* (Alpha-CoV), *Betacoronavirus* (Beta-CoV), *Deltacoronavirus* (Delta-CoV) and *Gammacoronavirus* (Gamma-CoV). Coronaviruses can infect avian species (Delta-CoVs and Gamma-CoVs) and numerous mammalian species (mainly Alpha-CoVs and Beta-CoVs). To date, seven coronaviruses that are able to infect humans have been identified. They can be separated into two categories, namely the endemic human coronaviruses and the highly pathogenic coronaviruses [280]. The endemic human coronaviruses consist of human coronavirus-OC43 (HCoV-OC43), human coronavirus-HKU1 (HCoV-HKU1), human coronavirus-229E (HCoV-229E), and human coronavirus NL63 (HCoV-NL63) [281–284]. Typically, infection with endemic human coronaviruses will only result in a mild respiratory tract illness. However, they can induce a more severe course of disease in infants and the immunosuppressed [285,286]. Some members of the endemic human coronaviruses have been circulating since the 1960s, hence their geographical origins are difficult to elucidate [280]. The first highly pathogenic coronavirus outbreak, caused by severe acute respiratory syndrome coronavirus (SARS-CoV), originated in the Guangdong province of China and lasted from late 2002 to 2004 [287]. Middle East respiratory syndrome coronavirus (MERS-CoV) was the causative agent of the second highly pathogenic coronavirus outbreak, the virus was first identified from a patient with acute pneumonia from Saudi Arabia in 2012 [288]. MERS-CoV is still circulating today, but the majority of its cases are localised within the Arabian Peninsula with only sporadic cases occurring elsewhere in the world [289]. The third and most recent highly pathogenic coronavirus outbreak began in late 2019, and originated from the city of Wuhan in China [290]. The last outbreak was caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the associated illness became known as coronavirus disease 2019 (COVID-19) [291]. Since its emergence, SARS-CoV-2 has surpassed SARS-CoV and MERS-CoV in regards to the number of infections and spatial range of endemic regions, and has rapidly escalated into a global pandemic [292]. The ongoing COVID-19 situation represents a huge threat to public health and has resulted in considerable economic ramifications [293].

COVID-19 cases can be categorised as asymptomatic, mild, moderate, severe and critical (Table 1.6). Asymptomatic individuals do not show any clinical symptoms and chest imaging findings, but can transmit the virus to others. Mild cases present with minor clinical symptoms such as fever, fatigue, cough, anorexia, malaise, myalgia, pharyngitis, dyspnea, nasal congestion and cephalgia, with no chest imaging findings. Moderate cases are defined by respiratory tract symptoms and fever, with chest imaging showing mild pneumonia. Severe cases are determined when patients have either respiratory distress (≥30 breaths/min), low resting pulse oxygen saturation (≤93%), hypoxemia (Horowitz index ≤300mmHg) or chest imaging showing significant lesion progression (>50%) within 1-2 days. Cases are considered to be critical when patients have either respiratory failure (ventilation required), septic shock or multiple organ failure [294,295]. SARS-CoV-2 associated hospitalisation rates are higher in the elderly (aged over 65) and individuals that have pre-existing conditions such as hypertension, diabetes and obesity [296]. COVID-19 patients with underlying health conditions have been shown to be six times more likely to require hospitalisation and 12 times more likely to succumb to their disease [297]. Further, males that have contracted COVID-19 have a higher risk of a worse prognosis compared to females [298,299]. A meta-analysis of studies from the UK and US indicated that people in Black, Hispanic and Asian ethnic minority groups are at a higher risk of COVID-19 infection, with Asians having an elevated risk of severe disease and death [300]. Prior to global vaccine rollouts, the case fatality rate (CFR) of SARS-CoV-2 (3.6–3.8%) was considerably lower than SARS-CoV (14–15%) and MERS-CoV (34.4%) [296,301,302]. However, SARS-CoV-2 is highly infectious and a large percentage of COVID-19 patients are

asymptomatic or present with mild symptoms [303]. This has enabled SARS-CoV-2 to effectively spread worldwide and cause a higher number of total fatalities than SARS-CoV and MERS-CoV combined. As of January 2022, there have been ~300 million SARS-CoV-2 infections and 5.5 million associated deaths reported to the World Health Organisation (WHO). Although, the lack of access to diagnostic testing in less developed communities and incomplete reporting of COVID-19 data means that these figures are likely an underestimate [304].

In most countries, pre-cautionary measures against the spread of SARS-CoV-2 were invoked such as social distancing, facemasks, travel restrictions, workplace closures and bans on public gatherings [305]. However, these restrictions were having a considerable impact on the livelihood and mental health of the global population, and the need for a more long-term mitigation was becoming apparent [306]. By March 2020, there were five COVID-19 vaccines in phase IV clinical trials as well as over 200 pre-clinical and clinical candidates [307]. These vaccines were developed and approved in record time, due to the coordination of government agencies with the scientific and medical community [308]. Consequently, regions such as Europe, the Middle East, and North America were able to start mass vaccination against COVID-19 by December 2020 [309]. Interpretation of COVID-19 vaccine effectiveness is complicated, this is because different vaccine candidates and prime-boost intervals have been implemented by various countries [310]. Further, since its emergence SARS-CoV-2 has undergone mutations and several circulating variants have been generated that differ to the ancestral strain presented via vaccination [311]. Additional factors such as prior infection status, time since vaccination and disease treatment can also obscure analyses of such data. Nonetheless, a meta-analysis of vaccination studies suggests that COVID-19 vaccines are safe and can effectively reduce infections, symptomatic cases, severe cases and death as a result of SARS-CoV-2 infection [312]. Management options for COVID-19 patients was initially limited, and relied on the use of experimental therapies and drug repurposing strategies. Over time treatment of this illness has improved and novel therapeutics have also been developed [313,314]. However, access to COVID therapeutics and vaccines is not universal with lower-income countries having a more limited supply [315].

*Table 1.6: COVID-19 disease severity classifications.*
*Shown here are the clinical characteristics used to classify disease severity in COVID-19 patients. Adapted from [294,295].*

| Type | Clinical characteristics | RT-PCR test for COVID-19 |
|---|---|---|
| **Asymptomatic** | • No clinical symptoms<br>• No abnormal chest imaging findings | Positive |
| **Mild** | • Mild clinical symptoms<br>• No abnormal chest imaging findings | Positive |
| **Moderate** | • Respiratory tract symptoms<br>• Fever<br>• Chest imaging showed mild pneumonia | Positive |
| **Moderate** | • Respiratory distress (≥30 breaths/min)<br>• Low resting pulse oxygen saturation (≤93%)<br>• Hypoxemia $Pa0_2/Fi0_2 \leq 300$ mmHg (1 mmHg = 0.133 kPa)<br>• Chest imaging showing significant lesion progression (>50%) within 1-2 days | Positive |
| **Critical** | • Respiratory failure (ventilation required)<br>• Septic shock<br>• Multiple organ failure | Positive |

The successful zoonotic transmission of disease from animals to humans is rare. However, coronaviruses have a wide distribution in animals, high genetic diversity and frequently recombine their genomes, increasing the likelihood of such an event. Further, infection and adaption within an intermediate host usually precedes and assists this process [291]. The endemic human coronaviruses HCoV-229E and HCoV-NL63 are believed to have originated from bats, whereas HCoV-OC43 and HCoV-HKU1 are more likely to have speciated from viruses found in rodents (Figure 1.27). The intermediate host of HCoV-OC43 is thought to be domestic animals such as cattle or swine, and it is assumed that HCoV-229E arose via transfer from camelids such as alpacas or camels. The zoonotic sources of HCoV-NL63 and HCoV-HKU1 are less fully understood [280,316]. Similarly to HCoV-229E and HCoV-NL63, the highly pathogenic coronaviruses are all considered to have originated from bats. Evidence suggests that SARS-CoV emerged through recombination of bat SARS-related coronaviruses (SARSr-CoVs) in the natural reservoir, which infected civets as an intermediate host before being transferred to humans. MERS-CoV likely spilled over from bats to dromedary camels and subsequently to humans, with human and camel strains sharing >99% sequence identity [316]. Early cases of SARS-CoV-2 could be epidemiologically connected to the Huanan seafood wholesale market of Wuhan, which participated in live wild animal trade [292]. Ancestral SARS-CoV-2 has been shown to be genetically related to bat SARSr-CoVs, particularly RaTG13 with which it shared >96% genetic identity [317]. Although bats are expected to be the natural reservoir of SARS-CoV-2, it is unlikely that the virus was directly transmitted to humans. This is supported by evidence such as bat RaTG13 not targeting the same host receptor as SARS-CoV-2, bats not being sold at the Huanan seafood market and the fact that bats normally hibernate throughout December in Wuhan. While snakes and turtles have also been suggested, pangolins appear to be the most likely intermediate hosts of SARS-CoV-2 available at the Huanan seafood market [291]. SARSr-CoVs isolated from pangolins show high sequence identity with SARS-CoV-2 (>90%) especially in the portion of the virus that interacts with host cell receptors [318]. However, considerable evolutionary gaps still exist between SARS-CoV-2 and the closest animal viral relatives. It remains possible that the progenitor virus circulates at a low prevalence or that the correct animal species has not yet been sampled [319].

Mutations occur naturally during viral replication, and RNA viruses tend to have higher mutation rates compared to DNA viruses [320]. However, coronaviruses generate fewer mutations than other RNA viruses because they express an exonuclease enzyme that lowers their replication error rate by ~15-20 fold [311]. Mutations can happen at any location in the coronavirus genome and the fates of mutant variants are determined by natural selection. Mutations that lead to better viral replication, transmission or immune escape will grow in frequency, whilst variants that display lower viral fitness are usually removed from circulation [320]. Although, in theory the majority of mutations are silent or mildly deleterious [321]. SARS-CoV-2 is highly transmissible and has a rapid replication rate which has resulted in evolution from the ancestral strain at a rate of ~0.0004-0.002 mutations per nucleotide per year [311]. Notably, prolonged SARS-CoV-2 infections in immunosuppressed patients can lead to the rapid generation of multi-mutational variants [322]. Moreover, recombination of multiple SARS-CoV-2 variants in a single individual can also lead to saltational evolution [323]. SARS-CoV-2 strains that have been shown to spread widely and be more contagious, induce more severe disease or escape immune responses have been categorised as variants of concern (VOC) by numerous public health organisations. Likewise, SARS-CoV-2 strains that bear similar mutations to VOCs but have not spread as widely are referred to as variants of interest (VOI) [311]. To date, there are five SARS-CoV-2 VOCs including B.1.1.7 (Alpha, α or UK variant), B.1.351 (Beta, β or South Africa variant), P.1 (Gamma, γ or Brazil variant), B.1.617.2 (Delta, δ or India variant) and B.1.1.529 (Omicron or o). With the exception of P.1 (γ) and B.1.351 (β), each of these VOCs have been the dominant globally circulating strain for a period of time (Figure 1.28) [324].

## 1.4.2 SARS-CoV-2 structure and genome

SARS-CoV-2 is an enveloped positive sense single-stranded RNA Beta-CoV, with a genome ~30kb in length (Figure 1.29a) [325]. The SARS-CoV-2 genome has a 5' cap and a 3' poly-A tail, and encodes 14 genes that can be translated to produce 29 different proteins. These include the non-structural genes ORF1a and ORF1b, the structural genes spike (S), envelope (E), membrane (M) and nucleocapsid (N), and the accessory genes ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8, ORF9b, ORF9c and ORF10 [326]. The genes ORF1a and ORF1b are translated as polyproteins pp1a and pp1ab, with pp1a encoding nsp1–

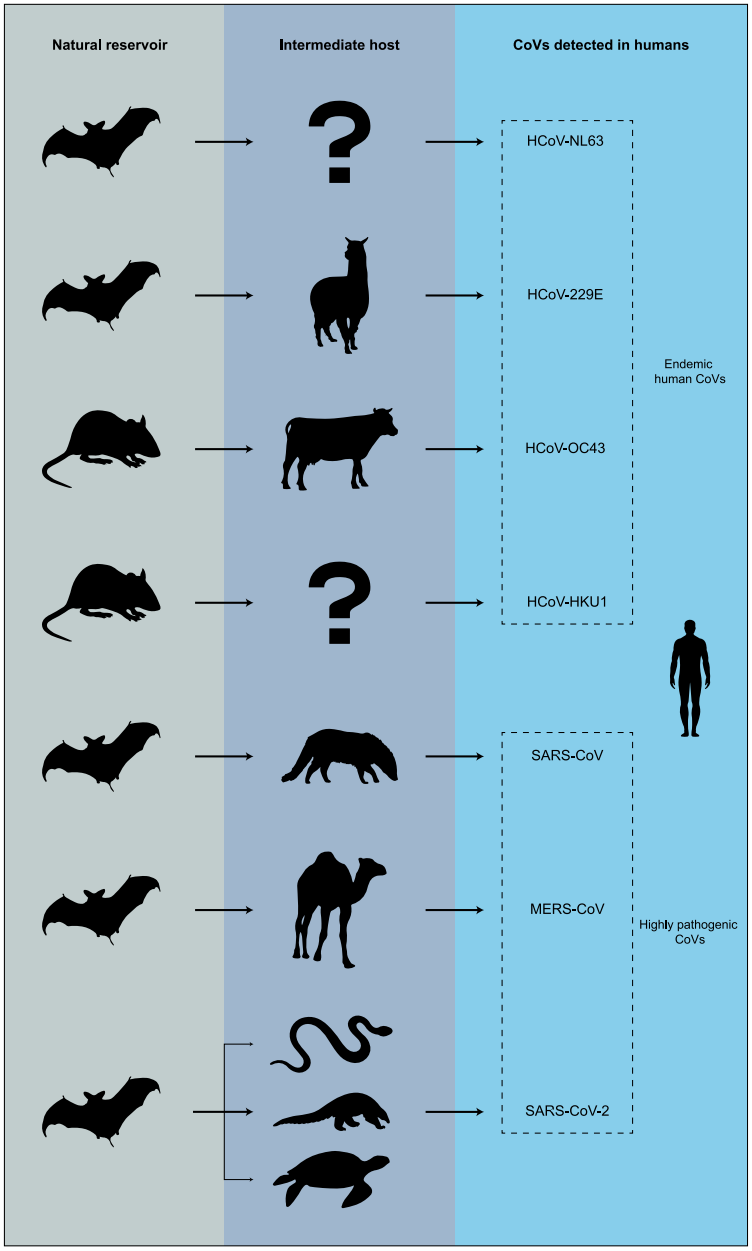



*Figure 1.27: Animal origins of human coronaviruses.*
*There are currently seven coronaviruses able to infect humans. They can be separated into two categories, the endemic human coronaviruses and the highly pathogenic coronaviruses. These viruses emerged as a result of cross-species zoonotic transfers, and typically required prior adaption within an intermediate host. The endemic human coronaviruses HCoV-229E and HCoV-NL63 originated from bats, whereas HCoV-OC43 and HCoV-HKU1 speciated from rodents. The intermediate hosts of HCoV-OC43 were domestic animals such as cattle or swine, and HCoV-229E arose via transfer from camelids such as alpacas or camels. The highly pathogenic coronaviruses are all considered to have originated from bats. The intermediate hosts of SARS-CoV were civets, whilst MERS-CoV spilled over from dromedary camels. The intermediate host of SARS-CoV-2 has not yet been fully ascertained, although snakes, turtles and pangolins have been suggested. Adapted from [291,316].*

nsp11 and pp1ab encoding nsp1–nsp16 [325]. To generate all 29 proteins the virus uses leaky scanning, frameshifts, discontinuous viral transcription and post-translational cleavage of polyproteins. The viral genome is flanked by stretches of UTRs, which have numerous cis-acting secondary RNA structural elements that are important for viral transcription and replication [327].

SARS-CoV-2 virions are ~100nm in diameter, spherical in appearance and have a dense viroplasm (Figure 1.29b). The virus is coated by a host derived lipid envelope, in which the viral proteins S, E and M are embedded. The S glycoprotein enables entry into host cells via interacting with the target receptor angiotensin-converting enzyme 2 (ACE2), and subsequently facilitating membrane fusion [328]. There are ~26 S glycoproteins on each particle, the majority of which are in a functional pre-fusion conformation [329]. E is the smallest of the structural proteins, it forms a homopentameric ion channel and is believed to play a role in viral budding. Although it is expressed copiously in infected cells, only a small amount of E proteins are incorporated into the viral membrane. The M protein is found in the highest abundance, it orchestrates viral assembly by forming interactions with all of the other structural proteins [326]. Inside the SARS-CoV-2 particles are numerous ribonucleoprotein (RNP) complexes (~26 per virion) made up of multiple N proteins and a single copy of the RNA genome. The RNPs have a bucket-like appearance and can be packaged into viruses with either a hexagonal or pyramidal morphology [329].

### 1.4.3 SARS-CoV-2 life cycle

SARS-CoV-2 can spread in the body via the generation of cell free particles or through cell-to-cell transmission [330]. The virus infects cells that express ACE2, and the major route of transmission is via the respiratory system. Cells susceptible to infection are found throughout human airways, in regions such as the nasal epithelium, the bronchial epithelia and lower lung. Further, ACE2 is also expressed in the small intestine, testis, kidney, heart muscle, colon and thyroid gland, representing other locations of vulnerability [331].

The S glycoprotein can be found on the surface of SARS-CoV-2 virions as a trimer of S1-S2 subunit heterodimers (Figure 1.30a) [332]. The S1 subunit can be separated into four domains including the

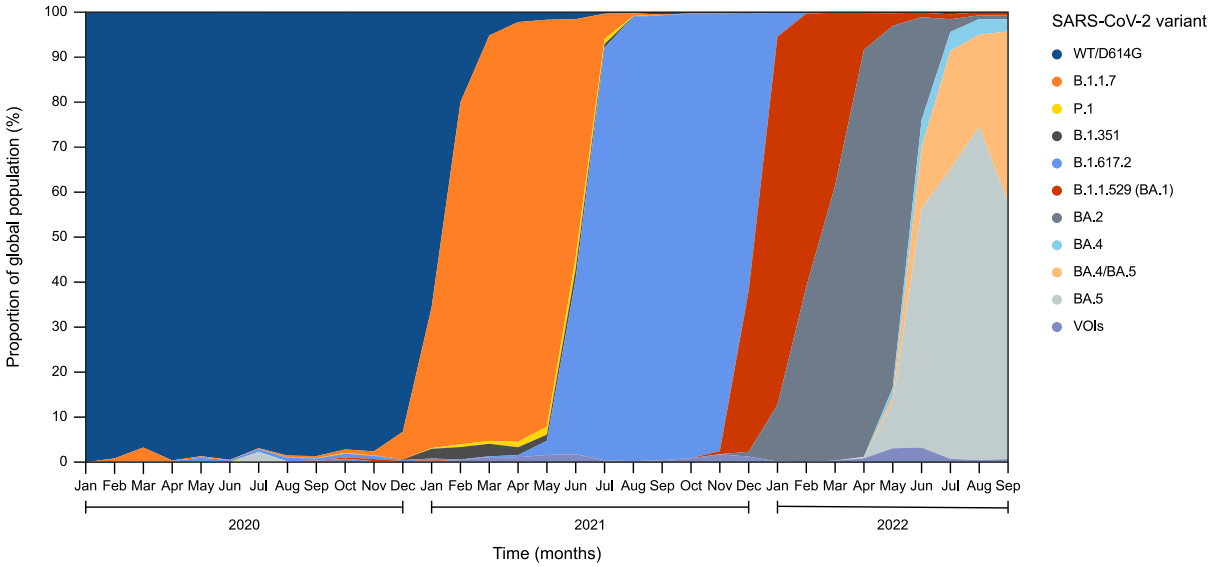



*Figure 1.28: Frequencies of SARS-CoV-2 VOCs.*
*SARS-CoV-2 strains that spread widely, are more contagious, induce more severe disease or escape immune responses have been categorised as VOCs by numerous public health organisations. There are five SARS-CoV-2 VOCs including B.1.1.7 (α), B.1.351 (β), P.1 (γ), B.1.617.2 (δ) and B.1.1.529 (o; BA.1 and sub-lineages named sequentially as they emerged). Similarly, SARS-CoV-2 strains that bear similar mutations to VOCs but have not spread as widely are referred to as VOIs. Shown above is the frequency of these VOCs and VOIs over time in the global population. Adapted from [324].*

amino-terminal domain (NTD), receptor binding domain (RBD) and two subdomains (SD1 and SD2; Figure 1.30b). The S2 subunit can be subdivided into the FP, HR1, central helix, connector domain, HR2, TMD and CT [333]. S requires cleavage by host cell proteases in order to facilitate viral entry, this occurs at the S1-S2 boundary and the S2' site that neighbours the FP. The RBDs are located at the apex of S, they exist in either a receptor-accessible (up) or receptor-inaccessible (down) conformation [331]. Similarly to HIV-1, S monomers are folded, oligomerised and glycosylated by the host cell apparatus as they pass through the secretory pathway. S is heavily glycosylated, with each protomer presenting 22 N-linked glycans [334].

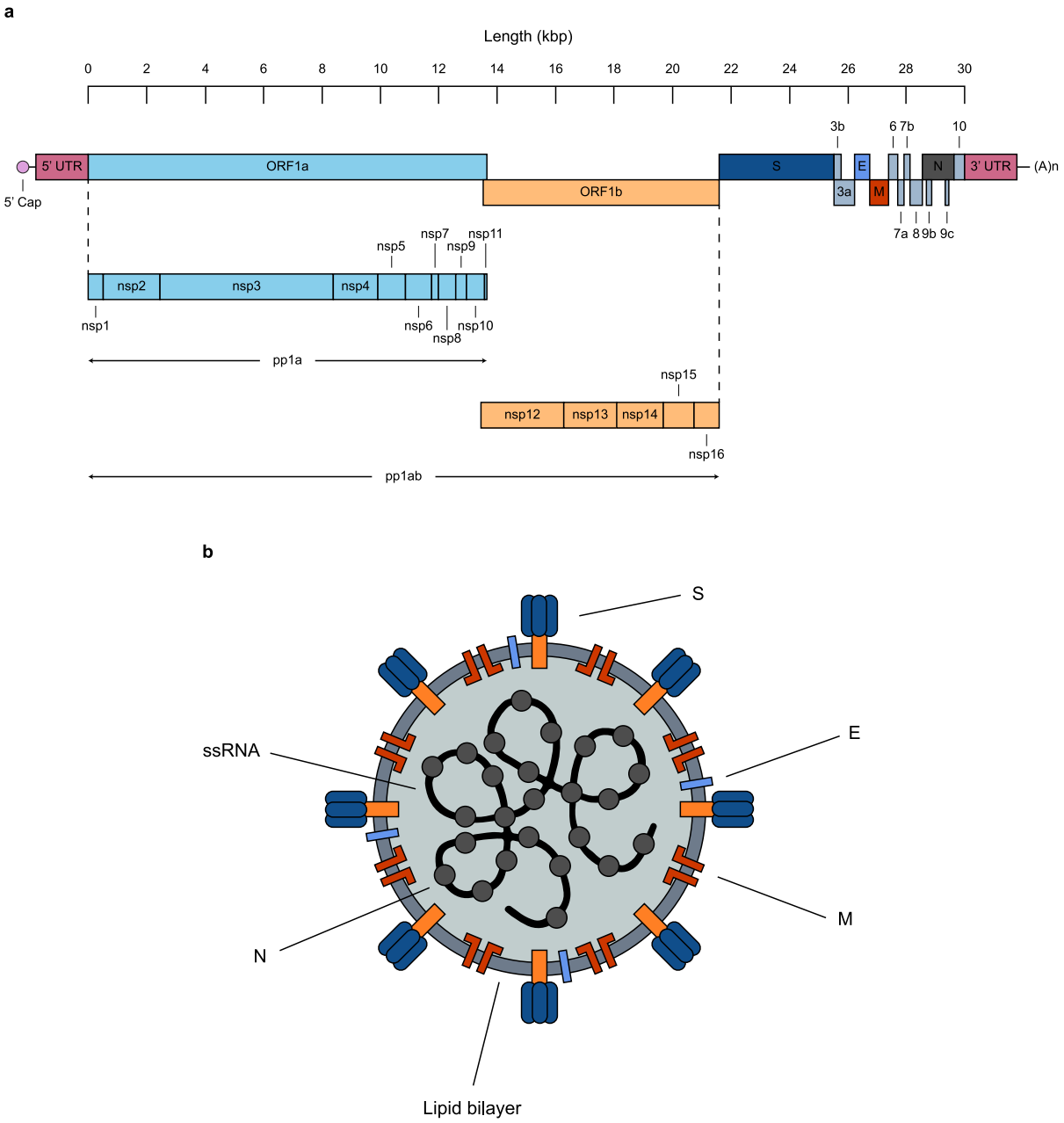



*Figure 1.29: Annotated SARS-CoV-2 genome and virion.*
*SARS-CoV-2 has a positive sense single-stranded RNA genome (~30kb) that encompasses 14 genes, is flanked by UTRs and has a 5' cap and a 3' poly-A tail (**a**). The genes include structural genes (S, E, M and N), non-structural genes (ORF1a and ORF1b) and accessory genes (ORF3a, ORF3b, ORF6, ORF7a, ORF7b, ORF8, ORF9b, ORF9c and ORF10). ORF1a and ORF1b are translated into polyproteins pp1a and pp1ab, with pp1a encoding nsp1–nsp11 and pp1ab encoding nsp1–nsp16. SARS-CoV-2 virions are spherical in shape (**b**) and have an envelope (derived from the host cell lipid bilayer), in which the structural proteins S, E and M are embedded. Inside SARS-CoV-2 particles there are numerous RNP complexes, consisting of multiple N proteins and a single copy of the RNA genome. Adapted from [326,331].*

Prior to initiation of the entry process, S must be cleaved at the S1-S2 boundary by host cell furin during production in an infected cell [335]. This yields mature trimers composed of S1-S2 monomers that are non-covalently linked. Entry begins with the binding of a mature S with ACE2 expressed on the

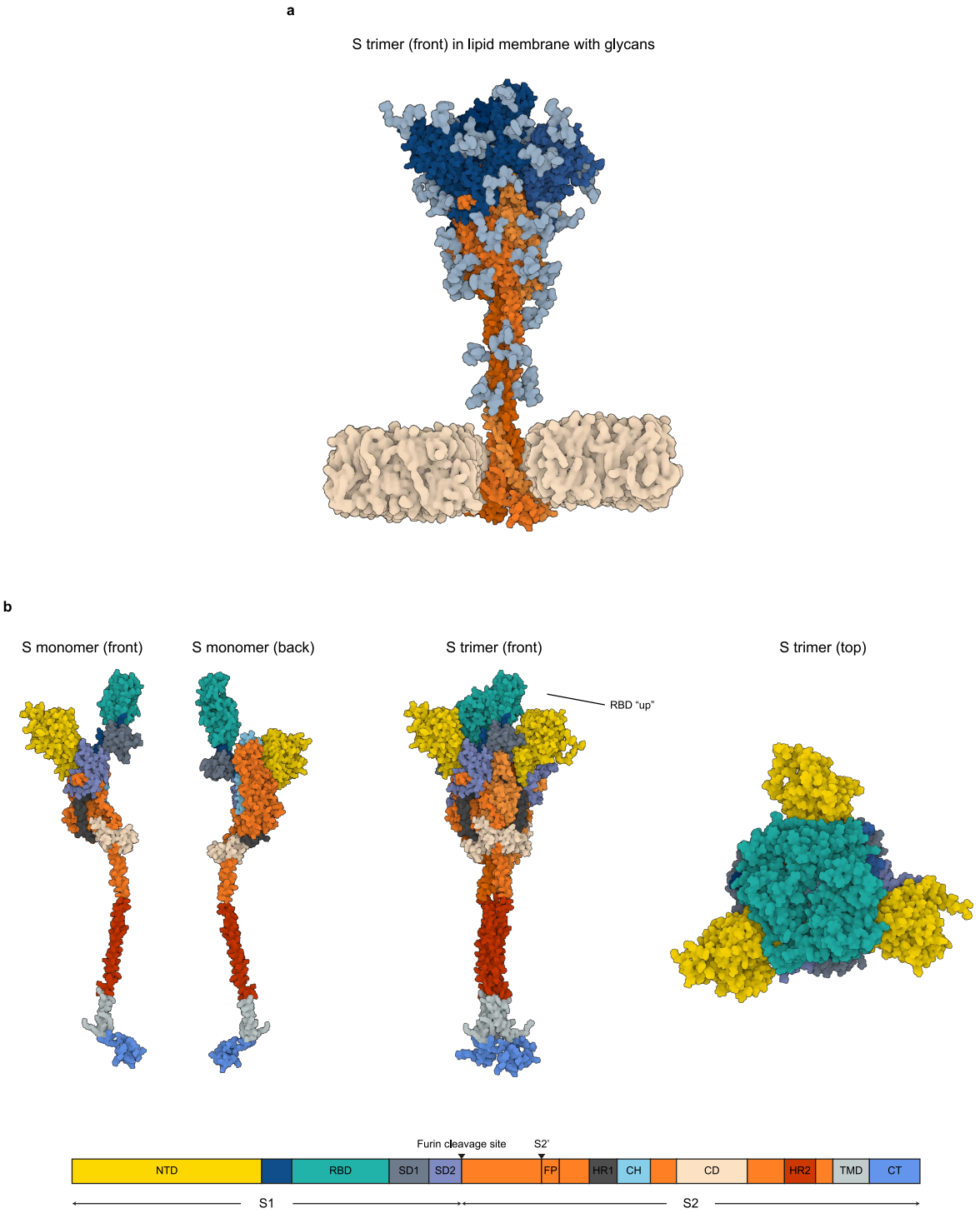



*Figure 1.30: Annotated structure of the SARS-CoV-2 S.*
*The SARS-CoV-2 S glycoprotein is a trimer of non-covalently linked S1 (blue) and S2 (orange) heterodimers expressed on the surface of virions (**a**). The S molecules are heavily glycosylated, with each monomer possessing ~22 PNGS (steel blue). The S1 subunit can be separated into four domains (**b**) including the NTD, RBD and two subdomains (SD1 and SD2). The S2 subunit can be segmented into the FP, HR1, central helix, connector domain, HR2, TMD and CT. S requires cleavage by host cell proteases in order to facilitate viral entry, this occurs at the S1-S2 boundary and the S2' site that neighbours the FP. RBDs located at the apex of S exist in either a receptor-accessible (up) or receptor-inaccessible (down) conformation. Adapted from [555,644]. Full model found downloaded from [644].*

surface of target cells. This causes a sustained conformational change that likely results in the dissociation of the S1 subunit, and exposure of the S2' site [331]. When transmembrane protease serine 2 (TMPRSS2) is present at the target cell membrane, it cleaves the S2' site [336]. However, when TMPRSS2 is not available the virus is internalised into endosomes via clathrin-mediated endocytosis, and the S2' site is cleaved by cathepsins (especially cathepsin L) [337]. In any case, shedding of the S1 subunit and cleavage at S2' unleashes the FP which proceeds to insert into the target cell membrane. The S2 subunit folds back on itself, forming a 6-helix bundle which leads to fusion of the viral and target cell membranes [331,338].

Successful membrane fusion results in the release of the SARS-CoV-2 genome into the host cell cytoplasm (Figure 1.31). Due to its 5' cap and 3' poly-A tail, the genome can act as mRNA and be immediately translated by host cell ribosomes into viral polyproteins [339]. ORF1a and ORF1b are translated to produce pp1a and pp1ab respectively, pp1ab is generated via a -1 programmed frameshift and is found 1.4-2.2 times less frequently than pp1a. The non-structural proteins (nsp) are

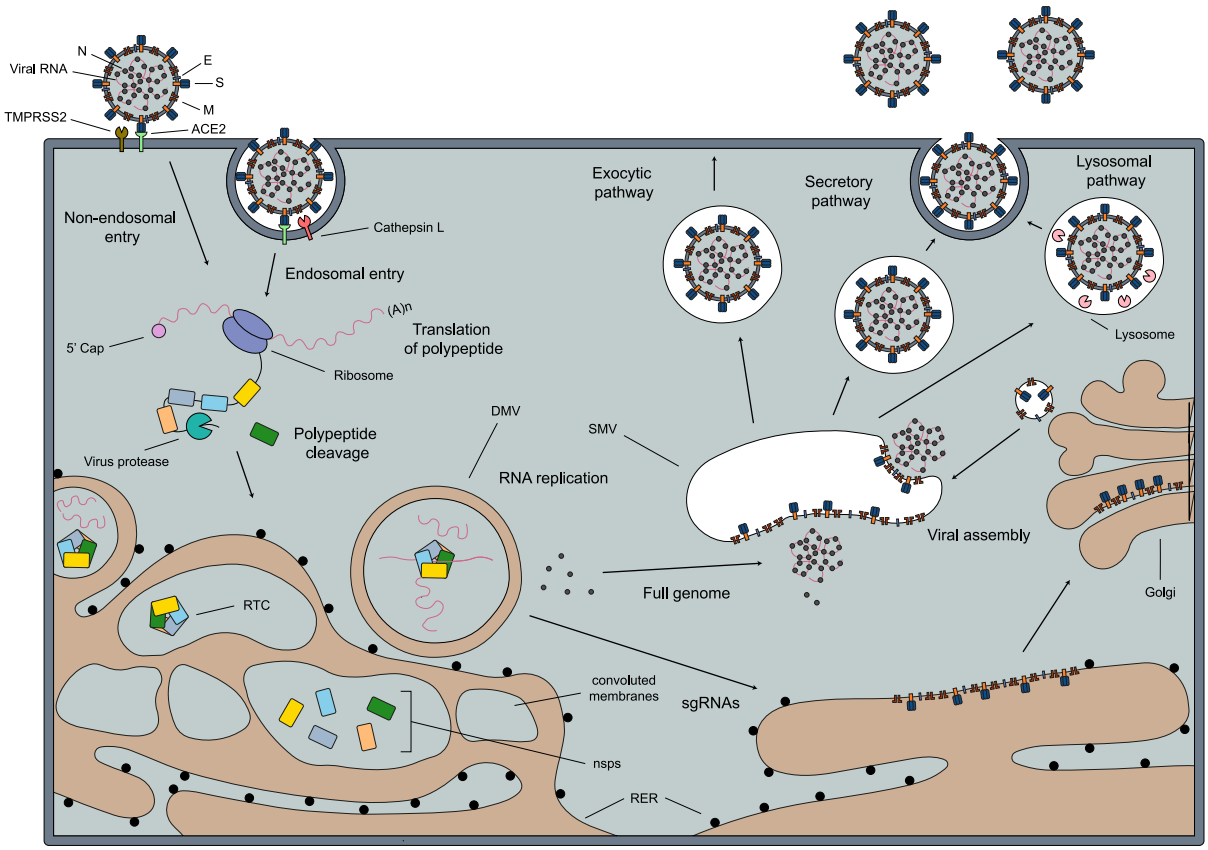



*Figure 1.31: SARS-CoV-2 life cycle.*
*SARS-CoV-2 entry begins when S interacts with ACE2 on the surface of target cells. Sustained conformational changes result in S1 subunit dissociation and exposure of the S2' cleavage site. S2' is digested by either TMPRSS2 at the host cell membrane or cathepsins in endosomes during clathrin-mediated endocytosis, enabling subsequent membrane fusion. In the cytoplasm, the SARS-CoV-2 genome is translated into viral polyproteins (pp1a and pp1ab). The non-structural proteins are released from the polyproteins through co-translational and post-translational proteolytic cleavage by nsp3 and nsp5. The RTC complex is formed and replication commences, producing new viral genomes that can be used as templates or packaged into new particles. A range of nested ssRNAs are generated via discontinuous viral transcription. Replication processes occur in DMV, convoluted membranes and double-membrane spherules, and viral transcripts are exported through molecular pores. The S, E and M structural proteins pass through the endoplasmic reticulum and Golgi apparatus, whilst N localises outside DMVs to package full-length viral genomes following their release. The membrane bound proteins are transported to SMVs, where viral assembly takes place and is orchestrated by M. The viruses bud into the lumen of SMVs and likely advance through the secretory pathway resulting in viral egress, however exit via the exosomal or lysosomal pathways have also been suggested. Adapted from [447].*

liberated from these polyproteins through co-translational and post-translational proteolytic cleavage. This is mediated by two cysteine proteases, namely the nsp3 papain-like protease and the nsp5 chymotrypsin-like protease [327]. Release of nsp1 occurs rapidly and it suppresses host gene expression by associating with the mRNA channels of ribosomes [340]. The rest of the non-structural proteins comprise the viral replication and transcription complex (RTC), and exert their actions at different subcellular locations. Playing a more supportive role, nsp2-11 have been shown to be important for intracellular membrane modulation, host immune evasion and to act as cofactors required for replication. On the other hand, nsp12-16 encompasses the core enzymes required for RNA synthesis, proof-reading and modification [327]. The synthesis of RNA is carried out by the nsp12 RNA-dependent RNA polymerase (RdRp), with nsp7-8 acting as cofactors and nsp13 as a helicase [341]. The coronaviruses capping machinery includes nsp10, nsp13, nsp14 and nsp16, which act as a cofactor, RNA 5'-triphosphatase, N7-methyltransferase and 2'-O-methyltransferase respectively. The protein nsp14 also has 3'–5' exonuclease activity and contributes to RNA synthesis by providing a proof-reading function [327].

Entire negative sense copies are made of the viral genome by the RTC, which in turn can be used as templates for the production of new positive sense genomes. These new genomes can be packaged into particles or employed as templates for the generation of more non-structural proteins and RTCs [327]. Like other coronaviruses, SARS-CoV-2 uses discontinuous viral transcription to make a range of nested 3' and 5' co-terminal subgenomic RNAs (sgRNA) [342]. In this process, the RTC transcription of negative sense RNA is halted when a transcription regulatory sequence (TRS) is encountered. The TRSs are located upstream of the majority of genes in the 3' portion of the viral genome. The interrupted transcription is later re-initiated at the TRS located near the leader sequence (TRS-L) on the positive sense template. This is possible due to the complementary nature of the nascent negative sense TRS and the positive sense TRS-L, resulting in the production of sgRNA with a leader sequence. A series of negative sense sgRNAs that vary in length are made, they act as templates for the generation of positive sense sgRNAs, which can then be translated into structural and accessory proteins. While coronavirus sgRNAs are structurally polycistronic, they are believed to be functionally monocistronic with only the first gene at the 5' end being translated [327].

As the expression levels of the non-structural proteins start to rise, infected cells begin to display viral replication organelles. These structures are derived from the endoplasmic reticulum and include double-membrane vesicles (DMV), convoluted membranes and double-membrane spherules [327]. The DMVs have been shown to be the site where viral genome replication and transcription occurs, potentially shielding this process from the host innate immune response [343]. These replication organelles are embedded with molecular pores, which can be used to export viral transcripts and genomes into the cytoplasm [344]. N proteins localise outside of DMVs and can therefore package full-length viral genomes following their release [345]. S, E and M undergo processing by the host cell machinery, passing through the endoplasmic reticulum and Golgi apparatus. These membranes bound proteins are transported by dense vesicles to single-membrane vesicles (SMV), the location where viral assembly and budding ensues. M can form interactions with S, E, N and itself, and mediates viral assembly when N packaged genomes are supplied [346]. Viruses bud into the lumen of SMVs, and membrane scission is expected to be facilitated by E in an ESCRT independent manner. Vesicles containing new particles likely advance through the secretory pathway and the viruses egress, however exit via the exosomal or lysosomal pathways has also been suggested [345].

### 1.4.4 SARS-CoV-2 pathogenesis

The multiciliated cells of the nasopharynx or trachea, or the sustentacular cells of the nasal olfactory mucosa, represent the first cells targeted by SARS-CoV-2 [347,348]. In the majority of COVID-19 cases, the

virus is cleared from the body by the induction of interferon (type I or III) and adaptive immune responses. However, in some instances the virus reaches the lower respiratory tract, either through direct infection or by gradually spreading through susceptible cells in the tracheobronchial tree [349]. Dissemination into the alveoli could result from the flow of mucus in the branching airways being disrupted, due to an infection induced loss of ciliation in the ciliated cells [350].

Most of the alveolar surface is covered by alveolar type 1 (AT1) cells, which are responsible for facilitating gaseous exchange (Figure 1.32). The alveolar type 2 (AT2) cells secrete pulmonary surfactants that lubricate the lung and reduce alveolar surface tension throughout respiration. The AT1 and AT2 cells form a complete epithelial lining in the adult human alveoli, with AT2 cells being the precursors of AT1 cells [351]. AT2 cells express ACE2 and TMPRSS2, and appear to be the main target of SARS-CoV-2 in the alveoli [352]. The alveolar epithelium can be damaged directly due to AT2 infection or indirectly as a result of local inflammatory responses [349]. The AT2 cells proliferate in an attempt to repair areas of damage, however they adopt a damage-associated transient progenitor phenotype, defined by a failure to fully differentiate into AT1 cells [353]. Stimuli such as hypoxia, cytokines and chemokines activate capillary endothelial cells, which may weaken their cell-cell bonds and trigger apoptosis [349]. SARS-CoV-2 can activate complement, inducing $CD16^+$ T-cells which promote the release of chemokines and microvascular endothelial cell damage [354].

Disruption of both the alveolar epithelium and the capillary endothelium leads to the induction of a leaky state, followed by a dysregulation of coagulation, fibrinolysis and inflammation. The exposed subendothelial extracellular matrix triggers the coagulation cascade, and results in fibrin deposition. Activated platelets are recruited to these sites of exposure, preventing leakage and secreting factors that sustain coagulation [349]. Platelets can stimulate neutrophils, which in cooperation with monocytes release neutrophil extracellular traps that drive coagulation and the formation of fibrin thrombi [355]. High levels of plasminogen activator inhibitor 1 in COVID-19 patients could further promote the generation of microthrombi by reducing fibrinolysis [356]. Hyaline membranes (fibrin-rich exudates) are generated inside the alveoli to stop fluid accumulation, but also limit gaseous exchange. Damage to the alveolar epithelium instigated by the virus leads to infiltration by immune cells, particularly monocytes and macrophages [353]. Once inside the alveoli, macrophages can adopt a pro-inflammatory and profibrotic phenotype, and undergo pyroptosis when infected. Overall, in severe COVID-19 cases the lung tissue of the patient becomes highly inflamed, flooded and scarred, impairing its function [349]. Despite being predominantly considered to act upon the respiratory system, COVID-19 is a systemic illness that can also have renal, cardiac, hematologic, gastrointestinal, hepatobiliary, endocrinologic, neurologic and cutaneous manifestations [357].

### 1.4.5 SARS-CoV-2 treatment and prevention

At this time, there are multiple therapeutic choices available for the treatment of COVID-19 including anti-inflammatory, immunomodulatory, anticoagulation and antiviral drugs. Additionally, anti-SARS-CoV-2 mAbs and convalescence plasma infusions have been investigated (Table 1.7) [358]. COVID-19 can be clinically separated into two stages, an early phase where SARS-CoV-2 replication is at its peak and a late phase where illness is driven by the establishment of hyperinflammatory and states. Hence, antiviral drugs and treatments based on antibody delivery are likely to be more effective during the early phases of COVID-19. On the other hand, anti-inflammatory, immunomodulatory and anticoagulation drugs are expected to be more effective in the later phase of infection [359].

Since the emergence of SARS-CoV-2, numerous novel and repurposed compounds have been shown to possess antiviral activity against coronaviruses [313,314]. In its active state, the antiviral agent remdesivir functions as a nucleoside analog that incorporates with the SARS-CoV-2 RdRp and causes

stalling following the addition of three extra nucleotides [360]. Studies conducted early into the pandemic showed remdesivir was superior compared to placebo, decreasing recovery time in hospitalised adults with moderate or severe COVID-19 infection [361,362]. However, a large study conducted later suggested remdesivir had little or no impact on the overall mortality, mechanical ventilation requirement and length of hospitalisation of COVID-19 inpatients, making its effectiveness disputed [363]. Since then, another nucleoside analog called molnupiravir was shown to effectively

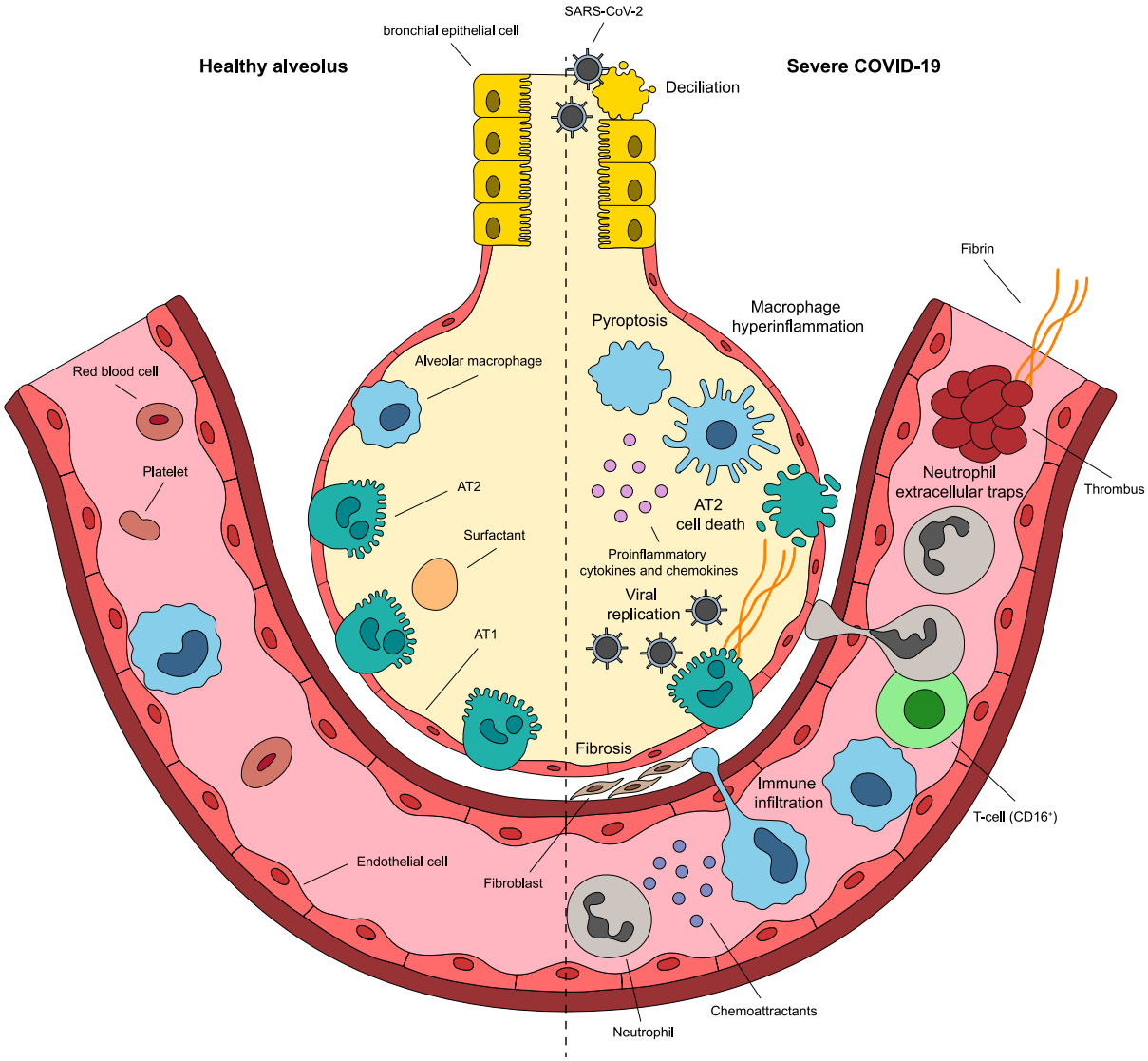



*Figure 1.32: COVID-19-associated acute respiratory distress syndrome.*
*SARS-CoV-2 initially infects ciliated cells in the upper conducting airways, gradually spreading through susceptible cells in the tracheobronchial tree. Dissemination into the alveoli occurs because of disrupted mucus flow, following infection induced loss of ciliation in ciliated cells. The alveoli are lined with AT1 and AT2 cells, responsible for gaseous exchange and surfactant secretion respectively. AT2 cells are precursors of AT1 cells, and express ACE2 and TMPRSS2 making them susceptible to SARS-CoV-2 infection. The alveolar epithelium becomes damaged due to AT2 infection and inflammatory responses. The AT2 cells proliferate but are unable to fully differentiate into AT1 cells. Stimuli such as hypoxia, cytokines and chemokines activate capillary endothelial cells, weakening cell-cell bonds and triggering apoptosis. Complement is activated, inducing CD16+ T-cells that promote the release of chemokines and microvascular endothelial cell damage. The exposed subendothelial extracellular matrix triggers the coagulation cascade, and results in fibrin deposition. Activated platelets are recruited to exposed sites, preventing leakage and secreting factors that sustain coagulation. Platelets stimulate neutrophils, which in cooperation with monocytes release neutrophil extracellular traps that drive coagulation and the formation of fibrin thrombi. Damage to the alveolar epithelium instigated by the virus leads to infiltration by immune cells, particularly monocytes and macrophages. Once inside the alveoli, macrophages adopt a pro-inflammatory and profibrotic phenotype, and undergo pyroptosis when infected. Adapted from [349].*

inhibit SARS-CoV-2 replication in human lung tissue [364]. Molnupiravir also targets the RdRp but instead functions as a mutagenising agent, causing an error catastrophe during the replication process [365]. A phase III clinical trial has reported that molnupiravir can decrease the risk of death or hospitalisation in unvaccinated adults with mild-moderate COVID-19 infection by ~50% compared to placebo [366]. A novel antiviral candidate has been recently developed called paxlovid, it contains a combination of nirmatrelvir (chymotrypsin-like protease inhibitor) and a low dose of ritonavir (pharmacokinetic boosting agent) [367]. The interim analysis of phase II-III data has shown that paxlovid can reduce COVID-19 associated hospitalisation or all-cause mortality by ~89% compared to placebo, when taken within three days of the onset of symptoms [368]. Interestingly, both molnupiravir and paxlovid can be taken orally, allowing the possibility of self-administration by patients at home to lower the risk of hospitalisation [369].

Immune dysregulation in COVID-19 cases can lead to the release of multiple inflammatory mediators, resulting in severe complications and a worse clinical outcome [358]. Corticosteroids are stress hormones secreted by the adrenal cortex that inhibit the release of several cytokines [370]. One such corticosteroid (dexamethasone), given at a dose of 6mg daily for up to 10 days, has been shown to reduce 28-day mortality in patients requiring oxygen or invasive ventilation support. However, patients that did not require respiratory assistance did not benefit from this treatment [371]. Drugs such as tocilizumab and sarilumab function by interacting with the receptor of the proinflammatory cytokine IL-6, inhibiting classic signalling and trans-signalling pathways [358]. Meta-analysis has shown that administration of anti-IL-6 receptor mAbs was associated with a lower 28-day all-cause mortality [372]. Imbalances between the Janus kinase (JAK) and signal transducer and activator of transcription pathways have been observed in severe cases of COVID-19. It has been suggested that this is associated with immune dysregulation [373]. The JAK 1 and JAK 2 inhibitor baricitinib was tested, in combination with the antiviral remdesivir, on hospitalised adults with COVID-19. Baricitinib and remdesivir were more effective compared to remdesivir alone, decreasing time to recovery and accelerating improvements in patient clinical status. These findings were most notable in patients that required respiratory assistance [374].

In severe cases of COVID-19, the coagulation cascade is activated and clotting factors become consumed [358]. Patients can experience coagulopathy and subsequent thromboembolic disease, which can be prevented and treated by anticoagulation therapy. Hence, anticoagulation therapy is a crucial

*Table 1.7: SARS-CoV-2 therapeutics.*
*Shown here are examples of SARS-CoV-2 therapeutics, with drug class and manufacturer information.*

| Name | Drug class | Manufacturer |
|---|---|---|
| **Remdesivir** | Nucleoside inhibitor | Gilead Sciences |
| **Molnupiravir** | Nucleoside inhibitor | Merck and Ridgeback Biotherapeutics |
| **Paxlovid** | Protease inhibitor | Pfizer |
| **Dexamethasone** | Corticosteroid | Various |
| **Tocilizumab** | IL-6 receptor antagonist | Roche |
| **Sarilumab** | IL-6 receptor antagonist | Regeneron and Sanofi |
| **Baricitinib** | JAK inhibitor | Eli Lilly and Incyte Corporation |
| **Heparin** | Anticoagulant | Pfizer |
| **REGN-COV2** | Monoclonal antibody cocktail | Regeneron and Roche |
| **Bamlanivimab-etesevimab** | Monoclonal antibody cocktail | Eli Lilly and Company |
| **Sotrovimab** | Monoclonal antibody | GlaxoSmithKline and Vir Biotechnology |

part of the treatment plan for hospitalised COVID-19 patients. However, there is a lack of clinical trials directly investigating the impact of anticoagulants on the clinical outcome of COVID-19 patients. Nonetheless, a meta-analysis of retrospective studies suggests that heparin can reduce 28-day mortality in patients showing signs of coagulopathy onset [375].

Treatment with convalescent plasma was assessed throughout the SARS-CoV, MERS-CoV and Ebola epidemics, but there was a lack of randomised control trials to endorse its efficacy [376]. A meta-analysis in COVID-19 patients has specified that convalescent plasma therapy does not decrease all-cause mortality, and should not be administered as part of standard care [377]. However, important considerations may have been overlooked in certain studies, such as the donor plasma having an adequate neutralisation titre against the variant the recipient is infected with [378].

A number of highly potent anti-SARS-CoV-2 mAbs have been isolated, some of which have undergone investigation as a treatment option for COVID-19 infection [379]. REGN-COV2, a cocktail of two non-competing neutralising mAbs that target the RBD of the SARS-CoV-2 S glycoprotein, were shown to function well prophylactically and therapeutically in non-human primates [380]. The interim analysis of a study has found that REGN-COV2 can decrease the viral load of non-hospitalised subjects infected with COVID-19, with a safety profile comparable to placebo [381]. A phase III clinical trial has shown that REGN-COV2 can reduce the risk of hospitalization or death in COVID-19 patients by ~71% compared to placebo. Additionally, REGN-COV2 treatment lowered viral load more rapidly, with the median time to resolution of symptoms reduced by four days [382]. Bamlanivimab and etesevimab are another pair of potent anti-S neutralising mAbs that target the RBD. Experiments have shown that etesevimab interacts with an epitope that is different to bamlanivimab, and can neutralise variants that bear escape mutations to bamlanivimab [383]. In phase II clinical trial co-administration of bamlanivimab-etesevimab to non-hospitalised COVID-19 patients with mild-moderate symptoms led to a significantly reduced viral load at day 11 [384]. A phase III clinical trial has shown that bamlanivimab-etesevimab can reduce the risk of hospitalization or death in COVID-19 patients by ~70% compared to placebo [385]. Unfortunately, mutations acquired by the B.1.1.529 (o) variant completely abolish the neutralisation capability of REGN-COV2 and bamlanivimab-etesevimab in vitro [386]. Sotrovimab is an engineered mAb, originally isolated from a SARS-CoV infected donor, which can neutralise multiple strains of SARS-CoV-2 (including B.1.1.529; o) [387,388]. The interim analysis of a phase III clinical trial has shown that sotrovimab can reduce the risk of hospitalization or death in COVID-19 patients by ~85% compared to placebo [389]. This highlights the importance of identifying novel mAbs that target conserved epitopes of the SARS-CoV-2 S glycoprotein as the virus evolves in nature.

### 1.4.6 SARS-CoV-2 novel vaccine technologies

The world required an effective SARS-CoV-2 vaccine, and this led to a development race in which multiple vaccine platforms were examined (Table 1.8). Some vaccines utilised to more traditional methodologies using live-attenuated virus, inactivated virus and viral protein subunits. However, the requirement for a fast production and rollout resulted in more modern platforms being considered, such as viral vector and mRNA vaccines. These newer candidates had an advantage over other technologies because they simply required design of the nucleic acid sequence being delivered prior to synthesis. Hence, viral vector and mRNA vaccines gained emergency use authorisation more quickly than other candidates [390].

The theory of mRNA vaccines was first conceived in the 1989, but would not achieve major clinical success until 2020 [391]. This came with the global deployment of the Pfizer-BioNTech (BNT162b2) and Moderna (mRNA-1273) vaccines. The mRNA in these vaccines mimic the structure of molecules found endogenously in humans, bearing a 5’ cap, 5’ UTR, ORF, 3’ UTR and poly-A tail. Following genome

sequencing of a pathogen, a desired antigen can be designed and implanted as the ORF into a DNA plasmid construct [392]. For both BNT162b2 and mRNA-1273, this antigen is a membrane-anchored stabilised (via S-2P mutations) pre-fusion S glycoprotein [393]. Generated plasmids are transcribed in vitro by bacteriophage polymerases, enabling a fast, scalable and cost-effective synthesis. The mRNA transcripts are purified via high performance liquid chromatography (HPLC), eliminating any contaminants and reactants. Due to its negative charge, mRNA is unable to pass through the anionic lipid bilayer of a cell. Further, once inside the body mRNA can be targeted by the innate immune system, and is subject to degradation by nucleases. Thus, the mRNA payloads are encapsulated by lipid nanoparticles (LNPs), which are comprised of four components including an ionisable lipid, cholesterol, helper phospholipid and polyethylene glycol (PEG) modified lipid [392]. Of note, the ionisable lipid utilised by BNT162b2 and mRNA-1273 vaccines are ALC-0315 and SM-102 respectively [394]. In a microfluidic mixer, the purified mRNA and lipid components are rapidly mixed in an acidic buffer. Ionisable lipids are positively charged in this setting and are able to attract their mRNA cargo, resulting in the precipitation of self-assembled LNPs. The LNPs are dialysed or filtered to get rid of non-aqueous solvents and any unencapsulated mRNA, finally being stored in sterilised vessels [392]. Results from large multinational clinical trials have shown that two doses of BNT162b2 (21 day interval) and mRNA-1273 (28 day interval) were able to achieve vaccine efficacies of ~95% and ~94.1% respectively [395,396].

Despite their overall success, there are some limitations associated with mRNA vaccines. Although rare, it has been documented that the BNT162b2 and mRNA-1273 vaccines can cause severe anaphylactic reactions [397]. The source of these reactions has not yet been ascertained, however it has been suggested that they may be the result of allergies to the PEG used in the LNPs [398]. Manufacturing these vaccines requires multiple expensive components, making them significantly more costly compared to alternative products. BNT162b2 and mRNA-1273 require a temperature of −70°C for long term storage, making distribution to regions that lack a cold-chain infrastructure problematic [399].

Adenoviruses were first discovered in 1953, and have since become promising vectors for the delivery of genetic material [400]. Adenoviruses have a characteristic icosahedral shape with an external protein encapsulation, and typically only cause mild self-limiting respiratory and ocular infections in humans [399]. They carry a compact double stranded DNA genome, expressing either early (E1 or E3) or late genes. The early genes are required for viral replication, whilst the late genes are important for the release of new particles [401]. Adenoviruses can be engineered to function as vectors for antigen gene delivery by replacing the E1 or E3 genes with a desired ORF. This prevents the virus from being able to undergo genome replication, whilst the gene encoding the antigen of interest is transcribed and released into the cytoplasm. Adenovirus vectors do not significantly integrate into the host genome, remaining episomal following entry into the nucleus. SARS-CoV-2 vaccines have been developed using human (Ad26.COV2-S by Janssen-Johnson) and chimpanzee (ChAdOx1-nCoV by Oxford-AstraZeneca)

*Table 1.8: Types of SARS-CoV-2 vaccine.*
*Shown here are examples of SARS-CoV-2 vaccines, with information regarding vaccine type and manufacturer.*

| Name | Type | Manufacturer |
|---|---|---|
| **ChAdOx1-nCoV** | Viral vector | Oxford-AstraZeneca |
| **Ad26.COV2-S** | Viral vector | Janssen-Johnson |
| **Ad5-nCoV-S** | Viral vector | CanSino Biologics |
| **mRNA-1273** | RNA or DNA based | Moderna |
| **BNT162b2** | RNA or DNA based | Pfizer-BioNTech |
| **INO-4800** | RNA or DNA based | Inovio |
| **NVX-CoV2373** | Protein subunit | Novavax |
| **CoronaVac** | Inactivated whole virus | Sinovac Biotech |
| **BBIBP-CorV** | Inactivated whole virus | Sinopharm |

adenoviruses. Chimpanzee adenovirus vectors were introduced in the hopes to circumvent pre-existing immunity to human serotypes, which has been shown to be capable of mitigating gene delivery [399]. The ChAdOx1-nCoV carries an unmodified full-length S glycoprotein gene, whilst the Ad26.COV2-S gene is stabilised (S-2P) and carries mutations in the furin cleavage site [390,402]. When a suitable transgene cassette has been cloned into the adenoviral backbone, the viruses are grown in specialised cells [403]. The vectors undergo a purification process involving depth filter clarification, tangential flow filtration (TFF), anion exchange chromatography and an additional round of TFF [404]. Results from large multination clinical trials have shown that a single dose of Ad26.COV2-S and two doses of ChAdOx1-nCoV were able to achieve vaccine efficacies of ~66.9% and ~70.4% respectively [405,406].

Almost all of the elements in the adenovirus vectors aid in the generation of anti-adenovirus immunity. For human adenovirus vectors, regional seroprevalence and cross-reactivity must be considered prior to vaccine deployment. Anti-adenovirus immunity can limit the prospects of annual boosters for both human and chimpanzee adenovirus vaccines [399]. The Ad26.COV2-S and ChAdOx1-nCoV vaccines have been linked to rare blood clotting events [407,408]. Numerous countries paused their administration of these products, but the majority have since resumed. Patients that have developed blood clots post-vaccination have shown similarities to heparin-induced thrombocytopenia (HIT), a rare disorder. However, the connection between adenovirus vaccines and HIT are not fully understood at this time [399].

Despite their differences, administration of either mRNA or viral vector vaccines results in mRNA encoding S glycoprotein being released into the target cell cytoplasm. This mRNA is translated and processed by the host cell machinery, yielding S glycoproteins. These glycoproteins can be displayed on the cell surface and interact with circulating elements of the host immune system. Alternatively, the glycoproteins may be broken into antigen fragments by the proteasome complex. The antigen fragments can then be presented to cytotoxic T-cells on MHC class I molecules, leading to their activation. The activated cytotoxic T-cells can proceed to eliminate the transfected or infected cells, through the secretion of molecules such as perforin and granzyme. Moreover, antigens that have been secreted can be endocytosed and degraded within the endosomes of antigen-presenting cells. These proteins can then be displayed on MHC class II proteins, activating helper T-cells which can stimulate other immune cells [392].

The rapid generation and overall success of the SARS-CoV-2 vaccine is in strong contrast to what was observed during the HIV-1 efficacy trials [242,244,395,396]. Reasons for this include SARS-CoV-2 having a much slower antigenic drift, a surface glycoprotein more susceptible to immune responses and the lack of requirement for sterilising immunity [409]. Utilisation of these more modern vaccination technologies in the HIV-1 field will likely help to overcome some of the issues that are being faced, such as allowing re-introduction of the transmembrane domain into stabilised Env trimers and removing the highly immunogenic based displayed by artificial truncation [410]. However, the difficulties in targeting rare bnAb germlines and generating a broad and potent antibody response will remain [411]. As seen with the SARS-CoV-2 vaccine, a large limitation of mRNA and viral vector vaccines is the lack of potency and the requirement for numerous doses to achieve efficacy. Further improvements in these methods, such as the generation of self-replicating mRNA vaccines, could be of benefit to both the HIV-1 and SARS-CoV-2 field [412].

### 1.4.7 Comparison of SARS-CoV-2 with influenza and HIV-1

Aside from its similarities with other human coronaviruses, parallels can also be drawn between SARS-CoV-2 and other viruses that cause acute respiratory infections, such as Influenza A [413].

There have been four pandemics caused by influenza A, occurring in 1918 (50 million deaths), 1957 (~1.1 million deaths), 1968 (~1 million deaths) and 2009-2010 (151,700–575,400 deaths) [413–417]. Since then, the novel influenza A of the 2009-2010 pandemic has continued to spread as a seasonal flu virus [413]. Comparable to SARS-CoV-2, antibody responses are generated to the Influenza A surface glycoprotein and are known to be protective against infection. A lack of antibody-based immunity in the human population is a major factor that allows emerging pandemic viruses to spread rapidly. The selective pressure of these antibody responses, generally resulting from natural infection, drive changes in the viral surface antigen [418]. However, the influenza virus surface glycoprotein has been shown to be particularly capable of tolerating antigenic drift [419]. This coupled with a mutation rate of $1.5\times10^{-5}$ mutations per nucleotide per replication cycle, has resulted in influenza vaccines needing to be updated each year to match the currently circulating strains [420,421]. A similar approach has been taken recently with SARS-CoV-2 vaccination, with the introduction of bivalent doses to protect against newer VOCs [422]. However, achieving neutralisation breadth against influenza strains is more difficult than what has been seen with SARS-CoV-2 [423]. Further, a universal flu vaccine will require methods more comparable to what has been tested in the HIV-1 field, with a series of immunisations with distinct antigens to particular bnAb germlines needed to elicit the desired response [423,424]. As noted with chronic viral infections, although there are many parallels between SARS-CoV-2 and influenza A, each virus will still require differing approaches to yield the best vaccination and treatment plans.

SARS-CoV-2 and HIV-1 are both enveloped RNA viruses, which are responsible for the currently ongoing COVID-19 and AIDS pandemics respectively [425]. They express surface glycoproteins that enable the entry of the virus into host cells, and represent the main target of neutralising antibody responses. However, a clear distinction in the life cycle of these viruses is that HIV-1 can integrate its genetic material into the hosts genome, whilst SARS-CoV-2 cannot [426,427]. This enables HIV-1 to form latent viral reservoirs limiting the control that the host immune system can exert on the virus, and thereby preventing clearance of the pathogen [428]. Conversely, during early SARS-CoV-2 infection, neutralising antibody responses continue to be beneficial in controlling the virus and limiting disease severity [381,382]. The chronic nature of HIV-1 means that a successful antibody-based vaccine against this virus would need to provide sterilising immunity to be effective [429]. On the contrary, whilst sterilising immunity is an advantageous feature for a SARS-CoV-2 vaccine, its primary goal is to reduce disease severity and prevent hospitalisation [430].

# Chapter 2: Materials and methods

## 2.1 Bacterial work

### 2.1.1 Preparation of LB broth and LB plates

25g of Lysogeny broth (LB; Fisher BioReagents) and 40g of LB agar (Miller) were dissolved in 1L containers of double distilled water (dd$H_2O$). Mixtures were agitated until completely dissolved and sterilised by autoclaving. LB broth without selection marker was stored at room temperature. LB agar without selection marker was poured under a Bunsen burner flame into plates, allowed to set and stored at 4°C.

Ampicillin salt (Sigma Aldrich) was dissolved in dd$H_2O$ at a concentration of 100mg/mL and sterilised through a 0.22µm filter (Millex), stocks were stored at -20°C. Ampicillin stocks were added to LB broth and LB agar to a final concentration of 100µg/mL. LB broth with ampicillin were stored at 4°C. LB agar with was ampicillin was poured under a Bunsen burner flame into plates, allowed to set and stored at 4°C.

### 2.1.2 Preparation of chemically competent *E. coli*

A fresh stock of Escherichia coli (*E. coli*; Table 2.1), was streaked on a LB agar plate without selection marker, incubated at 30-37°C for 12-18 hours and stored at 4°C. A single bacterial colony was inoculated into 10mL of LB broth without selection marker and grown at 30°C in a shaking incubator for 12-18 hours. 500mL of LB broth without selection marker was inoculated with 5mL of the starter culture and grown at 30°C in a shaking incubator. The absorption of the culture was measured regularly via spectrophotometer (DeNovix) at an optical density of 550nm, until it reached a reading of 0.5 (normally ~3 hours). The culture was put on ice for 10 minutes and centrifuged at 3000xg for 10 minutes at 4°C. After discarding the supernatant, the pellet was resuspended in 100mL of TfB1 (Table 2.2) and incubated on ice for 5 minutes. The *E. coli* were centrifuged at 3000xg for 10 minutes at 4°C and the supernatant was discarded. The pellet was resuspended in 8mL of TfB2 and incubated on ice for 10 minutes. The chemically competent bacteria were then aliquoted, flash frozen on dry ice and stored at -80°C.

### 2.1.3 Transformation of DNA plasmids into chemically competent *E. coli*

5µL of ligation (2.2.5) or Gibson assembly (2.2.6) product was added to 50µL of chemically competent *E. coli* and incubated on ice for 20-30 minutes. The mixture was heat shocked at 42°C for 45 seconds and put back on ice for 2 minutes. 200µL of LB broth without selection marker was added to the bacteria and the culture was grown in a shaking incubator at 30-37°C for 45 minutes. The culture was spread on LB agar plates with ampicillin, incubated at 30-37°C for 12-18 hours and stored at 4°C. The same process was used to re-transform 1µL of plasmid, but the first incubation on ice could be circumvented and the time spent in the shaking incubator could be reduced to 30 minutes.

### 2.1.4 DNA plasmid expansion and extraction

Single bacterial colonies were inoculated into 5-10mL, 100-200mL, 400-500mL or 800mL-1L of LB broth with ampicillin, and were extracted using QIAprep Spin Miniprep kits (Qiagen), NucleoBond Xtra Midi/Maxi kits (Macherey-Nagel) or Plasmid Plus Giga kits (Qiagen) respectively. DNA was

*Table 2.1: Chemically competent bacteria.*

| Company | Catalogue number | Bacteria Strain |
|---|---|---|
| **NEB** | C3040I | NEB stable |
| **NEB** | C2987I | DH5α |
| **NEB** | C3019I | DH10β |

resuspended using nuclease-free water (Ambion) and concentrations were measured using a nano-drop (ThermoFisher).

## 2.2 DNA work

### 2.2.1 Agarose gel electrophoresis and DNA fragment extraction

0.8% agarose gels were prepared by adding 0.6-0.8g of agarose (Invitrogen) to 80-100mL of TAE buffer, the mixtures were microwaved until the agarose dissolved. 8-10µL of ethidium bromide (Sigma) was added and the solution was poured into either a small or large mould with the desired comb. Once set, the gel was placed in an electrophoresis tank and submerged in TAE buffer. To examine DNA fragments, samples were mixed with 6x DNA loading dye (New England Biolabs) and loaded into wells alongside the appropriate amount of 1kb ladder (New England Biolabs). A power pack was used to electrify the gel at 100-120V for 35-50mins to achieve an adequate separation of the DNA bands. Gels were visualised using an ultraviolet (UV) transilluminator and the desired DNA fragments were excised using a razor blade. The DNA was extracted using a QIAquick Gel Extraction kit (Qiagen) following the manufacturer's instructions.

### 2.2.2 Generic PCR

Primers were designed manually to be ~18-24 bases in length, have GC contents of ~45-55%, incorporate GC-locks if possible, have annealing temperatures between 50-65°C (ascertained using NEB Tm calculator) and not self-hybridise. When being used for Gibson assembly (2.2.6), 15-40 base overhangs were added to the primers. These overhangs were complementary to the assembly areas of the linearised plasmid that the fragments were going to be cloned into. Similarly, if the polymerase chain reaction (PCR) was being used for restriction enzyme cloning or site-directed mutagenesis, restriction sites or mutations would be added to the primers as overhangs. The PCR reaction was conducted using the Phusion High-Fidelity PCR kit (New England Biolabs; Table 2.3). 5µL of PCR product was examined on a 0.8% agarose gel to confirm the presence of a DNA band at the correct size (2.2.1).

### 2.2.3 Site-directed mutagenesis

Primers were generated using the NEBaseChanger tool, but were re-designed manually if annealing temperatures were >72°C (2.2.2). The PCR reaction was conducted using the Q5 High-Fidelity PCR kit (New England Biolabs; Table 2.4). 5µL of PCR product was examined on a 0.8% agarose gel to confirm the presence of a DNA band at the correct size (2.2.1). If the PCR was successful, a digestion, phosphorylation and ligation reaction was carried out at room temperature for 1 hour (Table 2.5). 5µL of the ligation product was then transformed into chemically competent *E. coli* (2.1.3).

### 2.2.4 Overlap PCR

To swap large sections of glycoprotein, primers were designed to amplify recipient fragments and donor fragments from viral vectors, each with complementary overhangs to one another. PCRs (2.2.2), gel electrophoresis and gel extractions (2.2.1) were performed to amplify, confirm and extract the DNA fragments. The overlap PCR was carried out using equimolar amounts of the fragments as

*Table 2.2: Buffers used to prepare chemically competent bacteria.*

| | **TfB1** | **TfB2** |
|---|---|---|
| *KAc* | 10mM | - |
| *RbCl* | 100mM | 10mM |
| *CaCl2* | 10mM | 75mM |
| *MnCl2* | 50mM | - |
| *Glycerol* | 15% v/v | 15% v/v |
| *PIPES* | - | 10mM |

template DNA and the forward primer of the upstream recipient fragment and the reverse primer of the downstream recipient fragment. The overlap PCR product was then cloned into an expression vector using restriction enzyme digestion and ligation (2.2.5). 5µL of the ligation product was then transformed into chemically competent *E. coli* (2.1.3).

### 2.2.5 Restriction digestion and ligation of DNA

Plasmids were digested using restriction enzymes (New England Biolabs) according to the manufacturer's instructions, determined using the NEBcloner tool. When possible, double digests were performed and the restriction enzymes were heat inactivated (HI) upon completion of the reaction. The linearised fragments were dephosphorylated using shrimp alkaline phosphatase (rSAP) to prevent re-ligation (Table 2.6). The entire digestion product was examined on a 0.8% agarose gel to confirm the presence of fragments at the correct size, the desired bands were excised and the DNA was extracted (2.2.1).

Ligations were carried out between insert fragments and vector fragments with complementary sticky or blunt ends, through the use of T4 Ligase (New England Biolabs; Table 2.7). After calculating the molar concentrations manually, the insert fragment was used at a 3:1 ratio to the vector fragment. The reaction was incubated at room temperature for 2 hours or at 16°C for 12-18 hours. 5µL of the ligation product was then transformed into chemically competent *E. coli* (2.1.3).

### 2.2.6 Gibson assembly

Gibson reaction mix was generated using the appropriate enzymes, an isothermal reaction buffer and an enzyme storage buffer (Table 2.8). Plasmids were linearised using restriction enzyme digestion (2.2.5), and inserts were amplified using PCR (2.2.2). 0.1-0.5pmol of DNA, with the insert in 3-fold molar excess, was diluted in nuclease-free water to a total volume of 4µL and combined with 6µL of Gibson reaction mix. The Gibson assembly reaction was performed using a preheated thermocycler at 50°C for 45 minutes. 5µL of the assembled product was then transformed into chemically competent *E. coli* (2.1.3).

*Table 2.3: Phusion PCR recipe and thermocycler conditions.*

| | **Quantity** | **Final concentration** |
|---|---|---|
| *Template DNA (15ng/µL)* | 2.8µL | 2.1ng/µL |
| *GC buffer (5x)* | 4µL | 1x |
| *Nuclease-free H20* | 10µL | - |
| *Forward primer (10µM)* | 1µL | 0.5µM |
| *Reverse primer (10µM)* | 1µL | 0.5µM |
| *dNTPs (10mM)* | 0.5µL | 250µM |
| *DMSO* | 0.5µL | 2.5% v/v |
| *Phusion (2U/µL)* | 0.2µL | 0.02U/µL |

| Step | Temperature | Time |
|---|---|---|
| **Initial denaturation** | 98°C | 2 minutes |
| **45 cycles** | 98°C<br>56-72°C<br>72°C | 15 seconds<br>30 seconds<br>30 seconds per kb |
| **Final extension** | 72°C | 10 seconds |
| **Hold** | 4°C | ∞ |

### 2.2.7 Sequencing

PCR products and plasmid DNA samples were Sanger sequenced externally by Eurofins Genomics or GENEWIZ Germany GmbH. The sequencing results were examined using ApE [431].

## 2.3 General cell culture work

### 2.3.1 Cell lines and sub-culturing

Adherent cell lines were maintained in Dulbecco's Modified Eagle Medium (DMEM, Gibco), supplemented with 10% HI foetal bovine serum (FBS; Gibco or Sigma-Aldrich) and 1% penicillin (100 U/mL, Gibco) and streptomycin (100 µg/mL, Gibco) at 37°C with 5% CO2 (Table 2.9). Typically, the cell lines were sub-cultured using TrypLE Express (Gibco) every 2-3 days at a 1:10 ratio.

Non-adherent 293F cells were maintained in FreeStyle 293 Expression Medium (Gibco) at 37°C with 8% CO2 and 125 revolutions per minute (rpm) for agitation. Typically, this cell line was sub-cultured to retain a stocks at a concentration of $2x10^5$ cells/mL every 2-3 days.

### 2.3.2 Transient transfection of recombinant proteins

For small scale expressions of monoclonal antibodies, HEK293T or HEK293T/17 cells were pre-seeded in 12-well plates at a concentration of $5x10^5$ cells/mL and with a volume of 1mL per well. In an Eppendorf, 0.5µg of the $Ig_H$ and 0.5µg of the $Ig_L$ DNA plasmids were diluted in a total volume of 50µL in OptiMEM (Gibco). In another Eppendorf, 3µg of polyethyleniminely (PEI) Max (Polyscience) was diluted in a total volume of 50µL in OptiMEM. The two mixtures were combined and incubated at room temperature for 15 minutes, before being added dropwise to a single well. The supernatant was harvested 3 days later and sterilised through a 0.22µm filter.

For small scale expressions of viral glycoproteins, HEK293T or HEK293T/17 cells were pre-seeded in 12-well plates at a concentration of $5x10^5$ cells/mL and with a volume of 1mL per well. In an Eppendorf, 0.5µg of the viral glycoprotein DNA plasmid was diluted in a total volume of 50µL in OptiMEM. In another Eppendorf, 10µg of PEI Max was diluted in a total volume of 50µL in OptiMEM. The two mixtures were combined and incubated at room temperature for 15 minutes, before being added dropwise to a single well. The supernatant was harvested 3 days later and sterilised through a 0.22µm filter.

*Table 2.4: Q5 PCR recipe and thermocycler condition.*

| | **Quantity** | **Final concentration** |
|---|---|---|
| *Template DNA (15ng/µL)* | 2.8µL | 2.1ng/µL |
| *Q5 buffer (5x)* | 4µL | 1x |
| *Nuclease-free H20* | 11µL | - |
| *Forward primer (10µM)* | 0.5µL | 0.25µM |
| *Reverse primer (10µM)* | 0.5µL | 0.25µM |
| *dNTPs (10mM)* | 0.4µL | 200µM |
| *DMSO* | 0.5µL | 2.5% v/v |
| *Q5 (2U/µL)* | 0.3µL | 0.03U/µL |

| Step | Temperature | Time |
|---|---|---|
| **Initial denaturation** | 98°C | 30 seconds minutes |
| **25 cycles** | 98°C<br>50-72°C<br>72°C | 10 seconds<br>30 seconds<br>30 seconds per kb |
| **Final extension** | 72°C | 2 minutes |
| **Hold** | 4°C | ∞ |

For small scale expression of biotinylated proteins, HEK293T or HEK293T/17 cells were pre-seeded in 12-well plates at a concentration of $5x10^5$ cells/mL and with a volume of 1mL per well. A 10mM stock of biotin (Sigma Aldrich) was made in FreeStyle 293 Expression Medium, sterilised through a 0.22µm filter and added to each well at a final concentration of 2µM. In an Eppendorf, 0.4µg of the protein and 0.1µg of the pDisplay-BirA-ER (Addgene) DNA plasmids were diluted to a total volume of 50µL in OptiMEM. In another Eppendorf, 10µg of PEI Max was diluted to a total volume of 50µL in OptiMEM. The two mixtures were combined and incubated at room temperature for 15 minutes, before being added dropwise to a single well. The supernatant was harvested 3 days later and sterilised through a 0.22µm filter.

For large scale expressions, these procedures were scaled up and used to transfect 293F cells at a concentration of $1x10^6$ cells/mL. The supernatant was harvested 5-7 days later, centrifuged at 1200rpm for 5 minutes to remove suspension cells and centrifuged again at 10,000xg for 5 minutes to remove cell debris, prior to sterilisation through a 0.45µm filter.

### 2.3.2 Isolation of PBMCs from whole blood

Blood was collected in ethylenediaminetetraacetic acid (EDTA) treated tubes (Sigma Aldrich) and isolation of peripheral blood mononuclear cells (PBMC) was carried out within hours of receiving the samples. Blood was diluted at a 1:1 ratio with phosphate-buffered saline (PBS; Gibco). 15mL of Histopaque (Sigma Aldrich) was added to 50 mL falcon tubes and up to 30mL of diluted blood was layered on top. The falcon tubes were centrifuged at 600xg for 20 minutes with 0 deceleration. The diluted plasma layer was collected, aliquoted into screw cap Eppendorf tubes and stored at -80°C. The PBMC layer was collected and added to new 50mL falcon tubes that contained 30mL of PBS. The PBMCs were topped up to 50mL with PBS and centrifuged at 300xg for 10 minutes with 9 deceleration. The supernatant was discarded and the PBMCs were resuspended with 30mL of PBS. The PBMCs were centrifuged at 150xg for 20 minutes with 9 deceleration. The supernatant was discarded and the PBMCs were resuspended with 10mL of PBS, samples that had been separated were recombined at this point. The PBMCs were centrifuged at 300xg for 10 minutes with 9 deceleration. The supernatant was discarded and the PBMCs were resuspended in 5mL ACK (ammonium–chloride–potassium) Lysis Buffer (Lonza) and incubated for 5 minutes at room temperature in the dark. The PBMCs were topped up to 50mL with PBS and centrifuged at 300xg for 10 minutes with 9 deceleration. The supernatant was discarded and the PBMCs were resuspended in PBS for counting using a haemocytometer. The PBMCs were centrifuged at 300xg for 10 minutes with 9 deceleration. The supernatant was discarded and the PBMCs were resuspended in HI FBS with 10% dimethyl sulfoxide (DMSO; Sigma Aldrich), at a concentration of $>5x10^5$ cells/mL. The PBMCs were aliquoted into cryovials, stored at -80°C in a Mr. Frosty (Nalgene) for 12-18 hours and moved into liquid nitrogen for long term storage.

## 2.4 Virology cell culture work

### 2.4.1 HIV-1 pseudovirus production

HIV-1 pseudoviruses capable of single round infection were generated using a two-plasmid system.

*Table 2.5: Digestion, phosphorylation and ligation recipe.*

| | Quantity | Final concentration |
|---|---|---|
| *PCR product* | 1µL | - |
| *T4 DNA ligase buffer (10x)* | 1µL | 1x |
| *Nuclease-free H20* | 5µL | - |
| *T4 PNK (10U/µL)* | 1µL | 1U/µL |
| *DnpI (10U/µL)* | 1µL | 1U/µL |
| *T4 DNA ligase (400U/µL)* | 1µL | 40U/µL |

The pSG3Δenv (NIH AIDS reagents) plasmid encoded a HIV-1 backbone with an insertion mutation at the *env* gene, resulting in a stop codon. To enable infection, these pseudoviruses were supplemented by *env* genes encoded in either PSVIII or pcDNA3 LIC vectors. HEK293T or HEK293T/17 cells were pre-seeded in 12-well plates at a concentration of $5x10^5$ cells/mL and with a volume of 1mL per well. In an Eppendorf, 2µg of pSG3Δenv and 1µg of the env DNA plasmids were diluted to a total volume of 50µL in OptiMEM. In another Eppendorf, 9µg of PEI (Polyscience) was diluted to a total volume of 50µL in OptiMEM. The two mixtures were combined and incubated at room temperature for 15 minutes, before being added dropwise to a single well. The supernatant was harvested 3 days later and sterilised through a 0.22µm filter. The pseudovirus was used fresh or stored at -80°C.

For large scale virus production, this procedure was scaled up and used to transfect 10-20cm dishes of HEK293T or HEK293T/17 cells pre-seeded at a concentration of $5x10^5$ cells/mL and with a volume of 10-20mL per dish. When required, glycosidase inhibitors kifunensine (5mM stock) and swainsonine (5mM stock) were added to the reactions at final concentrations of 25µM and 20µM respectively.

When transfections did not produce adequate titres for downstream applications, the pseudovirus was concentrated over a sucrose cushion. In an ultracentrifuge tube (Beckman Coulter), 25mL of filtered pseudovirus was layered onto 5mL of 20% sucrose (Sigma Aldrich) in PBS. The pseudovirus underwent ultracentrifugation at 28,000rpm (4°C) for 75 min with 0 deceleration. The supernatant was discarded and the pseudovirus was resuspended with the desired volume of culture medium. The concentrated pseudovirus was used fresh or stored at -80°C.

### 2.4.2 SARS-CoV-2 pseudovirus production

SARS-CoV-2 pseudoviruses capable of single round infection were generated using a three-plasmid system. The HIV 8.91 gag/pol plasmid was used to produce lentiviral particles and pHIV-Luc plasmid was used to enable a luminescent readout upon pseudovirus infection. To enable infection, these pseudoviruses were supplemented by either full length or truncated (CT removed) S genes encoded in pcDNA3.1+ vectors. HEK293T or HEK293T/17 cells were pre-seeded in 12-well plates at a concentration of $5x10^5$ cells/mL and with a volume of 1mL per well. In an Eppendorf, 1.5µg of pHIV-Luc, 1µg of the gag/pol and 0.5µg of the S DNA plasmids were diluted to a total volume of 50µL in OptiMEM. In another Eppendorf, 9µg of PEI Max was diluted to a total volume of 50µL in OptiMEM. The two mixtures were combined and incubated at room temperature for 15 minutes, before being added dropwise to a single well. The supernatant was harvested 3 days later and sterilised through a 0.22µm filter. The pseudovirus was used fresh or stored at -80°C.

For large scale virus production, this procedure was scaled up and used to transfect 10-20cm dishes of HEK293T or HEK293T/17 cells pre-seeded at a concentration of $5x10^5$ cells/mL and with a volume of 10-20mL per dish. When required, glycosidase inhibitors kifunensine (5mM stock) and swainsonine (5mM stock) were added to the reactions at final concentrations of 25µM and 20µM respectively.

When transfections did not produce adequate titres for downstream applications, the pseudovirus was concentrated over a sucrose cushion. In an ultracentrifuge tube, 25mL of filtered pseudovirus was

*Table 2.6: Restriction enzyme digest and dephosphorylation recipe.*

| | Quantity | Final concentration |
|---|---|---|
| *Plasmid/insert* | 1µg | 33.3ng/µL |
| *Restriction enzyme buffer (10x)* | 3µL | 1x |
| *Nuclease-free H20* | To 30µL | - |
| *Restriction enzyme 1 (20U/µL)* | 1µL | 0.7U/µL |
| *Restriction enzyme 1 (20U/µL)* | 1µL | 0.7U/µL |
| *rSAP (1U/µL)* | 1µL | 0.03U/µL |

layered onto 5mL of 20% sucrose in PBS. The pseudovirus underwent ultracentrifugation at 28,000rpm (4°C) for 75 min with 0 deceleration. The supernatant was discarded and the pseudovirus was resuspended with the desired volume of culture medium. The concentrated pseudovirus was used fresh or stored at -80°C.

### 2. 4.3 Authentic SARS-CoV-2 strain and propagation

The SARS-CoV-2 Strain England 2 (England 02/2020/407073) was acquired from Public Health England. The virus was propagated by infecting a T75 flask of Vero E6 cells (60-70% confluent) at a multiplicity of infection of 0.005 in 3mL of DMEM supplemented with 10% HI FBS. The cells were incubated for 1 hour at 37°C with 5% CO2 prior to the addition of 15mL of the same medium. The supernatant was harvested following visible cytopathic effect, sterilised through a 0.22µm filter, aliquoted and stored at -80°C. The titre of the infectious virus was examined via plaque assay using Vero E6 cells.

### 2.4.4 Neutralisation assays and viral titrations

The TZM-bl cell line was used for neutralisation assays with HIV-1 pseudovirus. TZM-bl cells are a HeLa derived cell line that stably express the CD4 receptor and chemokine receptors CXCR4 and CCR5, therefore allowing infection with HIV-1 pseudovirus regardless of tropism. Further, TZM-bl cells have a HIV–1 Tat responsive firefly luciferase (Luc) and a β-galactosidase reporter gene. Serial dilutions of monoclonal antibodies or plasma/serum (HI for 30 minutes at 56°C) were prepared with DMEM media supplemented with 10% HI FBS and 1% penicillin/streptomycin in round bottom 96-well plates (Corning). 25µL of the dilutions was transferred to a 96-well half area plate (Corning) and 25µL of HIV-1 pseudovirus was added on top. The mixture was incubated for 1 hour at 37°C with 5% CO2, then 50µL of TZM-bl cells was added to the wells at a concentration of $2.5x10^5$ cells/mL. For pseudoviruses with low titres, DEAE-dextran (stock 10mg/mL in PBS) was added alongside the TZM-bl cells to a final concentration of 10µg/mL. After 3 days, the supernatant was aspirated and the TZM-bl cells were lysed using 20µL of Luc lysis buffer, and luminescence was assessed.

A HeLa cell line that stably expressed ACE2 (HeLa ACE2, a kind gift from James Voss) was used for the SARS-CoV-2 pseudovirus neutralisation assays. Serial dilutions of monoclonal antibodies or plasma/serum (HI for 30 minutes at 56°C) were prepared with DMEM media supplemented with 10% HI FBS and 1% penicillin/streptomycin in round bottom 96-well plates. 25µL of the dilutions was transferred to a 96-well half area plate and 25µL of SARS-CoV-2 pseudovirus was added on top. The mixture was incubated for 1 hour at 37°C with 5% CO2, then 25µL of HeLa ACE2 cells was added to the wells at a concentration of $5x10^5$ cells/mL. After 3 days, the supernatant was aspirated and the HeLa ACE2 cells were lysed using 20µL of Luc lysis buffer, and luminescence was assessed.

Vero E6 cells were used for the authentic SARS-CoV-2 neutralisation assays. Vero E6 cells were pre-seeded in flat bottom 96-well plates at a concentration of $2x10^4$ cells/mL and with a volume of 100µL per well. Serial dilutions of monoclonal antibodies or plasma/serum (HI for 30 minutes at 56°C) were prepared with DMEM media supplemented with 2% HI FBS and 1% penicillin/streptomycin in round bottom 96-well plates. 50µL of the dilutions was transferred to another round bottom 96-well plates and 50µL of authentic SARS-CoV-2 was added on top. The mixture was incubated for 1 hour at 37°C with 5% CO2. The supernatant was aspirated from the Vero E6 cells and the virus mixture was added.

*Table 2.7: T4 ligation recipe.*

| | Quantity | Final concentration |
|---|---|---|
| *Insert and vector mix* | 100ng | 10ng/µL |
| *T4 DNA ligase buffer (10x)* | 1µL | 1x |
| *Nuclease-free H20* | To 10µL | - |
| *T4 DNA ligase (400U/µL)* | 1µL | 40U/µL |

The assay was incubated at 37°C with 5% CO2 for 24 hours. The supernatant was aspirated, and the cells were fixed with 150µL of 4% formalin at room temperature for 30 minutes and topped up to 300µL using PBS. The supernatant was aspirated and the cells underwent intracellular staining of N (2.6.3).

To examine pseudovirus titres, the same TZM-bl or HeLa ACE2 procedures were carried out but the monoclonal antibodies or plasma/serum was replaced with DMEM media supplemented with 10% HI FBS and 1% penicillin/streptomycin. The titre of the authentic SARS-CoV-2 batches was determined via plaque assay using Vero E6 cells. All neutralisation assays incorporated virus only and cells only controls to enable calculation of percentage neutralisation.

## 2.5 Recombinant protein work

### 2.5.1 Purification of monoclonal antibodies

Filtered supernatants, collected from large scale transient transfections of monoclonal antibodies (2.3.2), were incubated with 1mL of Protein G sepharose beads (GE Life Sciences) per 1L of culture. The beads were washed 3 times with PBS before use and re-used up to 6 times. The mixtures were

*Table 2.8: Gibson assembly master mix recipe.*

| 5x Isothermal reaction buffer | Use 2 mL tube |
|---|---|
| **PEG-8000** | 0.5 g |
| **1M Tris/ HCl pH 7,5** | 1 ml |
| **1 M $MgCl_2$** | 100 µl |
| **1M DTT (Fermentas # BP17225, make fresh)** | 100 µl |
| **100 mM dATP** | 20 µl |
| **100 mM dCTP** | 20 µl |
| **100 mM dGTP** | 20 µl |
| **100 mM dTTP** | 20 µl |
| **50 mM NAD (NEB # B9007S, use fresh)** | 200 µl |
| **MQ** | ±120 µl (to 2 ml) |

| Enzyme storage buffer. | Use 15 mL tube |
|---|---|
| **MQ** | 4420 µl |
| **100% glycerol** | 5000 µl |
| **1 M Tris/HCl pH 7,5** | 500 µl |
| **0,5 M EDTA** | 20 µl |
| **1 M DTT** | 10 µl |
| **NaCl** | 5.8 mg |
| **20% Triton X-100** | 50 µl |

| 2x Gibson reaction mix | Makes 250 uL |
|---|---|
| **5X Isothermal reaction buffer** | 100 µl |
| **1 U/µl T5 Exonuclease (Epicentre # T5E4111K, diluted 1:10 in storage buffer)** | 2 µl |
| **2U/µl Phusion polymerase (NEB # M0530S)** | 6.25 µl |
| **40U/µl Taq DNA Ligase (NEB # M0208S)** | 50 µl |
| **MQ** | 91.75 µl |

incubated for 2 hours at room temperature or overnight at 4°C, the beads were agitated by placing the container on a roller. Econo columns (BioRad) were washed with 10mL $ddH_2O$ and equilibrated with 10mL of PBS. The mixtures were loaded through the columns, which were subsequently washed 2 times with 30mL of PBS. Amicon Ultra Centrifugal Filter Units (10 kDa cut-off, Merck) were equilibrated with PBS, and 80µL of 2M Tris base (pH 9) was added to each unit. The antibodies were eluted directly into the spin concentrators using 5mL of 0.1 M glycine (pH 2.8). The antibodies where concentrated and buffer exchanged into PBS by multiple rounds of centrifugation at ~4,000rpm, prior to storage at 4°C. The concentration of the antibodies was examined using a nano-drop and the IgG quality was determined via sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (2.5.3) and enzyme-linked immunosorbent assay (ELISA; 2.6.1). The beads were reconstituted with 5mL of 0.1M glycine (pH 2.8), washed with 30mL of PBS and stored at 4°C in 10mL of PBS with $NaN_3$ (0.01%). The columns were washed with 20mL of $ddH_2O$, 5mL of 2M NaOH, 20mL of $ddH_2O$ and 5mL of 70% ethanol.

*Table 2.9: Cell lines and culture media.*

| Cell line | Origin | Source | Media | Description |
|---|---|---|---|---|
| **293F** | Human embryonic kidney cells | ThermoFisher Scientific | FreeStyle 293 Expression Medium | Derived from HEK293T cell line, suspension cell system optimised for serum-free recombinant protein expression |
| **HEK293T** | Human embryonic kidney cells | ATCC | DMEM with HI FBS (10 %) Penicillin (100 U/mL) Streptomycin (100 µg/mL) | Derived from HEK293 cell line, highly transfectable and contains the SV40 T-antigen |
| **HEK293T/17** | Human embryonic kidney cells | ATCC | DMEM with HI FBS (10 %) Penicillin (100 U/mL) Streptomycin (100 µg/mL) | Derived from HEK293T cell line, selected specifically for its high transfectability. |
| **TZM-bl** | Human adenocarcinoma cells | ATCC | DMEM with HI FBS (10 %) Penicillin (100 U/mL) Streptomycin (100 µg/mL) | Derived from HeLa cell line, stably expresses CD4, CCR5 and CXCR4, expresses luciferase and β-galactosidase genes under control of the HIV-1 promoter |
| **HeLa ACE2** | Human adenocarcinoma cells | A kind gift from James Voss | DMEM with HI FBS (10 %) Penicillin (100 U/mL) Streptomycin (100 µg/mL) | Derived from HeLa cell line, engineered to stably express the ACE2 receptor |
| **Vero E6** | African green monkey kidney cells | ATCC | DMEM with HI FBS (10 %) Penicillin (100 U/mL) Streptomycin (100 µg/mL) | Derived from Vero cell line, highly susceptible to SARS-CoV-2 infection |

### 2.5.2 Purification of viral proteins

An Äkta Pure System was used to perform a two-stage protein purification, which involved affinity purification and a size exclusion step. 2M imidazole (pH 7.4) was added to supernatants, collected from large scale transient transfections of viral proteins (2.3.2), to a final concentration of 5mM (pH 7). HiTrap IMAC Sepharose FF 5mL columns (GE Life Science) were washed with 5 column volumes (CV) of $ddH_2O$ and charged with 5 CV of 0.1M $NiSO_4$. The columns were then washed with 5 CV of $ddH_2O$, 5 CV of wash buffer (PBS with 5mM imidazole), 5 CV of elution buffer (PBS with 500mM imidazole) and 5 CV of wash buffer. The supernatants were loaded onto the columns (at 1-5mL per minute) and the columns were washed with 5 CV of wash buffer. The proteins were eluted with a linear gradient of wash buffer to elution buffer, over a span of 5 CV. The fractions containing protein (determined by UV trace) were collected, concentrated and buffer exchanged into PBS using spin concentrators, prior to storage at 4°C. The nickel was stripped from the columns using 5 CV of 50mM EDTA (pH 7.5). The columns were washed with 5 CV of $ddH_2O$, 5 CV of 2M NaOH, 5 CV of $ddH_2O$ and 5 CV of 20% ethanol.

A HiLoad 16/600 Superdex 200 pg column (GE Life Science) was washed with 1.5 CV of $ddH_2O$ and 1.5 CV of PBS. The 1mL loop was washed 5 times with PBS, and the affinity purified proteins were loaded onto the column. PBS was run through the column and fractions containing protein (determined by UV trace) were collected. The size and purity of the different peaks were examined using SDS-PAGE, and the correct protein size was estimated using the Compute PI/MW tool. Proteins of the desired size were concentrated and buffer exchanged into PBS using spin concentrators, prior to storage at 4°C. The column was washed with 1.5 CV of $ddH_2O$ and 1.5 CV of 20% ethanol. The concentration of the proteins was examined using a nano-drop and their quality was determined via SDS-PAGE (2.5.3) and ELISA (2.6.1).

### 2.5.3 SDS-PAGE

Fixed percentage gels were made by adding the desired resolving gel mixture to a 0.75mm glass plate (Mini-PROTEAN Short Plates, Bio Rad) held by a casting frame (BioRad), isopropanol was layered on top. The stacking gel mixture (Table 2.10) was added to the set resolving gel and a well-comb was put in place. Once set, the gel was either used immediately or stored at 4°C and used within 7 days. The gel was placed in a Mini-PROTEAN Tetra cell (Bio-Rad), submerged in running buffer and the comb was removed. 3µg of purified protein or 10µL of crude protein was mixed with 4x Laemmli buffer (BioRad) and added to the gel alongside Precision Plus Protein Dual Colour standard (BioRad). A power pack was used to electrify the gel at 100V for ~120mins to achieve an adequate separation of the protein bands. The gel was stained with SimplyBlue SafeStain (Invitrogen) for ~15-30mins and destained using $ddH_2O$. The gel was imaged with a fluorescence/chemiluminescence imager (Odyssey, Licor).

### 2.5.4 Generation of $F(ab')_2$

Purified monoclonal antibodies (2.5.1) were incubated with 4µg of immunoglobulin G-degrading enzyme of *Streptococcus pyogenes* (IdeS) per 1mg of IgG, in PBS for 1 hour at 4°C. The Fc of the antibodies and IdeS was removed using 250µL of Protein A Sepharose Fast Flow beads (GE Healthcare Life Sciences) per 1mg of IgG and 50µL of Ni Sepharose beads (GE Healthcare Life Sciences) per 1mg of IgG. The beads were washed twice with PBS before being added to the reaction. After 10 minutes of incubation at room temperature, the beads were removed from the reaction using a Spin-X tube filter (Costar). The filtrate containing the $F(ab')_2$ was concentrated and buffer exchanged into PBS using a spin concentrator, prior to storage at 4°C. The concentration of the $F(ab')_2$ was examined using a nano-drop and quality was determined via SDS-PAGE (2.5.3) and ELISA (2.6.1).

## 2.6 Immunohistochemical work

### 2.6.1 ELISA

96-well half area plates (Corning) were coated with 25µL of viral protein (such as S, N and gp120) or goat anti-human Fc IgG antibody at a concentration of 3µg/mL for 2 hours at 37°C or 12-18 hours at 4°C. The plates were washed 5 times using PBS with 0.05% Tween-20 (Sigma Aldrich). The plates were blocked with 5% skimmed milk (Sigma Aldrich) in PBS with 0.05% Tween-20 for 1 hour at room temperature. The wells were emptied and 25µL of monoclonal antibody or plasma/serum serial dilutions, diluted in blocking solution, were added and incubated for 2 hours at room temperature. The plates were washed 5 times using PBS with 0.05% Tween-20. Alkaline-phosphatase conjugated with goat anti-human IgG Fc antibody was diluted 1:1000 in blocking solution and 25µL was added for 1 hour at room temperature (Table 3.1). The plates were washed 5 times using PBS with 0.05% Tween-20. The plates were developed by adding 25µL of alkaline phosphatase substrate diluted in alkaline phosphatase staining buffer (9.8 pH). Optical density readings were taken at 5 minute intervals with a spectrophotometer at 405nm.

ELISA was used to confirm complete biotinylation had been achieved with regards to the in vivo biotinylated proteins. The amount of protein needed to coat an ELISA plate was incubated with 50µL of Pierce Avidin Agarose beads (ThermoFisher Scientific) for 1 hour at room temperature. The beads were washed twice with PBS before being added to the reaction. The beads were removed from the reaction using a Spin-X tube filter and the filtration was used to coat the ELISA plate. The same ELISA procedure was then carried out and the level of biotinylation could be determine by comparing the results from the depleted plate with a non-depleted plate.

### 2.6.2 Competition ELISA

96-well half area plates were coated with 25µL of S protein at a concentration of 3µg/mL for 2 hours at 37°C or 12-18 hours at 4°C. The plates were washed 5 times using PBS with 0.05% Tween-20. The wells were emptied and 25µL of $F(ab')_2$ serial dilutions in blocking solution, starting at 100-fold molar excess of their S binding $EC_{80}$, were added and incubated for 1 hour at room temperature. The plates were washed 5 times using PBS with 0.05% Tween-20. 25µL of the competing IgGs, diluted in blocking solution, were added at their S binding $EC_{80}$ and incubated at room temperature for 1 hour. The plates were washed 5 times using PBS with 0.05% Tween-20. Alkaline-phosphatase conjugated with goat anti-human IgG Fc antibody was diluted 1:1000 in blocking solution and 25µL was added for 1 hour at room temperature. The plates were washed 5 times using PBS with 0.05% Tween-20. The plates were developed by adding 25µL of alkaline phosphatase substrate diluted in alkaline phosphatase staining buffer (9.8 pH).

To reverse the competition direction the same procedure was performed, however IgGs at their S binding $IC_{80}$ were added first, followed by $F(ab')_2$ at 100-molar excess of their S binding $IC_{80}$. The secondary used in this instance was alkaline-phosphatase conjugated with goat anti-human IgG $F(ab')_2$

*Table 2.10: SDS-PAGE gel recipe.*

| | 4% | 8% | 10% | 12% | 20% | Stacking |
|---|---|---|---|---|---|---|
| *H20 [mL]* | 6.5 | 5.5 | 5 | 4.5 | 2.5 | 3.75 |
| *Tris pH 8.8 (1.5M) [mL]* | 2.5 | 2.5 | 2.5 | 2.5 | 2.5 | - |
| *Tris pH 6.8 (0.5) [mL]* | - | - | - | - | - | 1.25 |
| *Acrylamide (40%) [mL]* | 1 | 2 | 2.5 | 3 | 5 | 0.47 |
| *TEMED [µL]* | 10 | 10 | 10 | 10 | 10 | 5 |
| *SDS (10%) [µL]* | 100 | 100 | 100 | 100 | 100 | 50 |
| *APS (10%) [µL]* | 50 | 50 | 50 | 50 | 50 | 25 |

antibody. The percentage competition was calculated as the reduction in IgG binding in the presence of $F(ab')_2$ as a percentage of the maximum IgG binding in the absence of $F(ab')_2$, or vice versa.

### 2.6.3 Intracellular N staining

Following fixation as part of the neutralisation assay procedure (2.4.3), the cells were washed once with PBS and permeabilised with 100µl of 0.1% Triton-X (Sigma Aldrich) in PBS at room temperature for 15 minutes. The cells were washed 2 times with PBS and blocked using 3% skimmed milk in PBS at room temperature for 15 minutes. The blocking solution was discarded and 50µl of murinised-CR3009 at 2µg/mL, diluted using 1% skimmed milk in PBS, was added at room temperature for 45 minutes. The cells were washed twice with PBS and 100µl of horseradish peroxidase conjugated with horse anti-mouse IgG (Cell Signaling Technology) was added at a 1:2000 dilution in 1% skimmed milk in PBS at room temperature for 45 minutes. The cells were washed twice with PBS, developed using TMB substrate (Thermo Scientific) for 30 minutes and quenched using 2M $H_2SO_4$. Optical density readings were taken with a spectrophotometer at 450nm.

### 2.6.4 Western blot

Following separation by SDS-PAGE (2.5.3), the proteins were transferred onto nitrocellulose membranes (Amersham Protran 0.45) using a Mini-PROTEAN Tetra Cell in transfer buffer at 16V for 12-18 hours on ice. The membranes were blocked with 5% skimmed milk (Sigma Aldrich) in PBS with 0.05% Tween-20 for 1 hour at room temperature. The blocking solution was discarded and primary antibodies, diluted in blocking solution, were added and incubated at room temperature for 2 hours. The membranes were washed 5 times using PBS with 0.05% Tween-20. The secondary antibodies, diluted in blocking solution, were added and incubated at room temperature for 1 hour (Table 2.11). The membranes were washed 5 times using PBS with 0.05% Tween-20. The blots were imaged with a fluorescence/chemiluminescence imager.

### 2.6.5 Flow cytometry

PBMCs were thawed in a 37°C water bath and transferred into falcon tubes containing 10mL of Roswell Park Memorial Institute (RPMI) medium supplemented with 10% HI FBS, 1% penicillin/streptomycin and DNaseI (104 U/mL final concentration). The cells were centrifuged at 1500rpm for 5 minutes at 4°C and the supernatant was discarded. The cells were resuspended in 15mL of PBS supplemented with 2% HI FBS and 2mM EDTA, known as fluorescence-activated cell sorting (FACS) buffer. The cells were counted using a haemocytometer and diluted or separated accordingly.

*Table 2.11: List of antibodies used for ELISA and western blot.*

| Target species | Antibody | Manufacturer | Host species | Working dilution |
|---|---|---|---|---|
| **Human** | IgG-Fc | ThermoFisher Scientific | Goat | 3µg/mL |
| **Human** | IgG-Fcγ-AP | Jackson ImmunoResearch | Goat | 1:1000 |
| **Human** | IgG, F(ab')2-AP | Invitrogen | Goat | 1:1000 |
| **Human** | IgM-HRP | Sigma | Goat | 1:1000 |
| **Human** | IgA-HRP | Sigma | Goat | 1:1000 |
| **Mouse** | IgG-HRP | Cell Signaling Technology | Horse | 1:2000 |
| **Biotin** | Streptavidin-AP | ThermoScientific | Conjugate | 1:10,000 |
| **Human** | IgG- IRDye 800CW | Licor | Goat | 1:1000 |
| **Biotin** | Streptavidin-PE-Cy7 | BD | Conjugate | 1:1000 |

The cells were centrifuged at 1500rpm for 5 minutes at 4°C and the supernatant was discarded. The cells were resuspended in 25µL of FACS buffer and 5µL of Fc block was added at 4°C for 15 minutes. The staining panel and antigen baits, diluted in FACS buffer, was added to a final volume of 100µL per sample and incubated at 4°C for 1 hour (Table 2.12). The cells were washed 2 times with 10mL of FACS buffer and centrifuged at 1500rpm for 5 minutes at 4°C. The supernatant was discarded and the cells were resuspended in 300µL of FACS buffer and strained into FACS tubes (Fisher Scientific). The samples were examined using either a FACS Melody cell sorter (BD Bioscience) or a FACSCanto II (BD Bioscience).

When adherent cell lines were being analysed using flow cytometry, cells were detached using PBS supplemented with 5mM EDTA. Cell lines were stained using the same procedure, but were not transferred to RPMI and Fc block was not used. When staining with biotinylated proteins, conjugation was carried out prior to the experiment. The biotinylated proteins (150nM final concentration in 100µL staining mix) were mixed in FACS buffer with a 4-fold molar excess of the desired streptavidin conjugated fluorochrome. The mixtures were incubated at 4°C for 1 hour.

## 2.7 mAbs isolation work

### 2.7.1 FACS

PBMCs were stained (2.6.5) and single IgG B-cells were sorted, using a FACS Melody cell sorter, into 96-well PCR plates containing 4µL of lysis mix (Table 2.13), prepared previously in a RNA/DNA-free hood. The plates were covered with a foil seal and centrifuged at 400g for 1 min at 4°C. The RNA was stored at -80°C for up to 6 months.

### 2.7.2 RT-PCR

All mixtures and reactions were prepared in a RNA/DNA-free hood. 7µL of reverse transcription polymerase chain reaction (RT-PCR) mix 1 (Table 2.14) was added to wells containing lysed B-cells. The plates were covered with a foil seal and centrifuged at 400g for 1 min at 4°C. The plates were incubated at 65°C in a preheated thermocycler for 5 minutes and 7µL of RT-PCR mix 2 was added. The plates were covered with a foil seal and centrifuged at 400g for 1 min at 4°C. The RT-PCR reaction was performed using a preheated thermocycler. The plates were covered with a plastic seal and centrifuged at 400g for 1 min at 4°C. The cDNA was stored at 80°C for up to 3 years.

### 2.7.3 Nested PCR

All mixtures and reactions were prepared in a DNA-free hood. 22.5µL of $Ig_H$, $Ig_L\kappa$ and $Ig_L\lambda$ PCR mix 1 (Table 2.15) was added to separate 96-well PCR plates, and 2.5µL of the cDNA was added to each plate. The plates were covered with a plastic seal and centrifuged at 400g for 1 min at 4°C. The PCR reaction was performed using a preheated thermocycler. 22.5µL of $Ig_H$, $Ig_L\kappa$ and $Ig_L\lambda$ PCR mix 2 (Table

*Table 2.12: List of antibodies used for flow cytometry.*

| Target species | Antibody | Manufacturer | Host species |
|---|---|---|---|
| **Human** | CD3-APC/Cy7 | Biolegend | Mouse |
| **Human** | CD8-APC-Cy7 | Biolegend | Mouse |
| **Human** | CD14-BV510 | Biolegend | Mouse |
| **Human** | CD19-PerCP-Cy5.5 | Biolegend | Mouse |
| **Human** | IgM-PE | Biolegend | Mouse |
| **Human** | IgD-Pacific Blue | Biolegend | Mouse |
| **Human** | IgG-PeCy7 | BD Biosciences | Mouse |
| **Human** | Streptavidin-APC | Thermofisher Scientific | Conjugate |
| **Human** | Srteptavidin-Alexa 488 | Thermofisher Scientific | Conjugate |
| **Human** | Streptavidin-PE | Thermofisher Scientific | Conjugate |

2.16) was added to separate 96-well PCR plates, and 2.5µL of the corresponding PCR1 product was added to each plate. The plates were covered with a plastic seal and centrifuged at 400g for 1 min at 4°C. The PCR reaction was performed using a preheated thermocycler. PCR2 products underwent agarose gel electrophoresis (2.2.1) to determine amplification success. Amplified products were purified using QIAquick PCR Purification kit (Qiagen).

The first PCR used a forward primer mixture specific for separate heavy and light leader regions and a reverse primer specific for the respective constant regions (Table 2.17). The second PCR used a forward primer mixture that annealed to the corresponding V genes and a reverse primer mixture specific for the respective J genes and constant regions.

### 2.7.4 Bulk ligation transfection

PCR products and linearised human expression vectors were diluted in nuclease-free water to final concentrations of 40ng/µL and 120ng/µL respectively [432]. 2µL of PCR product, 2µL of linearised plasmid and 6µL of Gibson reaction mix was added to a 96-well PCR plate. The plate was covered with a plastic seal and centrifuged at 400g for 1 min at 4°C. The Gibson assembly reaction was carried out (2.2.6) and the product was stored on ice.

Chemically competent *E. coli* was thawed on ice and 25µL was added to a round bottom 96-well plate. 4µL of the assembly product was added to the bacteria and incubated on ice for 10 minutes. A plastic seal was added to the plate and it was submerged in a 42°C water bath for 50 seconds. The plate was incubated on ice for 2 minutes and 175µL of LB broth without selection marker was added to the bacteria. The cultures were grown in a 37°C incubator for 1 hour, spread on LB agar plates with ampicillin, incubated at 30-37°C for 12-18 hours and stored at 4°C.

HEK293T/17 cells were pre-seeded in 12-well plates at a concentration of $5x10^5$ cells/mL and with a volume of 1mL per well. 50µL of OptiMEM, 6µL of assembled $Ig_H$ product and 6µL of assembled $Ig_L$ product was added to a round bottom 96-well plate. 3µL of PEI Max and 47µL of OptiMEM, per transfection reaction, were mixed in a separate round bottom 96-well plate. The two plates were combined and incubated at room temperature for 15 minutes, before being added dropwise to the HEK293T/17 cells. The supernatant was harvested 3 days later, sterilised through a 0.22µm filter and examined using ELISA (2.6.1).

### 2.7.5 mAb sequence analysis

Successfully binding $Ig_H$ and $Ig_L$ plasmid pairs were expanded (2.1.4) and sent for sequencing (2.2.7). The V(D)J arrangements of these mAbs was analysed using the IMGT database [433].

## 2.8 Data analysis work

### 2.8.1 Statistics

Statistics, alongside linear and non-linear regressions were performed using GraphPad Prism 9. The specifics of the statistics performed can be found in the figure legends.

*Table 0.13: Lysis mix recipe.*

| | Concentration | 1x well | 75x wells |
|---|---|---|---|
| *Nuclease-free water* | - | 3.2µL | 240 µL |
| *DPBS (no Ca/Mg)* | 10x | 0.2µL | 15 µL |
| *DL-DTT* | 100mM | 0.4µL | 30 µL |
| *RNaseOUT* | 40U/µL | 0.3µL | 22.5 µL |

### 2.8.2 Data visualisation

Graphs were generated using GraphPad Prism 9 or R studio. Figures and basic graphics were prepared using Adobe Illustrator 26.5. Structural images were made using pymol 2.5 and visualised with Mol* Viewer [434].

*Table 2.14: Recipes for RT-PCR mixes and thermocycler conditions.*

| | Concentration | 1x well | 75x wells |
|---|---|---|---|
| *Nuclease-free water* | - | 4.4µL | 330µL |
| *IGEPAL* | 10% | 0.5µL | 37.5µL |
| *Random hexamer primers* | 50ng/µL | 3µL | 225µL |
| *RNaseOUT* | 40U/µL | 0.2µL | 15µL |

| | Concentration | 1x well | 75x wells |
|---|---|---|---|
| *Nuclease-free water* | - | 2.5µL | 187.5µL |
| *First-strand buffer* | 5x | 3.5µL | 262.5µL |
| *DTT* | 100mM | 1.1µL | 82.5µL |
| *dNTP* | 25mM | 0.6µL | 45µL |
| *RNaseOUT* | 40U/µL | 0.1µL | 7.5µL |
| *SuperScript III RT* | 200U/µL | 0.3µL | 22.5µL |

| Step | Temperature | Time |
|---|---|---|
| **Step 1** | 42°C | 10 |
| **Step 2** | 26°C | 10 |
| **Step 3** | 50°C | 60 |
| **Step 4** | 94°C | 5 |
| **Hold** | 4°C | ∞ |

# Chapter 3: Neutralising antibody responses in a HIV-1 superinfected donor

## 3.1 Antibody responses to HIV-1 infection and mechanisms of escape

### 3.1.1 Introduction

One of the main goals of HIV-1 vaccine development research is to generate an antibody response that can neutralise a broad range of HIV-1 variants [435]. These bnAbs have been shown to arise in 10-30% of individuals chronically infected with HIV-1, typically after 2-3 years of infection [194,195]. While the presence of a broadly neutralising response does not usually lead to control of the disease, passive transfer of bnAbs in animal models has been shown to prevent infection [436–438]. The bnAbs target the Env glycoprotein, which is a trimer of non-covalently linked gp120-gp41 heterodimers [164]. The bnAbs target numerous conserved epitopes located across the HIV-1 Env. These include the CD4bs, V1/V2 apex, V3-glycan domain, FP/gp120-gp41 interface, silent face and MPER [200]. Notably, the circulating virus tends to be resistant to the bnAbs present, implying HIV-1 can escape such responses during chronic infection [201]. Co-evolution studies have shown that bnAbs are produced as a result of multiple rounds of viral escape and B-cell affinity maturation, with antibodies gradually targeting more conserved regions on Env [208,439]. However, bnAbs have not yet been elicited in response to immunisation with modern soluble recombinant Env trimers [440,441].

The SPARTAC trial began in 2003, and aimed to examine whether short-term ART during primary HIV-1 infection could increase the time before long-term ART was needed. The subjects were recruited into the trial within 6 months of HIV-1 infection. The groups were either given ART for 12 weeks, 48 weeks or not at all, with no ART being the standard of care at the time of the trial. Longitudinal sera and PBMC samples were taken until the CD4 levels of the subjects reduced to 350 cells/µl (after which ART was initiated), or the trial ended. Subjects were recruited from Australia, Brazil, Ireland, Italy, South Africa, Spain, Uganda, and the UK, therefore many HIV-1 clades were encompassed by the study [259]. Clinical guidelines have since adopted early ART following a HIV-1 diagnosis, making historical longitudinal clinical samples a valuable resource for bnAb development research [442,443].

Previous work by the Doores research group investigated the control arm of the SPARTAC cohort, identifying subjects with bnAb responses, the clinical factors associated with bnAb induction, the kinetics of bnAb development and the epitope specificity of these responses [444]. 50 subjects from the

*Table 2.1: Nested PCR1 recipe and thermocycler conditions.*

| | Concentration | 1x well | 75x wells |
|---|---|---|---|
| *Nuclease-free water* | - | 9µL | 675µL |
| *PCR Master Mix* | 2x | 12.5µL | 937.5µL |
| *Forward Primer* | 100µM | 0.0625µL | 4.7µL |
| *Reverse Primer* | 100µM | 0.0625µL | 4.7µL |

| Step | Temperature | Time |
|---|---|---|
| **Initial denaturation** | 94°C | 5 minutes |
| **45 cycles** | 94°C<br>58°C<br>72°C | 30 seconds<br>30 seconds<br>1 minute |
| **Final extension** | 72°C | 1 minute |
| **Hold** | 4°C | ∞ |

SPARTAC trial control arm were tested in neutralisation assays against a HIV-1 indicator panel, and each was assigned a neutralisation score based on their breadth and potency. This led to the identification of donor SJU027003 (SJU) with a bnAb response, and categorised this donor as one of the top ten neutralisers from the cohort. Epitope mapping showed that the bnAb response of SJU was targeted at the V3-glycan domain. Autologous viruses from SJU were isolated and sequenced from 13 different timepoints, using single genome amplification (SGA). This revealed that ~2 years after primary infection with a clade C virus, SJU had become superinfected with another clade C virus and showed increased neutralisation breadth after this event (Figure 3.1). Remarkably, aside from being of the same subtype, the superinfecting virus also encoded the N332 glycan [444].

Autologous neutralising antibody responses, occurring early in HIV-1 infection, are a driving factor of viral evolution and are responsible for moulding the viruses that generate bnAbs [445]. The *envs* isolated from SJU and longitudinal sera samples provide an excellent resource, enabling examination of the co-evolution between the autologous viruses and the neutralising antibodies. Through the use of a panel of previously isolated V3-glycan domain bnAbs, insights can be gained into how specific regions of the Env landscape developed and the mechanisms employed to escape the bnAb response.

Overall, the specific aims of this section were to:

- Confirm heterologous neutralisation using a 12-virus global panel.
- Analyse neutralising antibody responses to early primary and superinfecting viruses, as well as a recombinant strain.
- Examine the neutralising antibody response of sera and V3-glycan domain bnAbs to autologous viruses.
- Investigate the impact of glycan switching between N332 and N334 in autologous viruses on the neutralising antibody response.
- Assess the impact of autologous virus V1 loop length and disulfide bond usage on the neutralising antibody response.

*Table 2.2: Nested PCR2 recipe and thermocycler conditions.*

| | Concentration | 1x well | 75x wells |
|---|---|---|---|
| *Nuclease-free water* | - | 16µL | 1200µL |
| *GC Buffer* | 5x | 5µL | 375µL |
| *dNTPs* | 10mM | 0.5µL | 37.5µL |
| *Forward Primer* | 100µM | 0.0625µL | 4.7µL |
| *Reverse Primer* | 100µM | 0.0625µL | 4.7µL |
| *DMSO* | - | 0.25µL | 18.75µL |
| *Phusion* | 2U/µL | 0.25µL | 18.75µL |

| Step | Temperature | Time |
|---|---|---|
| **Initial denaturation** | 94°C | 5 minutes |
| **45 cycles** | 94°C<br>54°C (IgH) or 62°C (IgL)<br>72°C | 30 seconds<br>30 seconds<br>1 minute |
| **Final extension** | 72°C | 1 minute |
| **Hold** | 4°C | ∞ |

### 3.1.2 Human subjects and ethics

Peripheral blood mononuclear cells (PBMCs) were obtained from the King's Biobank. All donors signed a written informed consent and ethical approval for usage of these samples was granted by The London Multicentre Research Ethics Committee (MREC 04/2/025).

### 3.1.3 Heterologous neutralisation

Previous work examined the heterologous neutralisation capability of donor SJU, by testing longitudinal sera samples in neutralisation assays against a 6-virus cross-clade indicator panel (Figure 3.2a) [446]. This showed that heterologous neutralisation activity was first detectable at ~50 weeks post-infection. Between the start of heterologous neutralisation and the onset of superinfection, SJU could neutralise 50% of the cross-clade indicator panel. Following superinfection, this donor showed improved breadth and could neutralise ~83% of the viruses tested [444].

*Table 2.3: Nested PCR primer mixes.*

| | | |
|---|---|---|
| 5' L-VH 1 | ACAGGTGCCCACTCCCAGGTGCAG | IgGHV forward primer mix 1 |
| 5' L-VH 3 | AAGGTGTCCAGTGTGARGTGCAG | |
| 5' L-VH 4/6 | CCCAGATGGGTCCTGTCCCAGGTGCAG | |
| 5' L-VH 5 | CAAGGAGTCTGTTCCGAGGTGCAG | |
| 3' CH1 | GGAAGGTGTGCACGCCGCTGGTC | IgGHV reverse primer mix 1 |
| 5' LVL 1 | GGTCCTGGGCCCAGTCTGTGCTG | IgGLVλ forward primer mix 1 |
| 5' LVL 2 | GGTCCTGGGCCCAGTCTGCCCTG | |
| 5' LVL 3 | GCTCTGTGACCTCCTATGAGCTG | |
| 5' LVL 4/5 | GGTCTCTCTCSCAGCYTGTGCTG | |
| 5' LVL 6 | GTTCTTGGGCCAATTTTATGCTG | |
| 5' LVL 7 | GGTCCAATTCYCAGGCTGTGGTG | |
| 5' LVL 8 | GAGTGGATTCTCAGACTGTGGTG | |
| 3' CL | CACCAGTGTGGCCTTGTTGGCTTG | IgGLVλ reverse primer mix 1 |
| 5' LVK 1/2 | ATGAGGSTCCCYGCTCAGCTGCTGG | IgGLVκ forward primer mix 1 |
| 5' LVK 3 | CTCTTCCTCCTGCTACTCTGGCTCCCAG | |
| 5' LVK 4 | ATTTCTCTGTTGCTCTGGATCTCTG | |
| 3' CK 543 | GTTTCTCGTAGTCTGCTTTGCTCA | IgGLVκ reverse primer mix 1 |

| | | |
|---|---|---|
| AgeI_VH1 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCCCAGGTGCAGCTGGTGCAG | IgGHV forward primer mix 2 |
| AgeI_VH1/5 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCCGAGGTGCAGCTGGTGCAG | |
| AgeI_VH3 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCTGAGGTGCAGCTGGTGGAG | |
| AgeI_VH323 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCTGAGGTGCAGCTGTTGGAG | |
| AgeI_VH4 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCCCAGGTGCAGCTGCAGGAG | |
| AgeI_VH434 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCCCAGGTGCAGCTACAGCAGTG | |
| AgeI_VH118 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCCCAGGTTCAGCTGGTGCAG | |
| AgeI_VH124 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCCCAGGTCCAGCTGGTACAG | |
| AgeI_VH333 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCTCAGGTGCAGCTGGTGGAG | |
| AgeI_VH39 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCTGAAGTGCAGCTGGTGGAG | |
| AgeI_VH439 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCCCAGCTGCAGCTGCAGGAG | |
| AgeI_VH61 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCCCAGGTACAGCTGCAGCAG | |
| AgeI_VK15 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCTGACATCCAGATGACCCAGTC | IgGHLκ forward primer mix 2 |
| AgeI_VK19 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCAGACATCCAGTTGACCCAGTCT | |
| AgeI_VK1D43 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTGTGCCATCCGGATGACCCAGTC | |
| AgeI_VK224 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATGGGGATATTGTGATGACCCAGAC | |
| AgeI_VK228 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATGGGGATATTGTGATGACTCAGTC | |
| AgeI_VK230 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATGGGGATGTTGTGATGACTCAGTC | |
| AgeI_VK311 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCAGAAATTGTGTTGACACAGTC | |
| gel_VK315 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCAGAAATAGTGATGACGCAGTC | |
| AgeI_VK320 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCAGAAATTGTGTTGACGCAGTCT | |
| Age1_VK41 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTGTACATTCGGACATCGTGATGACCCAGTC | |
| AgeI_VL1 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTTCCTGGGCCCAGTCTGTGCTGACKCAG | IgGHLλ forward primer mix 2 |
| AgeI_VL2 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTTCCTGGGCCCAGTCTGCCCTGACTCAG | |
| AgeI_VL3 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTTCTGTGACCTCCTATGAGCTGACWCAG | |
| AgeI_VL4/5 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTTCTCTCTCSCAGCYTGTGCTGACTCA | |
| AgeI_VL6 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTTCTTGGGCCAATTTTATGCTGACTCAG | |
| AgeI_VL7/8 | CATCCTTTTTCTAGTAGCAACTGCAACCGGTTCCAATTCYCAGRCTGTGGTGACYCAG | |
| 3SalI.JH1/4/5 | GGAAGACCGATGGGCCCTTGGTCGACGCTGAGGAGACGGTGACCAG | IgGHV reverse primer mix 2 |
| 3SalI.JH2 | GGAAGACCGATGGGCCCTTGGTCGACGCTGAGGAGATGGTGATTGGG | |
| 3SalI.JH3 | GGAAGACCGATGGGCCCTTGGTCGACGCTGAAGAGAC GGTGACCMK | |
| 3SalI.JH6 | GGAAGACCGATGGGCCCTTGGTCGACGCTGAGGAGAC GGTGACGACG | |
| 3BsiWI_JK1/4 | GAAGACAGATGGTGCAGCCACCGTACGTTTGATYTCCA CCTTGGTC | IgGHLκ reverse primer mix 2 |
| 3BsiWI_JK2 | GAAGACAGATGGTGCAGCCACCGTACGTTTGATCTCCAGCTTGGTC | |
| 3BsiWI_JK3 | GAAGACAGATGGTGCAGCCACCGTACGTTTGATATCCA CTTTGGTC | |
| 3BsiWI_JK5 | GAAGACAGATGGTGCAGCCACCGTACGTTTAATCTCCA GTCGTGTC | |
| 3XhoI_CL | GTTGGCTTGAAGCTCCTCACTCGAGGGYGGGAACAGA GTG | IgGHLλ reverse primer mix 2 |
| 3XhoI_CL | GTTGGCTTGAAGCTCCTCACTCGAGGGYGGGAACAGA GTG | |

A 12-virus global panel has since been developed, and has been shown to be adequately representative of a larger 219-virus panel [447]. To build upon earlier data, sequential sera samples from SJU were tested in neutralisation assays against this global panel (Figure 3.2b). This assay used HIV-1

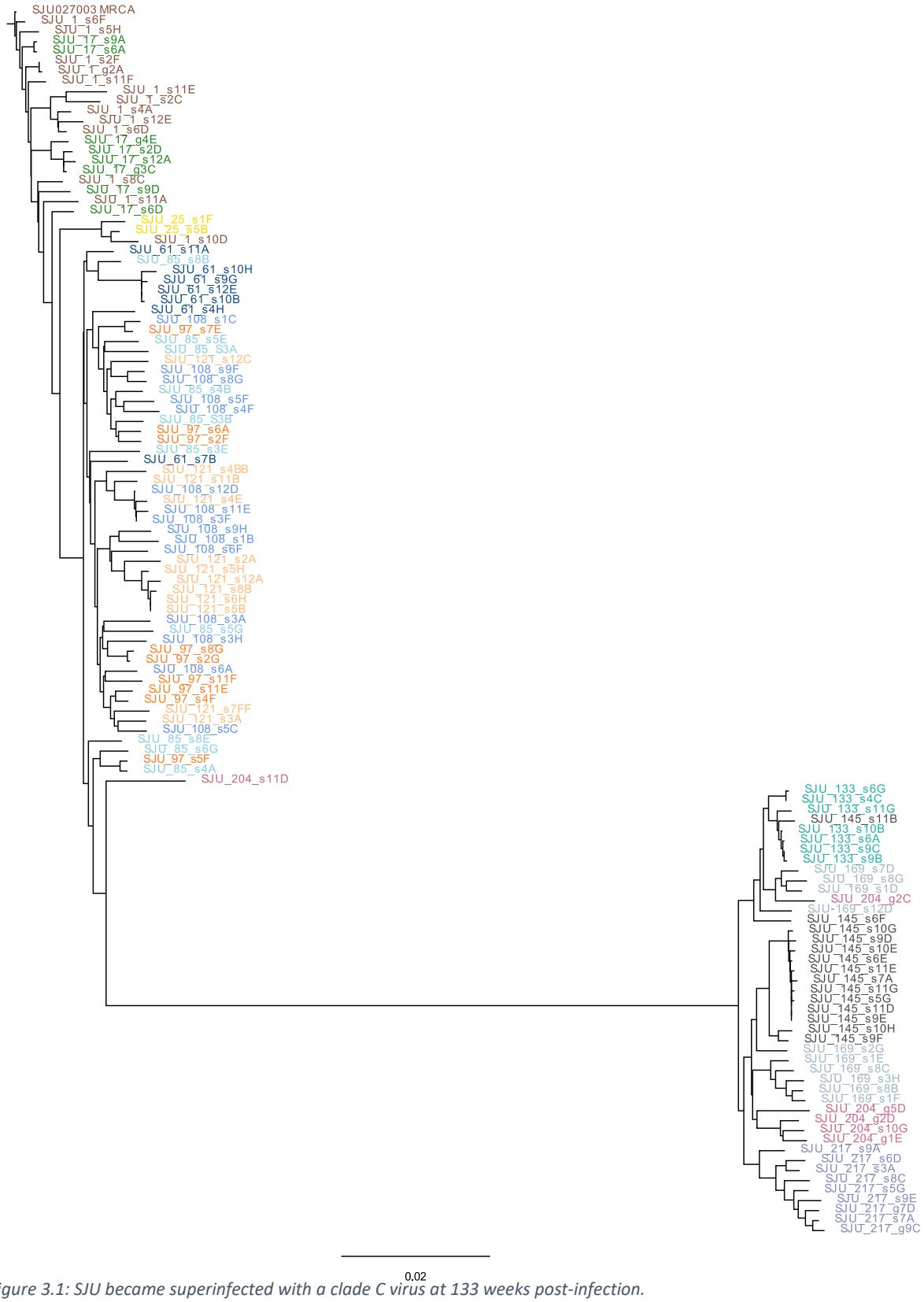



*Figure 3.1: SJU became superinfected with a clade C virus at 133 weeks post-infection.*
*A rooted uncorrected pairwise distances phylogenetic tree of autologous viruses isolated donor SJU between 1-217 weeks post-infection. Adapted from data by Luke Granger.*

based viral particles (PSG-3Δenv) pseudotyped with full-length Env against TZM-bl cells, which stably express CD4, CXCR4 and CCR5, and contain a Tat-responsive Luc reporter gene [448,449]. Neutralisation potency was quantified by the sera inhibitory dilution at which 50% viral neutralisation was attained, commonly referred to as the $ID_{50}$.

Between the start of heterologous neutralisation and the onset of superinfection, SJU could neutralise ~42% of the global panel viruses. Similar to what was seen with the cross-clade indicator panel, after superinfection this donor displayed increased neutralisation breadth and could neutralise ~58% of the viruses tested. The potency of the sera against each virus tended to increase with time to a certain level and then become relatively stable. However, there was a noticeable drop in neutralisation potency at week 121 against 398F1 (clade A). Overall, the neutralisation breadth was lower against the global panel compared to the cross-clade indicator panel. This was expected as the global panel incorporated more subtypes such as clade G, CRF07_BC and AC recomb. Nonetheless, this work confirmed that SJU displayed good neutralisation breadth and potency against a range of HIV-1 strains.

### 3.1.4 Autologous neutralisation

Previous work had isolated and sequenced autologous viruses from SJU via SGA. Analysis of the *env* sequences using DIVEIN enabled the identification of the most recent common ancestors of the primary infecting virus (1_s5H) and the superinfecting virus (133_s9C), as well as one recombinant virus (204_s11D) [450]. To investigate how long it took for neutralising antibodies to develop against each divergent strain, these viruses were tested in neutralisation assays using longitudinal sera samples from SJU (Figure 3.3a). Neutralisation activity was detectable against 1_s5H and 133_s9C weeks after their emergence, specifically from weeks 17 and 145 respectively. This indicated that 133_s9C was not susceptible to sera collected at the time of superinfection, or beforehand. Interestingly, 204_s11D was neutralised by sera taken prior to its emergence, starting from week 168. This may have occurred as a result of the recombination incorporating a susceptible epitope from a previously circulating virus. For all three viruses, neutralisation capability was maintained until the final samples with only 133_s9C showing a gradual reduction in potency over time.

To explore how the sera neutralisation developed in response to chronic infection, sera samples were then examined via neutralisation assay against a wider range of autologous viruses (Table 3.1). It was found that as the infection progressed, the sera became capable of neutralising more autologous variants. For example, serum from week 168 could neutralise all 22 viruses it was tested against. Conversely, with the exception of 108_s5C, serum from week 61 could not neutralise any of the 14 viruses it was tested against. Remarkably, between weeks 85-121 post-infection, SJU was able to

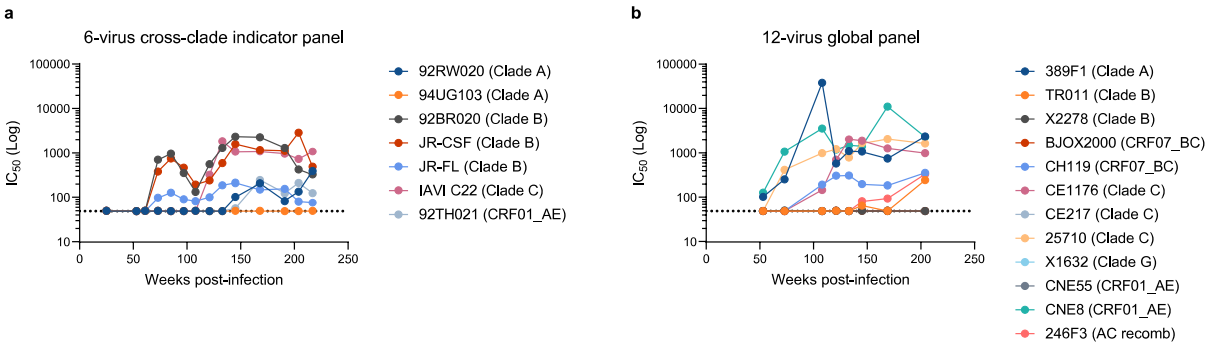



*Figure 3.2: Neutralisation of heterologous viruses by SJU sera.*
*Line graphs showing the longitudinal neutralisation potency of SJU sera over time, against a 6-virus cross-clade indicator panel, adapted from data by Luke Granger (**a**) and a 12-virus global panel (**b**). The black dotted lines indicate the cut-off serum concentration used in the neutralisation assays (1:50).*

contemporaneously neutralise numerous viral variants. The resistant viruses from week 85 timepoints failed to assume dominance and incomplete viral escape persisted until superinfection occurred. After superinfection the contemporaneous neutralisation ceased, but by week 145 neutralising activity could be detected against the superinfecting viruses.

The period of contemporaneous neutralisation between weeks 85-121 could be facilitated by either bnAbs, autologous neutralising antibodies or a combination of both. To ascertain their susceptibility at the V3-glycan domain, the autologous viruses were tested in neutralisation assays against a panel of HIV-1 bnAbs (Figure 3.3b). This panel was composed of bnAbs specific for the V3-glycan domain, such as PGT-121, PGT-128, PGT-130 and PGT-135, but each adopted a different angle of approach [451–453]. Notably, PGT-130 is a somatic variant of PGT-128 that shows more tolerance to N334 glycans [56]. PG9, a bnAb that targets the V1/V2 apex, was used as a positive control [454]. All viruses tested were sensitive to neutralisation by at least one member of the V3-glycan domain bnAb panel, except for 121_s4BB. This indicated that the V3-glycan domain was largely accessible in the strains circulating throughout this period. As 121_s4BB was neutralised by the contemporaneous serum, this indicated that autologous neutralising antibodies were playing a role to some extent. The strains resistant to sera neutralisation showed susceptibility to PGT-128 and the majority were also neutralised by PGT-

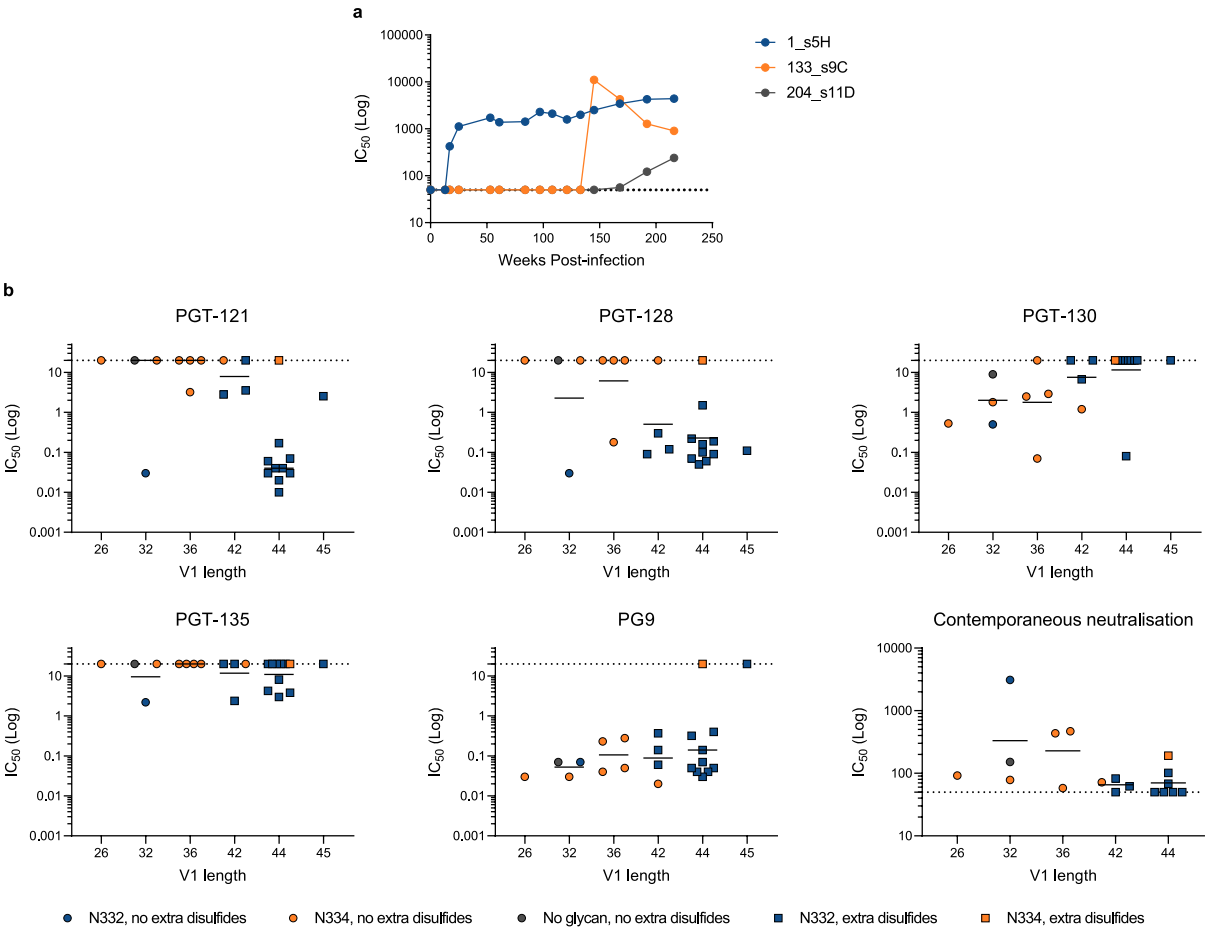



*Figure 3.3: Neutralisation of autologous viruses by SJU sera and a bnAb panel.*
*(**a**) Line graph showing the longitudinal neutralisation potency of SJU sera over time, against the most recent common ancestors of the primary infecting virus (1_s5H) and the superinfecting virus (133_s9C), as well as one recombinant virus (204_s11D). (**b**) Neutralisation potency of a panel of bnAbs against autologous viruses isolated between 85-121 weeks post-infection. V3-glycan domain bnAbs PGT-121, PGT-128, PGT-130 and PGT-135 were employed, alongside PG9 (V1/V2 apex) as a positive control. Important features of the Envs are displayed, including V1 loop length, presence of extra V1 loop disulfide bonds and V3 loop glycan usage. The black dotted lines indicate the cut-off antibody concentration used in the neutralisation assays (20µg/mL).*

121. Therefore, if V3-glycan domain bnAbs generated by SJU were responsible for the contemporaneous neutralisation witnessed during this period, they likely had a different mode of neutralisation to PGT-121 and PGT-128. The neutralisation profiles of PGT-130 and PGT-135 matched that of the sera for some variants but differed for others. Collectively, these findings suggested that V3-glycan domain bnAbs could have been facilitating contemporaneous neutralisation in this donor, however until such antibodies are isolated a definitive answer cannot be provided (see 3.3). It remains likely that the overall response was the result of strain specific antibodies and bnAbs acting synergistically.

### 3.1.5 N332 to N334 glycan shifting

The viral population was very heterogeneous during the period of contemporaneous neutralisation. For instance, the N332 glycan was shifted to position 334 in a large subset of viruses. To determine whether this characteristic impacted serum neutralisation several N-glycan shifted mutants were constructed using site directed mutagenesis. When the N-glycan was shifted from position 334 to 332, a mild increase in neutralisation was seen for the majority of variants (Figure 3.4a). However, 121_s11B demonstrated a large increase in potency following this mutation. When the glycan was shifted from position 332 to 334 a mild reduction was seen for 121_s2A, whilst mild increases in

*Table 3.4: Neutralisation of autologous viruses by longitudinal SJU serum samples.*

| Virus | ID50 (weeks post-infection) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 53 | 61 | 85 | 97 | 108 | 121 | 133 | 145 | 168 |
| 85_s3E | <50 | <50 | 403 | 331 | 893 | - | - | - | - |
| 85_s4A | <50 | <50 | <50 | <50 | 124 | - | - | - | - |
| 85_s4B | <50 | <50 | <50 | 123 | 157 | - | - | - | - |
| 97_s2F | - | <50 | <50 | <50 | 268 | 507 | - | - | - |
| 97_s2G | - | <50 | <50 | <50 | 188 | 180 | - | - | - |
| 97_s4F | - | <50 | 96 | 61 | 102 | 207 | - | - | - |
| 108_s1C | - | <50 | <50 | - | 58 | 70 | - | - | 499 |
| 108_s3F | - | <50 | 328 | - | 468 | 776 | - | - | 2099 |
| 108_s5C | - | 100 | 853 | - | 3097 | 1652 | - | - | 1954 |
| 108_s5F | - | <50 | <50 | - | <50 | 54 | - | - | 601 |
| 108_s6A | - | <50 | <50 | - | 151 | 121 | - | - | 498 |
| 108_s8G | - | <50 | <50 | - | 83 | 195 | - | - | 841 |
| 108_s9F | - | <50 | <50 | - | <50 | 79 | - | - | 390 |
| 108_s11E | - | <50 | 250 | - | 436 | 1014 | - | - | 1600 |
| 121_s2A | - | - | <50 | - | <50 | <50 | 106 | - | 594 |
| 121_s3A | - | - | <50 | - | <50 | 72 | 107 | - | 338 |
| 121_s4BB | - | - | 80 | - | 104 | 191 | 191 | - | 331 |
| 121_s7FF | - | - | <50 | - | <50 | 78 | 244 | - | 839 |
| 121_s11B | - | - | <50 | - | 60 | 92 | 195 | - | 931 |
| 121_s12C | - | - | <50 | - | <50 | <50 | 71 | - | 646 |
| 133_s4C | - | - | - | - | | <50 | <50 | 798 | 441 |
| 133_s6C | - | - | - | - | | <50 | <50 | 156 | 170 |
| 133_s6G | - | - | - | - | | <50 | <50 | 2291 | 1026 |
| 133_s9B | - | - | - | - | | <50 | <50 | 118 | 175 |
| 133_s9C | - | - | - | - | | <50 | <50 | <50 | 73 |
| 133_s10B | - | - | - | - | | <50 | <50 | 180 | 219 |
| 133_s11G | - | - | - | - | | <50 | <50 | 979 | 607 |

neutralisation potency were seen for the rest of the viruses tested (Figure 3.4b). These findings showed that for some viruses, shifting the N332 glycan to position 334 was a mechanism of escape. However, in most cases glycan shifting was insufficient for complete escape, and in some viruses it increased susceptibility. The incomplete viral escape could have occurred for several reasons. It is possible that the SJU bnAbs exhibited glycan promiscuity, a helper lineage specific for the N334 glycan was present or autologous neutralising antibodies were responsible for the neutralisation of N334 bearing virus.

### 3.1.6 V1 loop interference

The V1 length also varied greatly during the period of contemporaneous neutralisation and some of the viruses with longer V1 loops also harboured an uncommon extra disulfide bond (Figures 3.5a-b). Mutants that removed this additional disulfide bond were produced to determine their impact on serum and V3-glycan domain bnAb neutralisation. Employing the same bnAb panel as before revealed that removal of the extra disulfide bond resulted in a large increase in neutralisation activity for all of the antibodies except PGT-128, which neutralised both Env conformations potently (Figure 3.5c). Furthermore, 108_s9F became more susceptible to sera after the removal of the extra disulfide bond, seen in samples collected as early as 61 weeks post-infection (Figure 3.5d). This indicated that the introduction of an extra disulfide bond in the V1 loop of 108_s9F enabled escape from neutralisation by donor sera and V3-glycan domain bnAbs.

It has been shown previously that longer V1 loops (with more PNGS) can decrease the neutralisation potency of V3-glycan domain bnAbs [455–457]. A range of V1 lengths were observed in the autologous viruses isolated from donor SJU (Figure 3.5a). Between weeks 1-17 post-infection, the V1 loop lengths were very long (>45 amino acids). Between weeks 25-121 post-infection, the V1 loop lengths became bimodal with longer and shorter populations coexisting. After superinfection, the V1 loop lengths became very short (>20 amino acids) and remain this way until the final timepoint sampled. Four autologous viruses bearing the N332 glycan were selected from these periods to investigate the impact of V1 length on V3-glycan domain mediated neutralisation. The autologous viruses 1_s5H, 108_s5C, 108_s9F and 133_s4C were selected to represent the early long V1, middle short V1, middle long V1 and later short V1 periods respectively. Of note, only 108_s9F incorporated the aforementioned extra disulfide bond in the V1 loop. The V3-glycan domain bnAb panel were able to potently neutralise 1_s5H and 108_s5C. However, only PGT-128 and PGT-130 were able to neutralise 108_s9F and 133_s4C.

To investigate this further, a series of chimeric *envs* were developed using overlap PCR and tested in neutralisation assays against the V3-glycan domain bnAb panel (Table 3.2). When the long V1 of 1_s5H (46 amino acids) was replaced with the short V1 of 108_s5C (32 amino acids), there were minimal

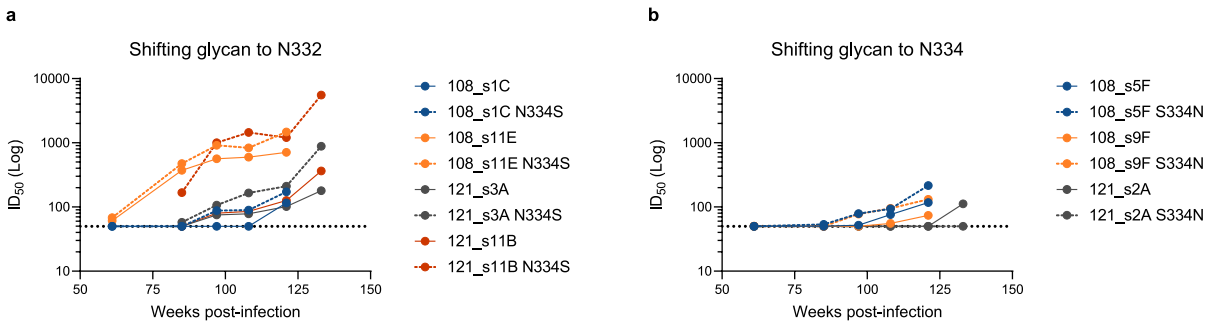



*Figure 3.4: Impact of glycan shifting in autologous viruses on neutralisation by SJU sera.*
*Line graphs showing the impact on SJU plasma neutralisation caused by shifting the N334 glycan to position 332 (**a**) and vice versa (**b**), in autologous viruses. The black dotted lines indicate the cut-off serum concentration used in the neutralisation assays (1:50).*

changes seen in the neutralisation potency of the bnAb panel. Unfortunately, when the V1 of 1_s5H was swapped with the other representative viruses, the products generated gave non-infectious pseudovirions. When the short V1 of 108_s5C (32 amino acids) was replaced with the shorter V1 of 133_s4C (23 amino acids), only mild changes in neutralising activity were seen for PGT-130 and PGT-135. When the long V1 of 108_s9F (42 amino acids) was replaced with either the short V1 of 108_s5C (32 amino acids) or the shorter V1 of 133_s4C (23 amino acids), a considerable increase in neutralisation was observed for most of the bnAbs tested. Unfortunately, when the V1 of 108_s9F was swapped with that of 1_s5H, the product generated gave non-infectious pseudovirions. Finally,

when the short V1 of 133_s4C (23 amino acids) was swapped with the long V1 of 1_s5H (46 amino acids), a reduction in neutralisation potency was found for PGT-121, but the rest of the bnAbs remain largely unchanged.

Overall, these results show that although long V1 loops can interfere with V3-glycan domain specific neutralising antibodies, the conformation of the V1 needs to also be considered. For instance, the V1 of 1_s5H is longer than 108_s9F but does not impact access to the V3-glycan domain. It appears that 108_s9F developed an extra disulfide bond in its V1 loop that enabled the obstruction of this epitope. Further, shorter V1 loops in association with the presence of the N332 glycan does not ensure V3-

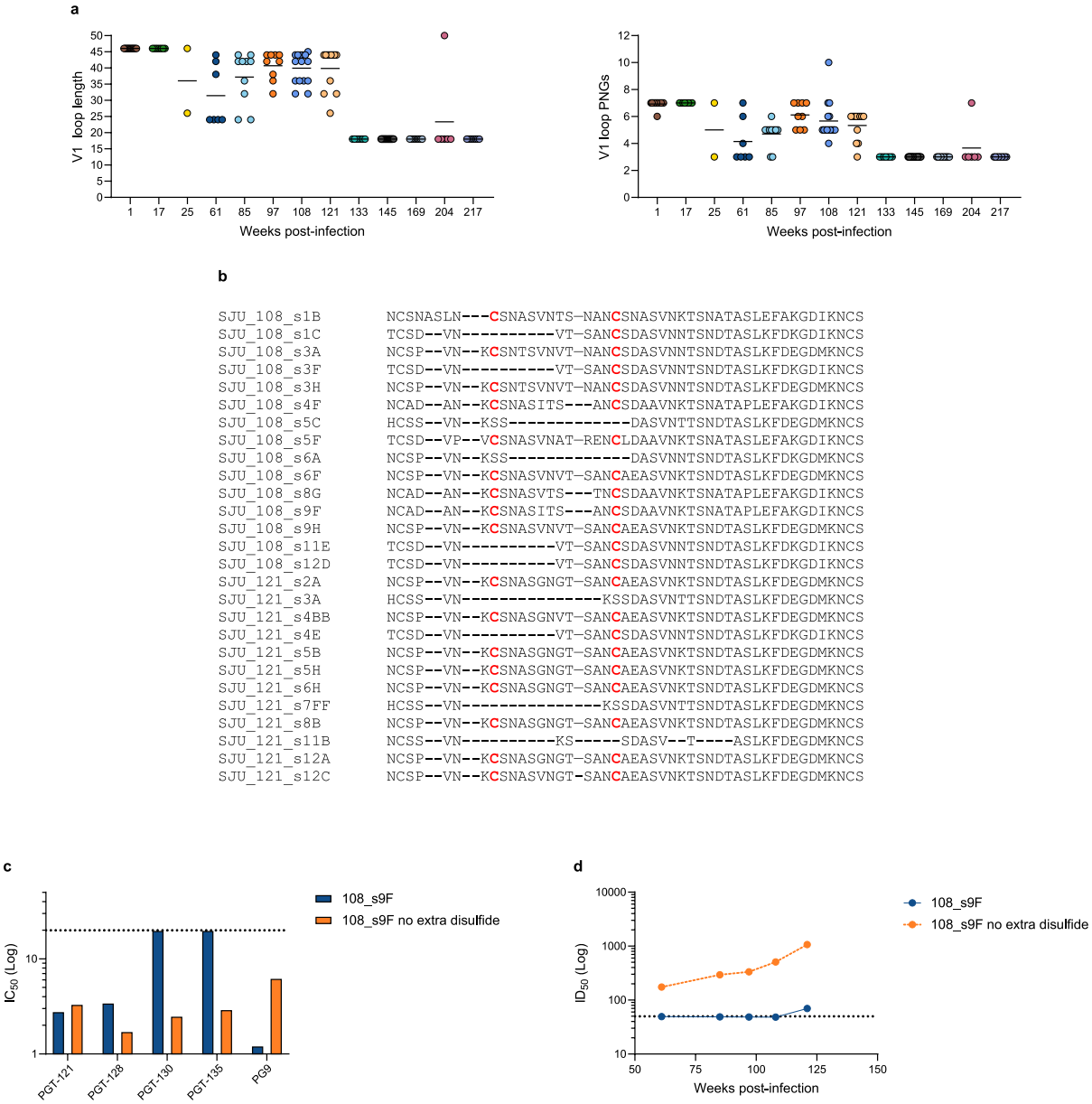



*Figure 3.5: Extra V1 loop disulfide bonds impact neutralisation by V3-glycan domain bnAbs and SJU sera.*
*(**a**) Graph showing variation in the V1 loop length and number of V1 PNGs over time, in autologous viruses. (**b**) Alignment of V1 loops from autologous viruses isolated between 108-121 weeks post-infection. Shown in red are the cysteine residues that form the additional disulfide bond. (**c**) Grouped bar chart showing the impact of removing the extra disulfide bond from an autologous virus, on the neutralisation potency of a V3-glycan bnAbs panel. The panel consists of PGT-121, PGT-128, PGT-130 and PGT-135, with PG9 (V1/V2 apex) acting as a positive control. The black dotted line indicates the cut-off antibody concentration used in the neutralisation assays (20µg/mL). (**d**) Line graph showing the impact of removing the extra disulfide bond from an autologous virus, on the neutralisation potency of the contemporaneous serum. The black dotted line indicates the cut-off serum concentration used in the neutralisation assays (1:50).*

glycan domain neutralisation. The virus 133_s4C had a very short V1 loop, which allowed enhanced V3-glycan domain neutralisation when swapped into 108_s9F, but was resistant to some bnAbs in its original form. Furthermore, despite SJU having broadly neutralising sera that was V3-glycan domain specific, superinfection with a short V1, N332 glycan bearing virus still occurred. This highlights that other regions of the superinfecting viruses may have been facilitating resistance from the neutralising antibody response of this donor.

### 3.1.7 Discussion

By 204 weeks post-infection, the serum from SJU could neutralise ~83% and ~58% of viruses from the cross-clade indicator panel and global panel respectively. This was less broad compared to the sera of HIV-1 donors from other V3-glycan domain co-evolution studies. Plasma tested from donors BF520 and PC039 were able to neutralise 85% and 65% of panel strains respectively [46,219]. Whilst BF520 was examined using the same 12 virus global panel, PC039 was tested using a different 37 virus panel preventing direct comparison to SJU. Although the serum from SJU was not as broad as BF520, it neutralised the viruses more potently. Furthermore, BF520 is an infant and prior research has shown that HIV-1 infected infants develop better bnAbs responses compared to adults, although it is not fully understood why [439]. A recent study isolated V3-glycan domain bnAbs from a superinfected donor, whose sera was able to neutralise 83% of the global panel. However, serum epitope mapping had not been carried out, thus bnAbs targeting other lineages may have also been present [458].

In HIV-1 infected individuals, it is understood that increased viral diversity is associated with the development of neutralisation breadth [459,460]. However, it is currently debated whether superinfection can drive increasing breadth of existing B-cell responses or merely induce de novo antibody responses against the superinfecting virus [210,458]. Thus, it remains unclear if the increase in serum breadth witnessed here resulted from increased bnAb coverage by the existing lineages targeting the V3-

*Table 3.5: bnAb IC50 values fold changes, resulting from swapping V1 loops in autologous viruses.*

| | | | bnAb (fold change) | | | | |
|---|---|---|---|---|---|---|---|
| Backbone | V1 | Length | PGT-121 | PGT-128 | PGT-130 | PGT-135 | PGV04 |
| 1_s5H (Early) (V1 =46) | 108_s5C | 32 | 8 | 1 | 2 | 1 | 1 |
| | 108_s9F | 42 | - | - | - | - | - |
| | 133_s4C | 23 | - | - | - | - | - |

| | | | bnAb (fold change) | | | | |
|---|---|---|---|---|---|---|---|
| Backbone | V1 | Length | PGT-121 | PGT-128 | PGT-130 | PGT-135 | PGV04 |
| 108_s5C (Sensitive) (V1 =32) | 1_s5H | 32 | NI | NI | NI | NI | NI |
| | 108_s9F | 42 | NI | NI | NI | NI | NI |
| | 133_s4C | 23 | 1 | 2 | 8 | -26 | 1 |

| | | | bnAb (fold change) | | | | |
|---|---|---|---|---|---|---|---|
| Backbone | V1 | Length | PGT-121 | PGT-128 | PGT-130 | PGT-135 | PGV04 |
| 108_s9F (Resistant) (V1 =42) | 108_s5C | 32 | - | - | - | - | - |
| | 108_s9F | 42 | 46 | 3 | >172 | >32 | 1 |
| | 133_s4C | 23 | 49 | 6 | >72 | >127 | 1 |

| | | | bnAb (fold change) | | | | |
|---|---|---|---|---|---|---|---|
| Backbone | V1 | Length | PGT-121 | PGT-128 | PGT-130 | PGT-135 | PGV04 |
| 133_s4C (Superinfection) (V1 =23) | 108_s5C | 32 | >-10 | 1 | 1 | 1 | 1 |
| | 108_s9F | 42 | NI | NI | NI | NI | NI |
| | 133_s4C | 23 | - | - | - | - | - |

glycan domain, or was caused by the development of new lineages. In study isolating V3-glycan domain bnAbs from a superinfected donor, it was found that multiple distinct B-cell lineages were activated in response to either the primary or superinfecting viruses. Whilst one lineage targeted the V3-glycan domain and neutralised multiple variants, the other lineages did not have their epitopes mapped and showed a lower heterologous neutralisation capacity. It was suggested that the V3-glycan domain bnAbs were elicited in response to the superinfecting virus [458]. To further understand how superinfection impacted V3-glycan domain bnAbs in SJU, numerous antibody lineages need to be isolated and characterised across multiple timepoints.

The neutralisation response to 1_s5H was purely autologous, as there was no detectable heterologous neutralisation observed until ~60 weeks post-infection. Conversely, neutralisation of 133_s9C could have been autologous or an expansion of the bnAb response. Interestingly, 204_s11D was neutralised pre-contemporaneously, and this could explain why other recombinant strains were not recovered via SGA. Similarly to 133_s9C, 204_s11D may have been neutralised by autologous or bnAb responses. Each of these viruses would have undergone vast diversification to evade the host immune response. Specific progeny from 1_s5H would have triggered the development of the V3-glycan domain specific bnAbs in SJU, and progeny from 133_s9C and 204_s11D may have expanded these lineages or generated new ones. Isolation of mAbs from this donor would allow identification of which autologous Envs stimulated the V3-glycan domain bnAb response.

The data from this project currently suggests that interplay between autologous and donor bnAbs led to the rare occurrence of incomplete viral escape. This phenomenon persisted for ~36 weeks and was disrupted by superinfection between weeks 121 and 133 weeks post-infection. The primary virus was no longer detected via SGA and the only recombinant virus isolated was sensitive to pre-contemporaneous neutralisation. Rebound or complete escape of the primary infecting virus never occurred during the 252-week sampling period, although this could be due to the low sampling capability of SGA. Contemporaneous neutralisation could have been due to the viral population becoming trapped between host immune responses and the fitness cost of resistance [461]. However, the viruses could have instead been replicating in ways which prevented antibody access, such as via cell to cell spread. The continued exposure to susceptible viral variants could have led to the high potency of SJU serum through continued B-cell activation. High potency serum responses following incomplete viral escape has been documented in other donors such as CAP256 [461,462]. Gaining a better understanding of this co-operation between strain specific and bnAbs could form the basis of an effective therapeutic vaccine.

It was demonstrated that most of the circulating viruses isolated during the period of contemporaneous neutralisation were susceptible to neutralisation by a panel of V3-glycan domain bnAbs. Hence, it was likely that the contemporaneous neutralising activity observed in sera was to some extent the result of V3-glycan domain bnAbs. However, between weeks 108 and 121 post-infection this donor was able to contemporaneously neutralise many N334 bearing autologous viruses. This was surprising because shifting the glycan from position 332 to 334 is a common mechanism of viral escape from V3-glycan domain bnAbs [463]. Although, some V3-glycan domain bnAbs show glycan promiscuity, for instance PGT-130 is more tolerant of N334 glycan bearing viruses [56,464]. Collectively, incomplete escape could have arisen as a result of the V3-glycan domain bnAbs exhibiting glycan promiscuity, a helper lineage specific for the N334 glycan or autologous neutralising antibodies targeting other regions on Env. This cannot be fully elucidated until autologous or bnAbs are isolated from donor SJU.

Many papers have described a correlation between increased V1 lengths and escape from V3-glycan domain bnAbs [455–457]. To investigate the impact of the V1 loop on the accessibility of the V3-glycan

domain epitope, chimeric *envs* were developed in which entire V1 loops were swapped between resistant and sensitive autologous viruses. It was shown that replacing the long V1 (42 amino acids) of a resistant virus (108_s9F) with a short V1 from either a sensitive (105_s5C, 32 amino acids) or resistant virus (133_s4C, 23 amino acids), led to a considerable increase in neutralisation sensitivity by a panel of V3-glycan domain bnAbs. This finding has been corroborated by others, and it is likely that the V1 can simply sterically block access to the V3 loop [445]. However, early viruses isolated from SJU (1_s5H) had longer V1 loops (46 amino acids) but were still highly sensitive to V3-glycan domain bnAbs. Further investigation showed that an extra disulfide bond was present in the V1 loop of the resistant autologous virus (108_s9F), and this was interfering with the V3 epitope. Additional disulfide bonds in the V1 loop are uncommon but have been documented in other donors [455,456]. Therefore, long V1 loops can interfere with V3-glycan domain specific neutralising antibodies, but the conformation of the V1 needs to also be considered. These structural changes in the V1 loop were likely a mechanism used to escape neutralisation by the V3-glycan domain targeted bnAbs.

The presence of the N332-glycan and a short V1 have been demonstrated as being important for neutralisation by V3-glycan domain bnAbs. However, SJU became superinfected by a virus of the same subtype (clade C), with an N332 glycan and a short V1 loop, whilst V3-glycan domain bnAbs were circulating in their sera. This indicates that other regions of the Env are important mediators of neutralisation by V3-glycan domain bnAbs. Notably, the superinfecting strains showed some mutations in the V3 loop and insertions in the V4 loop that require further investigation. Isolating bnAbs would allow examination of the Env alterations that enabled resistance to the neutralising antibody response of this donor.

## 3.3 Isolating mAbs from a HIV-1 superinfected donor

### 3.3.1 Introduction

The first mAbs were generated in 1975, through the use of hybridoma technology [465]. This method involved fusing splenocytes from immunised mice with nonsecretory myeloma cells. The generated hybridomas were maintained in vitro and continually secreted the mAb encoded by the fused B-cell [432]. Although used extensively, this method was very time consuming and had a low throughput due to a low fusion success rate. Hence, isolating mAbs using hybridomas did not adequately sample the diversity of the immune repertoire [466]. By 1985, phage display technology was introduced and enabled the screening of billions of distinct antibodies [432,467]. This process involved the generation of genetic diversity, the coupling of genotype to phenotype via a display platform, the screening of displayed antibodies and the amplification of enriched candidates [468]. However, the $Ig_H$ and $Ig_L$ pairings were created randomly, and were therefore not representative of the naturally occurring immune repertoire. Regardless of these technical limitations, a first-generation of HIV-1 bnAbs was successfully discovered in the early 1990s [469]. However, these antibodies showed unusual characteristics such as auto-reactivity, required high doses to provide protective immunity and could not be induced through vaccination despite repeated attempts [198,470]. As a result, vaccination efforts became almost exclusively focused on T-lymphocyte mediated immunity [471]. Almost a decade later, the availability of single-cell B-cell cloning techniques and large donor cohorts (for example protocol G) enabled characterisation of a second-generation of bnAbs against HIV-1. This new generation showed remarkably improved neutralisation breadth and potency which reignited confidence that an antibody-based vaccine could be achieved [218]. The bnAbs identified targeted numerous conserved epitopes located across the HIV-1 Env. These include the CD4bs, V1/V2 apex, V3-glycan domain, FP/gp120-gp41 interface, silent face and MPER [200].

Unfortunately, bnAbs are slow to develop and it can take several years before HIV-1 infected adults display sera with desirable neutralisation breadth [195]. Furthermore, bnAbs typically present unusual characteristics such as high levels of SHM, indels and a long CDRH3 [55]. The elevated level of SHM found in bnAbs suggests that they have undergone multiple rounds of affinity maturation to acquire breath. However, it remains uncertain if the extended CDRH3s displayed also arise through this gradual process [35]. Indeed, it has been shown that bnAbs can develop from germlines with pre-existing long CDRH3s produced via VDJ recombination [197]. An additional concern is that bnAbs often display auto-reactivity to endogenous factors such as phospholipids [198]. This would imply that re-elicitation of these antibodies would be disfavoured due to host tolerance mechanisms. However, this is not always the case and promising bnAb lineages with no significant auto reactivity have been successfully isolated [199].

HIV-1 manipulates host glycosylation machinery adding glycans to Env at PNGS, shielding conserved epitopes from immune responses. The location of N-linked glycans is encoded within the viral genome at N-X-S/T (where X is any amino acid other than proline) sequons. The high density of sugars present in the V3-glycan epitope sterically hinders the accessibility of glycoprocessing enzymes, resulting in high-mannose glycans which can be recognised as non-self [166,213]. Indeed, V3-glycan specific bnAbs target continuous amino acid residues (such as the GDIR motif) and high-mannose glycans, located on the V3 loop of Env. Although the V3-glycan epitope centres on the glycan at amino acid position 332, other high-mannose glycans (such as N301) can be involved. V3- glycan specific bnAbs exhibit lower levels of SHM and a less restricted germline usage compared to bnAbs targeting other epitopes [220,439]. Furthermore, this class of bnAbs have been shown to develop relatively quickly following HIV-1 infection [219]. Collectively, these features make V3-glycan specific bnAbs of great interest for vaccine

development. Examining the co-evolution of new bnAb lineages targeting the V3-glycan domain would help to inform such ventures.

Overall, the specific aims of this section were to:

- Develop biotinylated gp120 sorting probes for antigen specific B-cell sorting.
- Perform antigen specific B-cell sorting to isolate gp120 specific B-cells from donor SJU.
- Clone, sequence and produce a library of functional mAbs.
- Sequence variable regions of mAbs, and examine features such as germline gene usage, level of SHM and CDRH3 length.
- Characterise the neutralisation capability and epitope binding of this library, identifying any bnAb lineages present.

### 3.3.2 Human subjects and ethics

Peripheral blood mononuclear cells (PBMCs) were obtained from the King's Biobank. All donors signed a written informed consent and ethical approval for usage of these samples was granted by The London Multicentre Research Ethics Committee (MREC 04/2/025).

### 3.3.3 Probe production and optimisation

The HIV-1 clade B strains JR-CSF and 92BR020 were selected to be made into gp120 sorting probes because they were potently neutralised by SJU serum. Neutralisation of JR-CSF was shown to be dependent on the presence of the N332 glycan, whilst neutralisation of 92BR020 was less reliant on this feature (Figure 3.6). A JR-CSF N332A mutant was generated via site-directed mutagenesis, and it would be used to prevent selection of B-cells that were not specific for the V3-glycan domain. A 92BR020 N332A mutant was also produced as an alternative negative selection probe. The gp120 cDNA was PCR-amplified from pSVIII vectors and sub-cloned into pHLsec vectors using EcoRI-Kpnl restriction sites. The gp120s were cloned into the pHLsec vector to incorporate a 6x histidine tag and an AviTag downstream of the constructs. This would allow purification by nickel affinity chromatography and in vivo biotinylation when co-expressed with pDisplay-BirA-ER in the presence of biotin.

Following production and purification, the probes were subjected to quality control via western blot analysis and ELISA. The western blot revealed that gp120 monomers had been successfully purified, although the JR-CSF probe had some aggregation indicated by the second band in the non-reduced sample (Figure 3.7c). Further, visualisation with streptavidin conjugated to PE-Cy7 showed that the in vivo biotinylation had been successful. ELISAs with bnAbs targeting different epitopes (V3-glycan domain and CD4bs) verified that the monomers were correctly folded (Figure 3.7a). Additionally, the

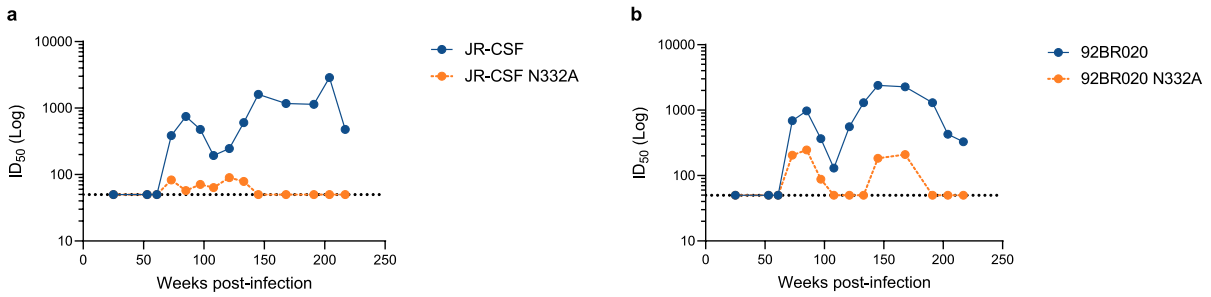



*Figure 3.6: SJU sera neutralisation of particular heterologous strains is dependent on the presence of the N332 glycan. Line graphs showing the longitudinal neutralisation potency of SJU sera over time, against JR-CSF (**a**) and 92BR020 (**b**), as well as their associated N332A mutants. The black dotted lines indicate the cut-off serum concentration used in the neutralisation assays (1:50). Adapted from data by Luke Granger.*

binding of streptavidin conjugated to alkaline phosphatase confirmed biotinylation (Figure 3.7b). SJU sera samples from 121-217 weeks post-infection were found to bind well via ELISA, verifying the probes were suitable for antigen specific B-cell sorting (Figure 3.8).

293T cells which had been transduced to stably express a PGT-121 bnAb precursor (3H3L cells, a kind gift from Laura Mccoy), were used to test the gp120 probes in a flow cytometry setting [472]. 3H3L did not bind to the JR-CSF, JR-CSF N332A or 92BR020 N332A probes, therefore optimisations were only performed with 92BR020 gp120. The probe was conjugated to PE, APC and Alexa-488 (conjugated to streptavidin) and binding to 3H3L cells was investigated through flow cytometry (Figure 3.9a). When conjugated with APC or PE, 92BR020 enriched ~60% of the population validating the probe. However, due to the high auto-fluorescence exhibited by 3H3L cells in the Alexa-488 channel enrichment was not observed. To determine the optimal probe concentration for B-cell sorting, whilst minimising the

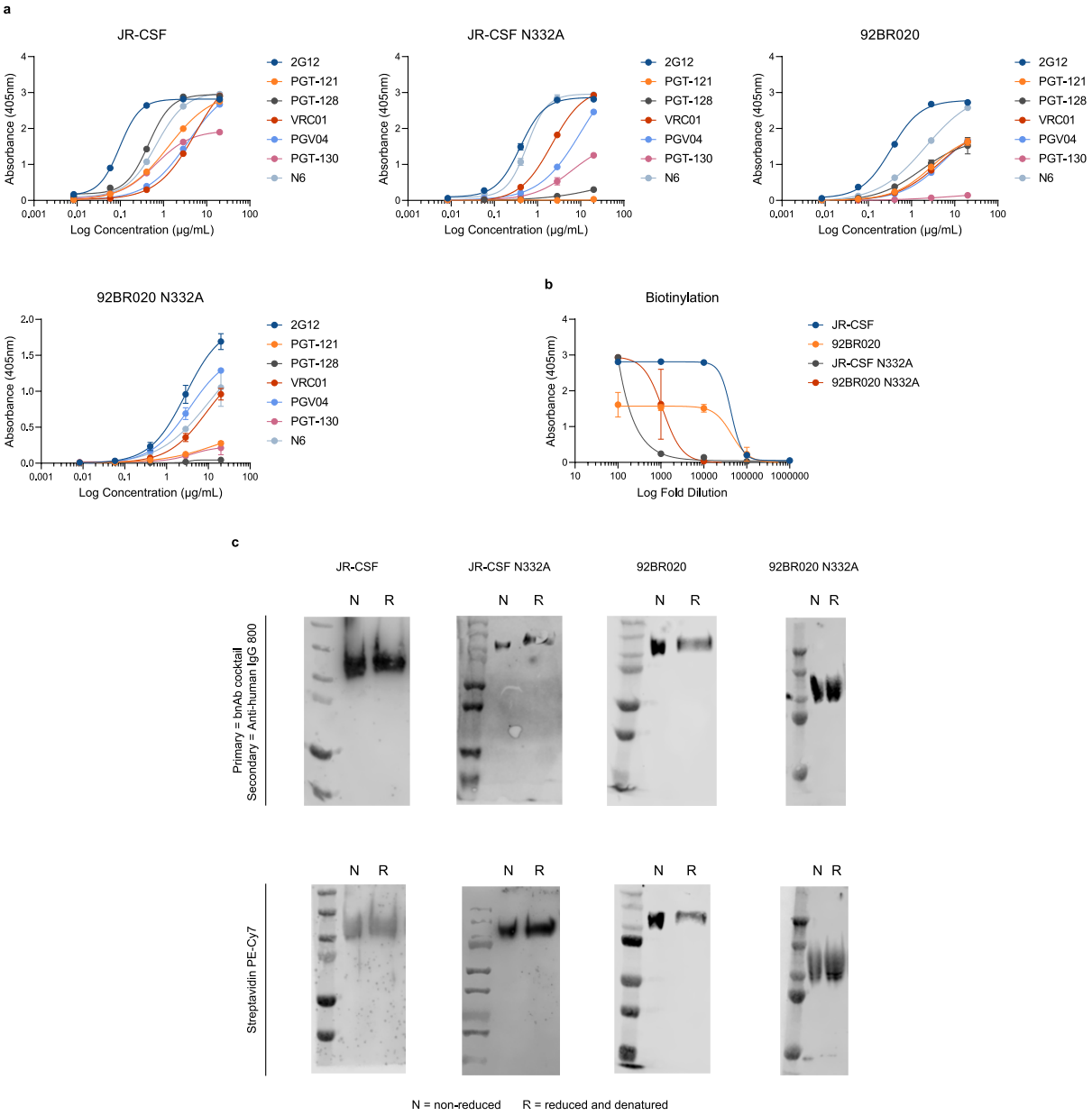



*Figure 3.7: Quality control of JR-CSF and 92BR020 biotinylated sorting probes.*
*ELISAs showing the binding of a panel HIV-1 bnAbs (**a**), and streptavidin conjugated to AP (**b**), to JR-CSF, JR-CSF N332A, 92BR020 and 92BR020 N332A recombinant gp120s. The panel included V3-glycan domain (PGT-121, PGT-128, PGT-130 and 2G12) and CD4bs (VRC01, PGV04 and N6) bnAbs. (**c**) Western blots of the gp120 proteins in their non-reduced (N) and reduced (R) forms, developed using a cocktail of bnAbs (with an anti-human IgG as a secondary) or streptavidin conjugated to PE-Cy7.*

background binding to other cells, flow cytometry was performed on healthy donor PBMCs and 3H3L cells stained with varying concentrations of gp120 probe conjugated to APC. However, the binding of 3H3L to 92BR020 was sub-optimal and would likely not represent what would be expected with SJU PBMCs. Therefore, a SF162 gp120 probe that had greater affinity for 3H3L was used for these experiments (Figure 3.9b). Increasing the concentration of SF162 gp120 led to increased enrichment and separation of the 3H3L cells. However, increasing the concentration of the probe also increased the background binding to healthy PBMCs. Therefore, a concentration of 150nM was chosen because it could give a strong enrichment and separation of the desired cells, whilst not incurring high background binding.

### 3.3.4 Antigen specific B-cell sorting from SJU PBMCs collected 252 weeks post-infection

Antigen-specific B-cell sorting was used to isolate $CD19^{+}IgG^{+}$ B-cells from SJU PBMCs collected at 252 weeks post-infection (Figures 3.10a-b). Cells that were specific for both of the JR-CSF and 92BR020 probes were selected. Due to a reduced number of cells than expected (~650,000 cells stained) and a low enrichment (0.31% double positive), the negative selection probe JR-CSF was not used to differentiate the cell populations. 10 B-cells were sorted and the $V_H$ and $V_L$ regions were reverse transcribed and amplified via nested PCR, using gene-specific primer mixes. The PCR products were purified using PCR cleanup kits and ligated into $Ig_H$ or $Ig_L$ expression vectors using Gibson assembly. A total of 5 mAbs were isolated, the plasmids were purified using Miniprep kits and sent for Sanger sequencing. The $Ig_H$ and $Ig_L$ were co-transfected into 293T cells and the supernatant was tested via ELISA with the three sorting probes (Figure 3.10c). A single mAb (SJU_252_H8) was found to bind to JR-CSF and 92BR020 probes, however it also bound to the N332A mutant indicating a lack of V3-glycan domain specificity.

The $V_H$ and $V_L$ genes of SJU_252_H8 were analysed using the IMGT database, revealing the predicted germline gene usage, %SHM and CDRH3 length (Table 3.3) [433]. SJU_252_H8 showed moderate levels of SHM (~11% $V_H$ and ~8% $V_L$), a longer than average CDRH3 (18 amino acids) and used the germlines VH1-69 and VK4-1. When tested in neutralisation assays, SJU_252_H8 was not able to neutralise JR-CSF or 92BR020 (data not shown).

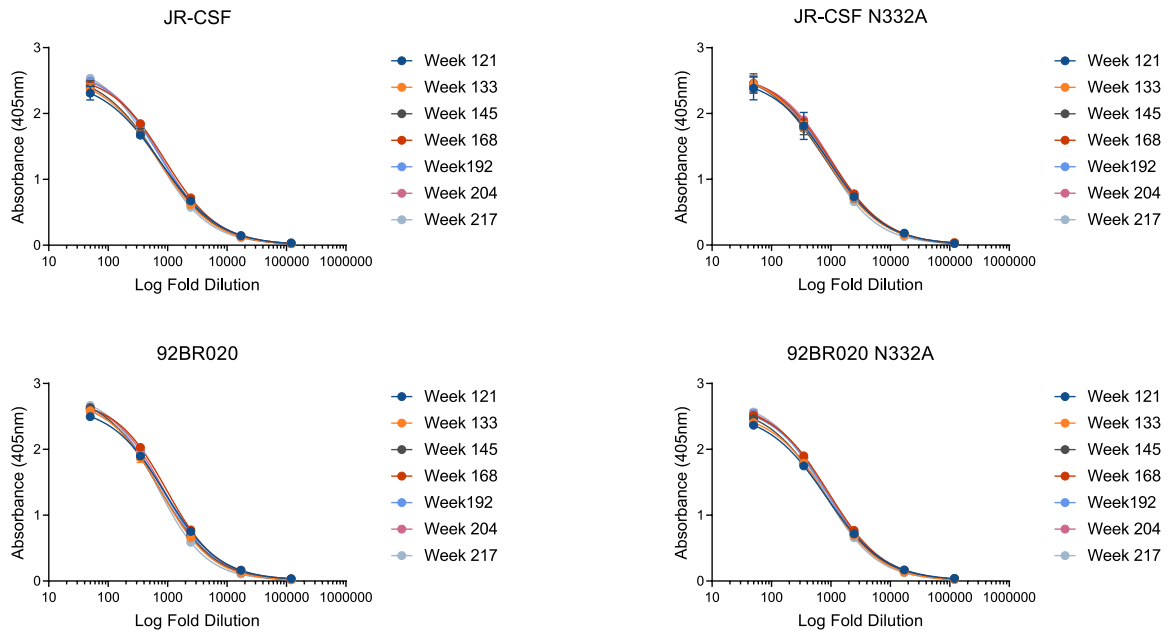



*Figure 3.8: Longitudinal binding of SJU serum samples to JR-CSF and 92BR020 biotinylated sorting probes. ELISAs showing the binding SJU sera collected between 121-217 weeks post-infection, to JR-CSF, JR-CSF N332A, 92BR020 and 92BR020 N332A recombinant gp120s.*

### 3.3.5 Modifying gp120 probes

Prior data had shown that using sorting probes at a concentration of 150nM could achieve strong enrichment and separation of the desired population (3.3.2). However, only a small enrichment was observed during the previous sort (3.3.3), indicating the presence of a population of cells that were competing with the gp120-specific B-cells.

In nature, the purpose of the gp120 subunit is to interact with CD4 and chemokine receptors. It has been estimated that there can be as many as ~$1.5x10^5$ CD4 receptors on the surface of a single $CD4^+$ lymphocyte, in cryopreserved PBMC samples [473]. Therefore, a flow cytometry experiment was carried out to investigate the interactions between $CD3^+CD4^+$ T-cells and the SF162 sorting probe (Figure 3.11). The T-cells from healthy PBMCs were shown to bind the sorting probe, and increasing the SF162 gp120 concentration from 50nM to 200nM resulted in higher levels of enrichment. This indicated that a higher concentration of sorting probe would be needed to overcome T-cell competition. However, earlier work showed that increasing the sorting probe concentration to 200nM would increase background measurements in the $CD19^+IgG^+$ B-cell channel. Processing a high number of B-cells as a result of elevated background would have greatly increased the workload of the project. Hence, new gp120 probes were developed with a D368R mutation which prevented interaction with the CD4 receptor.

D368R mutations were made to JR-CSF, 92BR020, JR-CSF N332A and 92BR020 N332A in pSVIII vectors via site-directed mutagenesis. The resulting gp120 cDNA underwent sub-cloning into pHLsec, transfection into 293F cells (with in vivo biotinylation), affinity chromatography and size exclusion chromatography. Despite multiple attempts, JR-CSF D368R N332A could not be produced and showed only low quantities of aggregate during size exclusion chromatography. The other proteins were examined in the same manner as the aforementioned sorting probes (ELISA and western blot). Western blot showed that the D368R gp120s were of the correct size with no aggregation observed and visualisation with streptavidin conjugated to PE-Cy7 showed that the in vivo biotinylation had been successful (Figure 3.12c). ELISA with bnAbs targeting different epitopes verified that the monomers were correctly folded (Figure 3.12a). Notably, the binding of some bnAbs specific for the V3-glycan domain (such as PGT-121) was weaker towards D368R gp120s compared to the wild-type (WT) probes. The binding of streptavidin conjugated to alkaline phosphatase conftirmed biotinylation (Figure 3. 12b). Overall, the D368R probes were considered suitable for use in antigen specific B-cell sorting.

*Table 3.6: Characteristics of the mAbs isolated from donor SJU at 252 and 108 weeks post-infection.*

| Name | V-germline | J-germline | D-germline | % SHM | CDRH3 length | CDRH3 |
|---|---|---|---|---|---|---|
| SJU_108_B10 | IGHV3-11 | IGHJ2 | IGHD5-18 | 9.4 | 16 | ARVNRGYSFRYFYFDF |
| SJU_108_C4 | IGHV1-69 | IGHJ5 | IGHD6-19 | 7.3 | 14 | ASAPMQWPYSWNDP |
| SJU_108_D6 | IGHV3-11 | IGHJ2 | IGHD5-18 | 5.2 | 16 | ARVNRGYSYYYYYFDL |
| SJU_108_E4 | IGHV1-69 | IGHJ4 | IGHD3-16 | 5.9 | 10 | ATLAYSDKDF |
| SJU_108_E9 | IGHV1-69 | IGHJ4 | IGHD5-18 | 6.6 | 13 | TRPGDNYGFEFDH |
| SJU_108_F7 | IGHV3-11 | IGHJ2 | IGHD5-18 | 8 | 16 | ARVNRGYSYHYHYLDV |
| SJU_108_G7 | IGHV1-69 | IGHJ4 | IGHD2-2 | 11.8 | 9 | ATPDAKMAY |
| SJU_252_H8 | IGHV1-69 | IGHJ5 | IGHD3-3 | 10.76 | 18 | ARDHTVVFGAVTDNWFDP |

| Name | V-germline | J-germline | % SHM | CDRL3 length | CDRL3 |
|---|---|---|---|---|---|
| SJU_108_B10 | IGKV3-20 | IGKJ4 | 5.3 | 9 | QLYDGNSVT |
| SJU_108_C4 | IGKV3-20 | IGKJ1 | 3.2 | 10 | QQYGSSPSVT |
| SJU_108_D6 | IGKV3-20 | IGKJ4 | 3.2 | 8 | QQYDSSVT |
| SJU_108_E4 | IGKV3-20 | IGKJ3 | 2.1 | 9 | QQYGSSPFT |
| SJU_108_E9 | IGKV3-20 | IGKJ3 | 3.5 | 9 | QQYGNSPFT |
| SJU_108_F7 | IGKV3-20 | IGKJ4 | 5.7 | 8 | QQYDSSVT |
| SJU_108_G7 | IGKV3-20 | IGKJ4 | 5.7 | 9 | QQYGSSLFT |
| SJU_252_H8 | IGKV1-39 | IGKJ2 | 7.89 | 9 | QQSYSTPRT |

### 3.3.6 Antigen specific B-cell sorting from SJU PBMCs collected 108 weeks post-infection

CD19⁺IgG⁺ B-cells were sorted from SJU PBMCs collected at 108 weeks post-infection using the D368R sorting probes (Figure 3.13a). PBMCs collected at 108 weeks post-infection were selected due to high sample availability. Cells that were positive for both of the JR-CSF D368R and 92BR020 D368R probes, but negative for the 92BR020 D368R N332A probe were selected. Due to a low enrichment (0.78% double positive), cells that were strongly positive for all three gp120s were also sorted. 60 B-cells were isolated from 5 million PBMCs, and 28 mAbs were generated using the cloning and production pipeline mention previously (3.3.3). When tested via ELISA, 7 mAbs were found to be reactive to JR-CSF, 92BR020 and JR-CSF gp120 (Figure 3.13b). Therefore, these mAbs displayed binding breadth but not V3-glycan domain specificity.

The sequences of the $V_H$ and $V_L$ genes were analysed using the IMGT database, allowing analysis of their predicted germline gene usage, %SHM and CDRH3 lengths (Table 3.3) [433]. The gp120-specific

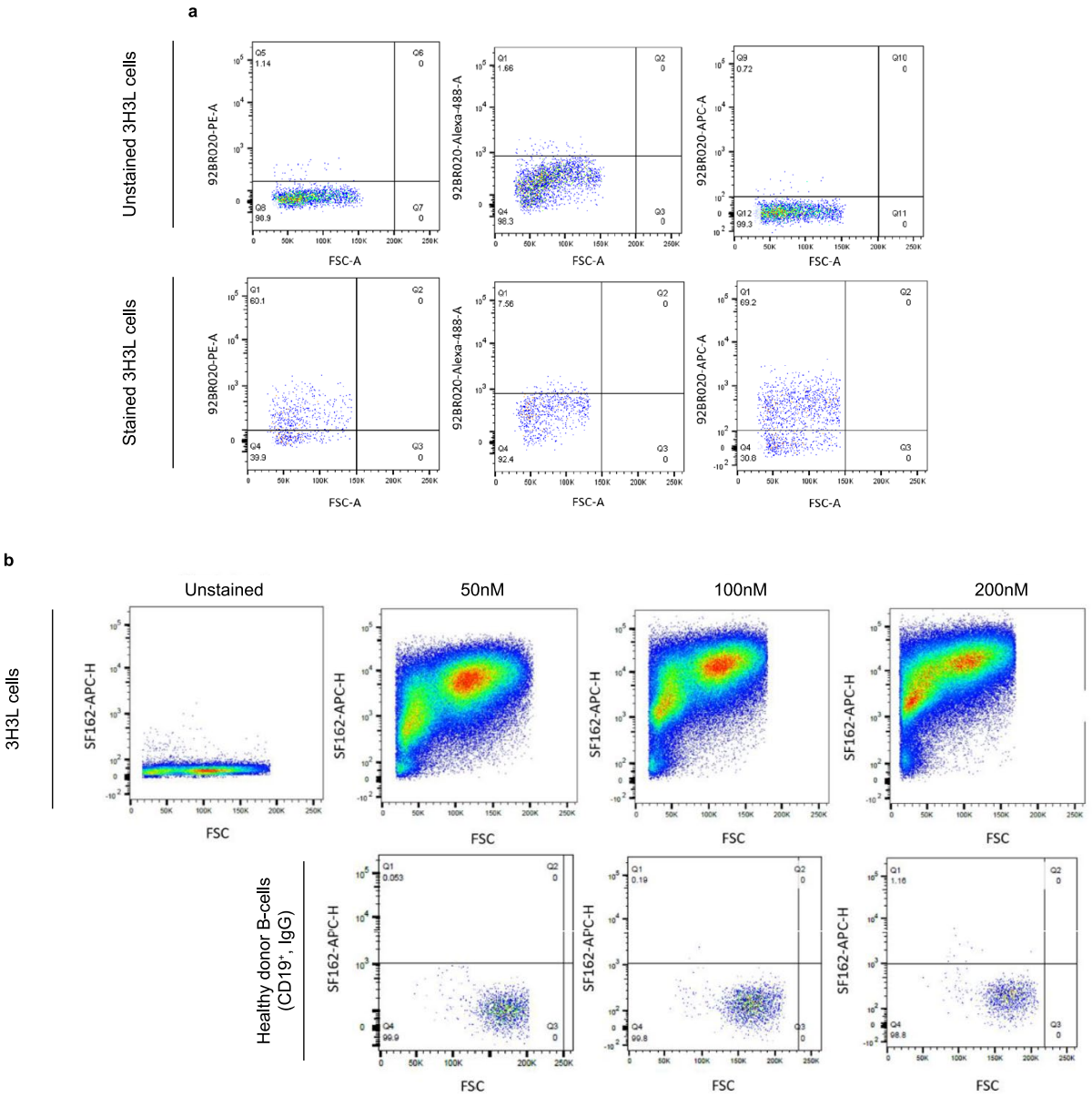



*Figure 3.9: Enrichment of 3H3L cells by biotinylated gp120 sorting probes.*
*(**a**) Flow cytometry of 3H3L cells, unstained or stained with 92BR020 gp120 conjugated to numerous fluorophores. (**b**) Flow cytometry of 3H3L cells and healthy B-cells, stained with varying concentrations of SF162 conjugated to APC. The PBMCs were also stained with a cocktail of antibodies (Table 2.11) enabling gating of the CD19+, IgG B-cell shown above.*

mAbs were found to have moderate SHM in the $V_H$ and $V_L$ genes (means of 8.1% and 4.4%, respectively), with SJU_108_G7 showing the highest level of mutation (11.8% $V_H$ and 5.6% $V_L$). However, the CDRH3 length of the mAbs was below average (mean length of 14 amino acids), with SJU_108_B10, SJU_108_D6 and SJU_108_F7 displaying the longest CDRH3s at 16 amino acids. Interestingly, these mAbs used either VH1-69 or VH3-11 in combination with the VK3-20 germline gene. SJU_108_D6 and SJU_108_F7 were found to belong to the same clonal lineage, using identical

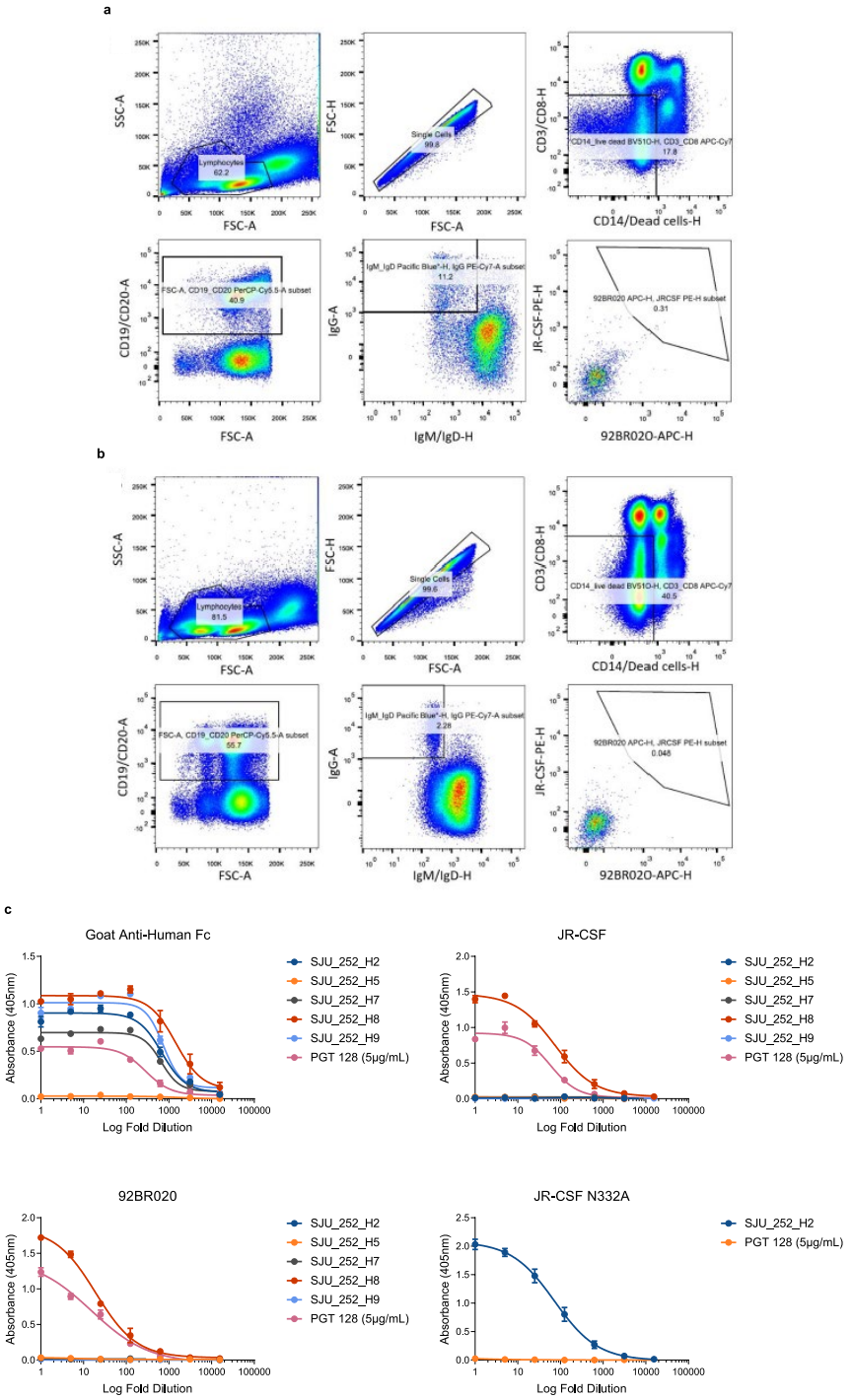



*Figure 3.10: Antigen specific B-cell sorting of SJU PBMCs collected 252 weeks post-infection.*
*Gating strategy used to sort N332-specific CD19⁺ IgG B-cells from SJU PBMCs collected at 252 weeks post-infection (**a**), with healthy donor PBMCs acting as a negative control to assess background (**b**). (**c**) ELISAs showing the binding of mAbs isolated from this timepoint to the gp120 sorting probes, as well as goat anti-human Fc allowing examination of antibody expression levels.*

V(D)J germline genes and having CDRH3s of the same length. When tested in neutralisation assays, none of the mAbs were able to neutralise JR-CSF or 92BR020 (data not shown).

### 3.3.7 Discussion

The sorting strategy applied here was unable to isolate bnAbs that were specific for the V3-glycan domain. There are numerous considerations as to why these experiments were unsuccessful. The low number of stained cells was likely a factor that limited the isolation of bnAbs in the 252 weeks post-infection sort. Additionally, it is possible that the competition of $CD3^{+}CD4^{+}$ T-cells for the g120 sorting probes further reduce enrichment. To counteract these issues, a higher number of PBMCs were stained in the 108 weeks post-infection sort and D368R gp120 sorting probes were used to prevent $CD3^{+}CD4^{+}$ T-cell interaction. This improved the enrichment of $CD19^{+}IgG^{+}$ B-cells positive for the JR-CSF and 92BR020 probes by more than 2-fold, and yielded 7 mAbs that bound to the sorting probes. However, none of these mAbs were specific for the V3-glycan domain or able to neutralise JR-CSF or 92BR020 in neutralisation assays. It was observed that some V3-glycan domain bnAbs were not able to bind as well to the D368R probes compared to the WT gp120s. Collectively, the data indicated that although the modified probes improved this process, they were still not optimal for isolating V3-glycan domain bnAbs from this donor.

Following cleavage by host cell furin, the gp120 and gp41 subunits of Env remain non-covalently linked [164]. The gp120 subunit can shed and immunodominant regions that are not displayed on trimeric Env can be presented to the host immune system [185]. The mAbs that were isolated during this work could be targeting such epitopes. The use of stabilised HIV-1 Env proteins such as SOSIPs, NFLs or UFOs would remove this problem, and would enable the isolation of mAbs that were dependant on quaternary structures [474]. D368R mutations could be made to these probes eliminating $CD3^{+}CD4^{+}$ T-

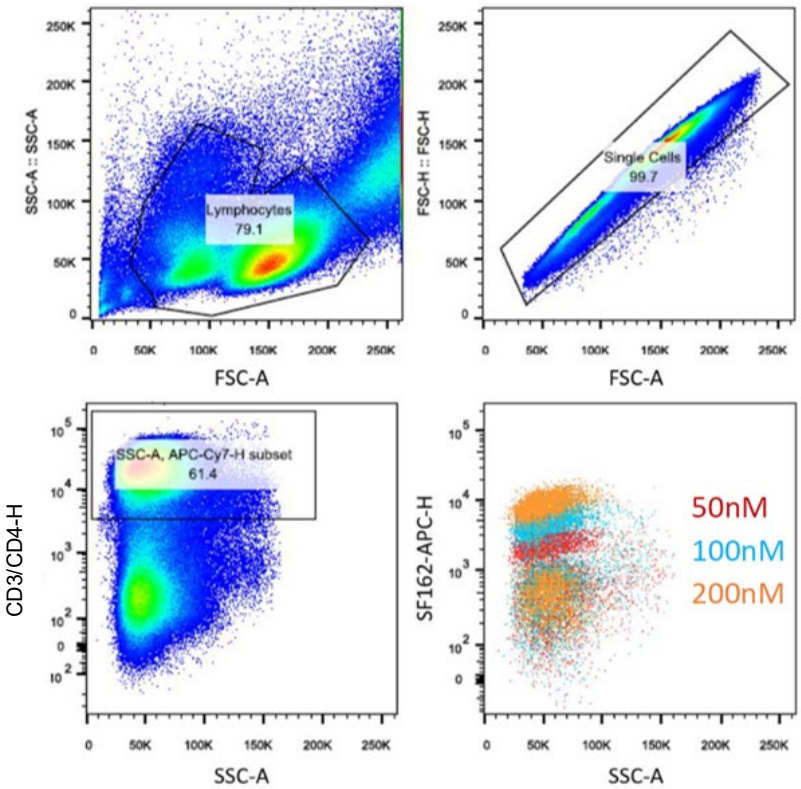



*Figure 3.11: Interactions between CD3+CD4+ T-cells and biotinylated gp120 sorting probes.*
*Flow cytometry of healthy PBMCs, stained with a cocktail of antibodies and varying concentrations of SF162 conjugated to APC.*

cell interaction, although this could inadvertently alter the V3-glycan domain as seen with the gp120 probes. Alternatively, B-cell isolation kits could be used to remove any competing cell populations prior to staining with WT stabilised Env. However, B-cell isolation kits typically require large blood volumes to isolate a suitable sample of B-cells, which may be incompatible with the limited quantity of SJU PBMCs held currently [475]. This could be overcome through the use of B-cell culturing techniques, with the caveat of altering the abundance of the natural repertoire [476]. A final alternative would be the use of V3-loops fused to non-human IgG1 Fc regions (V3-Fc). However, the lack of steric hindrance during the production of V3-Fc leads to the protein being produced with complex type glycans

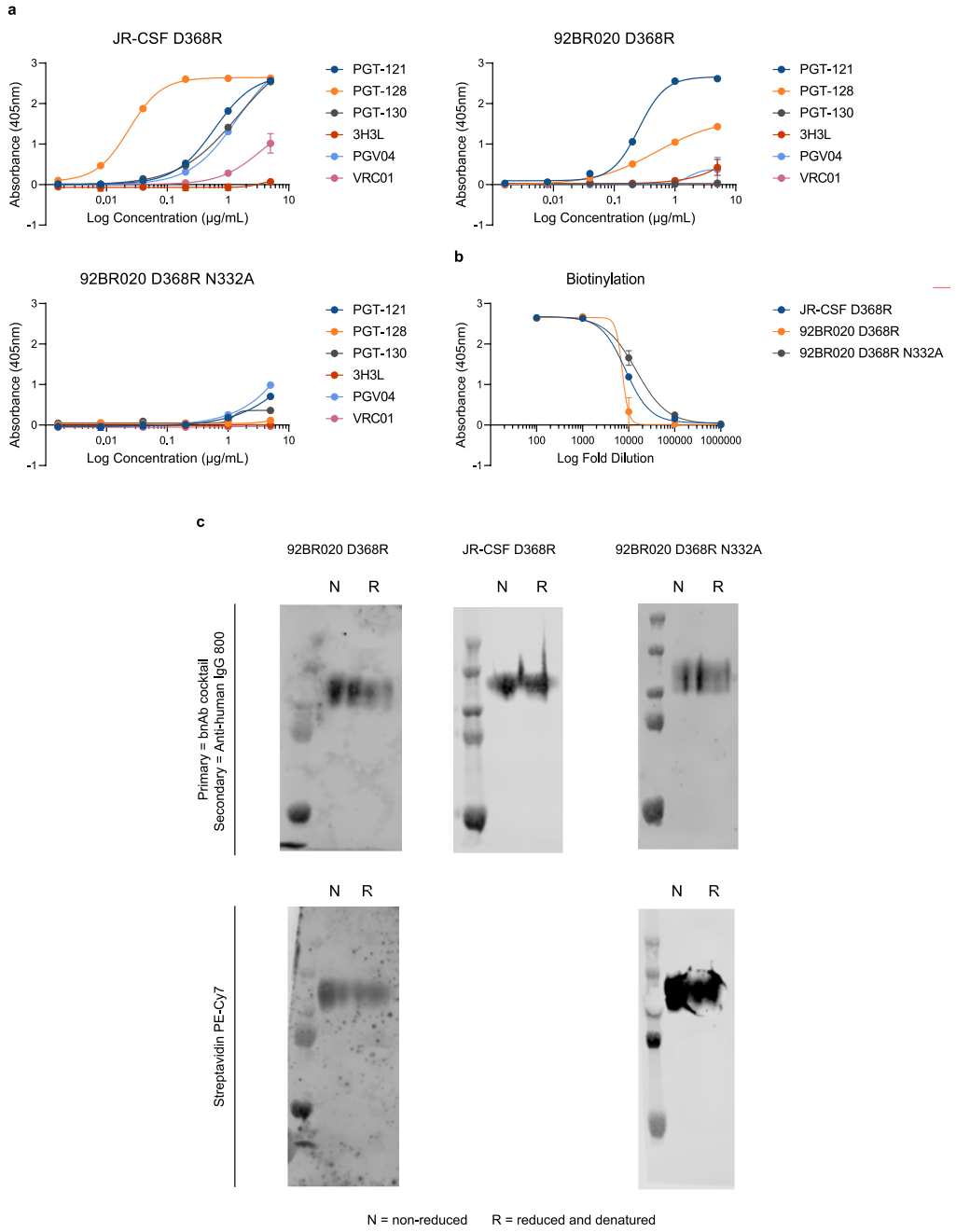



*Figure 3.12: Quality control of D368R JR-CSF and 92BR020 biotinylated sorting probes.*
*ELISAs showing the binding of a panel HIV-1 bnAbs (**a**), and streptavidin conjugated to AP (**b**), to JR-CSF D368R, 92BR020 D368R and 92BR020 D368R N332A recombinant gp120s. The panel included V3-glycan domain (PGT-121, PGT-128, PGT-130 and 3H3L) and CD4bs (VRC01 and PGV04) bnAbs. (**c**) Western blots of the gp120 proteins in their non-reduced (N) and reduced (R) forms, developed using a cocktail of bnAbs (with an anti-human IgG as a secondary) or streptavidin conjugated to PE-Cy7. JR-CSF D368R developed with streptavidin-PE-Cy7, data is not shown.*

opposed to high-mannose N-linked glycans [477]. To nullify this concern, the V3-Fc could be produced under a gradient of kifunensine and purified using affinity chromatography to a panel of V3-glycan domain bnAbs.

Despite being non-neutralising and not specific for the V3-glycan domain, 7 mAbs were isolated that were able to bind heterologous gp120s from different HIV-1 clades. As mentioned previously, these mAbs may have been targeting non-functional forms of Env arising from gp120 shedding. On the other hand, they may have been specific for non-neutralising epitopes located on trimeric Env that are conserved across the two strains. The mAbs showed moderate levels of SHM, indicating that they had undergone multiple rounds of affinity maturation. An enrichment for the VH1-69 and VH3-11 germlines was observed in the mAbs. VH1-69 is a germline that is commonly activated in response to infection with viruses such as HIV-1, influenza and hepatitis C. VH1-69 antibodies have been shown to share common genetic and structural features. For instance, the CDRH2 typically contains a hydrophobic residue at amino acid position 53, a crucial phenylalanine at 54 and a conserved tyrosine within the CDRH3 [478]. VH1-69 bnAbs isolated from other donors have been shown to interact with the hydrophobic regions of epitopes located on Env such as the MPER and CD4 binding site [479,480]. The

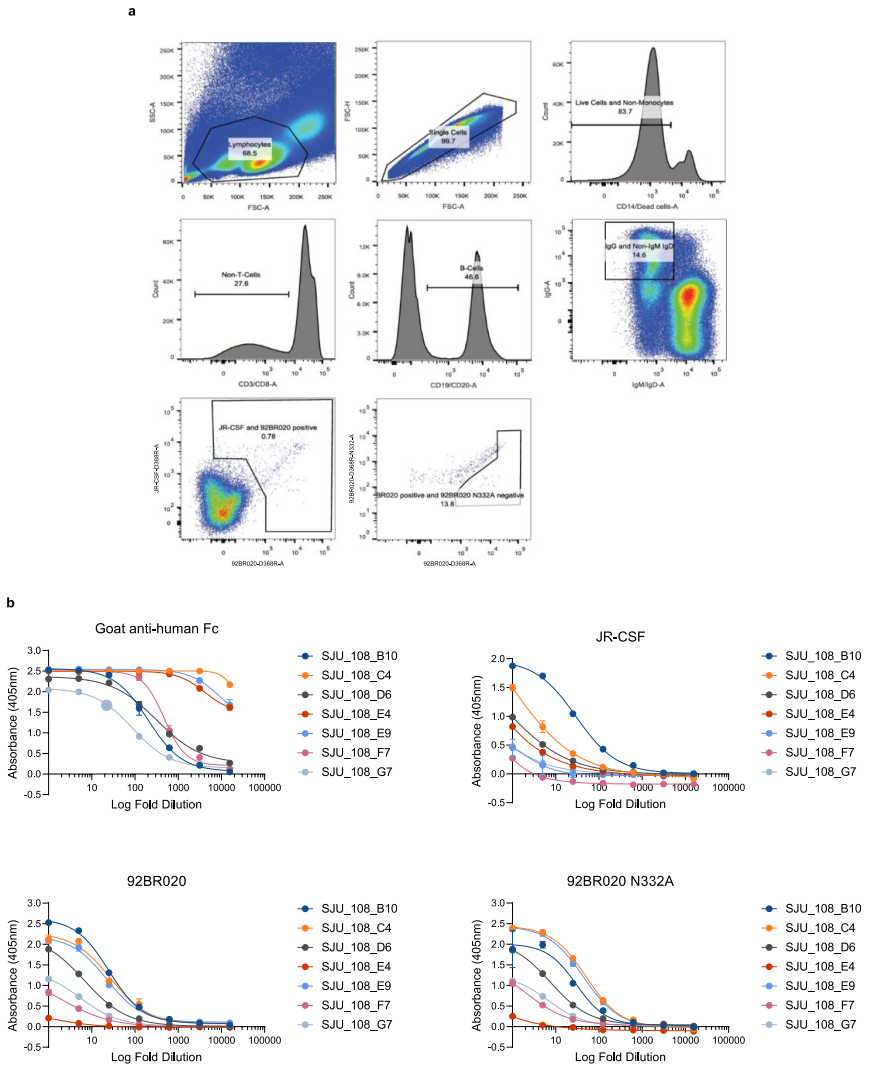



*Figure 3.13: Antigen specific B-cell sorting of SJU PBMCs collected 108 weeks post-infection.*
*(**a**) Gating strategy used to sort N332-specific CD19⁺ IgG B-cells from SJU PBMCs collected at 108 weeks post-infection. (**b**) ELISAs showing the binding of mAbs isolated from this timepoint to the gp120 sorting probes, as well as goat anti-human Fc allowing examination of antibody expression levels.*

VH3-11 germline is rarely found in HIV-1 bnAbs, but is used by PGDM12 which targets the V3-glycan domain [481]. There was also an enrichment of the VK3-20 germline, which is used by the MPER-specific bnAb PGZL1 [482]. The enrichment of these germlines, and the presence of clonally related mAbs, would imply the mAbs could be targeting similar regions on Env. These mAbs could be playing a role in as ADCC, and mAbs using the VH1-69 germline have been shown to be capable of mediating this effector function [483].

The next stage of this project was to begin making either stabilised HIV-1 Env or V3-Fc sorting probes, which would be used to isolate bnAbs from donor SJU. If the V3-Fc protein was chosen, this would require optimising a production pipeline that yielded a product with the correct glycosylation profile. Conversely, if stabilised HIV-1 Env proteins were to be employed, optimising a B-cell isolation and culturing method would be worth considering. Once bnAbs had been isolated, it would be beneficial to use a modern sequencing platform, such as 10x genomics, to bulk sequence B-cells from this donor across multiple timepoints. This would enable rapid identification of the bnAb lineage, and allow later investigation with autologous and heterologous viruses in a co-evolution study. However, at this time the UK had entered its first national lockdown as a result of the COVID-19 pandemic. Laboratories were closed unless working on a SARS-CoV-2 related project, and given our access to CL3 facilities and knowledge of serological techniques, it was decided that I would switch my research focus.

# Chapter 4: Neutralising antibody responses following SARS-CoV-2 infection

## 4.1 Antibody responses to SARS-CoV-2 infection

### 4.1.1 Introduction

The WHO declared COVID-19 a pandemic in March 2020, due to its continual global spread [484]. By June 2020, >6 million cases of SARS-CoV-2 infection had been confirmed worldwide, and there had been >370,000 associated deaths. The lack of vaccine and treatment options led to strict government enforced quarantines, in an attempt to limit the spread of the virus. This caused substantial economical disturbances, impacting quality of life and healthcare provision. The recommended diagnosis of SARS-CoV-2 involved RT-PCR of upper or lower respiratory samples (such as nasopharyngeal swabs), measuring the presence of N or RdRp genes [485]. These highly sensitive tests were able to detect vestigial viral RNA levels, but their effectiveness was dependant on the time the specimen was taken. This meant that RT-PCR was optimally applied for early detection of the virus, as the viral load reduced following the first week of symptoms [486,487]. Therefore, serological assays were developed that allowed measurement of antibody binding and neutralising responses, as well as the determination of seroconversion [485]. Although these assays were less suitable for identifying acute infection, they enabled examination of the humoral immune response to SARS-CoV-2 in a more qualitative and quantitative manner. Given that a proportion of COVID-19 cases were asymptomatic and diagnosis testing was not typically performed routinely in the early phases of the pandemic in the UK, serological assays facilitated a more precise estimation of the infection rate in an affected area [488].

At the time of this work, it had been shown that SARS-CoV-2 infected subjects developed IgM, IgA and IgG responses against S and N antigens in the first two weeks following infection, and these antibodies remained elevated after viral clearance [485,488,489]. Further, it was known that the S glycoprotein was the target of neutralising antibodies, and most of the highly potent neutralising mAbs isolated at the time targeted the RBD [490–492]. There had been a wide range of SARS-CoV-2 neutralising antibody titres reported in cross-sectional studies, and these varied depending on the length of time since infection and disease severity [493,494]. Data on the magnitude, kinetics and longevity of the neutralising antibody response against SARS-CoV-2, as determine via longitudinal studies, was limited and remained a somewhat debated topic [495,496]. Therefore, additional study of the neutralising antibody response following SARS-CoV-2 infection was required. Understanding the role these antibodies play in disease clearance and protection from reinfection was of great significance. Further, investigation of seroprevalence against SARS-CoV-2 in the community using serological methods was important. This would enable the definition of parameters in which these assays could be used to provide meaningful data in the absence of RT-PCR testing.

Neutralising antibody responses to other human coronaviruses had been shown to wane over time [497,498]. However, the rate of neutralising antibody decline varied between the endemic human coronaviruses and the highly pathogenic coronaviruses. For instance, neutralising antibody responses against endemic human coronaviruses had been shown to last as little as 12 weeks, but could persist following SARS-CoV and MERS-CoV infection for 12-34 months [499–501]. Therefore, it was crucial to ascertain the rate of neutralising antibody decline in SARS-CoV-2 infected subjects, and examine the impact of disease severity on such responses. These findings would have implications regarding protection against SARS-CoV-2 reinfection and vaccine durability, whilst also raising considerations for the convalescent plasma treatment trials that were being conducted at the time [495].

Overall, the specific aims of this chapter were to:

- Generate PCR and ELISA-confirmed SARS-CoV-2 infected cohorts, representing hospitalised and community populations respectively.
- Collect samples longitudinally and document the disease severity experienced by the subjects.
- Monitor IgM, IgA and IgG binding responses to S, RBD and N antigens following SARS-CoV-2 infection.
- Measure the neutralising antibody titres of these SARS-CoV-2 infected subjects, and consider correlations with antibody binding responses and kinetics.
- Examine the longevity of the neutralising and binding antibody responses following SARS-CoV-2 infection.
- Investigate the impact disease severity has on these findings.

### 4.1.2 Human subjects and ethics

Surplus serum from patient biochemistry samples taken as part of routine care was retrieved at the point of being discarded, aliquoted, stored and linked to a limited clinical dataset by the direct care team, before anonymization. Work was undertaken in accordance with the UK Policy Framework for Health and Social Care Research and approved by the Risk and Assurance Committee at Guy's and St Thomas' NHS Foundation Trust (GSTFT). Serum was collected from consenting HCWs with expedited approval from the GSTFT R&D office, occupational health department and medical director.

### 4.1.3 PCR-confirmed cohort

The neutralising antibody response, in the first three months following SARS-CoV-2 infection, was examined in a cohort of 65 real-time quantitative PCR (RT-qPCR) confirmed subjects (Table 4.1). The cohort was made up of 59 patients that had been admitted to GSTFT, and healthcare workers (HCW). The cohort was 78.5% male and the average age was ~55 years old, information regarding ethnicity was not collected. Each subject was assigned a severity score based on the maximal level of respiratory support needed throughout hospitalisation. The score ranged from 0-5, and was implemented to prevent underestimation of disease severity in subjects that did not require escalation above ward-based care. Disease severity was classified as follows:

- Asymptomatic or no requirement for supplemental oxygen.
- Requirement for supplemental oxygen (fraction of inspired oxygen ($FiO_2$) < 0.4) for at least 12 hours.
- Requirement for supplemental oxygen ($FiO_2 \geq 0.4$) for at least 12 hours.
- Requirement for non-invasive ventilation/continuous positive airway pressure or proning or supplemental oxygen ($FiO_2 > 0.6$) for at least 12 hours, and not a candidate for escalation above level one (ward-based) care.
- Requirement for intubation and mechanical ventilation or supplemental oxygen ($FiO_2 > 0.8$) and peripheral oxygen saturations <90% (with no history of type 2 respiratory failure; T2RF) or <85% (with known T2RF) for at least 12 hours.
- Requirement for extracorporeal membrane oxygenation (ECMO).

The cohort included 14, 10, 7, 2, 25 and 7 subjects with severity scores of 0, 1, 2, 3, 4 and 5 respectively. Numerous comorbidities were observed in this cohort, such as diabetes mellitus, hypertension and obesity (Appendix 7.5). Samples were collected sequentially at timepoints between 1 and 94 days post-onset of symptoms (POS). The number of samples was dependent on the availability of discarded serum taken as part of routine clinical care, or as part of a HCW study.

### 4.1.4 Antibody binding responses to SARS-CoV-2

ELISAs were performed using longitudinal samples to determine the level of serum IgM, IgA and IgG reactivity against S, RBD and N antigens (Figure 4.1). The optical density (OD) was measured for a total of 302 serum samples, collected from 65 subjects. The plasma samples were diluted 1:50, and a subject was considered seropositive when they had an OD that was four-fold above background. It was observed that 2/65 subjects had undetectable antibody responses against these antigens throughout their follow-up (Table 4.2). However, plasma samples were only available up until 2 and 8 days POS for these two individuals. The mean time to seroconversion against a single antigen in this cohort was 12.6 days POS. Therefore, it remained possible that these individuals could have seroconverted after being discharged from hospital. IgG reactivity against S, RBD and N antigens could be detected in 92.3%, 89.2% and 93.8% of subjects respectively (Table 4.2). Similarly, IgM reactivity against S, RBD and N antigens could be detected in 92.3%, 92.3% and 95.4% of subjects respectively. The frequency of subjects with a detectable IgA response to RBD and N was lower, 72.3% and 84.6% respectively. Conversely, the number of subjects generating an IgA response to S was similar to IgM and IgG (93.8%).

A cumulative frequency analysis was performed on positive IgG, IgA and IgM responses against S, RBD and N antigens. However, this did not indicate a more rapid elicitation of IgM and IgA responses against a particular antigen, possibly due to the sporadic nature in which the plasma samples had been collected (Figure 4.1a). Hence, a subset of subjects that had been sampled sequentially throughout early infection (<14 days POS) were examined independently, and different patterns of seroconversion were observed (Figures 4.2a-b). For instance, 51.6% of subjects seroconverted synchronously to IgM, IgA and IgG, whilst others seroconverted singularly to IgM (9.7%), IgA (9.7%) or IgG (9.7%). It was observed that 58.1% of subjects seroconverted synchronously to S, RBD and N antigens, whereas others seroconverted singularly to S (16.1%) or N (16.1%).

When analysed longitudinally, a rapid decline was seen for IgM and IgA responses to all three antigens following their peak OD at 20 and 30 days POS respectively (Figure 4.1b). This finding is comparable to what is typically observed during acute viral infections [497,502]. For certain individuals, IgM and IgA responses could be seen approaching baseline after 60 days POS (Figure 4.3b). On the other hand, IgG responses remained high in the majority of subjects and persisted up to 94 days POS in individuals that had been sampled until that timepoint (Figure 4.1b).

*Table 4.1: Cohort description: gender, severity, age and outcome.*

| | |
|---|---|
| **Gender** | |
| Male, no. (%) | 51 (78.5) |
| Female, no. (%) | 14 (21.5) |
| **Age** | |
| Mean (years) | 55.2 (23–95) |
| **Severity** | |
| 0 | 14 |
| 1 | 10 |
| 2 | 7 |
| 3 | 2 |
| 4 | 25 |
| 5 | 7 |
| **Outcome** | |
| HCW | 6 |
| Died | 12 |
| Discharged | 41 |
| Still in hospital | 5 |
| Transferred to local hospital | 3 |

### 4.1.5 Neutralising antibody responses to SARS-CoV-2

SARS-CoV-2 neutralisation capability was assessed using a surrogate viral inhibition assay. This assay employed HIV-1 based viral particles, which were pseudotyped with the ancestral S glycoprotein of SARS-CoV-2, and a HeLa cell-line stably expressing the ACE-2 receptor [503]. Neutralisation potency was quantified by the plasma inhibitory dilution at which 50% viral neutralisation was attained, commonly referred to as the $ID_{50}$. The neutralisation potency of the samples increased as the days POS increased, with each subject reaching a peak neutralisation titre on average 23.1 days POS (Figure 4.3a). The two individuals that were found to not seroconvert via ELISA also had undetectable neutralisation ($ID_{50}$ <50). The rest of the subjects could be categorised by their neutralisation potency. At peak, 7.7% had low ($ID_{50}$ 50-200), 10.8% medium ($ID_{50}$ 201-500), 18.5% high ($ID_{50}$ 501-2,000) and 60% potent ($ID_{50}$ >2,001) neutralisation titres. However, the percentage of subjects with potent neutralisation had reduced to 16.7% in plasma samples collected after 65 days POS. The $ID_{50}$ values of the plasma samples correlated well with OD values for the IgM, IgA and IgG responses against all three antigens (Figure 4.4a). Although, the best fits ($r^2$) were found between the $ID_{50}$ values and the OD values for IgM or IgA responses against S glycoprotein. Time to detectable neutralisation was on average 14.3 days POS, with some subjects developing neutralising activity before an IgG response against the S or RBD antigens could be detected by ELISA (Figure 4.2c). This emphasised that S and RBD-specific IgM and IgA antibody responses were capable of mediating neutralisation during acute infection in the absence of measurable IgG [504].

To examine the impact of disease severity on the neutralising antibody response, the peak $ID_{50}$ values were compared between subjects with severity scores of 0-3 and 4/5. It was observed that the magnitude of the peak neutralising antibody response was significantly higher in subjects with more

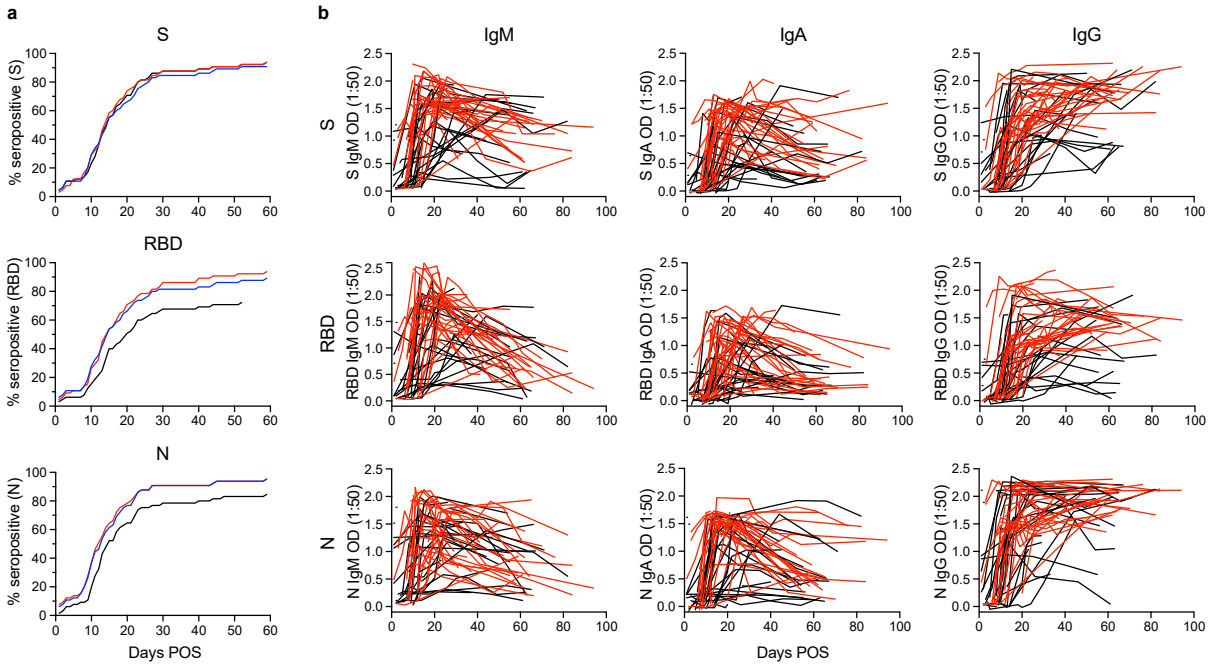



*Figure 4.1: Kinetics of antibody development against SARS-CoV-2 antigens over time following infection.*
*(**a**) Cumulative frequency analysis describing the point of seroconversion for each individual in the cohort. The graphs show the percentage of individuals in the cohort that became IgM, IgA or IgG positive to S glycoprotein, RBD and N protein, against days POS. A serum sample was considered positive when the OD was fourfold above background. (**b**) OD values at 1:50 serum dilution for IgM, IgA and IgG against S glycoprotein, RBD and N protein over time. Each line represents one individual (n = 65). Severities 0–3 are shown in black and severities 4/5 in red. More than 300 pre-COVID-19 healthy control samples and >100 sera from PCR-confirmed SARS-CoV-2-infected individuals were previously used to develop and validate the ELISA setup. The ELISAs were conducted once. Contributions to data collection from Jeffrey Seow, Sam Acors, Kathryn Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Isabella Huettner and Katie Doores.*

severe disease (Figure 4.5a). However, the mean time taken to develop a neutralising response and the mean time to reach peak neutralisation did not differ significantly between the two groups (Figures 4.5b-c). This suggested that more severe disease enhanced the magnitude of the neutralising antibody response, without altering its kinetics. The IgM, IgA and IgG OD values against S glycoprotein were compared in the plasma samples that corresponded with peak neutralisation. The OD values of IgM and IgA were shown to be significantly higher in subjects with more severe disease, whilst no significant difference was noted for IgG (Figures 4.5d–f). This further highlights the potential importance of IgM and IgA neutralising antibodies in acute infection [504]. Within the severity score 4/5 group, a proportion of subjects were undergoing treatment with immunomodulatory drugs in response to the development of a persistent hyperinflammatory state. It was initially hypothesised that individuals with this phenotype would have altered antibody responses. However, no significant difference was found between the peak $ID_{50}$ values of these subjects and the remainder of the severity 4/5 group (Figure 4.5g).

### 4.1.6 Longevity of the neutralising antibody response

Following the neutralisation peak, a waning in the $ID_{50}$ values was detected in subjects that had been sampled after 40 days POS. Examination of the $ID_{50}$ at peak neutralisation and the final timepoint collected revealed that a decrease in neutralisation potency occurred in almost all cases (Figure 4.6a). For some individuals, with a severity score of 0 and $ID_{50}$ values within the range of 100-500, neutralisation titres dropped below the limit of detection by the latest timepoint. For example, subjects 52 and 54 both produced lower peak neutralisation titres (174 and 434 respectively), which became undetectable 39 and 34 days later respectively (Figure 4.3b). A subset of subjects, that were representative of the range of neutralising antibody responses, were selected for neutralisation assay testing with authentic virus and Vero E6 cells. This was performed to establish whether similar neutralisation trends would be observed in comparison to the pseudovirus assays. In corroboration with other studies, it was found that neutralisation titres against authentic virus correlated very well ($r^2 = 0.9612$, $P < 0.0001$) with those ascertained using the pseudovirus assay (Figure 4.4b) [490,505,506]. Further, neutralisation of authentic virus was also undetectable ($ID_{50}$ <20) in the final samples collected from donors 52 and 54 (Figure 4.3b).

To get a more quantitative measure of the longevity of the IgG binding titre, additional ELISAs were carried out against all three antigens with the peak neutralisation and final timepoint samples. In this instance, the dilution effective in producing 50% of the maximal response was determined, commonly

*Table 4.2: Cumulative frequency of seropositivity for SARS-CoV-2 antigens following infection.*
*Contributions to data collection from Jeffrey Seow, Sam Acors, Kathryn Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Isabella Huettner and Katie Doores.*

| | Cumulative seropositivity (%) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | IgG | | | IgM | | | IgA | | | |
| Days POS | S | RBD | N | S | RBD | N | S | RBD | N | S, RBD or N |
| 7 | 10.8 | 10.8 | 15.4 | 12.3 | 10.8 | 15.4 | 12.3 | 6.2 | 9.2 | 21.5 |
| 14 | 47.7 | 47.7 | 58.5 | 52.3 | 49.2 | 61.5 | 49.2 | 32.3 | 48.7 | 66.2 |
| 21 | 67.7 | 69.9 | 76.9 | 75.4 | 72.3 | 80 | 70.8 | 52.3 | 64.6 | 86.2 |
| 28 | 83.1 | 80 | 90.8 | 84.6 | 81.5 | 90.8 | 86.2 | 64.6 | 76.9 | 93.8 |
| 35 | 84.6 | 81.5 | 90.8 | 87.7 | 86.2 | 90.8 | 87.7 | 67.7 | 78.5 | 93.8 |
| 42 | 86.2 | 83.1 | 90.8 | 89.2 | 89.2 | 90.8 | 89.2 | 69.2 | 80 | 93.8 |
| 49 | 89.2 | 86.2 | 93.8 | 90.8 | 90.8 | 93.8 | 90.8 | 70.8 | 83.1 | 95.4 |
| 56 | 90.8 | 87.7 | 93.8 | 92.3 | 92.3 | 93.8 | 92.3 | 72.3 | 83.1 | 95.4 |
| 63 | 92.3 | 89.2 | 93.8 | 92.3 | 92.3 | 93.8 | 93.8 | 72.3 | 84.6 | 96.9 |

known as the $ED_{50}$. Of note, a stronger correlation was witnessed between the $ID_{50}$ and $ED_{50}$ values compared to what was seen between the $ID_{50}$ and OD values (Figure 4.6e). Similar to what was observed with neutralisation potency, a reduction in the $ED_{50}$ values was measured between the timepoints tested, for all three antigens (Figures 4.6b–d). Individuals that had $ID_{50}$ values approaching baseline also showed IgG $ED_{50}$ values to the RBD and S glycoproteins that decreased in a similar manner. To examine if the reductions in IgG titres were reaching a plateau, $ED_{50}$ values were measured for all of the samples taken from four representative subjects. These individuals showed a steady decline in neutralisation and IgG $ED_{50}$ to all antigens, during the time frame that was studied (Figure 4.6f). However, additional observations of antibody binding and neutralisation after 94 days POS would be required to fully determine the longevity of the neutralising antibody response.

### 4.1.7 ELISA-confirmed HCW cohort

To further investigate antibody responses to SARS-CoV-2 infection, sequential plasma samples were analysed from 31 seropositive HCWs from GSTFT. The antibody responses in these subjects were likely to be more similar to those from SARS-CoV-2 individuals that had not required hospitalisation. Identification of acute SARS-CoV-2 infection via RT-qPCR was not performed routinely at the time in these subjects, therefore ELISA was used to confirm infection. Plasma samples were collected longitudinally every 1-2 weeks from subjects, between 13/03/2020 and 10/06/2020. Plasma samples were diluted 1:50 and assessed via ELISA, the HCWs were considered seropositive when IgG OD values against both S and N antigens were fourfold above negative control plasma. Symptoms related to COVID-19 infection were self-reported by the subjects, and days POS in seropositive individuals was concluded based on this information. In asymptomatic subjects, days POS was defined from the first sample in which SARS-CoV-2 antibodies could be detected.

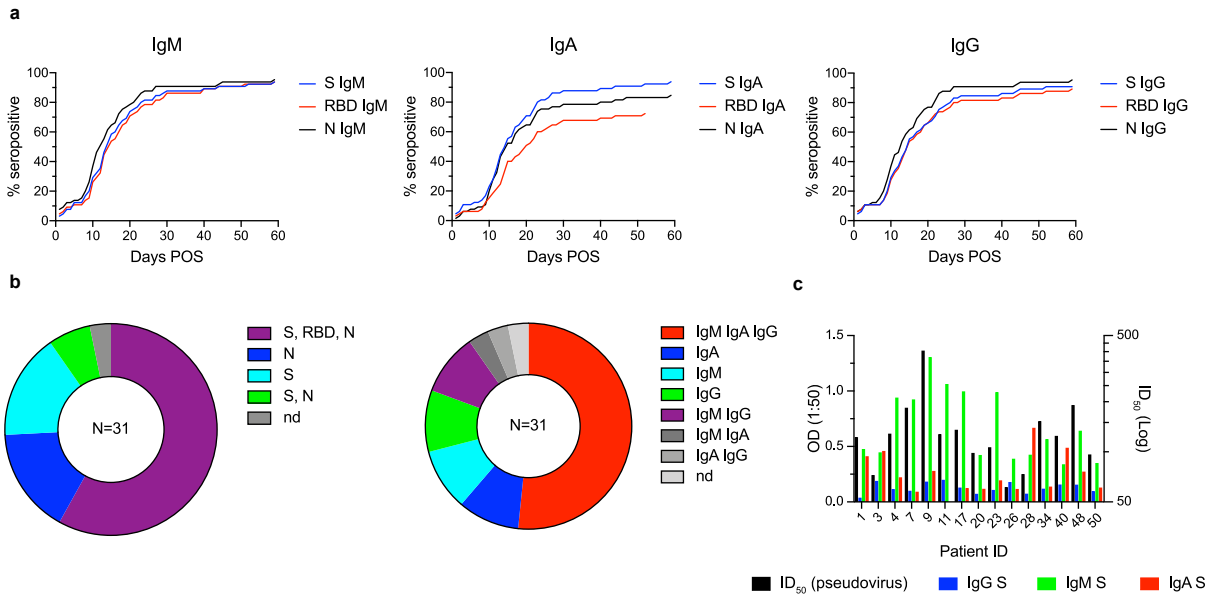



*Figure 4.2: Antibody isotype and specificity at the timepoint for which an individual became seropositive for SARS-CoV-2. (**a**) Cumulative frequency analysis describing the seroconversion in the hospitalised cohort. The graphs show the percentage of individuals in the cohort that become IgM, IgA or IgG seropositive against S glycoprotein, RBD and N protein, against days POS. (**b**) Patterns of seroconversion based on antigen and antibody isotype for a subset of donors from which sera was collected over sequential early time points (<14 days POS, n=31). For individuals with at least 2 sera collected <14 days POS, the first antigen(s) and antibody isotype(s) that gave an OD that was 4-fold above background is reported. (**c**) Pseudovirus neutralisation and ELISA binding at early time points POS for a subset of individuals (n=15) where there were measurable neutralisation titres but IgG binding to S was low (OD <0.2) and IgM and/or IgA to S is higher (OD >0.4). Each bar shows data for one patient only. Contributions to data collection from Jeffrey Seow, Sam Acors, Kathryn Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Isabella Huettner and Katie Doores.*

### 5.1.8 Antibody responses in HCW cohort

For each subject, the IgM and IgG binding to all three antigens and pseudovirus neutralisation titre was measured longitudinally (Figures 4.7a-b). Similarly to the PCR-confirmed cohort, $ID_{50}$ values correlated with the OD values for IgM and IgG binding to S and RBD antigens (Figure 4.7e). Although, the IgM and IgG responses to N protein correlated poorly with the $ID_{50}$ values ($r^2 = 0.030$ and 0.381 respectively) in the HCWs. The mean peak $ID_{50}$ values of asymptomatic HCWs (6/31) and symptomatic HCWs (25/31) were very similar, but lower compared to subjects with severity scores of 0-3 and 4/5 (Figure 4.7d). Interestingly, a few asymptomatic subjects developed neutralisation titres >1,000. Similarly to the PCR-confirmed cohort, a waning of neutralising antibody responses was observed in the HCWs (Figure 4.7c). For numerous subjects, with a peak $ID_{50}$ within the range of 100–500,

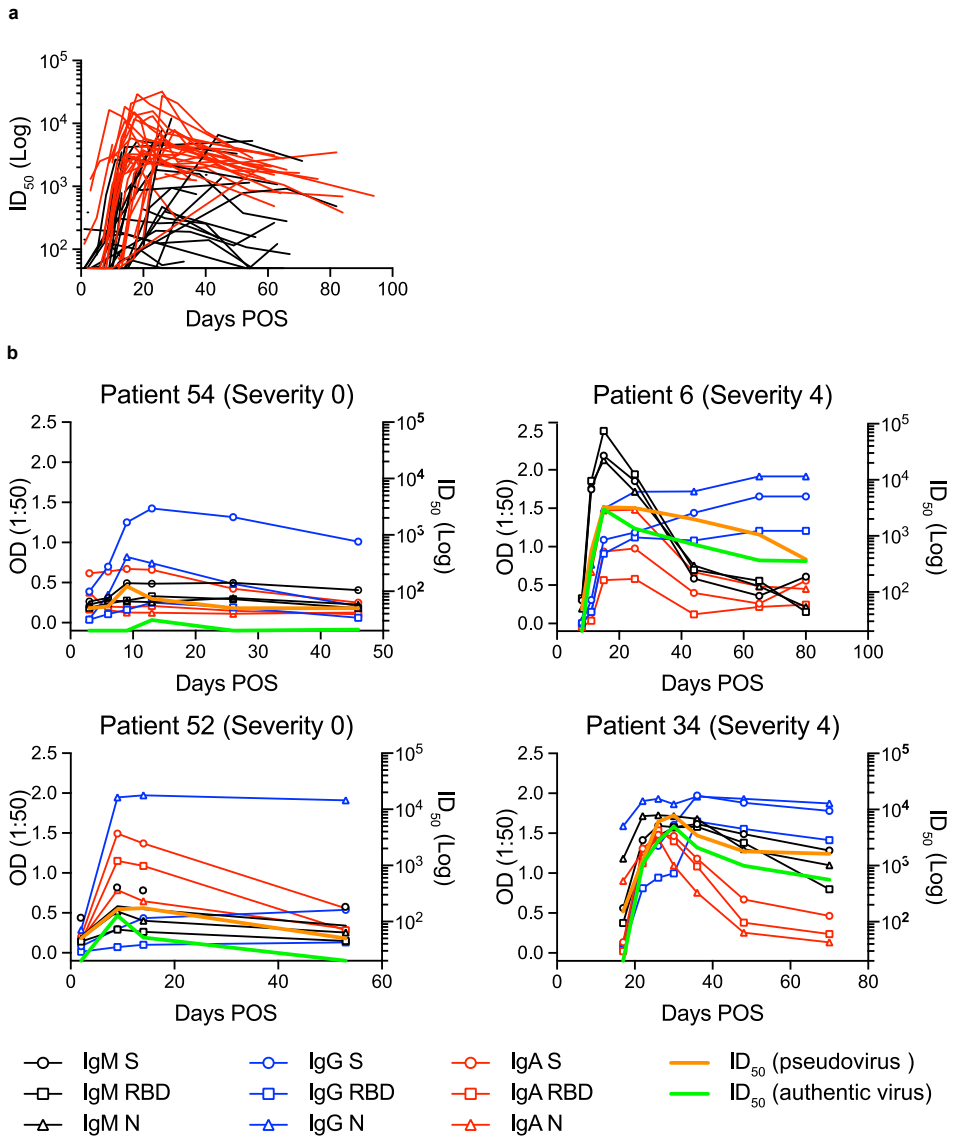



*Figure 4.3: Kinetics of neutralising antibody responses in SARS-CoV-2 infection.*
*(**a**) Neutralising antibody $ID_{50}$ changes related to days POS. Each line represents one individual (n = 65). Severities 0–3 are shown in black and severities 4/5 in red. A subset of neutralisation experiments (n = 25) was conducted twice, yielding similar results. $ID_{50}$ values for the remaining samples were measured once. (**b**) Example kinetics of antibody responses (IgM, IgA, IgG binding to S glycoprotein, RBD and N protein, and $ID_{50}$ against PVs and authentic virus) for four individuals during acute infection and the convalescent phase. The graphs show comparison between disease-rated severity 0 (left) and disease-rated severity 4 (right). The cut-off for the pseudovirus and authentic virus neutralisation assays are 1:50 and 1:20, respectively. Contributions to data collection from Jeffrey Seow, Sam Acors, Kathryn Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Isabella Huettner, Suzanne Pickering and Katie Doores.*

neutralisation reached the limit of detection after 50 days POS. Due to the lower mean peak $ID_{50}$ values in the HCWs, decline of neutralisation towards baseline was observed more frequently compared to the PCR-confirmed cohort.

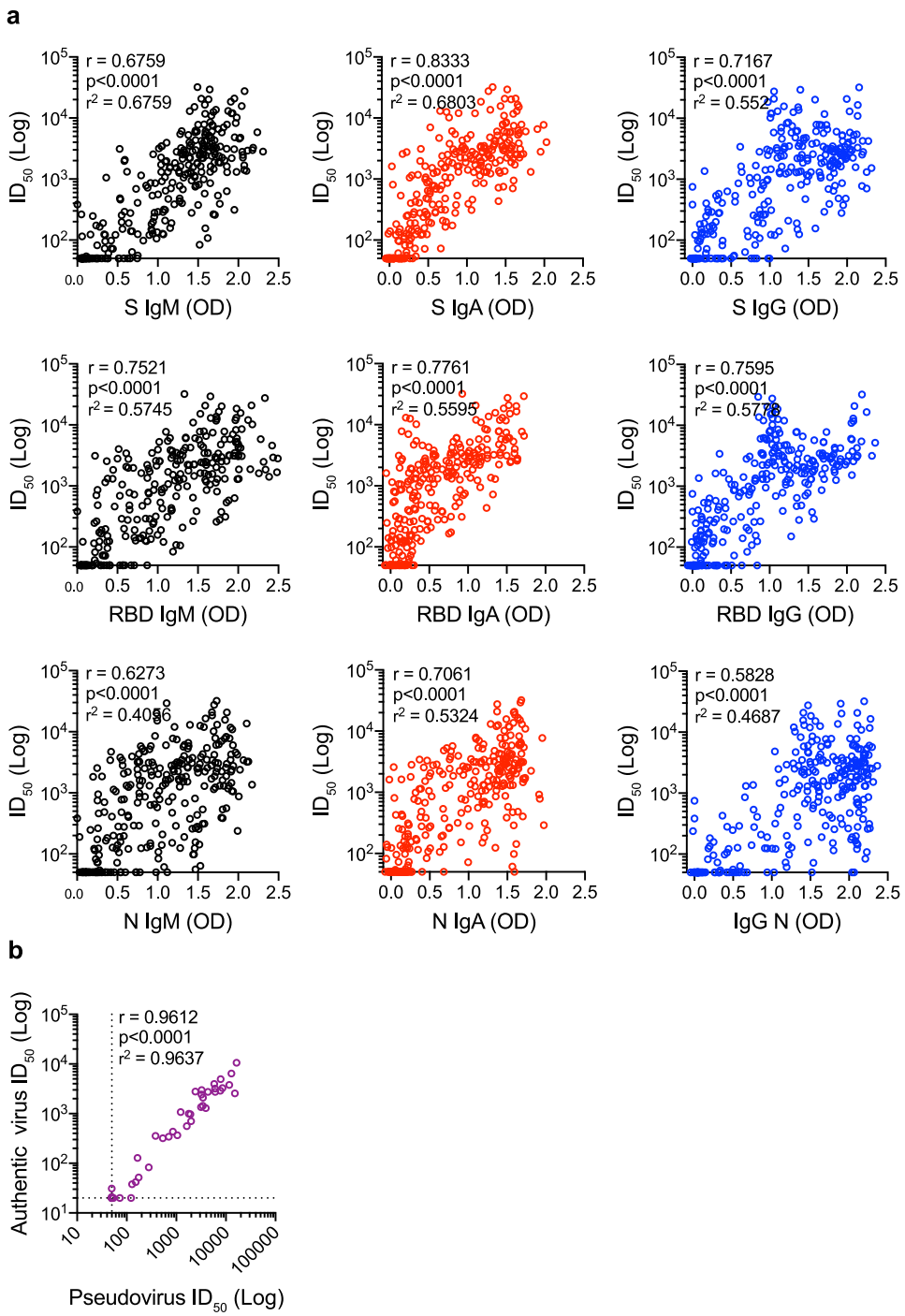



*Figure 4.4: Correlations between binding to different SARS-CoV-2 antigens and neutralisation potency following infection. (**a**) Correlation of $ID_{50}$ measurements against SARS-CoV-2 pseudovirus with IgG (blue), IgM (black), IgA (red) OD values (at 1:50) against S glycoprotein, RBD and N protein (Spearman correlation, r). (**b**) Correlation of $ID_{50}$ measured using authentic virus and $ID_{50}$ measured using pseudovirus neutralisation assays (Spearman correlation, r). A linear regression was used to calculate the goodness of fit ($r^2$). The dotted lines represent the lowest serum dilution used in each assay. The lowest dilution used for the pseudovirus and authentic virus neutralisation assays are 1:50 and 1:20, respectively. Contributions to data collection from Jeffrey Seow, Sam Acors, Kathryn Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Isabella Huettner, Suzanne Pickering and Katie Doores.*

### 4.1.9 Discussion

The collection and evaluation of sequential serum samples from SARS-CoV-2 infected subjects, enabled characterisation of the kinetics and longevity of the neutralising antibody response in greater detail than had been already reported. The kinetics of the antibody response against SARS-CoV-2 found here was typical of an acute viral infection [497,502]. The peak in neutralising antibodies occurs as a result of the rapid generation of short-lived antibody secreting plasma cells, these cells subsequently die and a there is a reduction in virus-specific antibodies [507]. A wide range of peak neutralising antibody titres were detected ($ID_{50}$ 98–32,000) at ~23 days POS, which was similar to what was seen by other cross-sectional studies [493,508]. Notably, more severe disease was associated with higher neutralising

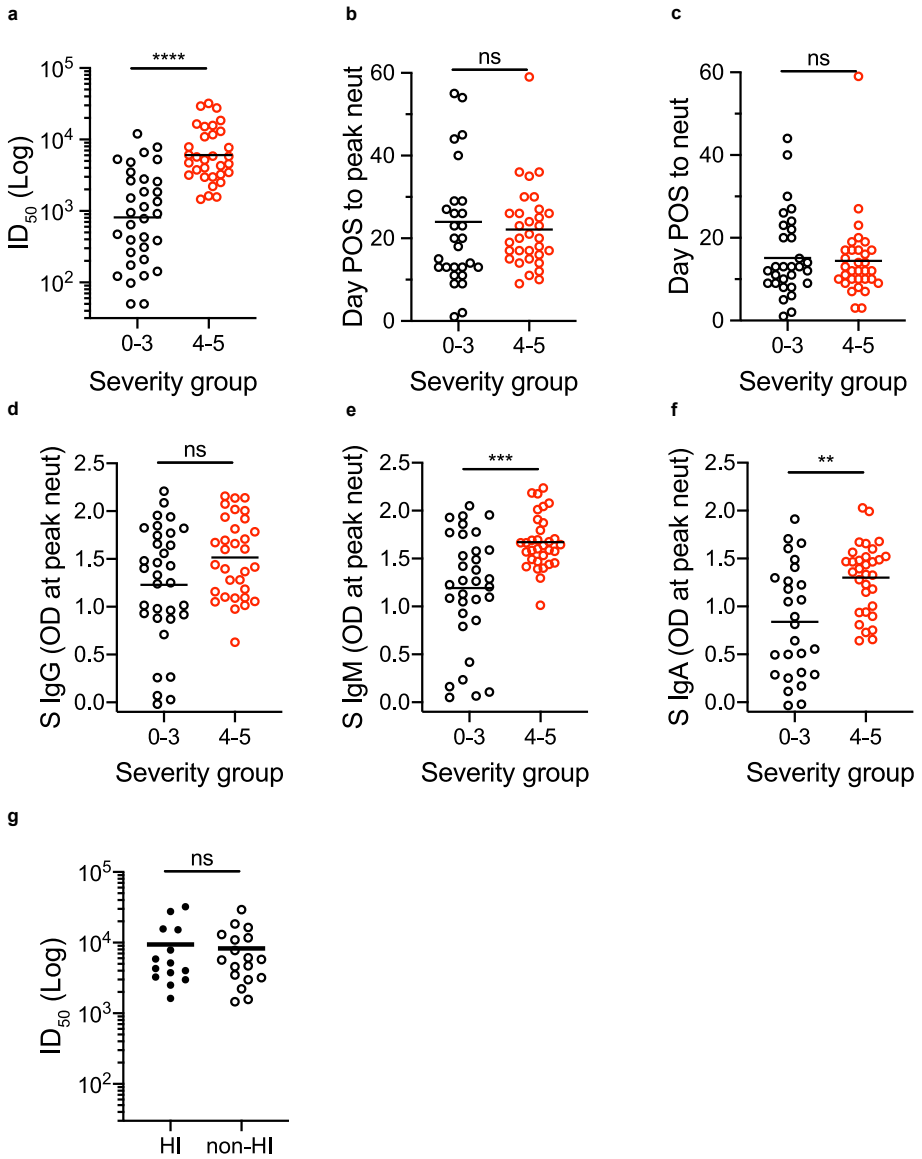



*Figure 4.5: Impact of disease severity on antibody responses in SARS-CoV-2 infection.*
*(**a–c**) Comparison for individuals with disease severity 0–3 (n = 33 individuals) or 4/5 (n = 32 individuals) for peak $ID_{50}$ of neutralisation assays, the time POS to reach peak $ID_{50}$ and the time POS to detect neutralising activity. $ID_{50}$ measured using HIV-1-based virus particles, pseudotyped with S glycoprotein of SARS-CoV-2. (**d–f**) Comparison of OD values for individuals with 0–3 or 4/5 disease severity for IgG, IgM and IgA against S glycoprotein measured at peak $ID_{50}$. (**g**) Comparison of the peak $ID_{50}$ value for individuals who were treated for hyperinflammation (HI; n = 14 individuals) or not treated (n = 18 individuals), and had 4/5 disease severity (P > 0.999). Statistical significance was measured using a Mann–Whitney two-sided test U-test. **P < 0.002, ***P < 0.0002, ****P < 0.0001 and ns (not significant). The line represents the mean $ID_{50}$ for each group. Contributions to data collection from Jeffrey Seow, Sam Acors, Kathryn Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Isabella Huettner, Suzanne Pickering and Katie Doores.*

antibody titres. Although this has been found in other studies, it remains unclear why this correlation exists [509,510]. It is possible that a higher viral load could cause more severe disease and in turn generate a stronger antibody response due to increased levels of viral antigen [510]. Alternatively, high antibody levels may play a causative role in disease severity, through mechanisms such as antibody-dependent enhancement [511,512].

By comparing the peak $ID_{50}$ from each subject with the $ID_{50}$ from the corresponding final timepoint, it was shown that the neutralisation titre declined irrespective of disease severity. There was also a decline in IgG-binding titres ($ED_{50}$) to S and RBD antigens, as well as IgM and IgA binding to S and RBD antigens (OD values) in the PCR-confirmed cohort. For some subjects with peak $ID_{50}$ values ranging from 100–500, neutralising titres reached or dropped below the level of detection ($ID_{50}$ <50) after only ~50 days from when the peak was measured. Although, IgG binding to S, RBD and N antigens was still detectable in these subjects at the later timepoints. This trend was also found in the HCW cohort,

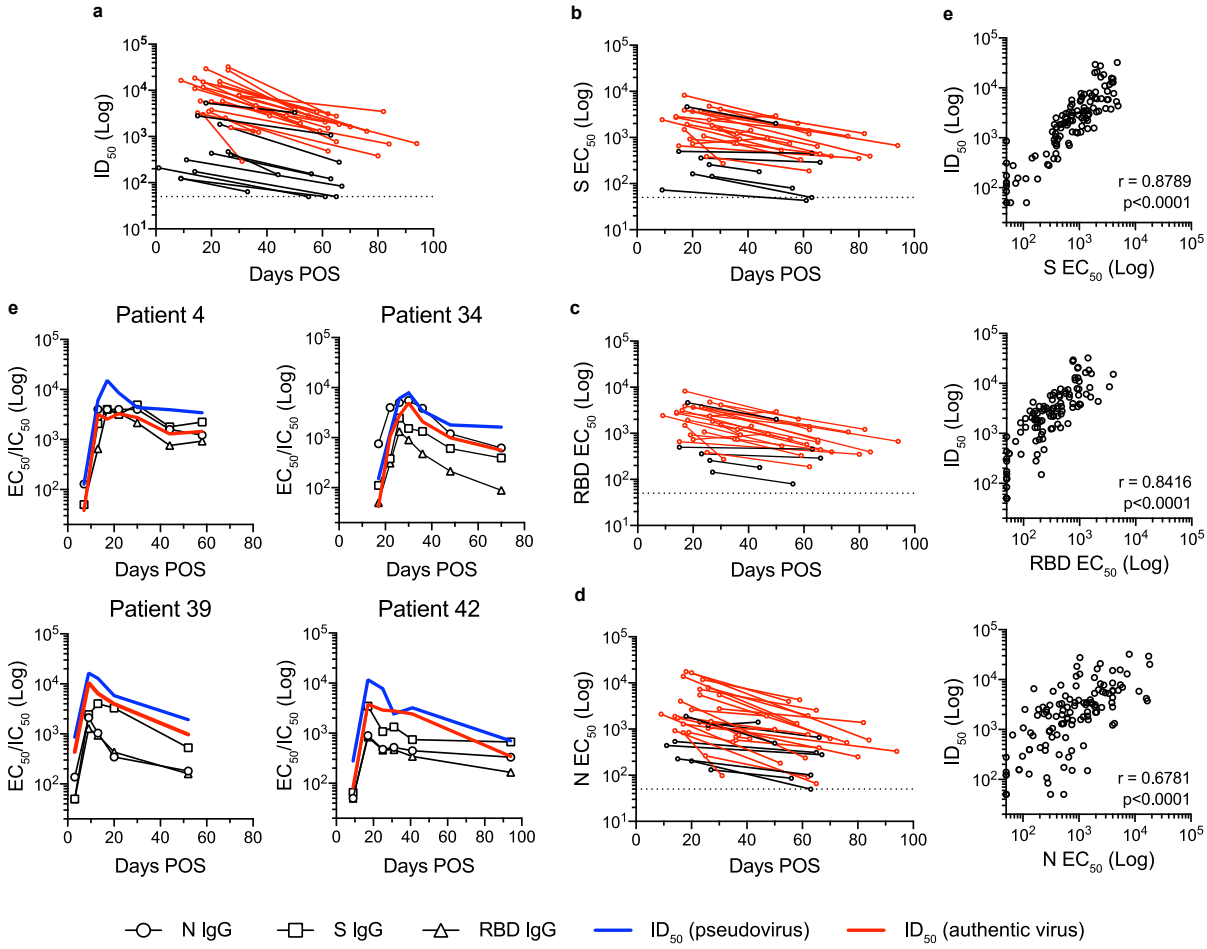



*Figure 4.6: Short-term longevity of the neutralising antibody response following SARS-CoV-2 infection.*
*(**a**) $ID_{50}$ at peak neutralisation is plotted with the donor-matched $ID_{50}$ at the last timepoint from which serum was collected. Only individuals with peak $ID_{50}$ occurring before the last timepoint and with a last timepoint >30 d POS were included in this analysis (n = 35). The dotted line represents the cut-off for the pseudovirus neutralisation assay. $EC_{50}$ values for IgG binding to S glycoprotein (**b**), RBD (**c**) and N protein (**d**) were calculated at timepoints with peak $ID_{50}$ and the final timepoint sera were collected. $EC_{50}$ at peak neutralisation is plotted with the donor-matched $EC_{50}$ at the last timepoint sera were collected. Individuals with a disease severity 0–3 are shown in black and those with 4/5 in red. The dotted line represents the cut-off for $EC_{50}$ measurement. The ELISAs to determine $EC_{50}$ values were conducted once. (**e**) Correlation of $ID_{50}$ with IgG $EC_{50}$ against S glycoprotein, RBD and N protein (Spearman's correlation, r). A linear regression was used to calculate the goodness of fit ($r^2$). (**f**) Change in IgG $EC_{50}$ measured against S glycoprotein, RBD and N protein, and $ID_{50}$ using pseudovirus and authentic virus over time for four example patients (all disease severity 4). The lowest dilution used for the pseudovirus and authentic virus neutralisation assays are 1:50 and 1:20, respectively. Contributions to data collection from Jeffrey Seow, Sam Acors, Kathryn Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Isabella Huettner, Suzanne Pickering and Katie Doores.*

suggesting that in some subjects, SARS-CoV-2 infection only yields a transient neutralising antibody response that rapidly wanes. However, for most subjects with a peak $ID_{50}$ value >4,000, even though a 2 to 23-fold decrease was witnessed over an 18 to 65 day period, neutralising antibody titres remained detectable at the final timepoint measured ($ID_{50}$ 1,000–3,500). Of note, whilst the lowest serum dilution applied in the pseudovirus neutralisation assay was relatively high (1:50), subjects that lacked detectable neutralisation via this method also lacked neutralisation when tested against authentic virus with at a lower dilution (1:20).

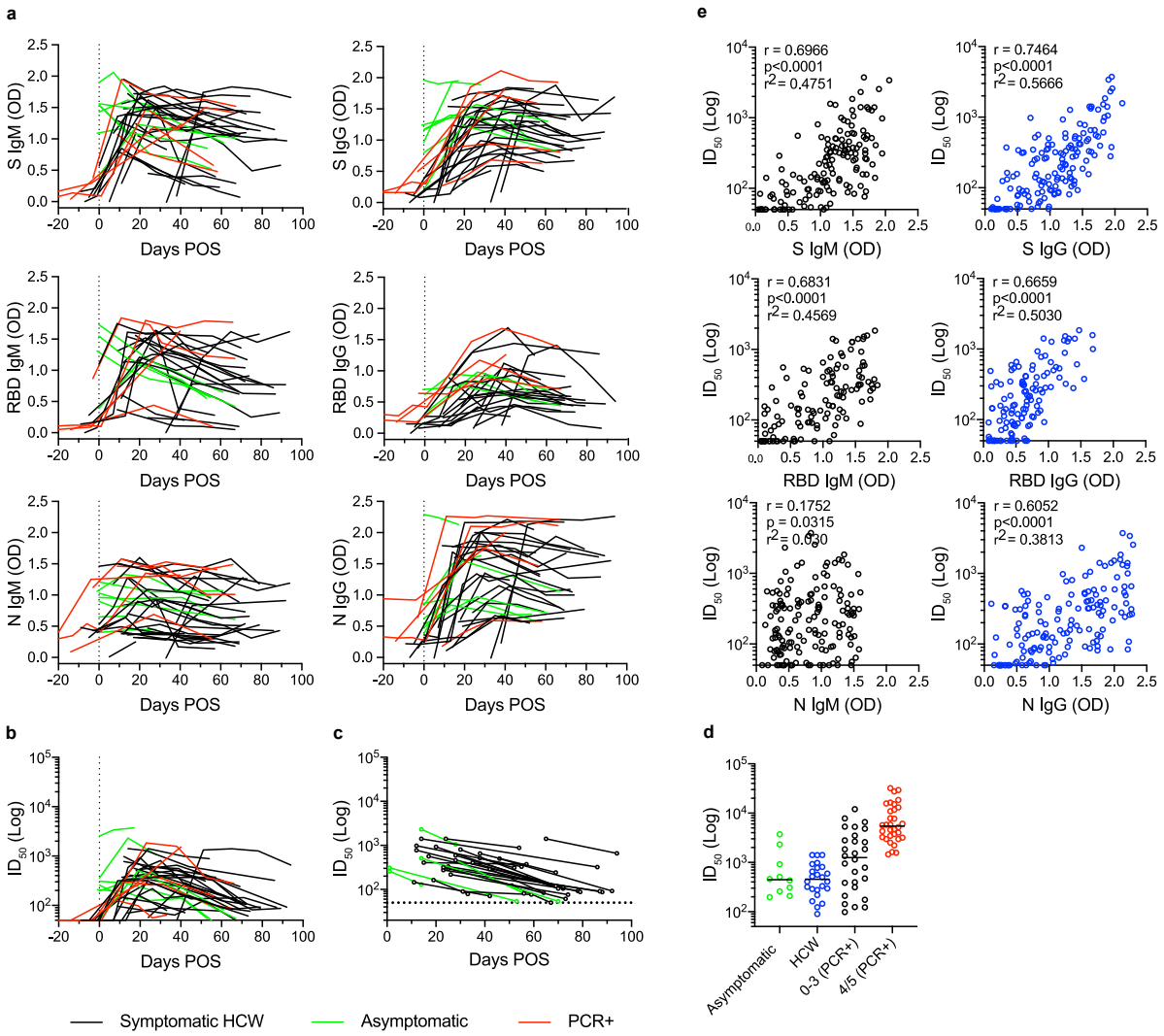



*Figure 4.7: Binding and neutralising antibodies responses against SARS-CoV-2 in seropositive HCWs.*
(**a**) *Antibody responses (IgG and IgM to S glycoprotein, RBD and N protein) over time. Each line represents one individual. Asymptomatic individuals shown in green, symptomatic individuals shown in black and PCR-confirmed SARS-CoV-2 infected individuals shown in red (for comparison). The dotted line represents day 0 POS.* (**b**) *$ID_{50}$ values plotted against the time POS at which sera were collected. Each line represents one individual (n = 37). Of seropositive individuals, 80.6% (25/31) recorded COVID-19-compatible symptoms (including fever, cough and anosmia) since 1 February 2020, and 19.4% (6/31) reported none. A subset of neutralisation experiments (n = 10) were conducted twice, yielding similar results. $ID_{50}$ values for the remaining samples were measured once.* (**c**) *$ID_{50}$ at peak neutralisation is plotted with the donor-matched $ID_{50}$ at the last timepoint sera were collected. The dotted line represents the cut-off (1:50) for the SARS-CoV-2-pseudovirus neutralisation assay.* (**d**) *Comparison of the peak $ID_{50}$ between asymptomatic individuals (n = 10, includes 7 HCW and 3 hospital patients), HCWs (n = 24 symptomatic HCWs with no PCR test), and PCR+ individuals with severity either 0–3 (n = 28) or 4/5 (n = 32). The two PCR+ individuals, who were sampled at early timepoints (<8 d POS) and did not seroconvert, were not included in this analysis.* (**e**) *Correlation of $ID_{50}$ with IgG and IgM OD values against S glycoprotein and N protein (Spearman correlation, r). A linear regression was used to calculate the goodness of fit ($r^2$). The ELISA assays were conducted once. A subset of neutralisation experiments were conducted twice yielding similar results. Contributions to data collection from Jeffrey Seow, Sam Acors, Kathryn Steel, Oliver Hemmings, Aoife O'Byrne, Neophytos Kouphou, Isabella Huettner, Suzanne Pickering and Katie Doores.*

These findings were novel at the time of this work, but subsequent reports were later released regarding early antibody responses to SARS-CoV-2 infection. In corroboration with this work, it has been shown that subjects with high $ID_{50}$ values early in infection retained high neutralisation titres regardless of the initial decline [303,495,513,514]. Further, in some subjects with lower $ID_{50}$ values and less severe disease, neutralising antibody titres were shown to decline to undetectable levels ($ID_{50}$ <20 or <50) at the final timepoint examined [495,513]. Conversely, some studies reported a sustained antibody response during the first 3 months following SARS-CoV-2 infection. Of note, these papers reported changes in antibody binding, or neutralising endpoint titres which are less quantitative compared to $ID_{50}$ values [496,515]. Although antibody binding correlates with neutralising antibody titres in some instances, this difference in measurement may account for the discrepancies in the kinetics described. Nonetheless, class-switched IgG memory B-cells against S glycoprotein have been found to persist in COVID-19 patients and vaccine recipients, verifying that long-lived memory responses can be mounted [516,517]. These memory B-cells, together with T-cells, can become activated in the event of re-exposure and may prevent severe disease [518,519].

To understand if the neutralising antibody titre declination would continue its trajectory or plateau to a steady state, a follow-up study was conducted using the same cohort [520]. This study, alongside others, found that although there was initially a rapid reduction in the neutralising antibody titres, the rate of decline slowed around 4-7 months POS [517,521–523]. This was attributed to the generation of long-lived plasma cells producing IgG antibodies specific for S glycoprotein [524]. Neutralising activity was detected in 18 of the 19 serum samples taken 257-305 days POS, with a geometric mean titre (GMT) of 640. Importantly, the follow-up cohort was also tested against the VOCs circulating at the time. It was shown that there was a 3.4-fold and 8.9-fold decrease in neutralisation potency against B.1.1.7 (α) and B.1.351 (β) respectively. At peak neutralisation, high GMTs of 3,331 and 1,303 were measured against B.1.1.7 (α) and B.1.351 (β) respectively. Further, neutralising activity could still be detected against these VOCs in the majority of subjects at 257-305 days POS. Remarkably, the fold reduction in neutralisation of these VOCs compared to WT decreased as days POS increased. This finding implied that the samples from later timepoints were able to tolerate S glycoprotein mutations better than those from earlier timepoints [520]. Indeed, it had been found that SARS-CoV-2 mAbs isolated at 6 months POS had higher levels of SHM and were more resistant to RBD mutations [525].

Previous studies have examined the longevity of neutralising antibody responses to other human coronaviruses [497–499]. Unlike SARS-CoV-2, infection with SARS-CoV generally results in more severe disease, with asymptomatic and low severity cases being less common [526]. In a hospitalised SARS-CoV infection cohort, neutralising antibody titres peaked at ~30 days POS (mean 1:590) with IgG binding and neutralising antibodies waning throughout a 3-year follow-up. Low neutralising antibody titres of 1:10 were measured in 17 of the 18 subjects tested, at 540 days POS [498]. Another study found low neutralising antibody titres (mean 1:28) could be detected in 89% of subjects, up to 36 months POS [500]. Initially, it was believed that the low neutralising antibody responses found in the 0-3 severity group could have been more akin to immune responses against the endemic human coronavirus. These responses have been reported to be more transient and reinfections are possible [501,527]. For instance, subjects experimentally infected with 229E produced high neutralising antibody titres after 2 weeks, which rapidly declined in the subsequent 11 weeks and reduced even further by 1 year. Later viral challenge resulted in reinfection (established by viral shedding), but the subjects were asymptomatic [497]. Meta-analysis of SARS-CoV-2 cases has estimated that the reinfection rate is ~0.65%, in which the reinfection can be symptomatic or asymptomatic. However, this report does acknowledge that viral mutations present in VOCs could be the cause of reinfection in many of these cases [528].

The neutralising antibody titre needed to prevent reinfection with SARS-CoV-2 is not fully understood. Neutralising mAbs isolated from SARS-CoV-2 infected subjects have been shown to provide protection from disease during challenge experiments in animal models. This occurs in a dose-dependent manner, emphasising neutralising antibody titres as a correlate of protection [491,492,529]. A study found that rhesus macaques infected with SARS-CoV-2 developed neutralising antibody titres of ~100, and displayed no clinical signs of illness when challenged 35 days after primary infection. However, virus could be detected in nasal swabs, and although the viral load was considerably lower compared to what was seen during primary infection, this indicated that there was immunological control instead of sterilising immunity [530]. However, another study was not able to detect virus after a second challenge, with the rhesus macaques developing neutralising antibody titres ranging from 8-20 [531]. Further insights into the neutralising antibody titres needed for protection from reinfection can be found in vaccine challenge studies. In rhesus macaques the BNT162b2 and mRNA-1273 vaccines was able to generate GMTs of 1,689 and 3,481 respectively, and this was shown to be capable of preventing viral replication in the upper and lower respiratory tracts [532,533]. The importance of T-cell responses generated through infection or vaccination also require consideration in regards to protection from reinfection [519]. While some progress has been made, more definitive correlates of protection are still needed [534,535]. This would enable a better comprehension of the significance of the neutralising antibody decline witnessed here. Follow-up investigations with sequential PCR testing and the use of standardised serological analysis in larger cohorts will be crucial for understanding the capacity of neutralising antibodies in preventing reinfections in humans.

In summary, examination of sequential samples from subjects that had been infected with SARS-CoV-2 revealed an antibody response that was typical of an acute viral infection, peaking at ~23 day POS and subsequently waning. Most subjects generated a robust neutralising response, with $ID_{50}$ values remaining >1,000 despite the initial decline. However, some individuals with less severe disease developed lower neutralising antibody titres, which declined below the limit of detection after ~50 days from when the peak was measured. Follow-up analysis of these subjects found that after this initial a rapid reduction, the rate of antibody decline slowed at around 4-7 months POS. The majority of these subjects were able to cross-neutralise VOCs, albeit at a lower neutralisation potency compare to the ancestral strain, and neutralising antibody titres remained detectable at 257-305 days POS. The fold reduction in neutralisation activity against VOCs compared to WT decreased as days POS increased. Still, further studies are required to understand the neutralising antibody threshold required for protection against reinfection with SARS-CoV-2.

# Chapter 5: Isolation of neutralising mAbs antibodies against SARS-CoV-2

## 5.1 Neutralising mAbs isolated from individuals infected with SARS-CoV-2

### 5.1.1 Introduction

The emergence of SARS-CoV-2 was still a relatively recent event by the time of this work, but much was known about the antibody response following infection at the monoclonal level. It had been shown that most SARS-CoV-2 infected individuals developed an antibody response 5-15 days POS, which peaked after ~3-4 weeks and subsequently declined [495,513,536]. The magnitude of the neutralising antibody response was associated with disease severity. Such that, individuals that experienced more severe disease generated a stronger antibody response, whilst those that were asymptomatic or had mild disease displayed lower neutralising antibody titres [509,510,536] SARS-CoV-2 neutralizing antibody levels are correlated with severity of COVID-19 pneumonia. It had been reported that antibodies targeting the RBD accounted for >90% of the neutralising activity measured in convalescent sera [537,538]. Several highly potent neutralising mAbs had been reported, however the majority were isolated from subjects with moderate to severe COVID-19 symptoms [490,492,539]. Most of these mAbs were RBD-specific, and numerous distinct epitopes had been found in this region [490,540]. A number of RBD-specific mAbs were being investigated as potential therapeutic options for COVID-19 [382,384,389]. Some NTD-specific mAbs had been identified with neutralising activity, indicating they may also be useful as a COVID-19 prophylaxis or treatment [541,542]. Therefore, more research was required on neutralising antibody responses at the monoclonal level in subjects that had experienced different levels of disease severity, examining neutralising epitopes located across the entire S glycoprotein.

The SARS-CoV-2 D614G variant became the globally dominant strain at an early stage in the pandemic [543]. This mutation was reported to increase infectivity by stabilising the RBD in the up conformation, but was not associated with neutralisation escape [544,545]. At the time of this work, the B.1.1.7 (α) variant had recently emerged and it contained eight mutations across the S glycoprotein (ΔH69/V70, ΔY144, N501Y, A570D, P681H, T716I, S982A and D1118H) [546]. B.1.1.7 (α) was associated with more efficient transmission and had become the dominant strain in London and the South East of England [547]. It was unknown if these mutations arose stochastically or if they had been selected as a result of their increased transmission capability [548]. It was suggested that these mutations could have been driven by neutralising antibodies, during longer term infection in immunocompromised patients undergoing passive immunotherapy [549–551]. It was also unclear if B.1.1.7 (α) would escape neutralising antibody responses that had been generated during wave 1 and/or through vaccination with the ancestral strain. Early studies had suggested that B.1.1.7 (α) was sensitive to polyclonal sera from vaccine recipients and patients infected with early circulating SARS-CoV-2 variants, but less was known about escape at the monoclonal level [552–554].

Overall, the specific aims of this section were to:

- Examine the plasma of subjects that had experienced different levels of COVID-19 disease severity, and identify good candidates for mAb isolation [507].
- Perform antigen-specific B-cell sorting on selected donors, generating a functional mAb library.
- Sequence variable regions of mAbs, and examine features such as germline gene usage, level of SHM and CDRH3 length.
- Characterise the regions of S glycoprotein bound by mAbs through ELISA, and measure neutralisation capability using pseudovirus and authentic virus neutralisation assays.

- Perform competition ELISAs on neutralising mAbs to identify distinct neutralising epitopes, then carry out further competitions with mAbs that have been previously described.
- Investigate the ability of the mAbs to inhibit interactions between S glycoprotein and ACE2.
- Examine the impact of S glycoprotein glycan heterogeneity on mAb neutralisation.
- Inspect the impact of S glycoprotein mutations from recent VOCs on the neutralising activity of donor plasma and mAbs.

### 5.1.2 Human subjects and ethics

This study used human samples from three individuals collected as part of the COVID-IP study [507]. All donors were male and P003, P008, P054 were 63, 29 and 62 years old, respectively. The study protocol for patient recruitment and sampling, out of the intensive care setting, was approved by the committee of the Infectious Diseases Biobank of King's College London with reference number COV-250320. The protocol for healthy volunteer recruitment and sampling was similarly approved by the same committee as an amendment to an existing approval for healthy volunteer recruitment with reference number MJ1-031218b. Both approvals were granted under the terms of the Infectious Disease Biobank's ethics permission (reference 19/SC/0232) granted by the South Central Hampshire B Research Ethics Committee in 2019. Patient recruitment from the ICU was undertaken through the ethics for the IMMERSE study approved by the South Central Berkshire Ethics Committee with reference number 19/SC/0187. Patient and control samples and data were anonymized at the point of sample collection by research nursing staff or clinicians involved in the COVID-IP project. We complied with all relevant ethical regulations.

### 5.1.3 Neutralising antibody responses following SARS-CoV-2 infection differ between donors with varied COVID-19 symptoms

In order to study a wide range of SARS-CoV-2 neutralising antibodies at the monoclonal level, three donors that had experienced different COVID-19 disease severities were examined that had been infected with the ancestral strain. P008 was asymptomatic, P054 was symptomatic but not hospitalised and P003 was hospitalised, during which they spent time in the intensive care unit (ICU). Of note, SARS-CoV-2 infection was only identified in P008 via serological screening [507]. ELISAs and pseudovirus neutralisation assays were used to investigate the longitudinal plasma samples taken from each subject, determining the binding and neutralisation titres. In corroboration with previous work, the highest neutralisation titre was detected in the subject that had the most severe disease ($ID_{50}$ 9,181) (Figure 5.1a) [509,510,536]. The symptomatic donor has the next highest neutralisation titre ($ID_{50}$ 3,936), whilst the lowest was found in the asymptomatic subject ($ID_{50}$ 820). The neutralising antibody response of these donors waned throughout the convalescent period, $ID_{50}$ values decreased to 25, and 258 in P008 and P054 respectively. Due to low a low sample size statistical analysis was not performed on this data. Levels of plasma IgG to the S and RBD antigens remained detectable in these donors, but a large decline from peak was observed. Notably, the levels of plasma IgM and IgA were still detectable in P054, but had reached baseline in P008. Unfortunately, there was a lack of sequential samples available from P003 to assess the longevity of their plasma antibody response.

### 5.1.4 Antibodies generated following SARS-CoV-2 infection target multiple epitopes across the S glycoprotein

Antigen-specific B-cell sorting was performed to isolate mAbs that were specific for the S glycoprotein of SARS-CoV-2. The PBMCs used for sorting were collected at 61, 48 and 20 days POS for P008, P054 and P003 respectively. An uncleaved and artificially trimerised (added C-terminal T4 fibritin trimerisation domain) S glycoprotein ectodomain (S2 truncated at 1138) that had been stabilised in

the pre-fusion confirmation (GGGG substitution at furin cleavage site and S-2P mutation) was employed as a sorting bait [555]. This configuration of the S glycoprotein was selected so that a full range of B-cells targeting neutralising and non-neutralising epitopes would be characterised. It was observed that 2.14%, 2.45% and 2.39% of $CD19^{+}IgG^{+}$ B-cells bound to the S probe in P008, P054 and P003 respectively, compared to 0.25% in a pre-COVID-19 healthy control sample (Figure 5.1b). Although P008 had a low $ID_{50}$ of 76 at day 61 POS, the frequency of S-specific B-cells measured was similar to that of the other two donors. Sorting yielded 150 B-cells that were reactive for S from P003 and P008, and 120 B-cells from P054. The $V_H$ and $V_L$ regions were reverse transcribed and amplified via nested PCR, using gene-specific primer mixes. The PCR products were purified using PCR cleanup kits and ligated into $Ig_H$ or $Ig_L$ expression vectors using Gibson assembly. Matching assembled products were co-transfected directly into 293T cells, and transformed into competent bacterial cells [492]. The supernatants from the transfection were tested via ELISA for IgG expression and S glycoprotein binding, and if confirmed the corresponding bacterial colonies were grown in small cultures. The plasmids were purified using Miniprep kits and sent for Sanger sequencing, enabling gene analysis. Pairs of $Ig_H$ and $Ig_L$ chains that could be cloned this way were co-transfected into 293T cells, and the supernatant was tested via ELISA with S, RBD, NTD and S1 antigens (Figure 5.2A). This led to a total of

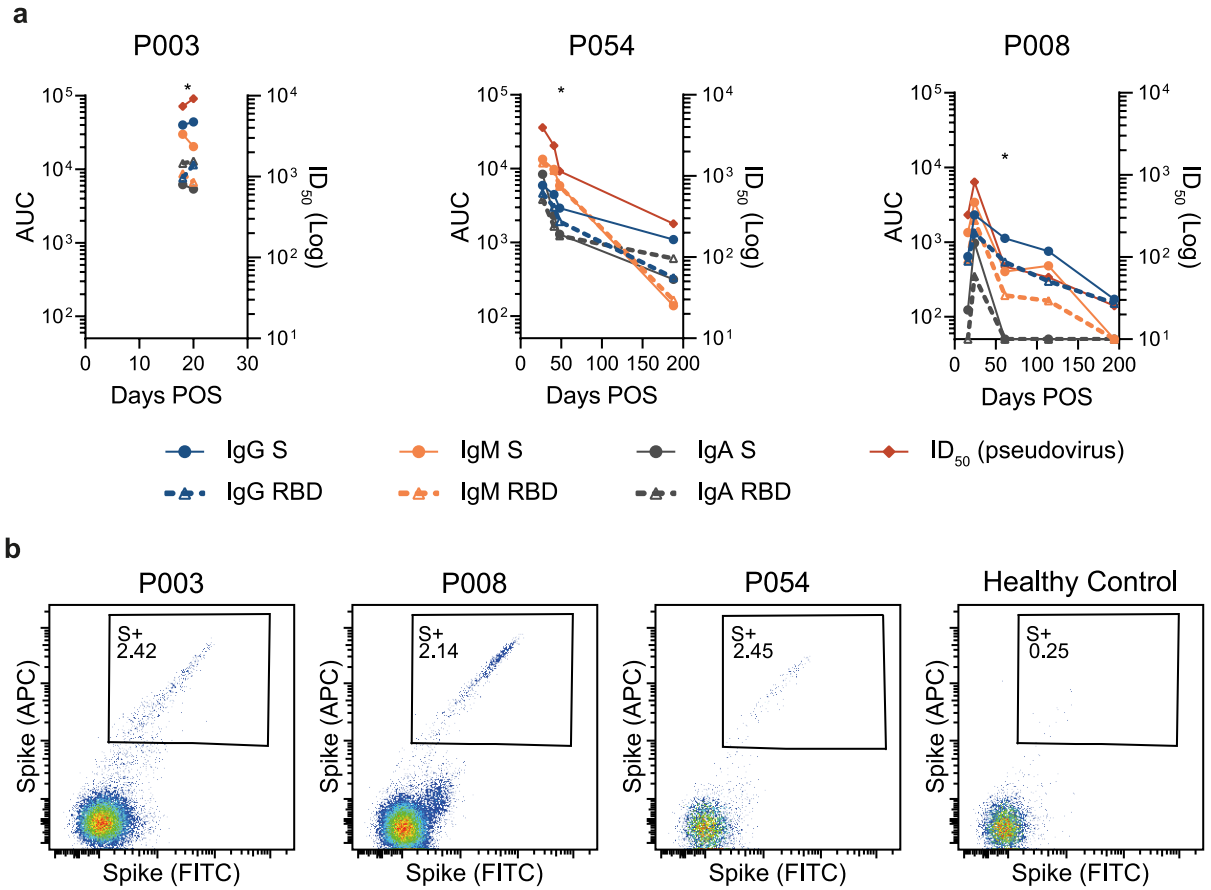



*Figure 5.1: Donor binding (IgG, IgA and IgM), neutralising antibodies and SARS-CoV-2 S reactive IgG+ B-cells.*
*(**a**) Kinetics of the antibody binding response (IgM, IgA, IgG against S and RBD) and neutralisation activity against SARS-CoV-2 pseudovirus for donors P003, P008, and P054 in the acute and convalescent phase. ELISA data is reported as area under the curve (AUC, left y axis) and neutralisation $ID_{50}$ is shown on the right y axis. The asterisk indicates the time point from which mAbs were cloned for each donor. Experiments were performed in duplicate and repeated twice where plasma was available. (**b**) Fluorescent activated cell sorting (FACS) showing percentage of $CD19^{+}IgG^{+}$ B-cells binding to SARS-CoV-2 S. A healthy control PBMC sample collected prior to the COVID-19 pandemic was used to measure background binding to S. Contributions to data collection from Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose Jimenez-Guardeño, Luke Muir, Weng Ng, Helen Duyvesteyn, Yuguang Zhao and Thomas Bowden.*

107 S-specific mAbs being identified and sequenced, with 64, 19 and 24 being isolated from P008, P054 and P003 respectively (Figure 5.2b). Of these mAbs, 35.5% and 32.7% bound to the S/S1/RBD

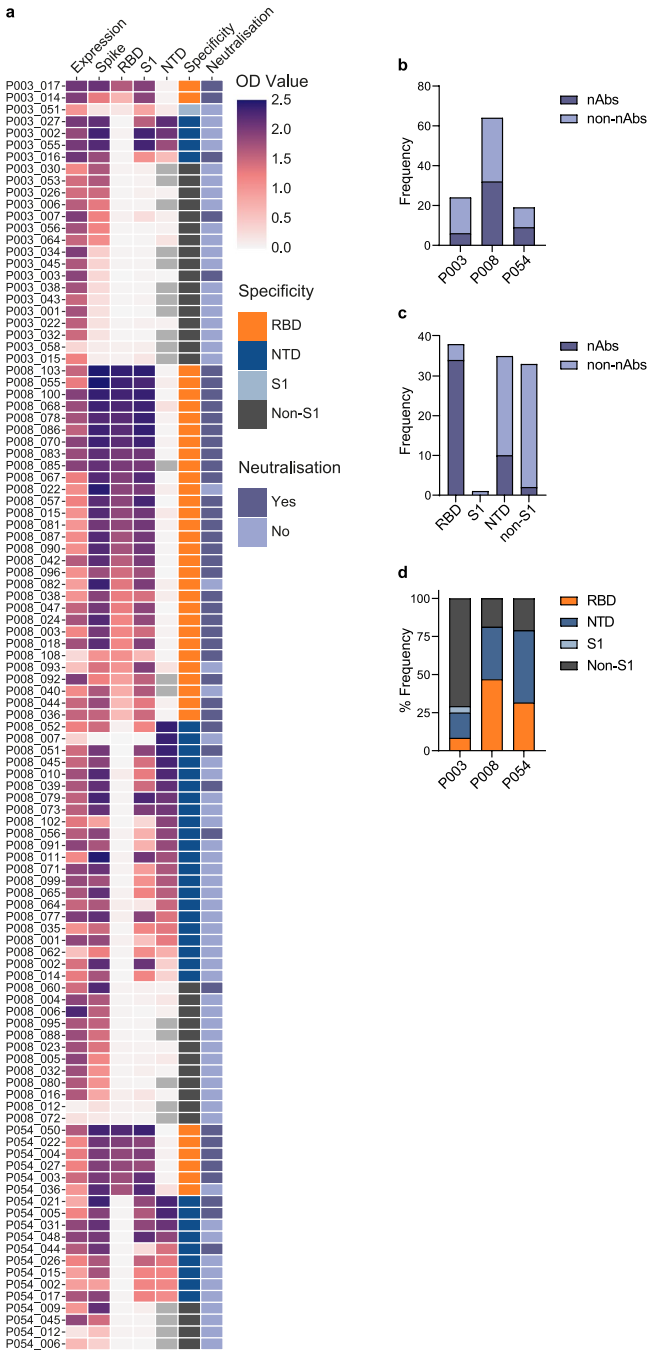



*Figure 5.2: SARS-CoV-2 S reactive mAbs binding to RBD, NTD, S1 and non-S1 epitopes.*
*(**a**) Heatmap showing IgG expression level and binding to SARS-CoV-2 S, S1, NTD, and RBD. Shown here are OD values from a single experiment (range 0–2.5) for undiluted supernatant from small scale expression of 107 cloned mAbs. Grey squares indicate samples that were not measured. Antigen binding was considered positive when OD at 405 nm was >0.3 after background was subtracted. SARS-CoV-2 S domain specificity for each antibody is indicated. Neutralisation activity was measured against pseudotyped virus using either small-scale purified IgG or concentrated supernatant. Antibodies were considered neutralising if at least 50% neutralisation was reached at the highest concentration (5 µg/mL for purified mAb) or concentrated supernatant (~30 times). (**b**) Bar graph shows frequency of neutralising and non-neutralising antibodies isolated from donors P003, P008, and P054. (**c**) Bar graph showing frequency of neutralising and non-neutralising antibodies targeting specific S sub-domains. (**d**) Bar graph showing the % of mAbs isolated from each donor targeting specific S sub-domains. Contributions to data collection from Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose Jimenez-Guardeño, Luke Muir, Weng Ng, Helen Duyvesteyn, Yuguang Zhao and Thomas Bowden.*

(RBD-specific) and S/S1/NTD (NTD-specific) antigens respectively, with 0.9% targeting S/S1 (S1-specific) (Figure 5.2c). The remaining 30.8% of mAbs were only reactive to S glycoprotein (non-S1-specific), suggesting that they were either specific for an epitope in the S2 or targeted a quaternary epitope [541]. The distribution of epitopes targeted by the mAbs differed between donors, for instance the majority of P003 mAbs were non-S1-specific (Figure 5.2d). Nonetheless, multiple mAbs were isolated from each donor that targeted RBD, NTD and non-S1 epitopes.

Next, the supernatants were either purified or concentrated, prior to being tested in pseudovirus neutralisation assays. This revealed that 43.9% of the mAbs isolated had neutralisation capability, highlighting the presence of non-neutralising epitopes on the S glycoprotein that are exposed and immunodominant. It was found that 91.9% of the RBD-specific mAbs were neutralising, whilst only 28.6% of the NTD and 8.8% of the non-S1 targeting mAbs had neutralising activity (Figure 5.2c). Thus, the RBD represented the dominant target of the neutralising mAbs, with the NTD acting as a subdominant epitope [537].

### 5.1.5 SARS-CoV-2 mAbs have diverse gene usage

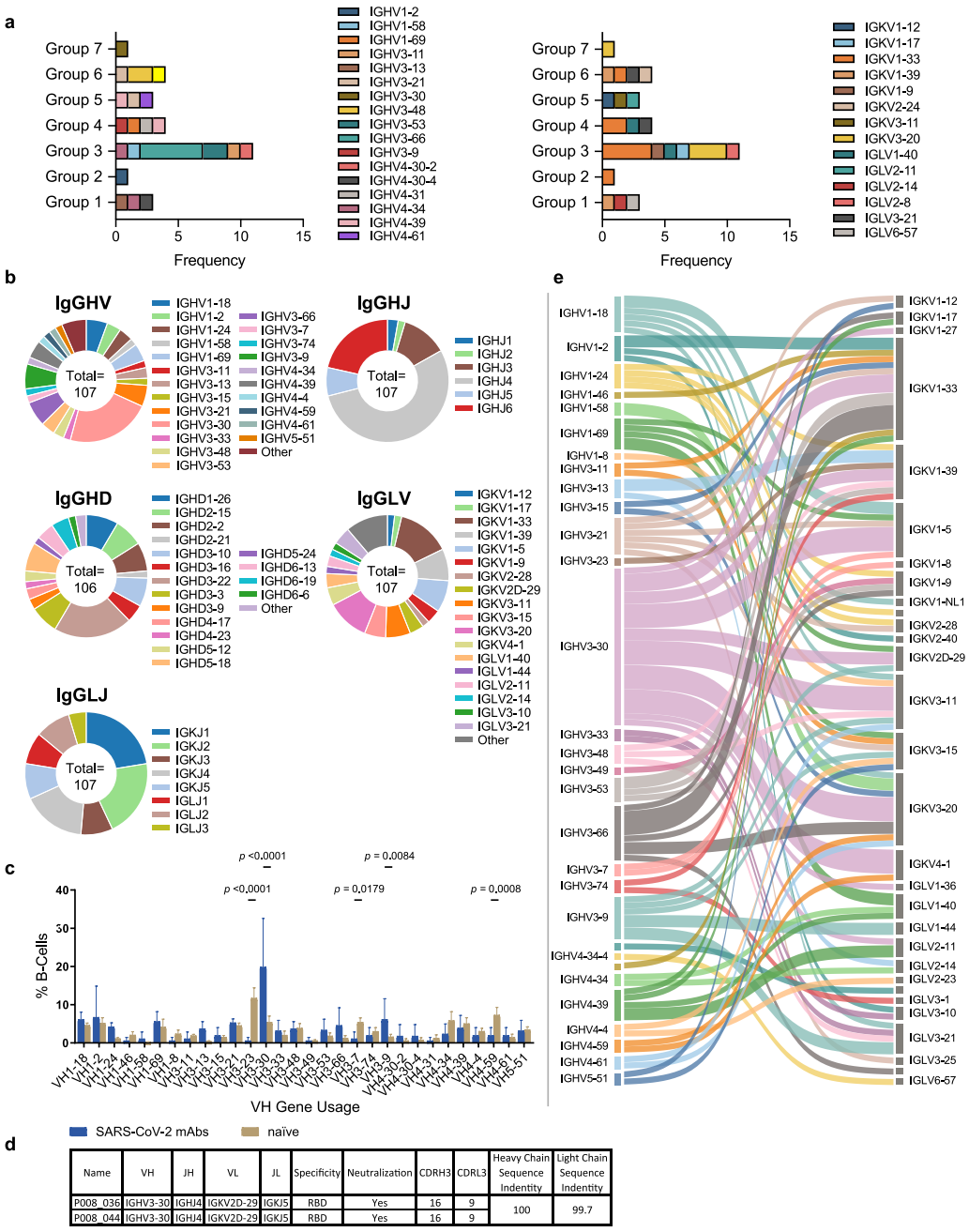
The sequences of the $V_H$ and $V_L$ genes were run through the IMGT database, which permitted analysis of their predicted germline gene usage, %SHM and CDRH3 lengths [433]. A wide range of $V_H$ and $V_L$ germline genes were used by the mAbs isolated from these subjects (Figure 5.3b). The $V_H$ germline gene usage of the 107 mAbs was compared to that of naïve B-cell repertoires, and showed an enrichment of the VH3-30 and VH3-9 genes (Figures 5.3c) [556]. On the other hand, it was found that there was a de-enrichment of the VH3-23, VH3-7 and VH4-59 genes. Although there was an enrichment in VH3-30 gene usage, ten different $V_L$ germline gene pairings were present, and they included both $Ig_L\kappa$ and $Ig_L\lambda$ genes (Figure 5.3e). There was only one example of clonal expansion identified amongst the mAbs, and this arose in P008 (Figure 5.3d). The mAbs that used the VH3-66 and 3–53 genes were shown to frequently target the RBD, similarly to what has been observed in other studies (Figure 5.3a) [540,557,558].

As expected following an acute viral infection, low levels of %SHM were found in the $V_H$ and $V_L$ genes (means of 1.9% and 1.4%, respectively). The mean %SHM in the $V_H$ and $V_L$ genes of mAbs from P008 (2.3%, isolated at 61 days POS) and P054 (2.0%, isolated at 48 days POS) were significantly higher compared to P003 (0.8%, isolated at 20 days POS) (Figure 5.4a). However, these differences were likely due to the PBMCs sorted from P003 being isolated at an earlier date post-infection. There was no significant difference in the %SHM of the $V_H$ and $V_L$ regions from mAbs specific for RBD, NTD and non-S1, or between neutralisers and non-neutralisers (Figures 5.4b-c). Comparison of the CDRH3 length distributions between the mAbs and naïve B-cell repertoires revealed that amino acid lengths of 21 and 22 were slightly enriched (Figure 5.4d). Overall, the SARS-CoV-2 targeted mAbs did not differ greatly from what can be observed in naïve repertoires, with diverse germline usage, low levels of %SHM and typical CDRH3 lengths. These findings have been corroborated by similar studies [490,492,541].

### 5.1.6 mAbs potently neutralise SARS-CoV-2

A representative sample of the neutralising (37/47) and non-neutralising (12/60) mAbs, with RBD, NTD or non-S1 specificities, were selected for further analysis. The $Ig_H$ and $Ig_L$ plasmids were amplified using Maxiprep kits and co-transfected into large cultures of 293F suspension cells. The mAbs were purified from the supernatant via filtration and affinity chromatography, after which the preparations were concentrated and buffer exchanged. Neutralisation potency was measured using pseudovirus and authentic virus neutralisation assays. A wide range of $IC_{50}$ values were observed for the pseudovirus (1.2–660ng/mL) and authentic virus (2.3–488ng/mL) (Appendix 7.6).

The RBD-specific mAb P008_108 had the highest potency, neutralising authentic virus with an $IC_{50}$ of 2.3ng/mL. At the time of its isolation, P008_108 was one of the most potent anti-SARS-CoV-2 mAbs discovered [559]. The $IC_{50}$ values determined using pseudovirus correlated well with those measured against authentic virus ($r = 0.7694$, $p < 0.0001$) (Figure 5.5a). However, $IC_{50}$ values against authentic virus were generally less potent than pseudovirus (5 to 10-fold lower), an observation seen previously




| Name | VH | JH | VL | JL | Specificity | Neutralization | CDRH3 | CDRL3 | Heavy Chain Sequence Indentity | Light Chain Sequence Indentity |
|---|---|---|---|---|---|---|---|---|---|---|
| P008_036 | IGHV3-30 | IGHJ4 | IGKV2D-29 | IGKJ5 | RBD | Yes | 16 | 9 | 100 | 99.7 |
| P008_044 | IGHV3-30 | IGHJ4 | IGKV2D-29 | IGKJ5 | RBD | Yes | 16 | 9 | | |

*Figure 5.3: Germline gene usage of SARS-CoV-2 S-specific mAbs.*
*(**a**) $V_H$ and $V_L$ germline gene usage for mAbs in the seven competition groups. (**b**) Pie charts showing percentage IGHV, IGHD, IGHJ, IGLV/IGLK and IGKJ/IGLJ usage for S reactive monoclonal antibodies. (**c**) Bar graph showing the mean $V_H$ germline gene percentage usage in SARS-CoV-2 specific mAbs (blue) compared to a representative naïve repertoire (grey). Error bars represent the standard deviation between donors used in the analysis (n = 3 for SARS-CoV-2 and n=10 for naïve repertoire). Differences between groups were determined using 2-way ANOVA with Šídák's multiple comparison test and p values <0.05 are shown. (**d**) Single example of clonal expansion observed from P008. (**e**) Sankey diagram showing the pairing between $V_H$ and $V_L$ germline genes for SARS-CoV-2 mAbs. Contributions to data collection from Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose Jimenez-Guardeño, Luke Muir, Weng Ng, Helen Duyvesteyn, Yuguang Zhao and Thomas Bowden.*

with patient serum samples [536]. Some of the mAbs that targeted epitopes outside of the RBD displayed weak neutralisation potencies (>10µg/mL) against pseudovirus (Figure 5.6b). These mAbs could only neutralise authentic virus at very high concentrations ($IC_{50}$ >50µg/mL), with some losing detectable neutralisation activity altogether (Figure 5.6c). However, the NTD-specific mAbs P008_056, P008_007 and P003_027 showed a ~10-fold higher potency against authentic virus (Figures 5.5c-d). Notably, P008_056 was able to neutralise authentic virus with an $IC_{50}$ of 14ng/mL, and was one of the most potent NTD targeted mAbs reported at the time of its isolation [560]. The neutralising mAbs that targeted the RBD were typically more potent compared to the non-RBD-specific mAbs (Figures 5.6b-c). Some mAbs displayed low neutralisation plateaus and shallow neutralisation curves, indicative of incomplete neutralisation (Figures 5.5b-c) [541]. These mAbs tended to be of a lower potency, with higher $IC_{50}$ values >1,000ng/mL. Seven non-S1-specific mAbs were shown to bind to S2 glycoprotein via ELISA, none of which had neutralisation capability (Appendix 7.6). Overall, highly potent neutralising mAbs against the RBD and NTD were identified from these donors.

### 5.1.7 Neutralising mAbs form seven binding competition groups

To understand the epitopes targeted by the neutralising mAbs further, competition ELISAs with S

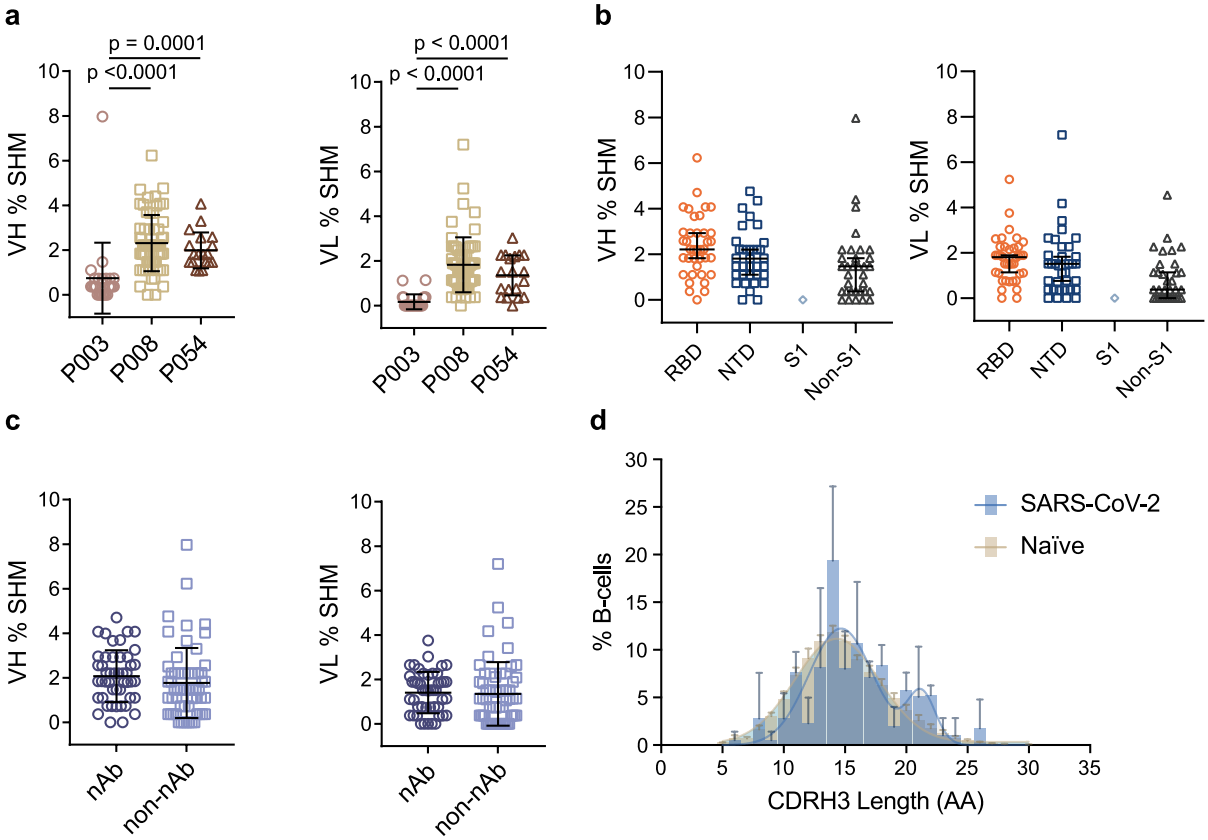



*Figure 5.4: Characteristics of SARS-CoV-2 S-specific mAbs: Germline gene usage, % SHM and CDRH3 length.*
*(**a**) % SHM in the $V_H$ and $V_L$ genes of S-reactive mAbs for donors P003, P008, and P054. Differences between groups were determined using Kruskal-Wallis multiple comparison test and p values <0.05 are shown. Black lines represent the mean SHM and error bars represent the standard deviation. (**b**) % SHM for mAbs targeting RBD, NTD, S1, or non-S1 epitopes (Kruskal-Wallis multiple comparison test). Black lines represent the mean SHM and error bars represent the standard deviation. (**c**) Percentage of $V_H$ and $V_L$ SHM for neutralising antibodies and non-neutralising antibodies (Mann-Whitney 2-sided U-test). Black lines represent the mean SHM and error bars represent the standard deviation. (**d**) Distribution of CDRH3 lengths for SARS-CoV-2 specific mAbs and representative naive B-cell repertoire. Error bars represent the standard deviation between donors used in the analysis (n = 3 for SARS-CoV-2 and n = 10 for naive repertoire). A bimodal distribution of CDRH3 length is observed for SARS-CoV-2 S reactive mAbs. Contributions to data collection from Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose Jimenez-Guardeño, Luke Muir, Weng Ng, Helen Duyvesteyn, Yuguang Zhao and Thomas Bowden.*

antigen were performed using 27 mAbs and their corresponding F(ab')$_2$ (Figure 5.6a). F(ab')$_2$ fragments were generated by incubating the mAbs with IdeS, the Fc fragements were removed via affinity chromatography and the preparations were concentrated and buffer exchanged. To relate this data to the literature, competition ELISAs were also carried out with numerous SARS-CoV-2 targeting

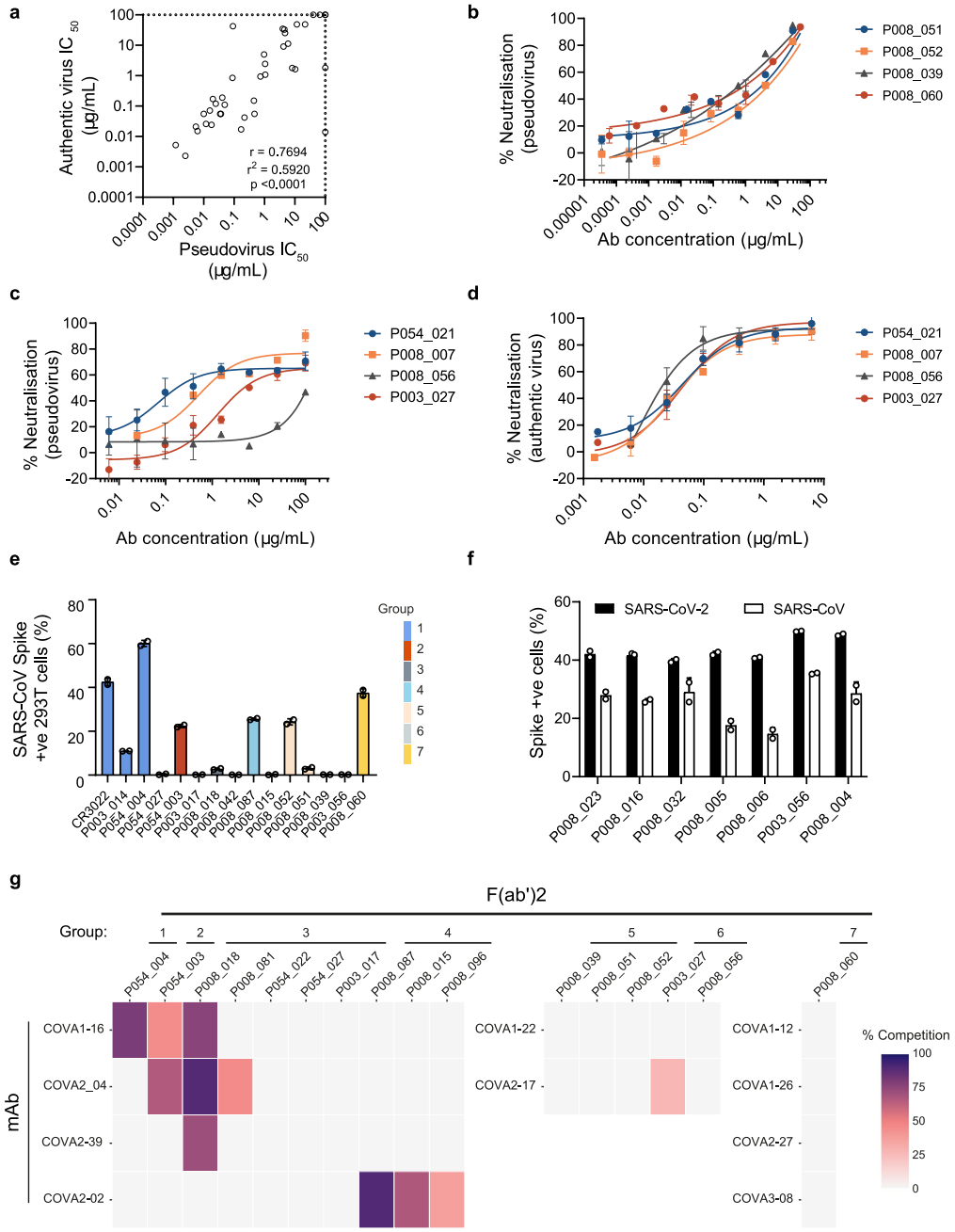



*Figure 5.5: mAb pseudovirus neutralisation correlates, authentic virus neutralisation and cross-reactivity with SARS-CoV. (**a**) Correlation of mAb neutralisation $IC_{50}$ against authentic virus and pseudotyped virus (Spearman correlation, r). A linear regression was used to calculate the goodness of fit, r2). (**b**) Examples of shallow neutralisation curves for Group 5 and Group 7 antibodies against SARS-CoV-2 pseudovirus. (**c-d**) Neutralisation curves of Group 6 neutralising mAbs against pseudotyped and authentic virus. (**e-f**) Binding of NTD and RBD nAbs and S2 non-neutralising mAbs to SARS-CoV-2 and SARS-CoV S glycoproteins expressed on the surface of HEK 293T cells measured by flow cytometry. Binding is reported as the % PE-positive cells. (**g**) Inhibition in binding of previously characterised SARS-CoV-2 neutralising antibodies to S by F(ab)2' fragments of representative members of each competition Group. The percentage competition was calculated using the reduction in IgG binding in the presence of F(ab')2 (at 100-molar excess of the IC80) as a percentage of the maximum IgG binding in the absence of F(ab')2. Contributions to data collection from Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose Jimenez-Guardeño, Luke Muir, Weng Ng, Helen Duyvesteyn, Yuguang Zhao and Thomas Bowden.*

neutralising mAbs that had been previously described (Figure 5.5g) [490]. Seven different competition

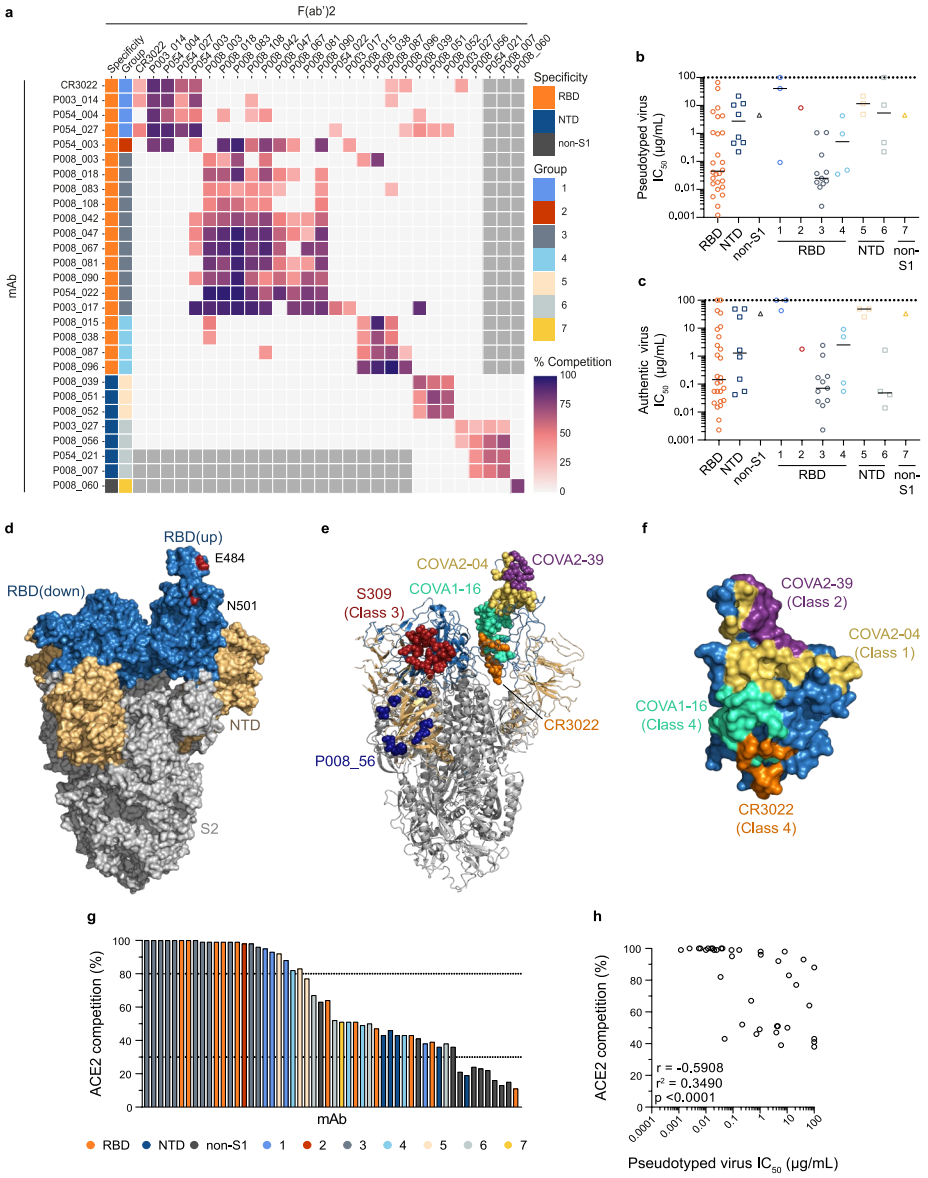



*Figure 5.6: SARS-CoV-2 S-specific mAb competition groups and neutralisation capabilities.*
*(**a-c**) Inhibition of IgG binding to SARS-CoV-2 S by F(ab)$_2$' fragments. Competition groups were determined using Ward2 clustering and clusters were then arranged according to binding epitopes. Experiments were performed in duplicate. Competition <25% is white. Grey boxes indicate competition not tested. IC$_{50}$ of mAbs targeting either RBD, NTD or non-S1, and competition Groups 1–7 against (**b**) SARS-CoV-2 pseudovirus and (**c**) authentic virus. Black lines represent the median IC$_{50}$ for each group. IC$_{50}$ values are the average of three independent experiments performed in duplicate. (**d–f**) Mapping of previously determined neutralising epitopes on RBD of SARS-CoV-2 S. (**d**) Surface rendered representation of SARS-CoV-2 S (side view) showing the RBD (blue), NTD (brown) and S2 (grey) domains. One RBD monomer is in the "up" conformation. Positions of S mutations relevant to neutralisation escape (N501Y and E484K) are indicated in red. (**e**) Cartoon representation of S showing antibody binding footprint for neutralising antibodies as coloured spheres. Epitopes were previously determined using crystal structures or cryo-electron microscopy of RBD or S-Fab complexes; COVA2-04 (yellow, RBD Class 1, COVA2-39 (purple, RBD Class 2), S309 (red, RBD Class 3), COVA1-16, and CR3022 (pale green and orange, respectively, RBD Class 4), and P008_056 (dark blue, NTD Group 6). (**f**) Surface representation of zoomed in RBD in "up" conformation showing footprint of RBD neutralising antibodies. (**g**) Ability of neutralising antibodies and non- neutralising antibodies to inhibit the interaction between cell surface ACE2 and soluble SARS-CoV-2 S. mAbs (at 600 nM) were pre-incubated with fluorescently labelled S before addition to HeLa-ACE2 cells. The percentage reduction in mean fluorescence intensity is reported. Experiments were performed in duplicate. (**h**) Correlation between IC$_{50}$ against pseudovirus and % ACE2 competition. (Spearman correlation, r). A linear regression was used to calculate the goodness of fit, r$^2$. Adapted from [388,555,563,568,645,646]. PDB = 6VSB. Contributions to data collection from Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose Jimenez-Guardeño, Luke Muir, Weng Ng, Helen Duyvesteyn, Yuguang Zhao and Thomas Bowden.*

groups were identified, mAbs in Groups 1-4 were RBD-specific, mAbs in Groups 5-6 were NTD-specific and there was one mAb in Group 7 that was only specific for S glycoprotein.

Group 1 mAbs competed with CR3022 (isolated from a SARS-CoV patient) and P054_004 (Group 1) competed with COVA1-16 (Class 4) (Figure 5.5g) [561]. Structural studies have shown that CR3022 and COVA1-16 bind to sites on the RBD that are distal to the receptor binding motif (RBM) (Figures 5.6d-f) [562]. Group 1 mAbs showed limited neutralisation potency, particularly when tested against authentic virus (Figures 5.6c). Group 3 was the largest and most potent competition group, it contained 57.8% of the neutralising mAbs tested via competition ELISA and the highly potent P008_108. P008_018 (Group 3) competed strongly with COVA2-04 (Class 1) and moderately with COVA2-39 (Class 2). On the other hand, P008_081 (Group 3) only competed weakly with COVA2-39. Class 1 mAbs bind directly to the RBM, whilst Class 2 mAbs target less or none of the residues in the RBM but maintain their ability to inhibit S-ACE2 interaction. It was observed that 63.6% of Group 3 members used the VH3-53 and VH3-66 germlines, a feature typical of both Class 1 and Class 2 mAbs (Appendix 7.6 and Figure 5.3a) [557,559]. P054_003 was the only Group 2 mAb and it competed with members of Group 1 and Group 3, as well as COVA1-16 and COVA2-04. This suggested that Group 2 had an epitope that overlapped with that of Group 1 and Group 3. Group 4 mAbs were of moderate neutralisation potency, and competed with COVA2-02 (Class 3) to varying extents. The neutralising epitope of COVA2-02 is believed to overlap with that of S309 (isolated from a SARS-CoV patient), which targets a site very distal on the RBD [388]. Some of the RBD-specific mAbs did not compete with any of the mAbs from previous studies, indicating that there could be differences in the epitope footprint and/or angle of approach.

Group 5 was made up of three NTD-specific mAbs that had limited neutralisation potency against pseudovirus ($IC_{50}$ 4.8–21.7µg/mL), and neutralised authentic virus more weakly ($IC_{50}$ 25.3–48.8µg/mL) (Figures 5.6b-c). Conversely, the four mAbs in Group 6 were able to neutralise authentic virus with greater potency compared to pseudovirus. Structural analysis has shown that P008_056 (Group 6) binds to an epitope on the NTD that is adjacent to the β sandwich fold (Figures 5.6d-f) [563]. P003_027 (Group 6) competed weakly with COVA2-17, but none of the other NTD-specific mAbs had detectable competition with COVA2-17 or COVA1-22 (Figure 5.3g). Group 7 contained only P008_060, which bound S glycoprotein in ELISA but not to S1 or S2 monomers. Additionally, P008_060 did not compete with the non-RBD targeting mAbs COVA1-26, COVA1-12, or COVA3-08. Originally, it was considered that P008_060 must target a quaternary epitope that spanned multiple protomers, similarly to 2-43 [541]. However, subsequent structural studies have revealed that P008_060 binds to an epitope on the SD1 of the RBD, which was truncated off of the S1 monomers used in this work [564].

### 5.1.8 mAbs inhibit interaction between S glycoprotein and ACE2 to differing extents

To investigate the potential mechanisms of neutralisation in these mAbs, flow cytometry-based assays were performed. The assays measured the ability of each mAb to prevent interaction between fluorescently conjugated S glycoprotein and HeLa ACE2 cells (Figure 5.6g). It was shown that Group 3 mAbs were able to inhibit S binding to ACE2 by >99%, indicating that their target epitope was located directly on the RBM. Of note, the mAbs displaying the best ACE2 competition typically had the highest neutralistion potency (Figure 5.6h). Resembling CR3022, the Group 1 mAbs had less complete ACE2 competition (88.2%–95.1%), whilst the Group 4 mAbs showed only partial ACE2 competition (43.1%–82.2%). The Group 5-7 members also showed partial to less complete ACE2 competition (38.4%–91.8%), whereas non-neutralising S1- and S2-specific mAbs showed negligible ACE2 competition. The mAbs that did not bind directly to the RBM but inhibited interaction between S and ACE2 may have done so sterically, or caused conformational changes in S that limited ACE2 binding. For instance, neutralising mAbs have been reported that lock the RBD in the down conformation, occluding the

ACE2 binding site [541]. Neutralisation mechanisms beyond receptor binding inhibition should also be considered, such as inhibiting membrane fusion and S1 shedding [537,560].

### 5.1.9 Glycan heterogeneity influences neutralisation potency

As stated previously, some mAbs presented shallow neutralisation curves that plateaued below 100% and this was generally accompanied by NTD specificity. Comparable neutralisation profiles have been seen with HIV-1 bnAbs, particularly in those that accommodate or bind to N-linked glycans on Env [565]. This phenotype is believed to arise due to glycan heterogeneity, and can be rescued by changing the composition of Env glycans via glycosidase inhibitors like kifunensine and swainsonine. Kifunensine inhibits the endoplasmic reticulum mannosidase I enzyme, resulting in only $Man_9GlcNAc_2$ glycans. Swainsonine inhibits the Golgi-α-mannosidase II enzyme, resulting in truncated complex-type glycans and naturally occurring high-mannose glycans.

The NTD is heavily glycosylated, therefore it was tested whether altering the glycan structure of S glycoprotein with kifunensine or swainsonine could impact neutralisation activity (Figure 5.7) [334]. As expected, RBD-specific mAbs P008_015, P008_087, P008_090, and P008_108 were not affected by this intervention. Conversely, mAbs from Group 5 and Group 7 (P008_039, P008_051, P008_052 and P008_060) displayed enhanced neutralisation against pseudovirus prepared with swainsonine. There was no change in the neutralisation profile of these mAbs when tested against pseudovirus made with kifunensine, although a lower infectivity was noted and had been previously reported [566]. Overall, it would appear that glycan structures can impact neutralising mAb epitope recognition through modulating S glycoprotein conformation or changing the accessibility of neutralisation sites. This raises important considerations about the glycan landscape of vaccination candidates.

### 5.1.10 Some mAbs display cross-reactivity with SARS-CoV

The S glycoprotein and RBD of SARS-CoV and SARS-CoV-2 share 73% sequence homology [567]. Hence,

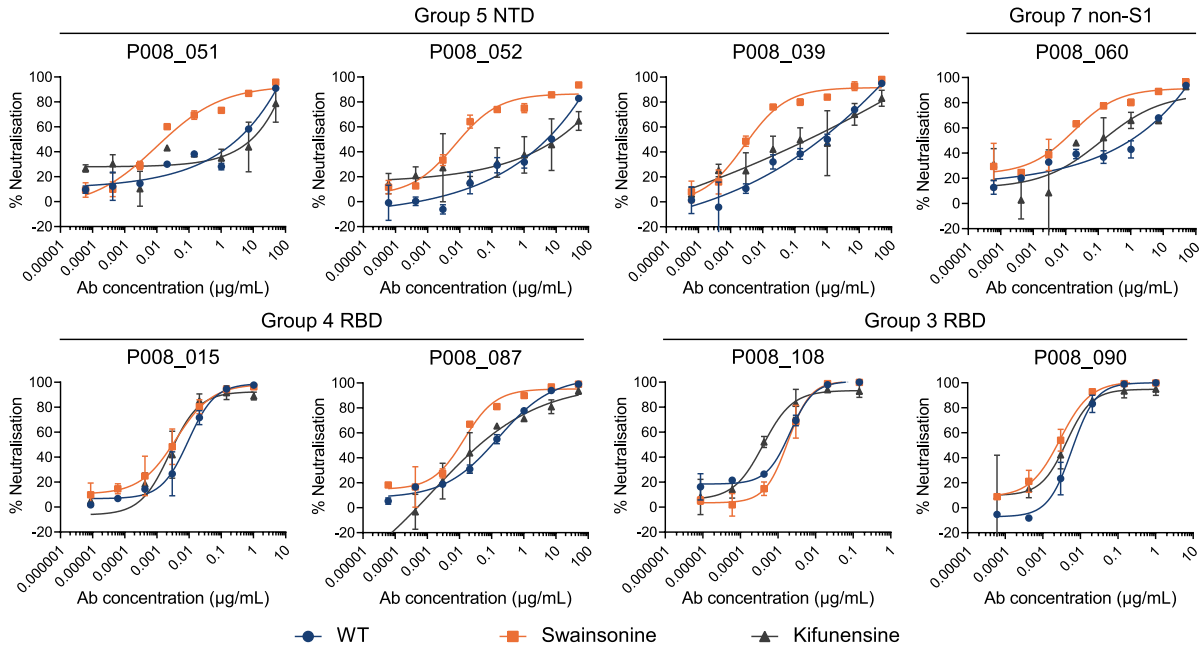



*Figure 5.7: Impact of changes in S glycosylation on Group 5 mAb neutralisation.*
*SARS-CoV-2 pseudovirus was expressed in the presence of glycosidase inhibitors kifunensine or swainsonine. Neutralisation potency of RBD and NTD neutralising antibodies against S-modified pseudoviruses was measured. Neutralisation assays were performed three times in duplicate and a representative experiment is shown. Contributions to data collection from Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose Jimenez-Guardeño, Luke Muir, Weng Ng, Helen Duyvesteyn, Yuguang Zhao and Thomas Bowden.*

additional pseudovirus neutralisation assays were carried out to determine whether the mAbs could cross-neutralise particles expressing the S glycoprotein of SARS-CoV. Although neutralising activity against SARS-CoV was observed for some mAbs, it was typically at a much reduced potency (3 to 65-fold) compared to SARS-CoV-2 (Appendix 7.6). CR3022 was isolated from a SARS-CoV infected donor, and has been shown to bind an epitope that is conserved between SARS-CoV and SARS-CoV-2 [568]. Despite competing with CR3022, only a single mAb from Group 1 displayed cross-neutralisation. Further, only one antibody from Group 4 and Group 5 were able to neutralise SARS-CoV pseudovirus. This suggested that there was a range of molecular contacts in each of the competition groups. Interestingly, P008_060 (Group 7) was able to neutralise SARS-CoV more potently (7-fold) than SARS-CoV-2. Binding to the S glycoprotein of SARS-CoV, expressed on the surface of HEK293T cells, was detected for the SARS-CoV neutralising mAbs but not for the other non-neutralising members of the competition groups (Figure 5.3e). In contrast, S2-specific mAbs that were unable to neutralise either SARS-CoV or SARS-CoV-2, were shown to bind to S glycoprotein of SARS-CoV in this context (Figure 5.3f). This suggested that there was a non-neutralising S2 epitope conserved between SARS-CoV and SARS-CoV-2. Whether mAbs targeting such an epitope could facilitate effector functions such as ADCC in vivo remains unknown [569]. Overall, there are conserved neutralising epitopes shared between SARS-CoV-2 and SARS-CoV located on the RBD and NTD.

### 5.1.11 Neutralising mAbs from distinct competition groups are differentially impacted by emerging SARS-CoV-2 variants

The neutralisation potency of the patient plasma, and mAbs from the seven competition groups was examined using pseudovirus particles bearing S glycoprotein with various mutations (D614G, N501Y, D614G+ΔH69/V70 and D614G+ΔY144), and the B.1.1.7 (α) variant with all eight mutations. In line with previous reports, the D614G mutation was found to have a limited impact on the neutralisation activity of most RBD-specific mAbs (Figure 5.8a) [570]. However, it was shown that Group 4 mAbs had a 3 to 22-fold decrease in neutralisation potency, and a Group 1 member (P054_027) had a 25-fold reduction. Further, Group 5 mAbs displayed a considerable reduction in neutralisation potency against D614G (19 to 450-fold), and P008_060 decreased 105-fold. Despite the N501Y mutation being located in the RBD, the neutralisation potency of the majority of RBD-specific mAbs was unaffected. Yet, the neutralisation activity of mAbs P003_017 and P008_003 (Group 3) reduced by 25 and 9-fold respectively, P054_004 and P054_027 (Group 1) decreased by 14 and 8-fold respectively, and P008_087 (Group 4) reduced by 15-fold. The neutralisation potency of NTD-specific mAbs was largely unaffected by the N501Y mutation, with the exception of P054_021 and P008_007 (Group 6) which decreased 8 and 17-fold respectively.

The ΔH69/V70 and ΔY144 mutations are located in the N1 loop of NTD, and have been linked to viral evolution in immunocompromised SARS-CoV-2-infected patients [549,550]. D614G+ΔH69/V70 had a limited impact on the neutralisation potency of RBD and NTD-specific mAbs (Figure 5.8a). Only P003_027 displayed a decrease in neutralisation activity (16-fold) compared to the D614G variant. On the other hand, D614G+ΔY144 abolished Group 6 mAb neutralisation and reduced the potency of Group 5 members by 7 to 15-fold compared to the D614G variant (Figure 5.8b). Particular RBD-specific mAbs showed a 6 to 11-fold decrease in neutralisation potency.

Similarly to the D614G+ΔY144 mutant, Group 6 mAbs were found to be unable to neutralise the B.1.1.7 (α) variant and Group 5 members showed only weak neutralisation (Figure 5.8a) ($IC_{50}$ 30-100 µg/mL). The most potent RBD-specific mAbs in Group 3 and Group 4 displayed no reduction in neutralisation potency when tested against the B.1.1.7 (α) variant. P008_081 and P054_022 (Group 3) showed a 6 and 18-fold decrease in neutralisation activity against B.1.1.7 (α) respectively, but maintained $IC_{50}$ values between 0.34-2.8µg/mL. Conversely, P003_017 (Group 3) and P054_027

(Group 1) presented the greatest reduction in neutralisation potency against B.1.1.7 (α), with weak neutralisation measured at 50µg/mL.

The E484K mutation has been observed in the B.1351 variant, which was prevalent in South Africa at the time of this work [546,571]. This mutation had also been found in combination with the B.1.1.7 (α) mutations in the UK [552]. To investigate if this extra mutation would result in escape from RBD-specific mAbs, neutralisation potency was measured against pseudovirus with the B.1.1.7+E484K S glycoprotein (Figure 5.8a). As expected, the neutralisation of the NTD-specific mAbs was largely

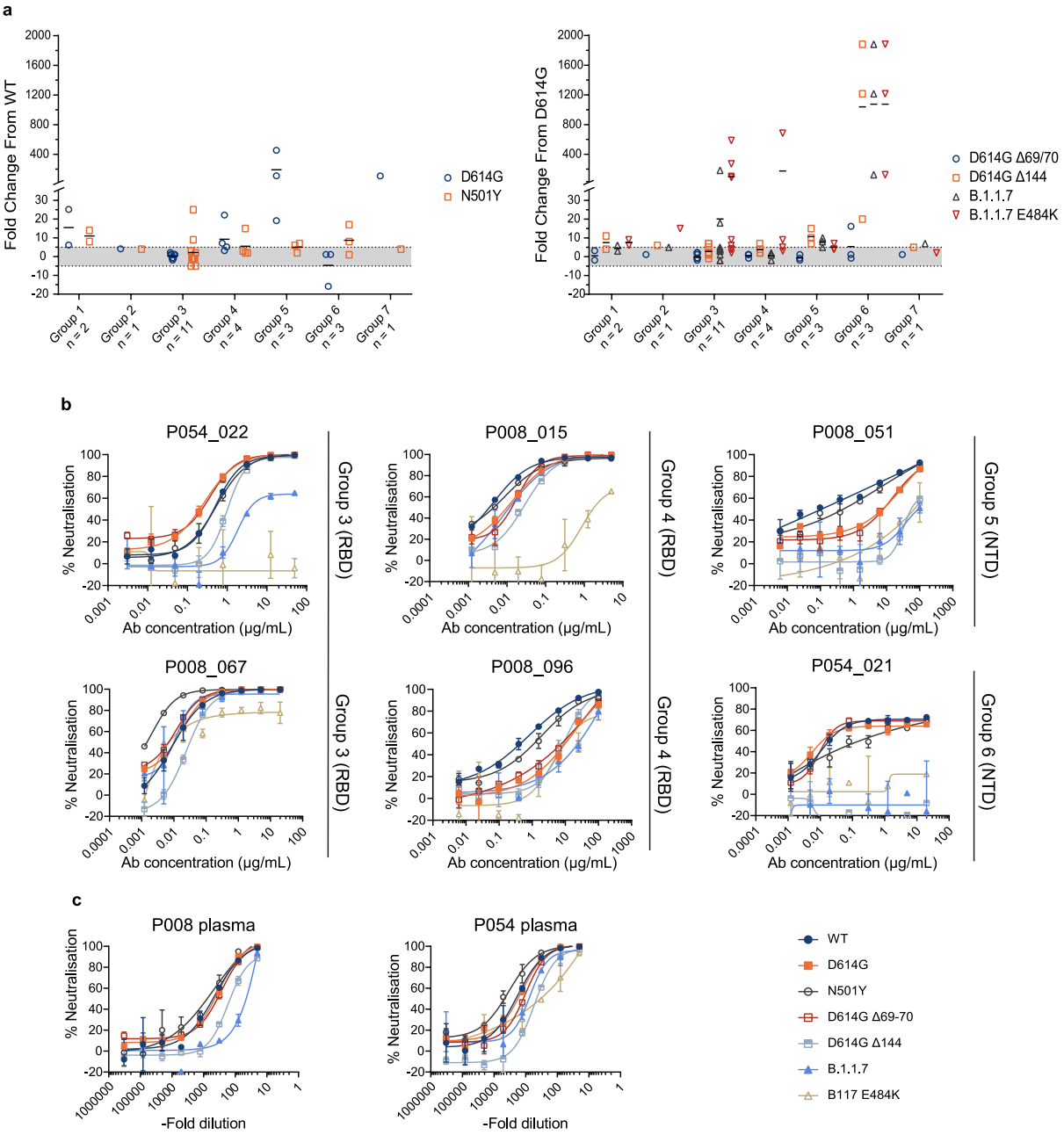



*Figure 5.8: Impact of B.1.1.7 (α) and S related variants on the neutralisation potency of mAbs.*
*(**a**) Fold change in neutralisation potency for D614G and N501Y mutations compared to authentic S, and D614G+ΔH69/V70, D614G+ΔH69/V70, B.1.1.7 (α) and B.1.1.7+E484K variants compared to D614G S. Black lines represent the mean fold change for each competition Group. $IC_{50}$ values were calculated from two independent experiments and used to calculate fold change. (**b**) Example neutralisation curves for Group 3, 4, 5, and 6 nAbs against S variants. (**c**) Neutralisation activity of P008 and P054 plasma against S variants at B-cell sorting time point. Contributions to data collection from Jeffrey Seow, Isabella Huettner, Hataf Khan, Neophytos Kouphou, Sam Acors, Helena Winstone, Suzanne Pickering, Rui Pedro Galao, Liane Dupont, Maria Jose Lista, Jose Jimenez-Guardeño, Luke Muir, Weng Ng, Helen Duyvesteyn, Yuguang Zhao and Thomas Bowden.*

unchanged compared to the B.1.1.7 (α) variant. Group 3 mAbs, which competed directly with ACE2 and were therefore the most likely to be impacted by the E484K mutation, showed a variety sensitivity to B.1.1.7+E484K. P054_022, P003_017, and P008_003 could not neutralise B.1.1.7+E484K, whilst the rest of Group 3 had slightly decreased neutralisation potencies compared to D614G (4 to 9-fold) (Figure 5.8b). The Group 4 mAbs showed modest reductions in neutralising activity compared to D614G (3 to 9-fold), except for P008_015 which had a >500-fold decrease in neutralisation. Of note, some of the RBD-specific mAbs were still able to neutralised B.1.1.7+E484K with $IC_{50}$ values as low as 0.008µg/mL.

Despite the reductions or losses of neutralisation potency for particular mAbs described above, P008 plasma (61 days POS) only showed an 8-fold decrease in neutralising activity against B.1.1.7 (α; Figure 5.8c). For P054 (48 days POS), no such reduction was observed against B.1.1.7 (α) but an 8-fold decrease in neutralisation potency was found against B.1.1.7+E484K. Unfortunately, there was not enough plasma available to test P003 with these mutants or variants, or P008 with B.1.1.7+E484K. Due to the low sample number, statistical analysis was not performed on this data. Nonetheless, the samples examined suggested that the polyclonal nature of the neutralising antibody response is able to overcome the mutations found in the B.1.1.7 (α) variant. Overall, the NTD-specific mAbs displayed the largest reductions in neutralisation activity against B.1.1.7 (α), which was mediated by the ΔY144 mutation. Although the neutralisation potency of some RBD-specific mAbs was impacted, potent neutralisation was still observed, even against B.1.1.7+E484K.

### 5.1.12 Discussion

Antigen specific B-cell sorting enabled the isolation of SARS-CoV-2 S-specific neutralising and non-neutralising mAbs from three convalescent subjects, each of which had experienced a different level of COVID-19 illness severity. In corroboration with previous studies, it was found that the antibody response to SARS-CoV-2 was highly diverse, not restricted to specific germline genes and did not require much SHM to enable neutralisation [490,492,541]. Antibodies targeting RBD, NTD and non-S1 epitopes were identified in each of the donors. Interestingly, the neutralising mAb with the highest potency (P008_108, with $IC_{50}$ 2.3ng/mL against authentic virus) was isolated from an asymptomatic subject that showed very low serum neutralising activity. Similar findings have been reported, indicating that the low neutralisation potency observed in asymptomatic or low severity individuals is due to the low quantity of neutralising antibodies in plasma, opposed to the sub-optimal potency of individual mAbs [557]. Further, this suggests that memory responses are not proportional to the antibodies present in the plasma of a patient, following the initial plasmablast burst. In the event of a re-exposure, subjects would be expected to generate highly potent neutralising antibodies [518].

The S glycoprotein was selected as the sorting bait to allow isolation of mAbs that target epitopes beyond the RBD, and enable the examination of their importance in regards to viral evolution. The RBD-specific neutralising mAbs bound to epitopes similar to those described previously, with some directly inhibiting ACE2 interactions via targeting the RBM and others binding epitopes more distally [490,492,541]. Of the mAbs isolated, ~1:3 were NTD-specific, of which only 28.5% were capable of neutralisation. The neutralising NTD-specific mAbs formed two distinct competition groups. The NTD-specific Group 6 mAbs were able to neutralise authentic virus more potently than pseudovirus. Assuming the structural conformations and dynamics of the S glycoproteins found on the authentic and pseudovirus particles were the same, the dissimilarities in neutralisation could have arisen for a number of reasons. These include differences in S glycoprotein density, S glycoprotein glycan heterogeneity and differences in the levels of ACE2 expressed on the surface of the target cells employed for the neutralisation assays. At the time of this work, P008_056 was one of the most potent NTD-specific neutralising mAbs reported ($IC_{50}$ 14ng/mL against authentic virus), and it was in line with

some of the most potent RBD-specific mAbs identified here [560]. Analysis of the structure of P008_056 Fab in complex with S glycoprotein showed that this mAb binds to the distal face of the NTD, inducing considerable conformational changes within this domain. Interestingly, biliverdin (a haem metabolite) has been shown to occupy a cleft in the vicinity of the epitope of P008_056. It was revealed that biliverdin supresses P008_056 binding to S glycoprotein, and substantially reduces its neutralising activity against authentic virus [563]. Therefore, the differences in the Group 6 neutralisation profiles against pseudovirus and authentic virus could have resulted from varying levels of biliverdin occupation in the S glycoprotein. The NTD-specific Group 5 mAbs were less potent compared to Group 6 and displayed atypical neutralisation curves. However, the neutralising activity of these mAbs was enhanced by reducing the size and/or the composition of the N-glycans found on S glycoprotein via preparation of pseudovirus in the presence of swainsonine. The NTD is heavily glycosylated and smaller glycans would be expected to increase the accessibility of the protein epitopes to mAbs, and therefore improve binding and neutralisation potency [334].

Many S-specific mAbs isolated here showed no neutralising activity, signifying the presence of immunodominant non-neutralising epitopes on the RBD, NTD and S2. Non-neutralising S2-specific mAbs were shown to cross-react with cells expressing the S glycoprotein of SARS-CoV. Further, the S2 subunit has been shown to be more conserved than S1 amongst coronaviruses [572]. Since these mAbs can target cell-surface expressed S glycoprotein, it was considered that they may play a role in Fc effector functions such as ADCC and aid in viral clearance [569]. Interestingly, one S2-specific mAb (P008_023) has been shown to be capable of triggering antibody-dependent NK cell activation (typically precedes ADCC) against SARS-CoV-2 infected cells, as well as causing the release of the proinflammatory cytokine TNFα. However, the NK cell responses generated by P008_023 were substantially lower compared to those generated by polyclonal serum from an infected donor, making the importance of such antibodies in this process uncertain [573]. Importantly, a pre-fusion S glycoprotein was used to isolate mAbs in this project, and this conformation may have not exposed the S2 neutralising epitopes that are found in nature. A study has shown that neutralising mAbs that are S2-specific can be generated through immunisation with recombinant S2 subunit immunogens [574].

At the time of this work, SARS-CoV-2 variants with multiple S glycoprotein mutations were rapidly emerging globally. There were concerns over whether antibodies produced against the ancestral strain would be able to provide protection against these new VOCs. The B.1.1.7 (α) variant could still be potently neutralised by most of the RBD-specific mAbs isolated here, but was resistant to NTD-specific mAbs. These findings were supported by other publications released around the time [546,575]. It was observed that the ΔY144 deletion of the B.1.1.7 (α) variant mediated escape from neutralisation by NTD-specific Group 6 mAbs. Analysis of the structure of P008_056 Fab in complex with S glycoprotein showed that Y144 resides in a loop that needs to undertake conformational rearrangement to enable access to the P008_056 epitope. The ΔY144 deletion has also been found to abrogate the binding of other NTD-specific mAbs such as S2M28, S2X28, S2X333, and 4A8 [542,560]. The Δ242-244 deletion of the B.1.351 (β) variant had also been shown to reduce the binding of NTD-specific mAbs 4A8 and 4-8 [546,571]. It could not be concluded from this data whether the NTD mutations were being selected as a result of NTD-specific neutralising mAbs and/or were enhancing viral functionality leading to increased transmission [548]. The NTD deletions of B.1.1.7 (α; ΔH69/V70 and ΔY144) had also been found in immunocompromised subjects receiving passive immunotherapy, whilst infected with SARS-CoV-2 for an extended period of time [549–551]. More research is required to understand what drives the accumulation of mutations in the S glycoprotein that are observed in circulating VOCs.

Although there was a loss in NTD-specific neutralisation against B.1.1.7 (α), neutralisation by RBD-specific mAbs was largely unchanged. When there was a RBD-specific reduction, the neutralisation potency remained in the range of 0.001–5µg/mL for most of the mAbs. The decrease in RBD-specific neutralising activity seen against B.1.1.7 (α) was lower than what had been observed for mAbs against the B.1.351 (β) variant [546,571]. The B.1.351 (β) variant has extra RBD mutations (K417N and E484K), and they have been linked to viral escape from RBD-specific mAbs [576,577] . Notably, P008 plasma neutralisation against the B.1.1.7 (α) variant was still detectable despite an 8-fold reduction, whereas P054 plasma neutralisation was unaltered. This suggested that although neutralisation of B.1.1.7 (α) was abolished for some mAbs targeting specific epitopes, more mutations would be needed for complete escape from neutralisation by the polyclonal plasma of these subjects. This finding has been supported by larger scale studies that have investigated neutralising responses in convalescent plasma [552,554,578]. Additional reductions in neutralisation potencies were observed in plasma and for some RBD-specific mAbs when the E484K mutation was added to the B.1.1.7 (α) variant, but this was not enough to invoke wide-spread resistance. However, more highly divergent VOCs have since emerged such as B.1.1.529 (ο), in which almost complete viral escape can be observed in convalescent donors and vaccine recipients [579].

In summary, potent neutralising mAbs were isolated from three convalescent subjects that had experienced different disease severities. The RBD was found to be the dominant target of the neutralising mAbs, with the NTD acting as a subdominant epitope. The neutralising mAbs were separated into seven distinct competition groups, and were shown to inhibit ACE2 interaction with S glycoprotein to varying degrees. The neutralisation potency of some NTD-specific mAbs was impacted by glycan heterogeneity, whilst others were found to neutralise authentic virus considerably more potently than pseudovirus. The B.1.1.7 (α) variant was resistant to NTD-specific mAbs, this highlighted the importance of researching the dominant and subdominant neutralising epitopes present on S glycoprotein when investigating viral evolution and antigenic drift.

## 5.2 Neutralising mAbs isolated from an individual vaccinated with ChAdOx1-nCoV

### 5.2.1 Introduction

By the time of this work, many mAbs had been isolated from subjects infected with SARS-CoV-2, and this had led to a better understanding of the neutralising responses generated following infection. It had been shown that there were neutralising epitopes located on the RBD, NTD and S2 [490,492,574,580]. The RBD was the dominant target of the neutralising mAbs, and the NTD acted as a subdominant epitope [580,581]. It was known that there were four distinct neutralising epitopes present on the RBD, some of which were able to inhibit the interaction of S glycoprotein with ACE2 [540,562]. Further, several non-overlapping neutralising epitopes had been identified in the NTD, and mAbs targeting this region had been shown to often be susceptible to viral mutations found in VOCs [560,580,582]. Many non-neutralising mAbs had also been isolated following SARS-CoV-2 infection, but the biological function of these molecules was not fully understood [573,583,584]. Collectively, the study of antibody responses at the monoclonal level had enabled the production of an antigenic map of the S glycoprotein [585,586]. However, it was not known whether COVID-19 vaccinations were able to generate neutralising antibodies targeting similar epitopes to those seen in convalescent individuals.

As a result of the emergence of SARS-CoV-2 in the human population, COVID-19 vaccines were rapidly developed and administered globally. The majority of licensed vaccines used the S glycoprotein as an antigen to stimulate humoral and cellular responses, and had shown promising efficacies in phase III trials [395,396,406]. However, concerns were raised on how vaccine efficacy would be impacted by newly emerging VOCs, such as B.1.1.7 (α), B.1.351 (β), P.1 (γ), B.1.617.2 (δ) and B.1.1.529 (ο), which carried several mutations in their S glycoprotein. Studies showed that neutralisation activity could still be detected in the sera of most double-vaccinated subjects, although at a decreased potency [552,553,587,588]. Real-world data suggested that the COVID-19 vaccines were still highly effective at preventing severe disease and hospitalisation in areas where VOCs were prevalent [589–591].

The ChAdOx1-nCoV vaccine is a replication-defective chimpanzee adenovirus-vectored vaccine that carries genetic material encoding the full-length ancestral S gene [406]. The ChAdOx1-nCoV vaccine was delivered via a prime-boost strategy, with some countries, including the UK, adopting an extended 8-12 week dose interval, as opposed to the original 3-4 week dose interval [592]. Vaccination with ChAdOx1-nCoV was paused in March 2021, due to reports of rare blood clotting events in vaccine recipients. By June 2021, it was decided that the benefits of the ChAdOx1-nCoV vaccine outweighed the potential risks, and many countries continued its administration. However, some nations recommended that individuals that had been given a single dose of ChAdOx1-nCoV, instead receive an mRNA-based vaccine boost [593]. At the time of this work, numerous studies had isolated mAbs generated by mRNA vaccines, but the antibody response to ChAdOx1-nCoV had not been assessed at the monoclonal level [594–596]. This was perhaps a result of ChAdOx1-nCoV having a lower efficacy, the aforementioned disruptions in vaccine rollout and the geographical location of laboratories that perform mAb isolation work. Therefore, research was required on how mAbs isolated from a ChAdOx1-nCoV vaccine recipient compared to those that had been identified in convalescent donors or individuals that had received an mRNA vaccine.

Overall, the specific aims of this section were to:

- Perform antigen-specific B-cell sorting on the available donor, generating a functional mAb library.
- Assess the S glycoprotein binding $IgG^+$ B-cells present in sequential PBMCs samples.
- Sequence variable regions of mAbs, and examine features such as germline gene usage, level of SHM and CDRH3 length.

- Characterise the regions of S glycoprotein bound by mAbs through ELISA, and measure neutralisation capability using pseudovirus and authentic virus neutralisation assays.
- Perform competition ELISAs on neutralising mAbs to identify which neutralising epitopes are targeted by mAbs
- Investigate the ability of the mAbs to inhibit interactions between S glycoprotein and ACE2.
- Inspect the impact of S glycoprotein mutations from recent VOCs on the neutralising activity of the ChAdOx1-nCoV recipient plasma and mAbs.
- Revert selected mAb variable regions to a germline configuration and examine the impact on binding, as well as neutralisation breadth and potency.

### 5.2.2 Human subject and ethics

This study used human samples from one donor collected as part of a study entitled ''Antibody responses following COVID-19 vaccination''. Ethical approval was obtained from the King's College London Infectious Diseases Biobank (IBD) (KDJF-110121) under the terms of the IDB's ethics permission (REC reference: 19/SC/0232) granted by the South Central Hampshire B Research Ethics Committee in 2019. VA14 was male and 23 years old.

### 5.2.3 Plasma binding and neutralisation potency after ChAdOx1-nCoV vaccination

Plasma and PBMCs were collected from donor VA14 (white male of 23 years) at 4 months (timepoint 1, TP1) and 9 months (TP2) post-vaccination with two doses of ChAdOx1-nCoV, given at a 12 week interval (Figure 5.9a). VA14 was considered to be SARS-CoV-2 naïve, based on regular PCR self-testing and N-specific IgG being undetectable in plasma samples via ELISA. Plasma IgG reactivity against S glycoprotein was confirmed by ELISA, with binding titres reducing between TP1 and TP2 (Figure 5.9b). Semi-quantitative ELISAs were also carried out, with 0.39µg/mL and 0.17µg/mL being measured at TP1 and TP2 respectively.

The neutralisation potency of the plasma samples was examined using neutralisation assays with ancestral and VOCs pseudoviruses, or authentic virus. Neutralising activity was low at TP1 with $ID_{50}$ values of ~100 against WT and P.1 (γ), and was reduced against B.1.1.7 (α), B.1.351 (β) and B.1.617.2 (δ; Figure 5.9c). Even though weak binding to S glycoprotein was measured at TP2, neutralisation was not detectable against any of the pseudoviruses tested ($ID_{50}$ <20) (Figure 5.9d). Similarly, plasma neutralisation potency at TP1 was low and was undetectable at TP2 ($ID_{50}$ <10) against authentic virus (Figure 5.9e).

### 5.2.4 S glycoprotein reactive B-cells detected up to 1 year following ChAdOx1-nCoV vaccination

The proportion of S glycoprotein binding IgG B-cells present at TP1 and TP2 was determined via flow cytometry. It was observed that 0.25% of the IgG B-cells at TP1 were S glycoprotein reactive (Figure 5.9f). Despite showing weaker IgG binding and undetectable neutralising activity, 0.27% of IgG B-cells were S glycoprotein reactive at TP2.

### 5.2.5 ChAdOx1-nCoV vaccination elicits antibodies targeting epitopes on S, NTD, RBD and S2

Antigen-specific B-cell sorting was used to select RBD or S glycoprotein reactive B-cells from TP1 PBMCs. Using the same mAb cloning and production pipeline described previously (section 4.1.3), a total 44 S glycoprotein targeting ChAdOx1-nCoV-elicited mAbs were isolated from VA14.

ELISAs were performed to measure mAb binding to S, RBD, NTD, S1 and S2 antigens (Figure 5.10a). Of the 40 mAbs sorted using S glycoprotein bait, 45% were RBD-specific, 35% were NTD-specific, 17.5% were S2-specific and 2.5% were only specific for S antigen (Figure 5.10b). An extra four mAbs were isolated from VA14 using a RBD bait, all of which were shown to be RBD-specific by ELISA. Of note, a

similar distribution of mAbs targeting S, RBD, NTD, and S1 antigens was observed in convalescent donors (P008 and P054) [580].

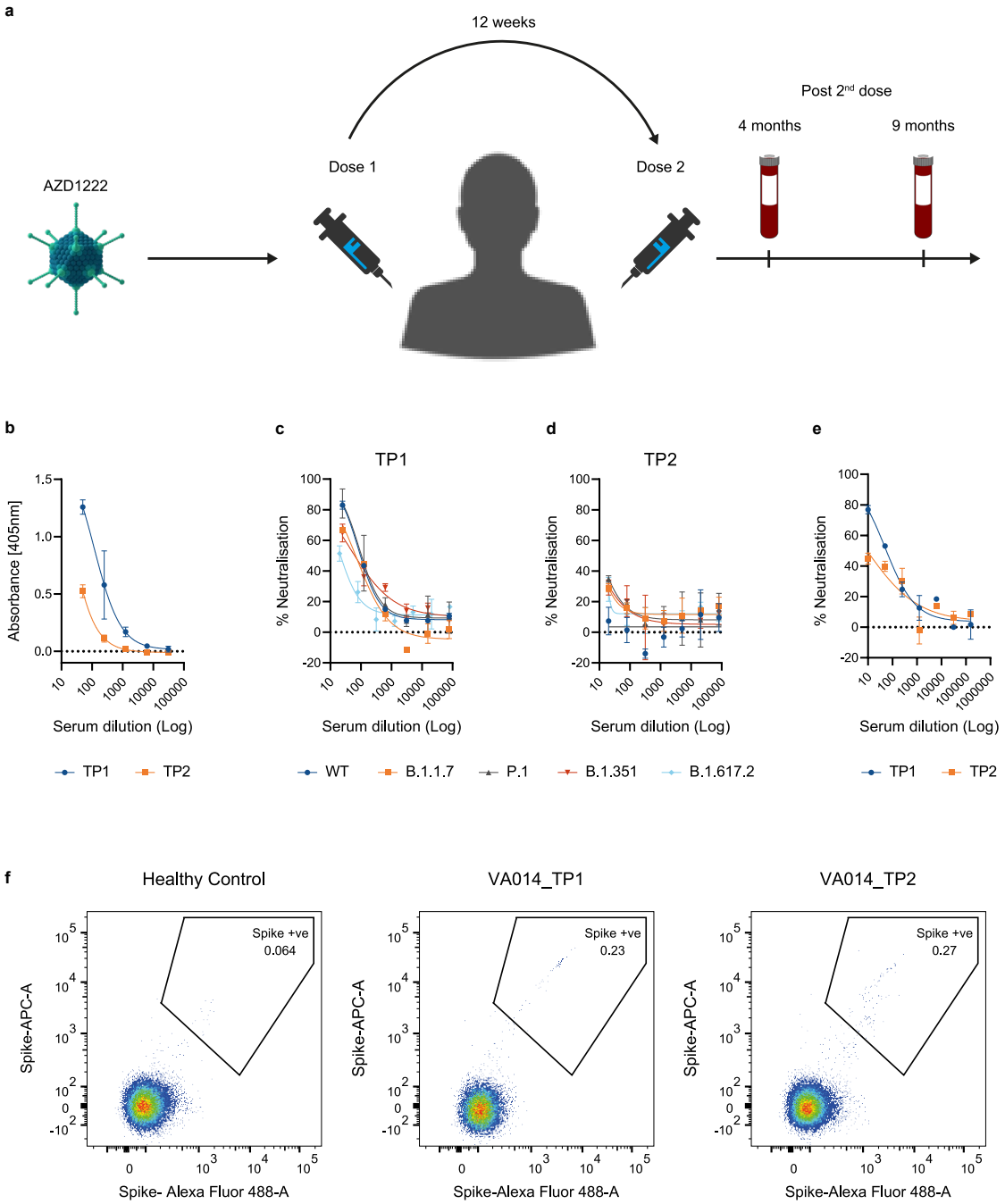



*Figure 5.9: Summary of VA14 vaccine doses and bleeds, plasma binding, plasma neutralisation and S-reactive B-cells. (**a**) Timeline of ChAdOx1 nCoV-19 vaccination, and blood sampling for donor VA14. (**b–d**) Plasma IgG binding to S at TP1 (4 months post-booster) and TP2 (9 months post-booster). Plasma-neutralising activity against HIV-1-based virus particles, pseudotyped with the Wuhan (WT), B.1.1.7 (α), P.1 (γ), B.1.351 (β) and B.1.617.2 (δ) S at (**c**) TP1 and (**d**) TP2. Experiments were performed in duplicate and repeated twice. A representative dataset is shown. Error bars represent the range of the value for experiments performed in duplicate. (**e**) Plasma neutralising activity against SARS-CoV-2 (England 02/2020/407,073) at TP1 and TP2. Experiments were performed in duplicate. (**f**) FACS showing percentage of CD19⁺ IgG⁺ B-cells binding to SARS-CoV-2 S at TP1 and TP2. A healthy control PBMC sample collected prior to the COVID-19 pandemic was used to measure background binding to S. Contributions to data collection from Jeffrey Seow, Sadie Hallett, Thomas Lechmere, Thomas Maguire, Isabella Huettner, Daniel Cox, Hataf Khan and Suzanne Pickering.*

### 5.2.6 ChAdOx1-nCoV vaccination elicits neutralising and non-neutralising antibodies that target epitopes across S glycoprotein

The neutralisation potency of the ChAdOx1-nCoV-elicited mAbs isolated from VA14 was determined using neutralisation assays, with pseudovirus expressing ancestral SARS-CoV-2 S glycoprotein. It was found that 59.1% of the mAbs had neutralising activity, of which 80.8% were RBD-specific, 15.5% were NTD-specific and 3.8% were only specific for S antigen (Figure 5.10a). Further, 95.5% of the RBD-

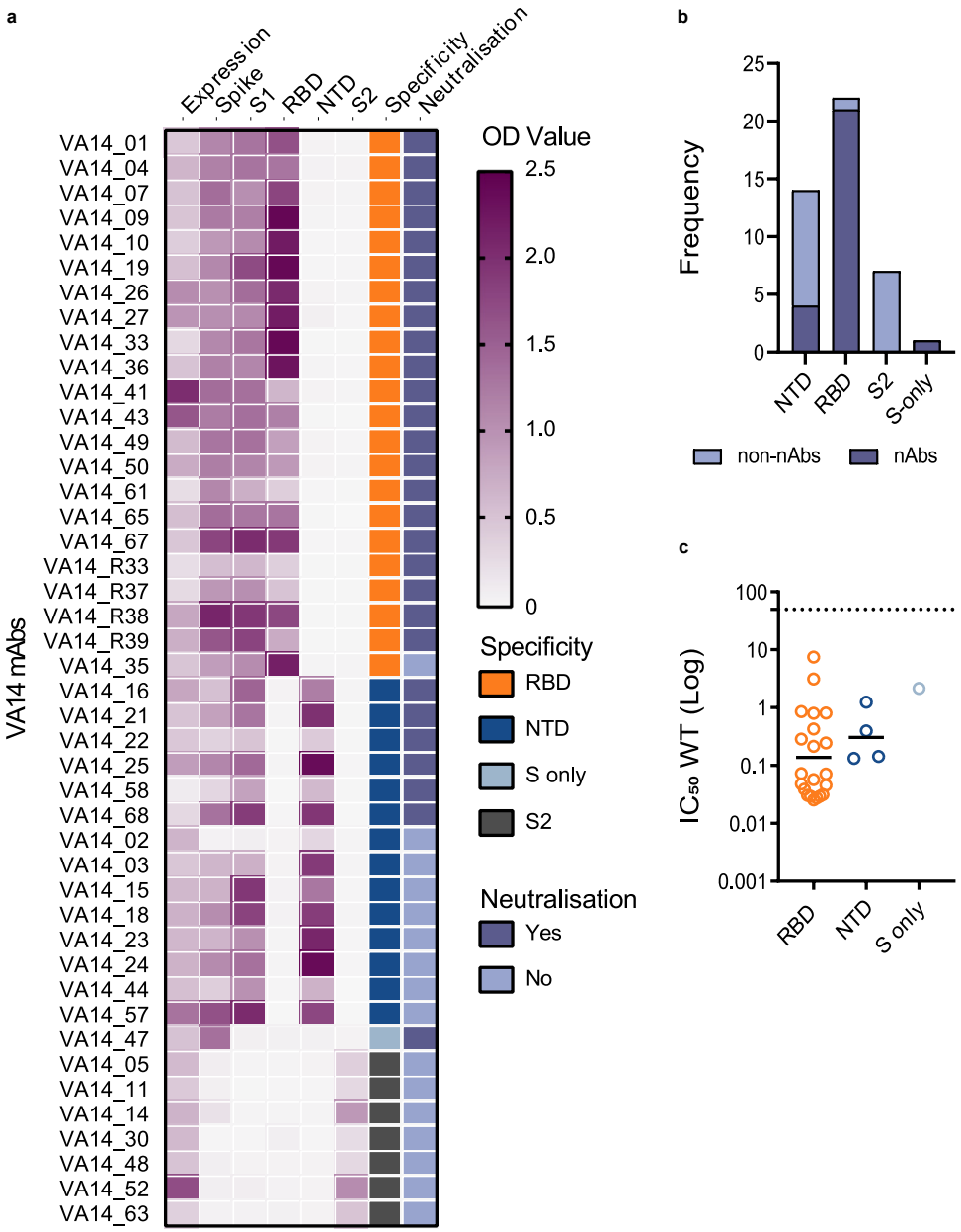



*Figure 5.10: ChAdOx1 nCoV-19 elicited neutralising and non-neutralising antibodies to RBD, NTD, S1, and S2 domains of S. (**a**) Heatmap showing IgG expression level and binding to SARS-CoV-2 S domains, RBD, NTD, S1, and S2. The figure reports OD values from a single experiment (range 0–2.5) for undiluted supernatant from small-scale expressions of 44 cloned mAbs. Antigen binding was considered positive when OD at 405 nm was >0.2 after background was subtracted. SARS-CoV-2 S domain specificity for each antibody is indicated. Neutralising activity was measured against WT pseudotyped virus using either small-scale purified IgG or concentrated supernatant. (**b**) Frequency of neutralising and non-neutralising antibodies targeting RBD, NTD, S-only, or S2. Graph includes only mAbs isolated using S as antigen bait for B-cell sorting. (**c**) Neutralisation potency ($IC_{50}$) against WT pseudotyped virus for mAbs targeting RBD, NTD, or non-S1. The black line represents the geometric mean $IC_{50}$. Neutralisation experiments were performed in duplicate and carried out at least twice. Contributions to data collection from Jeffrey Seow, Sadie Hallett, Thomas Lechmere, Thomas Maguire, Isabella Huettner, Daniel Cox, Hataf Khan and Suzanne Pickering.*

specific mAbs and 38.6% of the NTD-specific mAbs were neutralisers, whilst none of the S2-specific mAbs had neutralisation capability (Figure 5.10b). The neutralisation potencies of the mAbs against WT pseudovirus ranged from 0.01–7.3µg/mL, and the RBD-specific mAbs had a lower geometric mean $IC_{50}$ compared with NTD-specific mAbs (Figure 5.10c). Overall, in this subject the RBD was the dominant target of the neutralising mAbs and the NTD acted as the subdominant epitope, similarly to what was observed in convalescent donors [541,560,580].

### 5.2.7 ChAdOx1-nCoV-elicited mAbs are more highly mutated than mAbs from natural infection

The $V_H$ and $V_L$ genes of the mAbs were sequenced, and the predicted germline gene usage, %SHM and CDRH3 lengths were analysed using the IMGT database [433]. Mean SHM levels of 4.9% and 2.8% were found for the $V_H$ and $V_L$ genes respectively (isolated at 4 months post-vaccination), which was higher than what was observed in the convalescent donors (1.9% and 1.4% for $V_H$ and $V_L$ respectively) (Figure 5.11a) [580]. Notably, three pairs of clonally related mAbs were identified from VA14 (Figure 5.12a).

The germline gene usage and divergence from germline of neutralising and non-neutralising VA14 mAbs was compared to a database of SARS-CoV-2 targeting mAbs (n = 1,292) isolated from convalescent donors, and variable region sequences of paired IgG BCRs from healthy individuals (n = 862) [597,598]. Unfortunately, the SARS-CoV-2 mAb database only listed amino acid sequences, therefore the divergence from germline was examined at the amino acid level. The amino acid divergence from germline correlated well with the nucleotide %SHM ($r^2 = 0.8196$, $P < 0.0001$ and $r^2 = 0.7701$, $P < 0.0001$ for $V_H$ and $V_L$ respectively) (Figure 5.12b). The ChAdOx1-nCoV-elicited mAbs from VA14 showed statistically higher amino acid divergence from germline compared to the SARS-CoV-2 convalescent mAb database, but had a similar level to what was seen in the healthy donor BCRs (Figures 5.11b-c). With the exception of the $V_H$ genes from the SARS-CoV-2 convalescent mAb database, the levels of amino acid divergence from germline did not significantly differ between binding and neutralising mAbs (Figure 5.12c).

VH3-30 and VH3-53 germline genes were enriched in both SARS-CoV-2 infection and ChAdOx1-nCoV-elicited mAbs, and this has also been seen in SARS-CoV-2 mRNA vaccine-elicited mAbs (Figure 5.11d) [596]. It was found that 14.3% of the RBD-specific neutralising mAbs used the VH3-53/3-66 germlines, which have been associated with mAbs that bind directly to the RBM [540,568]. There was an enrichment of VH3-15, VH3-48, VH4-34, and VH4-59 germlines in the ChAdOx1-nCoV-elicited mAbs compared to the SARS-CoV-2 convalescent mAbs. The germlines VK3-20, VK1-39, VK3-15 and VK1-33 were enriched in the ChAdOx1-nCoV-elicited mAbs compared to the healthy donor BCRs (Figure 5.11e).

### 5.2.8 ChAdOx1-nCoV-elicited neutralising mAbs bind epitopes overlapping with mAbs generated in response to SARS-CoV-2 infection

To examine the epitopes targeted by the ChAdOx1-nCoV-elicited neutralising mAbs, competition ELISAs with S glycoprotein were performed against previously characterised neutralising mAbs isolated from convalescent donors [580]. ChAdOx1-nCoV-elicited mAbs IgG molecules were competed against $F(ab')_2$ fragments that were representative of the previously identified competition Groups 1-7. In addition, the ability of the neutralising mAbs to inhibit interaction between S glycoprotein and the ACE2 receptor was tested using flow cytometry.

It was found that ~19%, ~5, ~14% and ~62% of the RBD-specific neutralising mAbs isolated from VA14 competed with Group 1 (Class 4), Group 2 (overlaps Class 4 and Class 1), Group 3 (Class 1 and Class 2) and Group 4 (Class 3) respectively (Figures 5.13a and Figure 5.13e). The Group 1 competing mAbs displayed a wide range of neutralisation potencies and competed with ACE2 to varying degrees (Figures 5.13c-d). There was only one neutralising mAb (VA14_26) that competed with Group 2, it showed modest neutralisation potency and strong ACE2 competition. The Group 3 competing mAbs

used VH3-53/VH3-66 germlines, had the best mean neutralisation activity and competed strongly with ACE2. The majority of the VA14 RBD-specific mAbs competed with Group 4, with modest to potent neutralisation activity and a range of ACE2 competition levels.

The NTD-specific mAbs isolated from VA14 formed three competition groups (Figure 5.13b). The weakly neutralising mAbs VA14_21 and VA14_22 competed with Group 5, and were able to inhibit ACE2 interaction by 51%–58% (Figures 5.13c-d). The non-neutralising mAbs VA14_25 and VA14_58 competed with Group 6, and displayed weak ACE2 competition. Interestingly, VA14_16 and VA14_68 did not compete with any of the previously identified NTD-specific mAbs and could be considered as a novel group (Group 8). Group 8 showed better neutralisation potency against WT pseudovirus

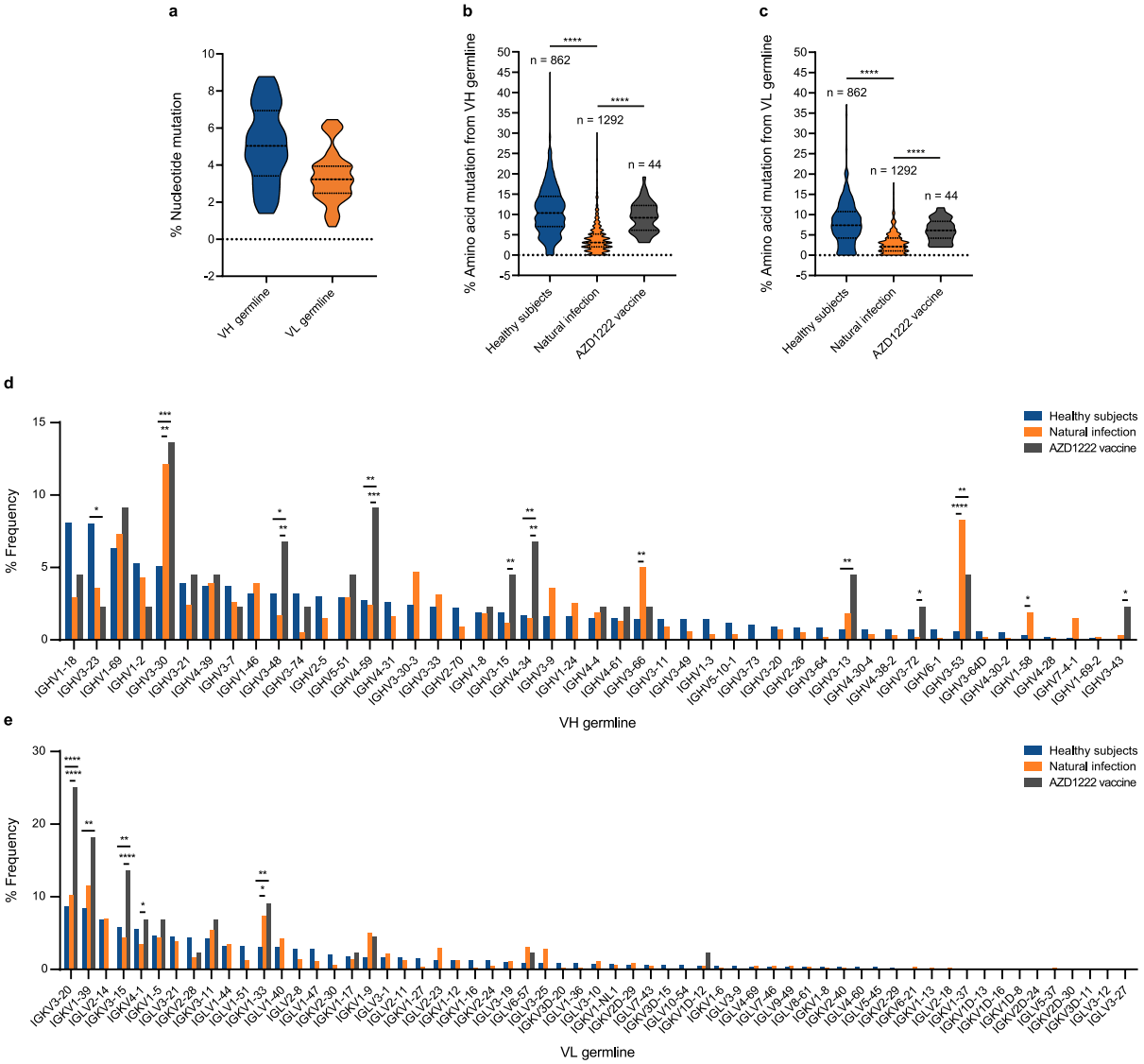


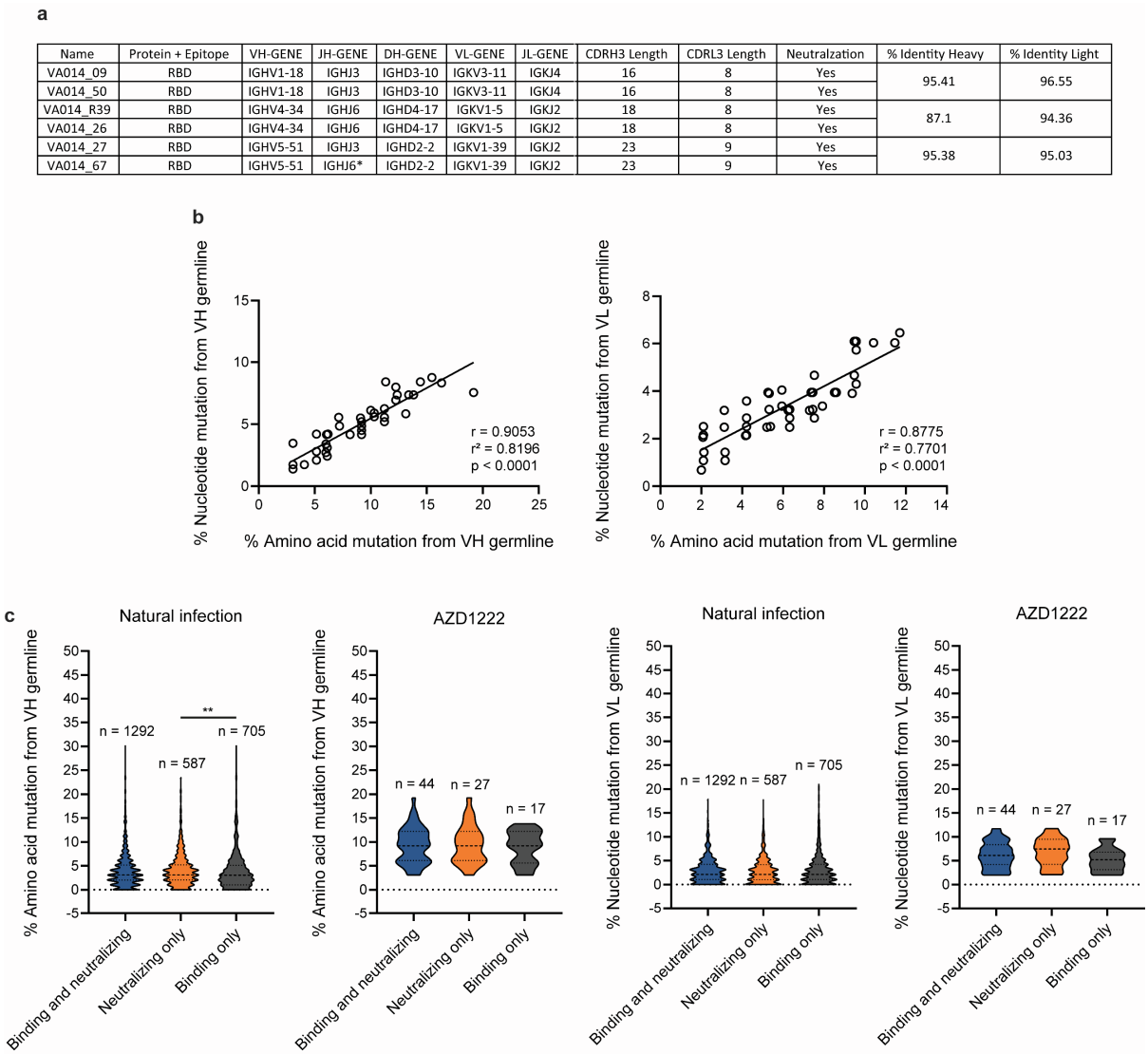

*Figure 5.11: ChAdOx1 nCoV-19 mAb SHM and germline usage compared to natural infection and healthy subjects.* ***(a–c)*** *Truncated violin plot showing the percentage of nucleotide mutation compared with germline for the $V_H$ and $V_L$ genes of S-reactive mAbs isolated following ChAdOx1 nCoV-19. Divergence from germline (based on amino acid alignments) for $V_H$ and $V_L$ genes for S-reactive mAbs from natural infection, ChAdOx1 nCoV-19 vaccination, and IgG BCRs from SARS-CoV-2-naive individuals. D'Agostino and Pearson tests were performed to determine normality. Based on the result, a Kruskal-Wallis test with Dunn's multiple comparison post hoc test was performed. *p < 0.0332, **p < 0.0021, ***p < 0.0002, and ****<0.0001.* ***(d-e)*** *Graph showing the relative abundance of $V_H$ and $V_L$ genes in mAbs elicited from ChAdOx1 nCoV-19 vaccination compared with SARS-CoV-2 infection mAbs and IgG BCRs from SARS-CoV-2-naive individuals. Two-sided binomial tests were used to compare the frequency distributions. *p < 0.0332, **p < 0.0021, ***p < 0.0002, and ****<0.0001. Contributions to data collection from Jeffrey Seow, Sadie Hallett, Thomas Lechmere, Thomas Maguire, Isabella Huettner, Daniel Cox, Hataf Khan and Suzanne Pickering.*

compared to the NTD-specific mAbs isolated previously from convalescent donors, and competed with ACE2 weakly.

Only one mAb (VA14_47) capable of competing with Group 7 was isolated from VA14 (Figure 5.13b). VA14_47 has low neutralisation activity and competed ACE2 interaction by 59% (Figures 5.13c-d). Antibodies targeting this epitope appear to be rarer, but have been found in multiple donors.

### 5.2.9 ChAdOx1-nCoV-elicited neutralising mAbs cross-neutralise SARS-CoV-2 VOCs

At the time of this work five VOCs had been identified, these included B.1.1.7 (α), B.1.351 (β), P.1 (γ), B.1.617.2 (δ) and B.1.1.529 (ο). Previous work had shown that mutations in B.1.1.7 (α) impacted the neutralisation capability of specific neutralising mAbs isolated from convalescent donors [546,580]. The S



a

| Name | Protein + Epitope | VH-GENE | JH-GENE | DH-GENE | VL-GENE | JL-GENE | CDRH3 Length | CDRL3 Length | Neutralzation | % Identity Heavy | % Identity Light |
|---|---|---|---|---|---|---|---|---|---|---|---|
| VA014_09 | RBD | IGHV1-18 | IGHJ3 | IGHD3-10 | IGKV3-11 | IGKJ4 | 16 | 8 | Yes | 95.41 | 96.55 |
| VA014_50 | RBD | IGHV1-18 | IGHJ3 | IGHD3-10 | IGKV3-11 | IGKJ4 | 16 | 8 | Yes | | |
| VA014_R39 | RBD | IGHV4-34 | IGHJ6 | IGHD4-17 | IGKV1-5 | IGKJ2 | 18 | 8 | Yes | 87.1 | 94.36 |
| VA014_26 | RBD | IGHV4-34 | IGHJ6 | IGHD4-17 | IGKV1-5 | IGKJ2 | 18 | 8 | Yes | | |
| VA014_27 | RBD | IGHV5-51 | IGHJ3 | IGHD2-2 | IGKV1-39 | IGKJ2 | 23 | 9 | Yes | 95.38 | 95.03 |
| VA014_67 | RBD | IGHV5-51 | IGHJ6* | IGHD2-2 | IGKV1-39 | IGKJ2 | 23 | 9 | Yes | | |



*Figure 5.12: ChAdOx1 nCoV-19 elicited mAbs, binding or neutralising, SHM compared to natural infection.*
*(**a**) Clonally related mAbs isolated from VA14. (**b**) Correlation between the % nucleotide mutation and % amino acid mutation for $V_H$ and $V_L$ germline for ChAdOx1 nCoV-19 elicited mAbs (Spearman correlation, two-tailed, r). A linear regression was used to calculate the goodness of fit, $r^2$. (**c**) Divergence from germline (based on amino acid alignments) for $V_H$ and $V_L$ genes for S-reactive mAbs arising from natural infection and ChAdOx1 nCoV-19 vaccination. S-reactive mAbs have been separated based on their binding and neutralising properties. D'Agostino & Pearson tests were performed on each dataset to determine normality. Based on the result, either a Kruskal-Wallis test with Dunn's multiple comparison post hoc test or an ordinary one-way ANOVA with Turkey's multiple comparison post hoc test was performed. *p<0.0332, **p<0.0021, ***p<0.0002 and ****<0.0001. Contributions to data collection from Jeffrey Seow, Sadie Hallett, Thomas Lechmere, Thomas Maguire, Isabella Huettner, Daniel Cox, Hataf Khan and Suzanne Pickering.*

glycoproteins of the VOCs carry mutations located across the RBD, NTD and S2 (Figure 5.14a). Some RBD mutations are found in multiple variants, for instance B.1.1.7 (α), B.1.351 (β), P.1 (γ) and B.1.1.529

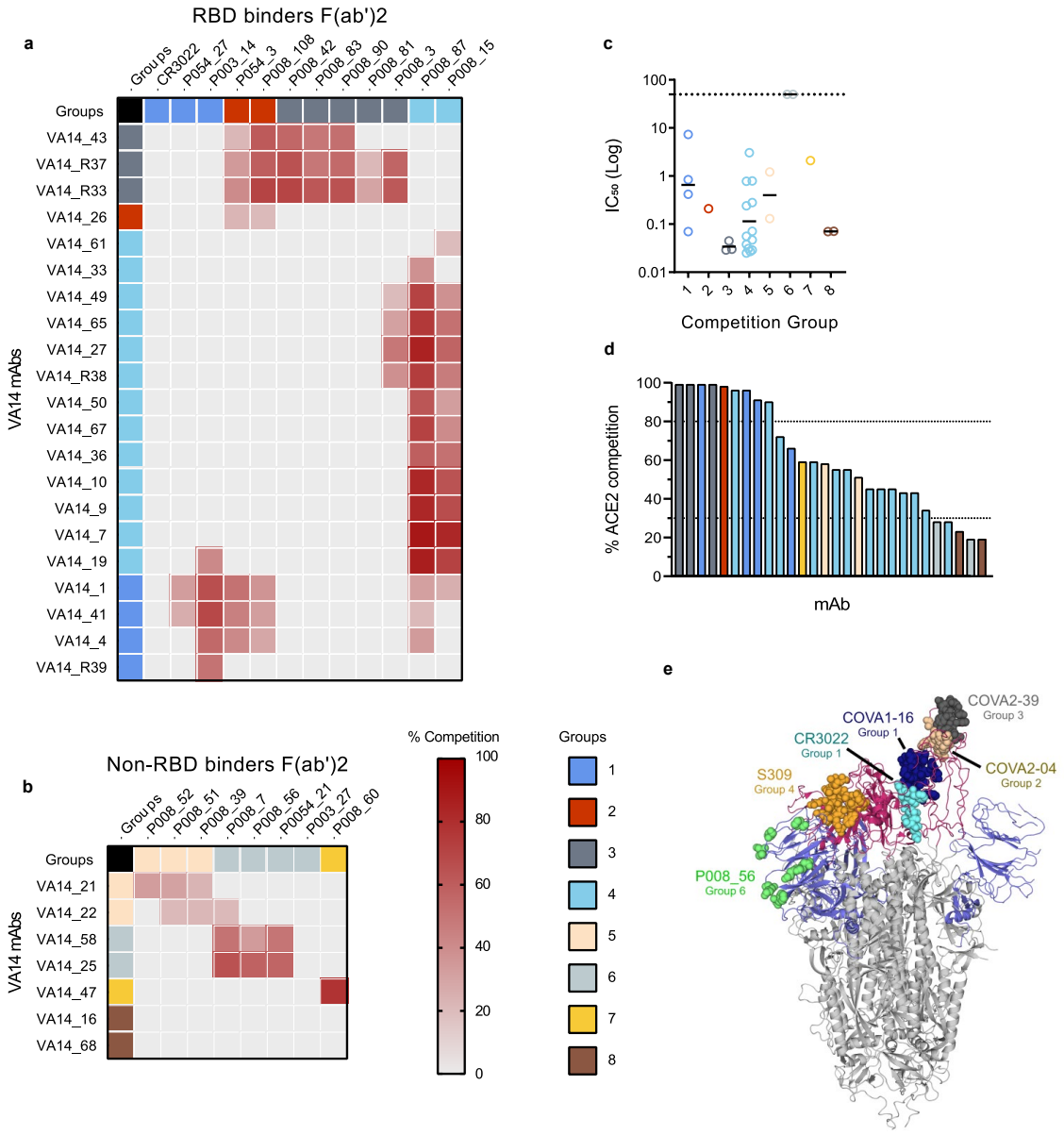



*Figure 5.13: ChAdOx1 nCoV-19 neutralising mAbs epitopes compared to those elicited following natural infection.* ***(a-b)*** *Competitive binding of ChAdOx1 nCoV-19 and SARS-CoV-2 infection-elicited neutralising antibodies. Inhibition of IgG binding to S by F(ab)$_2$' fragments was measured. The percentage of competition was calculated using the reduction in IgG binding in the presence of F(ab')$_2$ (at 100 molar excess of the IC$_{80}$) as a percentage of the maximum IgG binding in the absence of F(ab')$_2$. Competition was measured between RBD-specific and NTD-specific/S-only neutralising mAbs. Competition groups were determined according to binding epitopes. Experiments were performed in duplicate. Competition <25% is in white.* ***(c)*** *The IC$_{50}$ of mAbs targeting RBD, NTD or non-S1, and competition groups 1–8 against SARS-CoV-2 WT pseudotyped virus. The black lines represent the geometric mean IC$_{50}$ for each group. Neutralisation experiments were performed in duplicate and carried out at least twice.* ***(d)*** *Ability of neutralising mAbs to inhibit the interactions between cell surface ACE2 and soluble SARS-CoV-2 S. Neutralising mAbs (at 600 nM) were pre-incubated with fluorescently labelled S before addition to HeLa-ACE2 cells. The percentage reduction in mean fluorescence intensity is reported. Experiments were performed in duplicate.* ***(e)*** *Mapping of previously determined neutralising and non-neutralising epitopes on SARS-CoV-2 S. Cartoon representation of S showing antibody-binding footprint for neutralising antibodies used in competition ELISA as coloured spheres. Epitopes were previously determined using crystal structures or cryo-electron microscopy of RBD or S-Fab complexes; COVA2-04 (gold, group 2 [RBD Class 1]), COVA2-39 (grey, group 3 [RBD Class 2]), S309 (orange, group 4 [RBD Class 3]), COVA1-16 and CR3022 (dark blue and turquoise, group 1 [RBD Class 4]), and P008_056 (green, NTD group 6). Adapted from [388,555,563,568,645,646]. PDB = 6VSB. Contributions to data collection from Jeffrey Seow, Sadie Hallett, Thomas Lechmere, Thomas Maguire, Isabella Huettner, Daniel Cox, Hataf Khan and Suzanne Pickering.*

(o) all encode the N501Y mutation. However, mutations in the NTD vary substantially across VOCs and can include indels. The B.1.1.529 (o) has over >30 mutations, and had been shown to escape neutralisation by multiple SARS-CoV-2-specific mAbs and the sera of COVID-19 vaccine recipients

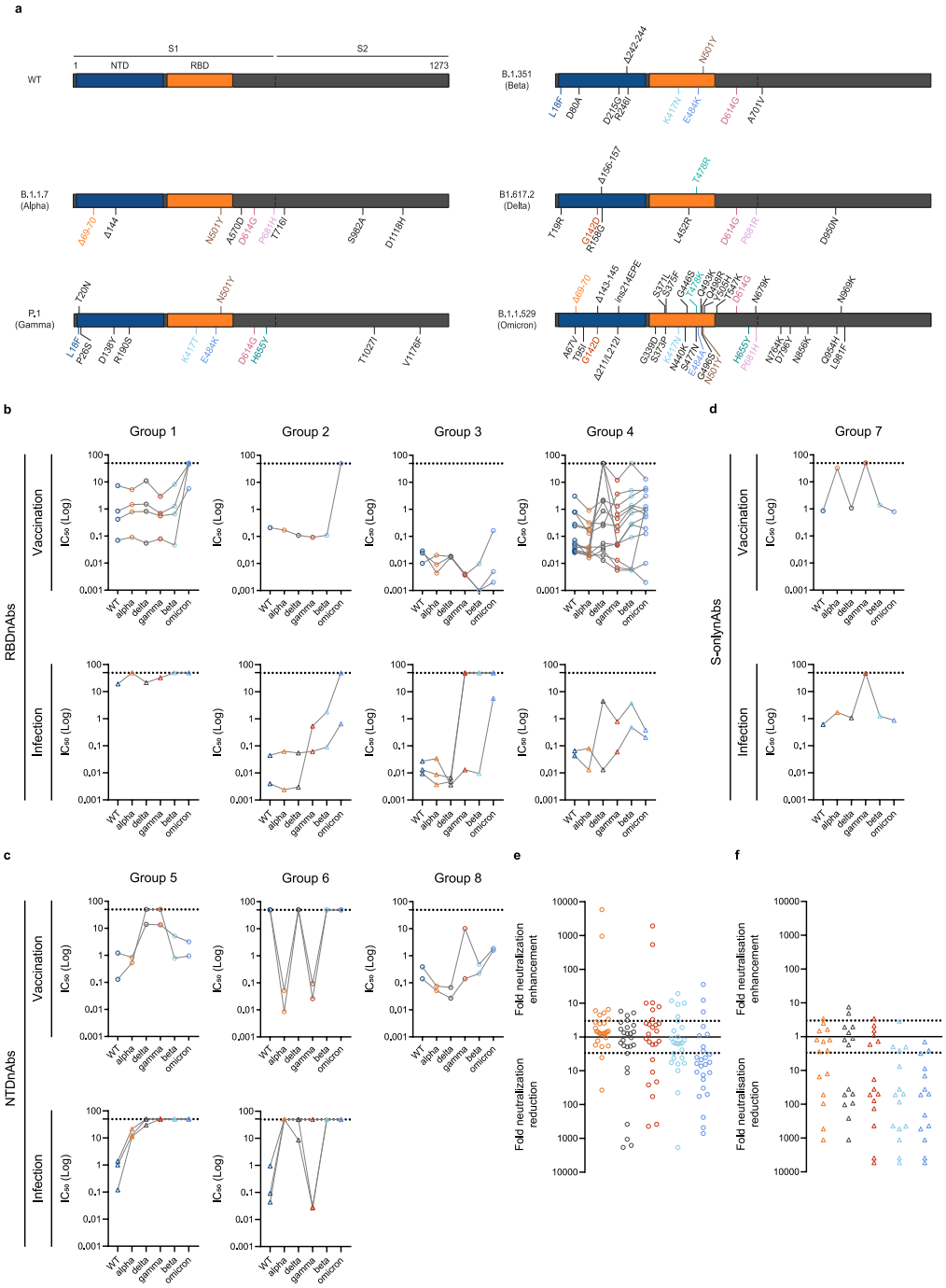



*Figure 5.14: Cross-reactivity of ChAdOx1 nCoV-19 generated mAbs against VOCs.*
*(**a**) Schematic showing mutations present in the S of SARS-CoV-2 viral VOCs. (**b**) Neutralisation by RBD-specific neutralising mAbs isolated following ChAdOx1 nCoV-19 vaccination or SARS-CoV-2 infection against main VOCs. Neutralising mAbs are separated by competition group (groups 1–4). (**c**) Neutralisation by NTD-specific neutralising mAbs isolated following ChAdOx1 nCoV-19 vaccination or SARS-CoV-2 infection against main VOCs. Neutralising mAbs are separated by competition group (groups 5, 6, and 8). (**d**) Neutralisation by S-only-specific neutralising mAbs isolated following ChAdOx1 nCoV-19 vaccination or SARS-CoV-2 infection against main VOCs. Neutralisation experiments were performed in duplicate and carried out at least twice. (**e-f**) Fold enhancement or reduction in neutralisation $IC_{50}$ against VOCs compared with the $IC_{50}$ against WT for ChAdOx1 nCoV-19-elicited mAbs and infection mAbs. The dotted line indicates a 3-fold reduction or enhancement in neutralisation. Contributions to data collection from Jeffrey Seow, Sadie Hallett, Thomas Lechmere, Thomas Maguire, Isabella Huettner, Daniel Cox, Hataf Khan and Suzanne Pickering.*

[386,599–601]. Hence, the neutralising mAbs from VA14, and representative mAbs from convalescent donors, were assessed using neutralisation assays with pseudovirus particles bearing the S glycoprotein of several SARS-CoV-2 variants.

All of the VA14 Group 1 mAbs were able to neutralise the B.1.1.7 (α), B.1.351 (β), P.1 (γ) and B.1.617.2 (δ) variants without losing neutralisation potency (Figure 5.14b). However, neutralisation activity against the B.1.1.529 (o) strain was reduced in all of the mAbs from this group, with total escape observed in three members. Similarly, VA14_26 (Group 2) was able to cross-neutralise all of the VOCs with the exception of B.1.1.529 (o), in which neutralisation was abolished. The VA14 Group 3 mAbs were able to potently neutralise all of the VOCs, with two members showing enhanced neutralisation of the B.1.351 (β), P.1 (γ) and B.1.1.529 (o) variants. The VA14 Group 4 mAbs were more varied in their ability to neutralise VOCs. It was observed that ~46% of this group were able to cross-neutralise all of the VOCs without a reduction in potency. Conversely, the remaining ~54% of this group displayed a >3-fold decrease in neutralising activity against at least one VOC, with the largest reductions being found against B.1.351 (β), B.1.617.2 (δ) and B.1.1.529 (o).

The VA14 Group 5 mAbs were able to cross-neutralise all of the VOCs, with the exception of the P.1 (γ) and B.1.617.2 (δ) variants where neutralisation activity was either reduced or abolished (Figure 5.14c). Interestingly, the VA14 Group 6 mAbs that were unable to neutralise WT pseudovirus were shown to neutralise the B.1.1.7 (α) and P.1 (γ) variants. Previous work found that some of the Group 6 mAbs isolated from convalescent donors were unable to neutralise WT pseudovirus due to the presence of biliverdin, but could potently neutralise authentic virus [563,580]. Therefore, neutralisation assays with authentic virus were carried out with the VA14 Group 6 mAbs and it was shown that these mAbs can potently neutralise the WT strain in this setting (Figure 5.15f). The VA14 Group 7 mAb (VA14_47) was able to neutralise B.1.351 (β), B.1.617.2 (δ) and B.1.1.529 (o), but showed reduced and abolished neutralisation activity against B.1.1.7 (α) and P.1 (γ) respectively (Figure 5.14d). The VA14 Group 8 mAbs maintained neutralisation capability against the VOCs, but a large decrease in potency against P.1 (γ) was observed for VA14_68 (Figure 5.14c). Conversely, VA14_16 potently neutralised B.1.1.7 (α), B.1.351 (β), P.1 (γ) and B.1.617.2 (δ; $IC_{50}$ values <0.22µg/mL), and neutralised B.1.1.529 (o) with an $IC_{50}$ of 1.59µg/mL. At the time of this work, VA14_16 was the only NTD-specific mAb that had been reported to cross-neutralise all of these VOCs [560].

When the neutralising mAbs from VA14 were compared to those from convalescent donors, the convalescent mAbs were found to be less sensitive to the S glycoprotein of the VOCs. The RBD-specific mAbs elicited from infection often showed greatly reduced neutralisation potencies against VOCs, particularly B.1.351 (β), P.1 (γ) and B.1.1.529 (o; Figure 5.14b). Of note, each of these variants contain RBD mutations at amino acid positions 417, 484, and 501. The NTD-specific mAbs isolated from the convalescent donors also had weaker cross-neutralising capability (Figure 5.14c). For instance, the infection Group 5 mAbs lacked neutralisation activity against B.1.351 (β), P.1 (γ) and B.1.1.529 (o). Further, the infection Group 6 mAbs were only able to neutralise WT and P.1 (γ), whilst the VA14 Group 6 mAbs were able to neutralise B.1.1.7 (α) and P.1 (γ) pseudovirus and WT authentic virus. Due to the low sample number, statistical analysis was not performed on this data. Overall, it would appear that despite competing for the same epitopes, the ChAdOx1-nCoV-elicited mAbs displayed a greater resistance S glycoprotein mutations compared to the infection-elicited mAbs (Figure 5.14e-f).

### 5.2.10 Role of SHM in neutralisation breadth

To investigate the impact of increased SHM on neutralisation breadth, four RBD-specific neutralising mAbs were expressed with their variable genes reverted to germline. The Group 1 mAbs VA14_01 and VA014_04 both used the VH3-13 (5.6% and 0.2% SHM respectively) and VK1-39 germlines (1.4% and

6.4% SHM respectively). The germline version of VA014_01 was not able to bind to WT S1 when tested in ELISA (Figure 5.15a). The germline version of VA014_04 bound weakly to WT S1, but did not show neutralisation activity against WT or any VOCs (Figures 5.15a-b). This demonstrated that SHM was crucial for antigen recognition and neutralisation in these Group 1 mAbs.

The Group 3 mAbs VA14R_33 and VA14R_37 also had their variable genes reverted to germline. VA14R_33 used VH3-66 (8.4% mutated) and VK1-33 (3.9% mutated) germlines, whereas VA14R_37 used VH3-53 (2.4% mutated) and VK3-20 (6.0% mutated). The reverted versions of both VA14R_33 and VA14R_37 retained the ability to bind to WT S1 (Figure 5.15c). This was expected as other studies have found that these germlines have amino acid motifs that are pre-configured to target the RBM of SARS-CoV-2 [602,603].

The germline version of VA14R_33 had reduced neutralisation potency against B.1.1.7 (α), B.1.351 (β) and B.1.1.529 (o), but neutralisation against WT and B.1.617.2 (δ) was not largely impacted (Figure 5.15d). The germline version of VA14R_37 was not able to neutralise B.1.1.529 (o) and had decreased potency against B.1.351 (β; Figure 5.15e). Notably, the germline version of VA14R_37 could neutralise WT and B.1.617.2 (δ) more potently compared to VA14R_37. These results further highlighted that SHM was important for neutralisation breadth and potency against VOCs.

### 5.2.11 Discussion

Gaining a more complete understanding of the of COVID-19 vaccines elicited neutralising antibodies function against emerging SARS-CoV-2 VOCs will be crucial for control of the current pandemic. The antibody response to two doses of the ChAdOx1-nCoV vaccine, given at a 12-week interval, was examined at the mAb level in donor VA14. Most studies investigating polyclonal immune sera from ChAdOx1-nCoV-vaccinated subjects have shown a lower neutralisation potency against B.1.1.7 (α;

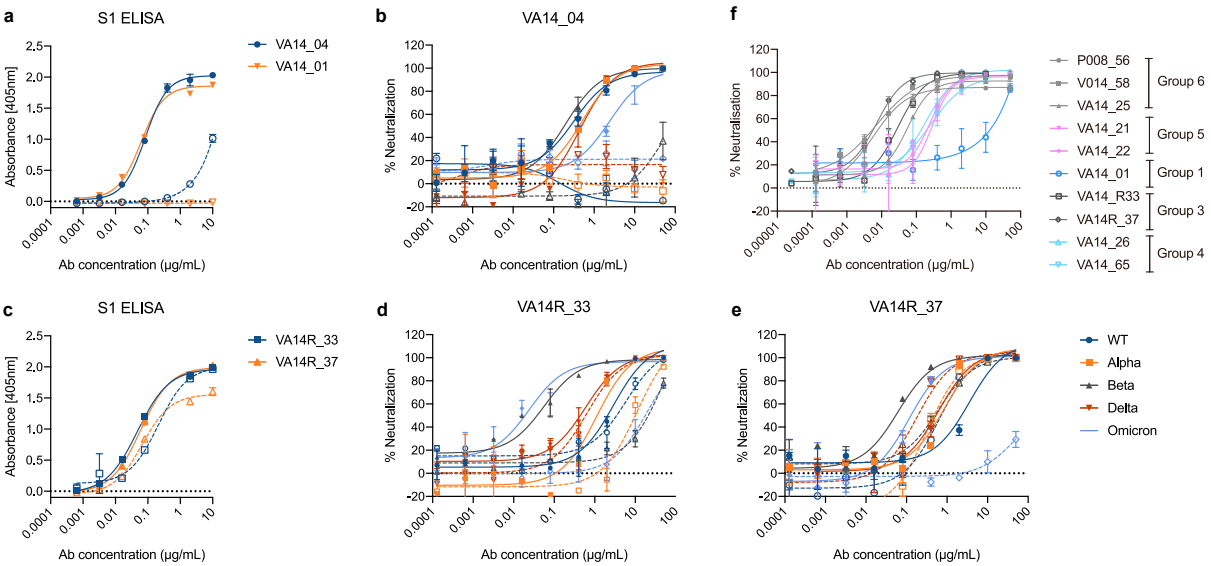



*Figure 5.15: Neutralising activity of reverted germline mAbs isolated from VA14.*
*(**a**) Binding of Group 1 reverted mAbs to WT S1 by ELISA. Reverted mAbs (VA14_01_rev and VA14_04 rev) are shown in open symbols and dotted lines. (**b**) Comparison of neutralisation activity for VA14_04 and germline reverted mAb against WT, Alpha, Beta, Delta, and Omicron. The reverted mAb is shown with open symbols and dotted line. (**c**) Binding of Group 3 reverted mAbs to WT S by ELISA. Reverted mAbs (VA14R_33_rev and VA14R_37 rev) are shown in open symbols and dotted lines. (**d**) Neutralisation of VA14R_33 and reverted mAb against VOCs. The reverted mAb is shown with open symbols and dotted line. (**e**) Comparison of neutralisation of VA14R_37 and germline reverted mAb against VOCs. The reverted mAb is shown with open symbols and dotted line. Experiments were performed in duplicate and repeated twice. A representative dataset is shown. Error bars represent the range of the value for experiments performed in duplicate. Contributions to data collection from Jeffrey Seow, Sadie Hallett, Thomas Lechmere, Thomas Maguire, Isabella Huettner, Daniel Cox, Hataf Khan and Suzanne Pickering.*

range 2.2 to 9.0-fold), B.1.351 (β; range 4.0 to 9.0-fold), P.1 (γ) (2.9-fold) and B.1.617.2 (δ; range 4.3 to 9.0-fold) compared to the ancestral strain or D614G variant [586,589,591,604–606]. Further, very limited neutralising activity has been observed in sera from ChAdOx1-nCoV-vaccinated individuals against the B.1.1.529 (o) variant [607]. The plasma of VA14 had a low neutralisation potency ($ID_{50}$ ~100) at 4 months post-vaccine booster, but ~59% of the S-specific mAbs isolated from this donor showed neutralising activity against the ancestral strain. Further, many of these mAbs were able to potently cross neutralise VOCs that were circulating at the time of this work. Similar to previous work with SARS-CoV-2 infected donors, the RBD was shown to be the dominant target of neutralising mAbs and the NTD acted as a subdominant epitope [580]. Notably, RBD-specific neutralising mAbs from competition groups 1 to 4, and NTD-specific mAbs were identified that could cross-neutralise all five VOCs tested, including the highly divergent B.1.1.529 (o) variant. Neutralisation of the B.1.1.529 (o) variant by RBD-specific mAbs elicited against the ancestral strain had been described previously, but was of a rare occurrence [387,608,609]. However, neutralisation of the B.1.1.529 (o) variant by NTD-specific mAbs elicited against the ancestral strain had not been reported at the time of this work. Collectively, these results suggested that the polyclonal nature of the neutralising response generated by ChAdOx1-nCoV vaccination would likely help limit complete vaccine escape in response to emerging VOCs.

Competition ELISAs showed that the ChAdOx1-nCoV-elicited mAbs bound to epitopes that overlapped with that of the mAbs generated by SARS-CoV-2 infection. Interestingly, despite targeting similar regions on the S glycoprotein, the ChAdOx1-nCoV-elicited mAbs had better neutralisation breadth compared to the convalescent mAbs. The greater neutralisation breath was likely the result of higher levels of SHM in the ChAdOx1-nCoV-elicited mAbs (isolated 4 months post-booster) compared with the convalescent mAbs (isolated 2–8 weeks POS), enabling improved tolerance of S glycoprotein mutations. It was shown that SHM was essential for antigen recognition by Group 1 mAbs VA14_01 and VA14_04. Binding of these mAbs to S glycoprotein was low or undetectable when their variable regions were reverted to germline, and both lacked neutralisation capability. Conversely, VA14R_33 and VA14R_37 (Group 3) were able to retain S glycoprotein binding when reverted to germline, but showed lower neutralisation breadth against the B.1.351 (β) and B.1.1.529 (o) variants. Structural analysis of RBM targeted neutralising mAbs CC12.1 and CC12.3 has revealed that a motif encoded by the VH3-53 germline (N32-Y33 and S53-G54-G55-S56) enables recognition of S glycoprotein [603]. Another VH3-53 RBM targeted mAb (CV30) has been shown to retain S glycoprotein reactivity but have decreased neutralisation potency, when reverted to germline [610]. Numerous other studies have found that higher levels of SHM improves neutralisation breadth against VOCs [525,611,612]. Structural analysis of the interactions between S glycoprotein and broad mAbs, as well as their germline reverted counterparts, would allow a better understanding of how higher levels of SHM enables neutralisation breadth. Collectively, these findings have shown that analysis of antibody-antigen interactions at the molecular level can provide insights into the mechanisms of neutralisation breadth.

At the time of this work, ChAdOx1-nCoV-elicited mAbs targeting the SARS-CoV-2 S glycoprotein had not been reported, but multiple studies have isolated mAbs from individuals following mRNA COVID-19 vaccination [594–596]. A higher number of RBM-specific mAbs were found in response to mRNA vaccination compared to what was witnessed in VA14 [525,610]. An enrichment of the VH3-53 and VH3-30 germlines had also been shown in subjects that have received mRNA vaccines [596,613]. Monoclonal antibodies have since been isolated from four subjects that received the ChAdOx1-nCoV vaccine, with an extended dose interval. An enrichment of VH3-30 was found but increased levels of other germline genes, such as VH3-53 and VH3-66, were not observed. Unfortunately, the author presented the total number of mutations in the variable regions instead of %SHM from germline, preventing comparison with the work discussed here. The ChAdOx1-nCoV-elicited neutralising mAbs were shown to target four distinct epitopes across the RBD, but whether the mAbs targeted separate epitopes on the NTD

was not investigated. Binding to B.1.351 (β), P.1 (γ), B.1.617.2 (δ) and B.1.1.529 (o) was assessed, and a good proportion of ChAdOx1-nCoV-elicited mAbs were able to cross-react with the VOCs examined [614]. Further studies investigating mAbs from ChAdOx1-nCoV vaccinated subjects is required to confirm this data and enable more powerful comparison with mAbs isolated from other vaccination platforms and natural infection.

The ChAdOx1-nCoV-elicited mAbs found here displayed greater cross-neutralisation capability compared to those described in previous studies, and there are a number of potential reasons for these differences [490,580,615]. Firstly, the time at which mAbs are isolated following S glycoprotein exposure has been shown to have an impact on SHM levels and neutralisation capability [525,580]. The length of vaccination booster intervals (3 weeks for mRNA reports and 12 weeks in this work), has been found to influence the SARS-CoV-2 neutralisation breadth of sera samples collected from mRNA vaccine recipients (manuscript in preparation). Lastly, mRNA-1273 and BNT162b2 encode a modified S glycoprotein, whilst ChAdOx1-nCoV carries unaltered S glycoprotein [390,393]. This, alongside the differences in vaccine technology platforms, could have had an impact on antigen presentation and the subsequent immune responses generated. Further examination of these factors will be critical for optimising vaccines strategies to generate the broadest neutralising antibody response, providing the best protection against emerging VOCs.

Unfortunately, plasma samples from VA14 were not available to investigate the peak neutralising response. Hence, the waning of neutralising antibodies following ChAdOx1-nCoV vaccination was not examined here. Although, the neutralising antibody responses following ChAdOx1-nCoV vaccination has been shown to wane over time in a similar manner to what is observed following SARS-CoV-2 infection [616,617]. The plasma neutralising antibody titre was low at 4 months post-booster, and it is unclear whether this level would provide sterilising or near-sterilising immunity. Although, the presence of B-cells that produce potent cross-neutralising antibodies against non-overlapping epitopes at 4 months post-booster, and the persistence of S-specific IgG B-cells at 9 months post-booster indicate that a rapid recall response would likely occur [518]. Such a response could be adequate to protect against severe disease and/or hospitalisation, in the event of reinfection with emerging VOCs. Nonetheless, an additional third dose of either BNT162b2 or mRNA1273, has been introduced in the UK since September 2021 and are now being administered globally. This has shown promising results in generating neutralising antibody responses against circulating VOCs, particularly the B.1.1.529 (o) variant [607,618,619].

In summary, mAbs isolated from a donor vaccinated with ChAdOx1-nCoV (12 week interval) were able to potently cross-neutralise SARS-CoV-2 VOCs, including the B.1.1.529 (o) variant. Similar to mAbs isolated from convalescent subjects, the ChAdOx1-nCoV-elicited neutralising mAbs targeted non-overlapping epitopes on the RBD and NTD. The ChAdOx1-nCoV-elicited mAbs had higher levels of SHM and greater neutralisation breadth compared to those isolated from convalescent donors. Although plasma neutralisation was undetectable at 9 months post-booster, S-specific IgG B-cells could still be measured via flow cytometry. This work provides key insights into long-term immunity and protection against SARS-CoV-2 emerging VOCs, following ChAdOx1-nCoV vaccination.

# Chapter 6: Discussion and outlook

It has been over 40 years since HIV-1 emerged into the human population, and despite extensive research, there is no licensed vaccine available [119,435]. Advances in ART transformed AIDS from a devastating fatal illness to a manageable chronic disease with an almost normal lifespan [620]. However, ~1.5 million people still contract HIV-1 each year, and with access to treatment being non-universal, AIDS remains a leading cause of death in some regions of the world [120,481]. Production of an effective HIV-1 vaccine is extremely challenging due to the considerable genetic diversity of the virus, its intricate mechanisms of immune evasion and ability to integrate itself into the host-genome. One of the primary goals of HIV-1 vaccine research is to elicit a bnAb response that can provide sterilising immunity [481]. Some HIV-1 positive individuals are capable of generating such a response, and are able to neutralise a wide range of globally circulating HIV-1 variants [194]. However, bnAbs typically arise following years of chronic infection and exposure to numerous autologous strains [196]. Consequently, bnAbs tend to have high levels of SHM, indicative of multiple rounds of affinity maturation [35]. This feature is unattractive for vaccine design, as re-elicitation would likely require multiple rounds of immunisation with a variety of immunogens. Further, bnAbs often display unusual features such as long CDRH3s and can sometimes be autoreactive, suggesting they may be difficult to generate via vaccination [481]. Nonetheless, many believe that within broadly neutralising HIV-1 infected donors, exists a blueprint to a successful HIV-1 vaccine [435]. Co-evolution studies, in which developing autologous viruses are examined in tandem with bnAb lineages, provide a key insight into this subject. However, co-evolution work requires longitudinal samples from untreated HIV-1 infected subjects, of which only a small quantity exist. Further, it is improbable that researchers will get ethical approval to repeat such studies [208]. Therefore, while co-evolution studies are highly useful, only a finite number have been carried out. Conducting more of these studies would enable a greater understanding of bnAb development across different donors. This in turn, could result in identification of recurring features that would provide the best chance of generating an effective immunisation regime.

Herein, the neutralising antibody responses of a single HIV-1 infected donor (SJU) were examined in detail. SJU was a control arm participant of the SPARTAC trial, in which longitudinal PBMC and serum samples were collected. Previous work by the Doores research group found that SJU was superinfected with a clade C virus, 2 years after primary infection with a clade C virus. Further, it was shown that SJU developed a bnAb response that targeted the V3-glycan domain, and numerous autologous viruses were isolated from the donor [444].

Building on this work, the neutralisation breadth of donor SJU was determined using neutralisation assays with representative global virus panels. Although not as broad as other donors reported in the literature, the SJU plasma samples were found to have good neutralisation breadth and potency [46,219,458]. As has been seen with other superinfected donors, the neutralisation breadth of SJU increased following superinfection [458,621,622]. Moreover, the autologous neutralisation capability of this donor was shown to increase over time. Interestingly, between weeks 85-121 post-infection SJU was able to contemporaneously neutralise numerous viral variants. This was unusual as during HIV-1 infection, circulating viruses tend to be resistant to the bnAbs present [201]. This raised the question of whether the bnAbs or autologous neutralising antibodies were responsible for this phenomenon. This was investigated by performing neutralisation assays with a panel of V3-glycan domain bnAbs against the autologous viruses from this period. It was determined that the incomplete escape could have arisen from V3-glycan domain bnAbs with glycan promiscuity, a helper lineage specific for the N334 glycan or autologous neutralising antibodies. In either event, the viral population was unable to completely escape the host immune response until superinfection occurred. Similar to other donors

that have been reported to exhibit contemporaneous neutralisation, SJU was found to have highly potent heterologous neutralisation [461,462].

The viral population was very heterogeneous during the period of contemporaneous neutralisation, prompting examination of the mechanisms employed by the virus to escape the bnAb response. Longitudinal serum samples were tested in neutralisation assays against autologous viruses that had been genetically altered to shift a critical N-glycan between positions 332 and 334. Surprisingly this had little impact on sera neutralisation, despite being a commonly reported mechanism of viral escape from V3-glycan domain bnAbs [463]. Instead, it was shown that alterations in the V1 loop, such as increasing length and introducing additional disulfide bonds, were responsible for disrupting V3-glycan domain based neutralisation. This escape mechanism has been corroborated by others, and likely involves the V1 loop sterically blocking access to the V3 loop [445,455–457]. This finding could be further examined through structural analysis via cryogenic electron microscopy (cryo-EM). Gaining a better understanding of how contemporaneous neutralisation occurs, and viral mechanisms of escape, could help to form the basis of a therapeutic vaccine.

The next phase of this work was to isolate bnAbs from donor SJU using antigen specific B-cell sorting. Biotinylated gp120 sorting probes were generated and used to sort $CD19^{+}IgG^{+}$ B-cells from SJU PBMCs collected at 252 weeks post-infection. A single mAb was isolated, which bound the gp120 of two heterologous viruses but was non-neutralising. The gp120 sorting probes were subsequently modified to incorporate a D368R mutation to prevent unwanted interactions with $CD3^{+}CD4^{+}$ T-cells in the PBMC samples. A second sort was carried out using SJU PBMCs collected at 108 weeks post-infection, and the new sorting probes greatly increased the enrichment of $CD19^{+}IgG^{+}$ B-cells. This yielded 7 mAbs, which also bound to two gp120 of two heterologous viruses, but were also non-neutralising. It was concluded that the sorting probes, although improved, were still not optimal for isolating bnAbs from this donor. Instead, stabilised Env proteins, in conjunction with B-cell isolation and culturing techniques, or V3-loops fused to non-human IgG1 Fc regions would be more suitable [474–477].

Although attempts to isolate bnAbs from donor SJU were largely unsuccessful, several protocols were optimised to carry out this project. These protocols included generating biotinylated sorting probes, antigen specific B-cell sorting and a mAb production pipeline, which were employed successfully and improved upon when used to investigate SARS-CoV-2. Indeed, many of the advancements seen in the HIV-1 field enabled research to be rapidly carried out on SARS-CoV-2. For instance, the pseudoviruses generated here were based on a 3 plasmids HIV-1 pseudovirus system. Further, the uncleaved and artificially stabilised S glycoprotein used as a sorting bait bear multiple similarities to stabilised Env trimers [247,490]. Moreover, stabling S-2P mutations in the S2 subunit of the S glycoprotein found in the mRNA based vaccines are reminiscent of the IP mutations found in SOSIP proteins [393].

Primarily, antibody responses to SARS-CoV-2 were examined in longitudinal serum samples from 65 individuals that had been infected with SARS-CoV-2. The use of sequential samples was based on principles from HIV-1 research, with most prior reports being cross-sectional [493,494]. As a result, these studies varied depending on the length of time since infection and disease severity, and often gave somewhat conflicting results. The work here showed that the early kinetics of the neutralising antibody response was typical of an acute viral infection, with declining neutralising antibody titres observed after an initial peak. By assigning the donors a severity score based on the maximal level of respiratory support needed throughout hospitalisation, this work was able to show that the magnitude of the neutralising antibody peak was dependent on disease severity [509,510]. Although the neutralising antibody titres of subjects with a high peak $ID_{50}$ remained relatively high, some with a lower peak $ID_{50}$ had neutralising antibody titres approaching baseline within the follow-up period. Correspondingly, a similar decline in neutralising antibody titres was seen in a cohort of 31

seropositive HCWs. These finding were largely corroborated by others, and suggested that vaccine boosters may be required to provide long-lasting protection [303,495,513,514]. Eventually, due to the waning of vaccine-elicited neutralising antibody levels, and the emergence of the B.1.1.529 (o) variant, a third vaccination became recommended.

A follow-up study was carried out to investigate whether the neutralising antibody titre declination would continue its trajectory or plateau to a steady state [520]. This ultimately showed there was initially a rapid reduction in the neutralising antibody titres, the rate of decline slowed around 4-7 months POS. Neutralisation against VOCs circulating at the time was also analysed, showing a 3.4-fold and 8.9-fold decrease in neutralisation potency against B.1.1.7 (α) and B.1.351 (β) respectively. Importantly, neutralising activity could still be detected against these VOCs in the majority of subjects at 257-305 days POS. Remarkably, the fold reduction in neutralisation of these VOCs compared to WT decreased as days POS increased. This finding implied that the samples from later timepoints were able to tolerate S glycoprotein mutations better than those from earlier timepoints [520]. This data had important connotations for public health, as there were global concerns at the time based around antibody longevity following infection and vaccination, as well as efficacy against VOCs.

Next, 107 anti-SARS-CoV-2 mAbs were isolated from three convalescent subjects, each of which experienced a different level of COVID-19 disease severity. Using an uncleaved and artificially stabilised S glycoprotein as a sorting bait allowed examination of antibodies that targeted regions beyond the RBD. The mAbs were shown to bind epitopes across S, including the RBD, NTD, and S2. As seen by others, the RBD was found to be the dominant target of the neutralising mAbs, with the NTD acting as a subdominant epitope. The mAbs were separated into seven distinct competition groups, and were shown to inhibit ACE2 interaction with S glycoprotein to varying degrees. Similar to V3-glycan domain bnAbs, neutralisation potency of some mAbs was impacted by glycan heterogeneity, whilst others were found to neutralise authentic virus considerably more potently than pseudovirus. The B.1.1.7 (α) variant was resistant to NTD-specific mAbs, this highlighted the importance of researching all neutralising epitopes present on S glycoprotein when investigating viral evolution and antigenic drift.

To investigate whether similar mAb responses were generated following SARS-CoV-2 vaccination, 44 anti-SARS-CoV-2 mAbs were isolated from an individual who received two doses of the ChAdOx1 nCoV-19 vaccine at a 12-week interval. Despite having a low neutralisation potency, it was found that S-specific $IgG^+$ B cells were detectable in this donor at 9 months post-boost. Interestingly, mAbs with potent neutralising activity against a panel of SARS-CoV-2 VOCs were isolated. The vaccine-elicited neutralising mAbs formed eight distinct competition groups and bound epitopes, seven of which overlapped with those identified in the convalescent donors. The mAbs found here were more mutated than mAbs isolated from convalescent donors at ~1–2 months post-infection. As the plasma neutralising antibody titre was low at 4 months post-booster, it was unclear whether this level would provide sterilising or near-sterilising immunity. Although, the presence of B-cells that produce potent cross-neutralising antibodies against non-overlapping epitopes at 4 months post-booster, and the persistence of S-specific IgG B-cells at 9 months post-booster indicate that a rapid recall response would likely occur [518]. Such a response may have protected against severe disease and/or hospitalisation, in the event of reinfection with emerging VOCs.

Researching antibody responses following viral infection, in polyclonal sera or at the monoclonal level, is important for a multitude of reasons. Analysis of sera using immunological techniques such as ELISA can aid as a confirmatory diagnostic test, and be used for surveillance purposes to better understand the disease prevalence in studied populations [623]. Further, collecting and testing longitudinal serum samples enables examination of the kinetics and longevity of the immune response following

infection. Neutralisation assays allow the protective properties of sera to be determined. Isolating mAbs can be used to map neutralising epitopes, as immunotherapy options and as molecular tools. Therefore, it would be strategically beneficial to develop a platform that would enable rapid examination of antibody responses, in the event of another pandemic. Whilst different methodologies can be repurposed to investigate a new pathogen, lack of standardisation across research groups limits comparison of results. Alternatively, introduction of standardised controls, to which results can be normalised, would help to remedy this problem. These controls would need to be introduced early into the pandemic and taken on board by the scientific community. For example, the WHO released an anti-SARS-CoV-2 antibody, intended for accurate calibration of assays to an arbitrary unit [624]. However, this did not happen until several months into the pandemic, and the reagent was not implemented by many researchers. Considerable importance needs to be placed on the defining correlates of protection from immunological research [625]. This would enable more effective extrapolation of serological data to real-world scenarios.

As the COVID-19 pandemic persists, research of the immune response against the virus has become more challenging to investigate. Factors such as vaccine status, differences in vaccine technology platforms, previous infections, disease severity and time since on-set of symptoms, complicate studies. Although not as diverse as HIV-1, highly divergent stains of SARS-CoV-2 have emerged, such as B.1.1.529 (o) [579]. It may be more beneficial to begin to search for broadly neutralising epitopes in SARS-CoV-2 research. This could lead to the production of immunogens that will better protect against future VOCs. In this work, it was shown that bnAbs against SARS-CoV-2 variants could be identified following just two immunisations with the ancestral strain S glycoprotein. These mAbs had a higher level of SHM compared to those isolated from convalescent donors, but were much less mutated compared to HIV-1 bnAbs. This would suggest that achieving neutralisation breadth will be a much less challenging task, in comparison to what has been experienced by the HIV-1 field. However, phenomenon such as immune imprinting and original antigenic sin would also need to be considered moving forward [626].

To review, this work examined antibody responses in a HIV-1 infected donor and SARS-CoV-2 infected individuals. In doing so, the antibody responses of donor SJU against autologous and heterologous viruses were examined, revealing some aspects of the complex co-evolutionary process that occurs during a HIV-1 infection. Longitudinal samples from SARS-CoV-2 infected subjects were examined, and key features regarding the kinetics and longevity of the neutralising response during early infection were ascertained. Multiple neutralising SARS-CoV-2-specific mAbs were isolated from convalescent donors and a vaccine recipient. This allowed the neutralising epitopes of the virus to be mapped, and mechanisms of viral escape to be determined. It is hoped that the worked carried out here, helped to answer important scientific questions and improved the understanding of immune responses towards the causative viruses of the AIDS and COVID-19 pandemic.

# Appendix

## 7.1 SJU envs

>SJU_1_g2A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPCNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCASNITGILLQYDGGNTNSSNGTETFRPGGGNIKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAAFLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLISLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGRDRLGRIEEEGGEQ
DKDRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s2C

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWAEAKTTLFCASNAKAYEKEMHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLTEGDIIIRSKDITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRRAHCNISKANWTATLQRVEKKLQRYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCASNITGILLQYDGGNTNSSNGTETFRPGGGNMKDIWRSELYKYKVVEIMPLGVAPTGAKRRVVEREKRAVGL
GAAFLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTTVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQ
DKDRPSRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
RSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s2F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCASNITGILLQYDGGNTNSSNGTETFRPGGGNIKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAAFLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLISLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGRDRLGRIEEEGGEQ
DKDRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s4A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRRAHCNISKANWTATLQRLEKKLQRYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLQHDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQ
DKDRSSRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGADRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s5H

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQRYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLQHDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQ
DKDRSSRLVNGFLAIVREDLRSLCLFSYHQLRDFILIAARAVELLGRNSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s6D

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSKATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRRAHCNISKANWTATLQRVEKKLQRYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNITKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLQYDGGNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGLGA
ALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQL
LGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDSWK
SLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQDK
DRSIRLVNGFLTIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKS
AISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s6F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNVTSANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILISCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGNIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLHYDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQ
DKDRSSRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s8C

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKELHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDAPVNKCSNASVNVTSANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTIHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLQHDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVSWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQ
DKDRSSRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s10D

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCRNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCASNITGVLLQYDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEARRRVVKREKRAVGL
GAVFLGFLGMAGNTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTTVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALDS
WKSLWSWLDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQ
DKDRSIRLVNGLLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGQEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s11A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDIPVNKCSNASVNVTSANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFNITNTTTATITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLHYDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGLGAAF
LGFLGTAGSTMGAASMALTIQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLG
LWGCSGKLICTTAVPWNSSWSNKSYTEIWDNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDSWKSL
WSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDR
SIRLVNGFLTIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKRSAI
SLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s11E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRRAHCNISKANWTATLQRVEKKLQRYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFNITNTTTATITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQYDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEARRRVVKREKRAVGLGAVF
LGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLG
LWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALDSWRSL
WSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDR
SIRLVNGFLTIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAI
SLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s11F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNATC
SDASVNNTSNDTASLKFDEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCASNITGILLHYDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAAFLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKNLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQ
DKDRSIRLVNGLLAIIWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_1_s12E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKELHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDAPVNKCSNASVNVTSANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLDGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRRAHCNISKANWTATLQRVEKKLQRYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLHYDGGNTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQ
DKDRSVRLVNGFLAIVWEDLRSMCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
RSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_17_g3C

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVPVNKCSNASVNVTRANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYVPPIAGN
ITCASNITGILLHYDGGNMNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTTVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQ
DKDRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_17_g4E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNVTSANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNSSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLQYDGGNMNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTTVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQ
DKDRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_17_s2D

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVPVNKCSNASVNVTRANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSHEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTSQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCASNITGILLHYDGGNMNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTTVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQ
DKDRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_17_s6A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNVTSANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKEQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQRYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLHYDGGNTNSSNETGTFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQ
DKDRSVRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_17_s6D

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATYACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNTSNANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLDGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFNIINTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CASNITGILLQHDGGDTNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGA
ALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQL
LGLWGCSGKLICTTAVPWNSSWSNKSYTEIWDNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDSWK
SLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDK
DRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKS
AISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_17_s9A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVSVNKCSNASVNVTSANC
SDTSVNKTSNATASLEFATGDIKNCSFNITTEIRDKEQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQRYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLHYDGGDTNSSNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQ
DKDRSVRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_17_s9D

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNVKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVPVNKCSNASVNVTSANC
SDASVNKTSNATASLEFATGDIKNCSFNITTDIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTRPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
LTCTSNITGILLQHDGGNTNSLNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTAVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEVSQNQQEKNEKDLLALDS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQWGLDRLERIEEEGGEQ
DKDRSIRLVNGLLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_17_s12A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVPVNKCSNASVNVTRANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCASNITGILLHYDGGNMNSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGL
GAALLGFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQ
QLLGLWGCSGKLICTTTVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALNS
WKSLWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQ
DKDRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSGLRGLQRGWEVLKYPGNLVLYWGLEIK
KSAISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_25_s1F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDASVNNTSNDTAPLKFDEG
DIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGNNSNEYILINCNTSTITQACPKVNFDPIPIHYCAPAGYAIL
QCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENTNTIIVHLNESVPIVCTRPSN
NTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWNTTLTRVKEKLQGYFPNKTIKFAPHSGGDLEVTTHSFNC
RGEFFYCNTTQLFKLNITNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASDITGILLQYDGGNSS
NGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEARRRVVKREKRAVGLGAVFLGFLGMAGNTMGAASLTL
TVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTTVPWNS
SWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALDSWKSLWSWFDITQWLWYIKIFIM
IVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLSIVWEDLRS
LCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTDRII
EVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_25_s5B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDAPVNKCSNASVNVTSANC
SDASVNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLDGNNSNEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKDITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWTTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFELNITNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCASNITGVLLQYDGGNTNSSETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEARRRVVKREKRAVGLGAV
FLGFLGMAGNTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLL
GLWGCSGKLICTTTVPWNSSWSNKSYTEIWGNMTWMQWDKEISNYTGTIYRLLEDSHIQQEKNEQELLALDSWKS
LWSWFDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKD
RSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSA
ISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_61_s4H

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNKCSNASVTSANCSDAS
VKKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYVLFYKPDVVPLNGENSNSTEYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQANCNISESNWTTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFTNTTKPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASN
ITGILLHYDGGNTNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLG
AAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCS
GKLICTTAVPWNSSWSNKSSTEIWDNMTWMQWDKEISNYTDTIYRLLEVSRNQQEKNEKDLLALDSWKSLWSWFD
ITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLV
NGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLNT
TAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_61_s7B

MRVKGTQRNWSQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSDVPVCSNASVNATRANCSD
ASVNKTSNATASLEFAEGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWNTTLTRVKEKLQGYFPNKT
ITFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKPNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNTNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSTEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSW
FDITQWLWYIKIFIMIVGGLIGLKIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFETALL*

>SJU_61_s9G

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEQEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCASVNKTSNATASLEFATGDI
KNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYCTPAGYAIL
QCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITKNTNTIIVHLNESVPIVCTRPSN
NTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSSWTTTLTRVTEKLQGYFPNKTIKFAPHSGGDLEVTTHSFSC
RGEFFYCNTTQLFELNITNTTNSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQHDGGGTN
SSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALLGFLGTAGSTMGAASM
ALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVPW
NSSWSNKSSTEIWDNMTWMQWDKEISNYTDTIYGLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWLWYIKIF
IMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWEDL
RSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTDR
IIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_61_s10B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEQEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCASVNKTSNATASLEFATGDI
KNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYCTPAGYAIL
QCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITKNTNTIIVHLNESVPIVCTRPSN
NTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSSWTTTLTRVTEKLQGYFPNKTIKFAPHSGGDLEVTTHSFSC
RGEFFYCNTTQLFELNITNTTNSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQHDGGDTN
SSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALLGFLGTAGSTMGAASM
ALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVPW
NSSWSNKSSTEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWLWYIKIF
IMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWEDL
RSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTDR
IIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_61_s10H

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEQEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCASVNKTSNATASLEFATGDI
KNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACPKVNFEPIPIHYCTPAGYAIL
QCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSKNITKNTNTIIVHLNESVPIVCTRPSN
NTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSSWTTTLTRVTEKLQGYFPNKTIKFAPHSGGDLEVTTHSFSC
RGEFFYCNTTQLFELNITNTTNSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQHDGGDTN
SSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALLGFLGTAGSTMGAASM
ALTVQTRQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVPW
NSSWSNKSSTEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWLWYIKIF
IMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWEDL
RSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTDR
IIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_61_s11A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVS
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSNASVNTSNANCSDASVKKT
SNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSTEYILINCNTSTITQACPKVNFDP
IPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVH
LNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISKSNWNTTLTRVKEKLQGYFPNKTIKFAPH
SGGDLEVTTHSFSCRGEFFYCNTTQLFELNITNTAEPNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASN
ITGILLQHDGGNTSSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALLG
FLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLW
GCSGKLICTTAVPWNSSWSNKSSTEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWS
WFDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSI
RLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISL
LDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_61_s12E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEQEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCASVNKTSNATASLEFATGDI
KNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYCTPAGYAIL
QCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITKNTNTIIVHLNESVPIVCTRPSN
NTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSSWTTTLTRVTEKLQGYFPNKTIKFAPHSGGDLEVTTHSFSC
RGEFFYCNTTQLFELNITNTTNSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQHDGGDTN
SSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALLGFLGTAGSTMGAASM
ALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVPW
NSSWSNKSSTEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWLWYIKIF
IMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWEDL
RSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTDR
IIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_85_S3A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNKCSNASVISANCSDAS
VNKTSNATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESIPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESNWNTTLTRVKEKLQGYFPNKTIN
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNLNITNTTTSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CTSNITGILLQRDGGNKDIPNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGA
AFLGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQL
LGLWGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWK
SLWSWFDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGGEQDK
DRSIRLVNGFLSIVWEDLRSLCLFSYHQLRDFILIAARALELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKS
AISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_85_S3B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNKCSNASVISANCSGAS
VNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISKLNWTTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFKLNITNTTKPNENEAITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLQRDGGNKNSSNETETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAV
GLGAAFLGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQ
DQQLLGLWGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLAL
DSWKSLWSWFNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLGRIEEEGG
EQDKDRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLE
IKKSAISLLDTTAIAVAGGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_85_s3E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEGHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHENIISIWDQGLKPCVKLTPLCVTLTCSDVPVCSNASVNVAKANCSD
ASVNRTSNATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSTEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITKNT
NTIIVHLNESVPIVCTRPSNNIRKSVRIGPGQAFYATNDIIGDIRQAHCNISKLDWTTTLIRVMEKLQGYFPNKT
ITFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSW
FDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPQNQRGLDRLKRIEEEGGEQDKDRSIRL
VNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLD
TTAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_85_s4A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCADVNKCSNASVISANCSDAA
VNKTSNATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNEENSNSSEYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESNWTTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNLNITNTTTPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CASNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALL
GFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGL
WGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLW
SWFDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRS
IRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEVKKSAIS
LLDTTAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_85_s4B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVPVCSNASVNATRTNCSD
AAVNKTSNATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISKSNWNTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNITNATEPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CTSNITGILLQRDGGSNRTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLG
FLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLW
GCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWS
WFNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSI
RLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISL
LDTIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_85_s5E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLVNVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNKCSNASVASANCSAAS
VNKTSNATASLEFATGDIKNCSFNITTEIRDKKQKAYALFYRPDVVPLNGENSNSTEYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESNWNTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNLNITNTTTPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CTSNITGILLQRDGGNKNSSETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFL
GFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGL
WGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTGTIYRLLEDSQNQQEKNEKDLLALDSWKSLW
SWFDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRS
IRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAIS
LLDTTAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_85_s5G

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSSVNKSSDASVNNTSNDTAS
LKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYC
APAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLNESVP
IVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESNWNTTLTRVKEKLQGYFPNKTIKFAPHSGGDLE
VTTHSFSCRGEFFYCNTTQLFNLNITNTTTPNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILL
QRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGSTM
GAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQTRVLAIERYLQDQQLLGLWGCSGKLICT
TAVPWNSSWSNKSYAEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLSLDSWKSLWSWFDITQWLW
YIKIFIMIVGGLIGLRIILGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAI
VWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVA
EGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_85_s6G

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCASVNKTSNATASLEFAKGDI
KNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYCAPAGYAIL
KCNNKTFNGTGPCQNVNTVQCTHGITPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLNESVPIVCTRPSN
NTRKSVRIGPGQTFYATNDIIGDIRQAHCNISKSNWNTTLTRVKEKLQGYFPNKTINFAPHSGGDLEVTTHSFNC
RGEFFYCNTTQLFNLNITNTTTSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQRDGGNPN
ETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALLGFLGTAGSTMGAASMALT
VQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVPWNSS
WSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNGKDLLALDSWKSLWSWFDITKWLWYIKIFIMI
VGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWEDLRSL
CLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTDRIIE
VLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_85_s8B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEQEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCASVNKTSNATASLEFAKGDI
KNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYCTPAGYAIL
QCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITKNTNTIIVHLNESVPIVCTRPSN
NTRKSVRIGPGQTFYATNDIIGDIRQAHCNISESNWNTTLTRVKEKLQGYFPNKTITFAPHSGGDLEVTTHSFSC
RGEFFYCNTTQLFTNTTETNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQRDGGNPNETET
FRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGSTMGAASLTLTVQAR
QLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVPWNSSWSNK
SSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWLWYIKIFIMIVGGL
IGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWEDLRSLCLFS
YHQLRDFILIAARVVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTDRIIEVLQG
IYRAFRNVPRRIRQGFEAALL*

>SJU_85_s8E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNVTSANCSDASVNRTSN
ATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACPKVNFDPIP
IHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLN
ESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISESNWTTTLTRVKEKLQGYFPNKTIKFAPHSG
GDLEVTTHSFSCRGEFFYCNTTQLFNLNITNTTTPNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNIT
GILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAA
GSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGK
LICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDIT
KWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNG
FLAIVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTA
IAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_97_s2F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDANKCSNASVISANCSNAS
VNKTSNATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITKNTNT
IIVHLNESIPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISKLNWTTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFKLNITNTTKPNENETITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLQRDGGNSSNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGA
AFLGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQL
LGLWGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWK
NLWSWFNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDK
DRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKS
AISLLDTTAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_97_s2G

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNTSVNVTNANCSD
ASVNNTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISESNWNTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTTPNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSIEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSW
FDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_97_s4F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEQEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSLVNKSSDASVNNTSNDTAS
LKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYC
APAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLNESIP
IVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAYCNISKSNWNTTLTRVKEKLQGYFPNKTIKFAPHSGGDLE
VTTHSFNCRGEFFYCNTTQLFTNTTKTNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQRDG
GNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGSTMGAAS
LTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVP
WNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWLWYIKI
FIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLSIVWED
LRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTD
RIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_97_s5F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCADVNKCSNASVISANCSDAA
VNKTSNATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNEENSNSSEYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESNWTTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNLNITNTTTPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CASNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALL
GFLGTAGSTMGAASMALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGL
WGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLW
SWFDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRS
IRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARVVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAIS
LLDTTAIAVAEGTDRIIEVLQGIYRAFRNVPRRIRQGFEAALL*

>SJU_97_s6A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDANKSVISANCSNASVNKT
SNATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDP
IPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVH
LNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISKSNWNTTLTRVKEKLQGYFPNKTITFAPH
SGGDLEVTTHSFSCRGEFFYCNTTQLFKLNITNTTKPNENGTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCT
SNITGILLQRDGGHSSNGTETFRPGGGNMRDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLG
FLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLW
GCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKNLWS
WFNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSI
RLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISL
LNTTAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_97_s7E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNVTSANCSDASVNNTSN
DTASLKFDKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDPIP
IHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLN
ESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNINKSNWNTTLTRVKEKLQGYFPNKTIIFAPHSG
GDLEVTTHSFNCRGEFFYCNTTQLFTNTTKTNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCTSNITGILL
QRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEARRRVVEREKRAVGLGAAFLGFLGAAGSTM
GAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICT
TAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKNLWSWFNITQWLW
YIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAI
VWEDLRSLCLFSYHQLRDFILIAARAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVA
EGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_97_s8G

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNTSVNVTNANCSD
ASVNNTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESNWNTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTTPNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSIEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSW
FDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_97_s11E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEQEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNTSVNVTNANCSD
ASVNNTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIKQAHCNISKSNWNTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKTNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSW
FDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLSIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_97_s11F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNTSVNVTNANCSD
ASVNNTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGNIIIRSKNITENT
NTIIVHLNESIPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESNWNTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTNSNETLTLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTGAKRRVVEREKRAVGLGAALLGF
LGTAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSW
FDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_108_s1B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLQPCVKLTPLCVTLNCSNASLNCSNASVNTSNANCS
NASVNKTSNATASLEFAKGDIKNCSFDITTEIRDKKQKAYALFYRPDVVPLNGENSNSTEYILINCNTSTITQAC
PKVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITEN
TNTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESTWNTTLTRVKEKLQGYFPNK
TIIFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITC
TSNITGILLQRDGGNKDNSTGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAA
FLGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLL
GLWGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEINNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKS
LWSWFNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKD
RSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSA
ISLLDTIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_108_s1C

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNVTSANCSDASVNNTSN
DTASLKFDKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDPIP
IHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLN
ESVPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNINKSNWNTTLTRVKEKLQGYFPNKTIIFAPHSG
GDLEVTTHSFNCRGEFFYCNTTQLFTNTTKTNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCTSNITGILL
QRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEARRRVVEREKRAVGLGAAFLGFLGAAGSTM
GAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICT
TAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKNLWSWFNITQWLW
YIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAI
VWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTIAIAVA
EGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_108_s3A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEQEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNTSVNVTNANCSD
ASVNNTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLNGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESMPIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISELSWNTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKTNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSAEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSW
FDITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARVVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_108_s3F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNVTSANCSDASVNNTSN
DTASLKFDKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDPIP
IHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLN
ESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNINESNWNTTLTRVKEKLQGYFPNKTIIFAPHSG
GDLEVTTHSFNCRGEFFYCNTTQLFTNTTTPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILL
QHDGGNMNSSTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGST
MGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLIC
TTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWL
WYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLA
IVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLRNLVLYWGLEIKKSAISLLDTAAIAV
AEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_108_s3H

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNTSVNVTNANCSD
ASVNNTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLNGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCTPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVPIECTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSNWNTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTEPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSW
FDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLSIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_108_s4F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGENLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCADANKCSNASITSANCSDAA
VNKTSNATAPLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSEYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNEIIGDIRQAHCNISKSNWNTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNITNATKPNENETITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CTSNITGILLQRDGGKSSNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAF
LGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLG
LWGCSGKLICTTAVPWNSSWSNKSHAAIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSL
WSWFNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDR
SIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAI
SLLDTIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_108_s5C

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLHCSSVNKSSDASVNTTSNDTAS
LKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYC
APAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLNESIP
IVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSNWNTTLTRVKEKLQGYFPNKTIKFAPHSGGDLE
VTTHSFNCRGEFFYCNTTQLFTNTTKTNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQRDG
GNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGSTMGAAS
LTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVP
WNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSPWSWFDITKWLWYIKI
FIMIVGSLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLSIVWED
LRSLCLFSYHQLRDFILIAARVVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTD
RIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_108_s5F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVPVCSNASVNATRENCLD
AAVNKTSNATASLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVSIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSNWNTTLTRVKEKLQGYFPNKT
IKFAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFTNTTEPSETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCT
SNITGILLQRDGGSSNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSHAAIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSW
FNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_108_s6A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKSSDASVNNTSNDTAS
LKFDKGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYC
APAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITKNTNTIIVHLNESVP
IVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNIRESNWNTTLTRVKEKLQGYFPNKTIKFAPHSGGDLE
VTTHSFNCRGEFFYCNTTQLFTNTTKPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQRDG
GNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGSTMGAAS
LTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVP
WNSSWSNKSSKEIWENMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITRWLWYIKI
FIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLSIVWED
LRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTD
RIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_108_s6F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNASVNVTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESIPIVCTRPNNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNISESNWNTALTRVKEKLQGYFPNKT
IEFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGNPNGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEARRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWDSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSW
FNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARALELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEGLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_108_s8G

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTETKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCADANKCSNASVTSTNCSDAA
VNKTSNATAPLEFAKGDIKNCSFNITTEIRDKKQKAYALFYRPDVVPLNGENSNSSEYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNEIIGDIRQAHCNISESTWNTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNLNITNTTTSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CTSNITGILLQRDGGSSNETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGF
LGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWG
CSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEINNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSW
FSITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIR
LVNGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLL
DTTAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_108_s9F

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCADANKCSNASITSANCSDAA
VNKTSNATAPLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKV
NFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNT
IIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNEIIGDIRQAHCNISESTWNTTLTRVKEKLQGYFPNKTIT
FAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNLNITNTTTSNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNIT
CTSNITGILLQRDGGNKNGSNETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAF
LGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLG
LWGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSL
WSWFNITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVRRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDR
SIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAI
SLLDTIAIAVTEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_108_s9H

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNASVNVTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESIPIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNISESTWNTTLTRVKEKLQGYFPNKT
IIFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCT
SNITGILLQRDGGNKDNSTGTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAF
LGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLG
LWGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWIQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSL
WSWFSITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDR
SIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARALELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAI
SLLDTTAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_108_s11E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNVTSANCSDASVNNTSN
DTASLKFDKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDPIP
IHYCAPAGYAILKCNNKTFNGTGPCQDVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLN
ESIPIVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNINESNWNTTLTRVKEKLQGYFPNKTIIFAPHSG
GDLEVTTHSFNCRGEFFYCNTTQLFTNTTTPNGTITLPCRIKQIINMWQEAGRAMYAPPMAGNITCASNITGILL
QHDGGNMNSSTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGST
MGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLIC
TTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWL
WYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLA
IVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLRNLVLYWGLEIKKSAISLLDTAAIAV
AEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_108_s12D

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNVTSANCSDASVNNTSN
DTASLKFDKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDPIP
IHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLN
ESIPIVCTRPSNNTRKSVRIGPGQAFYATNDIIGDIRQAHCNINESNWNTTLTRVKEKLQGYFPNKTIIFAPHSG
GDLEVTTHSFNCRGEFFYCNTTQLFTNTTTPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILL
QHDGGNMNSSTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGST
MGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLIC
TTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWL
WYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLA
IVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLRNLVLYWGLEIKKSAISLLDTAAIAV
AEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_121_s2A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGENLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLQPCVKLTPLCVTLNCSPVNKCSNASGNGTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVTIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSTWNTTLTRVKEKLQGYFHNKT
IEFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCT
SNITGILLQRDGGNENSSNETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAALLG
FLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLW
GCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEINNYTDPIYRLLEDSQNQQEKNEKDLLALDSWKSLWS
WFDITNWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSI
RLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARALELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISL
LNTTAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_121_s3A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLHCSSVNKSSDASVNTTSNDTAS
LKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDPIPIHYC
APAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLNESIP
IVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNINESNWNTTLTRVKEKLQGYFPNKTIKFAPHSGGDLE
VTTHSFNCRGEFFYCNTTQLFTNTTKTNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILLQHDG
GNMNSSTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGSTMGAA
SLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAV
PWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWLWYIK
IFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLSIVWE
DLRSLCLFSYHQLRDFILIAARVVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIRKSAISLLDTTAIAVAEGT
DRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_121_s4BB

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNASGNVTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVPIVCTRPNNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNINESNWNTTLTRVKEKLQGYFPNKT
IIFAPHSGGDLEVTTHSFNCRGEFFYCDTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCT
SNITGILLQHDGGNMNSSTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLG
FLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLW
GCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWS
WFDITKWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSI
RLVNGFLSIVWEDLRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLRNLVLYWGLEIKKSAISL
LDTAAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_121_s4E

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDVNVTSANCSDASVNNTSN
DTASLKFDKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDPIP
IHYCAPAGYVILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLN
ESIPIVCTRPNNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNINESNWNTTLTRVKEKLQGYFPNKTIIFAPHSG
GDLEVTTHSFNCRGEFFYCNTTQLFTNTTTPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGNITCASNITGILL
QHDGGNMNSSTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGST
MGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLIC
TTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWL
WYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLA
IVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLRNLVLYWGLEIKKSAISLLDTAAIAV
AEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_121_s5B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLQPCVKLTPLCVTLNCSPVNKCSNASGNGTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVTIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSTWNTTLTRVKEKLQGYFHNKT
IEFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGSNKTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFL
GAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGC
SGKLICTTAVPWDSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSWF
NITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRL
VNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLD
TIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_121_s5H

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLQPCVKLTPLCVTLNCSPVNKCSNASGNGTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVTIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSTWNTTLTRVKEKLQGYFHNKT
IEFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGSNKTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFL
GAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGC
SGKLICTTAVPWDSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSWF
NITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRL
VNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLD
TTAIAVAEGTDRIIEVLQGIYRAFRNVPTRIRQGFEAALL*

>SJU_121_s6H

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLQPCVKLTPLCVTLNCSPVNKCSNASGNGTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVTIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSTWNTTLTRVKEKLQGYFHNKT
IEFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGSNKTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFL
GAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGC
SGKLICTTAVPWDSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSWF
NITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRL
VNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLD
TIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_121_s7FF

MRVKGTQRNWPQWWIWGILGFWLIMICKGEGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLHCSSVNKSSDASVNTTSNDTAS
LKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSEYILINCNTSTITQACPKVNFDPIPIHYC
APAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLNESIP
IVCTRPSNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNINESNWNTTLTRVKEKLQGYFPNKTIKFAPHSGGDLE
VTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCTSNITGILLQRDG
GNPNETETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGSTMGAAS
LTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVP
WNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEVSQNQQEKNEKDLLALDSWKSLWSWFDITKWLWYIKI
FIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWED
LRSLCLFSYHQLRDFILIAARAAELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLDTTAIAVAEGTD
RIIEVLQGIYRAFRNVPTRIRQGLEAALL*

>SJU_121_s8B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLQPCVKLTPLCVTLNCSPVNKCSNASGNGTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVTIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSTWNTTLTRVKEKLQGYFHNKT
IEFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGSNKTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFL
GAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGC
SGKLICTTAVPWDSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSWF
NITQWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPKTQRGLDRLERIEEEGGEQDKDRSIRL
VNGFLAIVWEDLRSLCLFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLD
TIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_121_s11B

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSSVNKSSDASVTASLKFDEG
DMKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSNSSDYILINCNTSTITQACPKVNFDPIPIHYCAPAGYA
ILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENTNTIIVHLNESIPIVCTRP
SNNTRKSVRIGPGQTFYATNDIIGDIRQAHCNINESNWNTTLTRVKEKLQGYFPNKTIIFAPHSGGDLEVTTHSF
NCRGEFFYCNTTQLFTNTTKPNGTITLPCRIKQIINMWQEVGRAMYAPPIAGSITCASNITGILLQHDGGNMNSS
TETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFLGAAGSTMGAASLTLTV
QARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVPWNSSW
SNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSWFNITQWLWYIRIFIMIV
GGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWEDLRSLC
LFSYHQLRDFILIAARAGELLGRSSLRGLQRGWEVLKYLRNLVLYWGLEIKKSAISLLDTAAIAVAEGTDRIIEV
LQGIYRAFRNVPTRIRQGFEAALL*

>SJU_121_s12A

MRVKGTQRNWPQWWIWGILGFWLIMICKGGGNLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLQPCVKLTPLCVTLNCSPVNKCSNASGNGTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLSGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVTIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSTWNTTLTRVKEKLQGYFHNKT
IEFAPHSGGDLEVTTHSFNCRGEFFYCNTTQLFTNTTKSNETITLPCRIKQIINMWQEVGRAMYAPPIAGNITCA
SNITGILLQRDGGSNKTETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEAKRRVVEREKRAVGLGAAFLGFL
GAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQLLGLWGC
SGKLICTTAVPWDSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWKSLWSWF
NITQWLWYIRIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDKDRSIRL
VNGFLAIVWEDLRSLCLFSYHQLRDFILIAARALELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKSAISLLD
TIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_121_s12C

MRVKGTQRNWPQWWIWGILGFWLIMICKGGENLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWKNDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLNCSPVNKCSNASVNGTSANCAE
ASVNKTSNDTASLKFDEGDMKNCSFNITTEIRDKKQKAYALFYKLDVVPLNGENSNSSDYILINCNTSTITQACP
KVNFDPIPIHYCAPAGYAILQCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNITENT
NTIIVHLNESVTIVCTRPNNNTRKSVRIGPGQAFYATNDIIGDIRQAYCNISKSNWNTTLTRVKEKLQGYFPNKT
ITFAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFNLNITNTTTSNGTITLPCRIKQIINMWQEVGRAMYAPPIAGN
ITCTSNITGILLQRDGGNENSSNETFRPGGGNMKDNWRSELYKYKVVEIKPLGVAPTEARRRVVEREKRAVGLGA
AFLGFLGAAGSTMGAASLTLTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIKQLQARVLAIERYLQDQQL
LGLWGCSGKLICTTAVPWNSSWSNKSSKEIWDNMTWMQWDKEISNYTDTIYRLLEDSQNQQEKNEKDLLALDSWK
SLWSWFSITQWLWYIRIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQRGLDRLERIEEEGGEQDK
DRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAVRAVELLGRSSLRGLQRGWEVLKYLGNLVLYWGLEIKKS
AISLLDTIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_133_s4C

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKAYEREVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKANITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_133_s6A

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWENDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKANITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_133_s6G

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKAYEREVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKANITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAKGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_133_s9B

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKANITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_133_s9C

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKANITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_133_s10B

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKANITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_133_s11G

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSDYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKANITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s5G

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEVALL*

>SJU_145_s6E

MRVRGMLRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s6F

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEREVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPTKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKENITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWNWFNISSWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSIRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s7A

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCKIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s9D

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAICNIPSRIRQGFEAALL*

>SJU_145_s9E

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s9F

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDTKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKENITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYNLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s10E

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s10G

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s10H

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRRKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKENITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYNLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s11B

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTIGWNKTLEKVKRKLEEHFPNKIIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKEDITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s11D

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s11E

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIWLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_145_s11G

MRVRGMPRNWPQWWIWGILGFWMIIFCRGMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIVGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNISTEKWNTTLERVRKKLEEHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNTT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_169_s1D

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEREVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESINIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNINTIGWNKTLEKVKRKLEEHFPNKIIEFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNTE
VFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQA
RQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSN
KSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGG
LIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLF
SYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLKLKRNAIRLLDTIAIAVAEGTDRILEVIQ
RICRAIRNIPSRIRQGFEAALL*

>SJU_169_s1E

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEREVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNTTLERVRKKLEKHFPSKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNISEPNITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTSATEVFR
PGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQL
LSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSY
EKIWDNMTWMQWDREINNYPNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIG
LRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYH
QLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQRIC
RAIRNIPSRIRQGFEAALL*

>SJU_169_s1F

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGSCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNKSIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNTTLERVRKKLEEHFPGKTISFEPPSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNISEPNTTILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDNTTEVFRP
GGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQLL
SGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSYE
KIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIGL
RIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYHQ
LRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQRICR
AIRNIPSRIRQGFEAALL*

>SJU_169_s2G

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNTTLERVRKKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNTE
VFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQA
RQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSN
KSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGG
LIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLF
SYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQ
RICRAIRNIPSRIRQGFEAALL*

>SJU_169_s3H

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKAYEREVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNTTLERVRKKLEKHFPSKTISFEPPSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNISEPNITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNTEVFRP
GGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQLL
SGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSYE
KIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIGL
RIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWNDLRSLCLFSYHQ
LRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQRICR
AIRNIPSRIRQGFEAALL*

>SJU_169_s7D

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDAVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNINTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNTE
VFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQA
RQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSN
KSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGG
LIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLF
SYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQ
RICRAIRNIPSRIRQGFEAALL*

>SJU_169_s8B

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNTTLERVRKKLEEHFPGKTISFEPPSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNISEPNITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNTEVFRP
GGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQLL
SGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSYE
KIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIGL
RIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYHQ
LRDFILIVARTVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQRICR
AIRNIPSRIRQGFEAALL*

>SJU_169_s8C

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKGYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNTTLERVRKKLEEHFPGKTISFEPPSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNISEPNTTILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSDITGLLLTRDGGTNDNTTEVFRP
GGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQLL
SGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSYE
KIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIGL
RIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYHQ
LRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEIIQRICR
AIRNIPSRIRQGFEAALL*

>SJU_169_s8G

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGSLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKATFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCARPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNINTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNTE
VFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQA
RQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSN
KSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGG
LIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLF
SYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQ
RICRAIRNIPSRIRQGFEAALL*

>SJU_169_s12D

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEREVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
VTTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESININVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNINTIGWNKTLEKVKRKLEEHFPNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELFNKSKANITILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSNITGLLLTRDGGTNDTNTE
VFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQA
RQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSN
KSYEKIWDNMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGG
LIGLRIILGMLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLWSLCLF
SYHQLRDFILIVARAVELLGRSSLRGLRRGREALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQ
RICRAIRNIPSRIRQGFEAALL*

>SJU_204_g1E

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEDEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNTTLEKVRKKLEKYFPGKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSESNTTILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGPNENNNT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYERIWDNMTWMQWDREINNYTNTIYSLLEESQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLLFQTLIPNPRGLDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYEGLEIKRNAIRLQDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_204_g2C

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCNINTIGWNKTLEKVKRKLAGHFPKKIIEFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSVANTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSKITGLLLTRDGGPNENNNT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIAKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_204_g2D

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNTTLEKVRKKLEKYFPGKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSESNTTILPCRIRQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGPNENNNT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_204_g5D

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKINTEKWNTTLERVRKKLKGHFPGKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSVANTTTLSCRIKQIINMWQGVGRAMYAPPIAGNITCKSKITGLLLTRDGGPNENNDT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKSEQDLLALDSRKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEII
QRICRAIRNIPSRIRQGFEAALL*

>SJU_204_s10G

MRVRGMPRNWPQWWIWGILGFWMIIFCRVMGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKINTGKWNTTLERVRKKLEKHFSGKTISFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSESNTTILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGPNENNNT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYERIWDNMTWMQWDREINNYTNTIYSLLEESQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGLDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIGLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_204_s11D

MRVKGTQRNWPQWWIWGILGFWLIMICKGGENLWVTVYYGVPVWTEAKTTLFCASNAKAYEKEVHNIWATHACVP
TDPNPQEIVLANVTENFNMWENDMVDQMHEDIISIWDQGLKPCVKLTPLCVTLTCSDANVTRANCSGASVNKTNH
AANCSDASVNRTSNATVSLEFAKGDIKNCSFNITTEIRDKKQKAYALFYKPDVVPLNGENSSDYILINCNTSTIT
QACPKVNFDPIPIHYCAPAGYAILKCNNKTFNGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGDIIIRSKNI
TENTNTIIVHLNESVPIVCTRPSNNTRKSVRIGPGQAFYATNEIIGDIRQAYCTISKLNWTTTLTRVKKKLQGYF
NKTIEFAPHSGGDLEVTTHSFSCRGEFFYCNTTQLFKLNITNATKPNENETITLPCRIKQIINMWQEVGRAMYAP
PIAGNITCASNITGILLQRDGGETNETGGETNETGGETNGTETFRPGGGNMKDNWRSELYKYKVVELKPLGVAPT
EARRRVVEREKRAVGLGAALLGFLGTAGSTMGAASVALTVQARQLLSGIVQQQSNLLKAIEAQQHMLRLTVWGIK
QLQARVLAIERYLQDQQLLGLWGCSGKLICTTAVPWNSSWSNKSSKCIWDNMTWMQWDKEISNYTDTIYRLLEDS
QNQQEKNEQDLLALDSWKSLWSWFDITNWLWYIKIFIMIVGGLIGLRIIFGVLSIVKRVRQGYSPLSFQTPPQNQ
RGLDRLERIEEEGGEQDKDRSIRLVNGFLAIVWEDLRSLCLFSYHQLRDFILIAARALELLGRSSLRGLQRGWEV
LKYLGNLVLYWGLEIKKSAISLLDTIAIAVAEGTDRIIEVLQGIYRAFRNLPRRIRQGFEAALL*

>SJU_217_g7D

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSGYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIECTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCKINTEKWNKTLEKVKRKLEKHFPKKKIEFKPHSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNKSVENITTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSATEVFR
PGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQL
LSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSY
EKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIG
LRIILGVLSIVKKVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYH
QLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEIIQRIC
RAIRNIPSRIRQGFEAALL*

>SJU_217_g9C

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGIGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIECTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCKINTEKWNKTLEKVKRKLEKHFPKKKIEFKPHSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNKSVENITTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSATEVFR
PGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQL
LSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTTVPWNSSWSNKSY
EKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIG
LRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYH
QLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEIIQRIC
RAIRNIPSRIRQGFEAALL*

>SJU_217_s3A

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGIGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIECTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKISTEKWNKTLEKVKRKLEKHFPKKKIEFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSESNTTILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWENMTWMQWDREINNYTNTIYSLLEESQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIVLGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRFVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLKGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_217_s5G

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSGYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENLTDNVKTIIVHLNESIDIECTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKINTEKWNKTLEKVKRKLEKHFPKKKIEFKPHSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNKSVENITTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSATEVFR
PGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQL
LSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSY
EKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIG
LRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYH
QLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQRIC
RAIRNIPSRIRQGFETALL*

>SJU_217_s6D

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSGYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIECTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCKISTEKWNKTLEKVKRKLEKHFPKKKIEFKPHSGGDLEITTHSFNCRGEF
FYCNTTQLFNTTELRNKSESNTTILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWENMTWMQWDREINNHTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGCEAALL*

>SJU_217_s7A

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSSYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYTILKCNN
KTFSGIGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIECTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCKINTEKWNKTLEKVKRKLEKHFPKKKIEFKPHSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNKSVENITTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSATEVFR
PGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQL
LSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSY
EKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIG
LRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYH
QLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEIIQRIC
RAIRNIPSRIRQGFEAALL*

>SJU_217_s8C

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSGYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIVCTRPNNNTSK
GIRIGPGQTFYATGRVIGNIREAHCEISTEKWNKTLEKVKRKLEKHFPKKKIEFKPHSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNISGPNTTTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSATEVFR
PGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQL
LSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWSNKSY
EKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIG
LRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCLFSYH
QLRDFVLIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVIQRIC
RAIRNIPSRIRQGFEAALL*

>SJU_217_s9A

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPMWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSGYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIECTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKIRTEKWNTTLERVRKKLKEHFLNKTIEFEPHSGGDLEITTHSFNCRGEF
FYCNTTRLFNTTELRNKSESNTTILPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSAT
EVFRPGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQ
ARQLLSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTAVPWNSSWS
NKSYEKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVG
GLIGLRIIFGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDRDRSVRLVNGFLALAWDDLRSLCL
FSYHQLRDFVLIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRILEVI
QRICRAIRNIPSRIRQGFEAALL*

>SJU_217_s9E

MRVRGMPRNWPQWWIWGILGFWMIIFCRVVGNLWVTVYYGVPVWTDAKTTLFCASDAKAYEKEVHNVWATHACVP
TDPNPQEMVLGNVTENFNMWKNDMVDQMHEDIISIWDQSLKPCVKLTPLCVTLKCTNATNNNQSIIGEMTNCTFN
ITTEIRDKKKKVDALFYRSDVVPLGNQSSENNSGYYILINCNTSTITQACPKVTFDPIPIHYCAPAGYAILKCNN
KTFSGTGPCQNVSTVQCTHGIKPVVSTQLLLNGSLAEEEIIIRSENLTDNVKTIIVHLNESIDIECTRPNNNTSK
GIRIGPGQTFYATGRVIGDIREAHCKINTEKWNKTLEKVKRKLEKHFPKKKIEFKPHSGGDLEITTHSFNCRGEF
FYCNTSKLFNNNKSVENITTLPCRIKQIINMWQGVGRAMYAPPIAGNITCKSSITGLLLTRDGGKNDTSATEVFR
PGGGDMKDNWRSELYKYKVVEIKPLGIAPSKAKRRVVEREKRAVGMGAVIFGFLGAAGSTMGAAAVTLTVQARQL
LSGIVQQQSNLLRAIEAQQHMLQLTVWGIKQLQARVLAIERYLKDQQLLGLWGCSGKLICPTVVPWNSSWSNKSY
EKIWENMTWMQWDREINNYTNTIYSLLEKSQNQQEKNEQDLLALDSWKNLWDWFNISNWLWYIKIFIMIVGGLIG
LRIILGVLSIVKRVRQGYSPLSFQTLIPNPRGPDRLGRIEEEGGEQDKDRSVRLVNGFLAIAWDDLRSLCLFSYH
QLRDFILIVARAVELLGRSSLRGLQRGWEALKYLGALVQYGGLELKRNAIRLLDTIAIAVAEGTDRIIEIIQRIC
RAIRNIPSRIRQGFEAALL*

## 7.2 SJU mAbs

>SJU_108_B10_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASRFTFNDFYMSWIRQAPGKGLEWVSSISSAAGYTYYADSVRGRFTVSRDNG
KNSLFLQMNSLRVEDTAIYYCARVNRGYSFRYFYFDFWGRGSLVTVSS

>SJU_108_C4_heavy_chain

QVQLVQSGAEVKKPGSSVKVSCKASGGTLTTYNINWVRQAPGQGLEWMGRIIPFLGAANYAQNFQGRVTITADGS
RSTAYMELSRLTSDDTAVYYCASAPMQWPYSWNDPWGQGTLVTVSS

>SJU_108_D6_heavy_chain

QVQLVESGGGLVKPGESLRLSCEASGFTFSDYYMSWIRQAPGKGLEWVSEISSSSGYTKYADSVKGRFTMSRDNA
KKSLSLQMNSLRAEDTALYYCARVNRGYSYYYYYFDLWGRGTLVTISS

>SJU_108_E4_heavy_chain

QVQLVQSGAEVKKPGSSVKVSCKASGGTFNSYTINWVRQAPGQGLEWMGRLIPIFRTEDYAQDFQGRVTITADES
TSTSYMELRGLRSDDTAVYYCATLAYSDKDFWGPGTLVTVSS

>SJU_108_E9_heavy_chain

EVQLVQSGAEVKKPGSSVKVSCKASGGTFGSYAISWVRQAPGQGLEWMGRIIPMFALTNYAQRFQGRVTITADES
TSTAYMQLSSLKFEDTAVYYCTRPGDNYGFEFDHWGQGTLVTVSS

>SJU_108_F7_heavy_chain

EVQLLESGGGLVKPGGSLRLSCTASGFTFNDFYMTWIRQAPGKGLQWVSDISTSSGYSNYGESVKGRFTISRDNA
KKALYLQMNSLRADDTAVYYCARVNRGYSYHYHYLDVWGRGTLVTVSS

>SJU_108_G7_heavy_chain

EVQLVQSGAEVKPPGSSVRVSCKASEDPFTMKTISWVRQAPGQGLDWVGRIIPILGVADYAQKFQGRVTITADTS
TRTSYMELSSLTSEDTAIYYCATPDAKMAYWGQGSLVTVSS

>SJU_252_H8_heavy_chain

QVQLVQSGAEVKQPGSAVRLSCKASGGTFSTYAFSWVRQAPGQGLEWMARIVPIFGTTNYAQKFQDRVSVSADAS
TSTAYMEMSSLTSEDTAVYYCARDHTVVFGAVTDNWFDPWGQGTLVTVSS

>SJU_108_B10_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSFSSDYLAWYLQKPGQAPRLLIYAASTRATGIPDRFSGSGSGADFTLS
ISRLEPDDFAVYFCQLYDGNSVTFGGGTKVEIK

>SJU_108_C4_light_chain

DIVMTQSPGTLSLSPGETATLSCRASQSVTSSYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSPSVTFGQGTKVDIK

>SJU_108_D6_light_chain

DVVMTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQRAGQAPRLLIYGASSRAPAIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYDSSVTFGGGTKLEIK

>SJU_108_E4_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSVDSSYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTVT
ISRLEPEDFAVYYCQQYGSSPFTFGPGTKVEIK

>SJU_108_E9_light_chain

EIVLTQSPGILSLSPGERATLSCRASQSVSSSYLAWYQQTPGQAPRLLIYGASNRAAGVPDRFSGGGSGTDFTLT
ISRLAPEDFAVYYCQQYGNSPFTFGPGTKVEIK

>SJU_108_F7_light_chain

EIVLTQSPGILSLSPGERATLSCRASHSVNSNYLTWYQQKPGQAPRLLIYAASTRATGIPDRFSGSGSGTDFTLT
INRLEPEDFAVYFCQQYDSSVTFGGGTKVDIK

>SJU_108_G7_light_chain

EIVMTQSPGTLSLSPGERATLSCRASQSISSNYLAWYQQRPGQAPRLLIYDASSRAPGIPDKFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSLFTFGGGTKVEIK

>SJU_252_H8_light_chain

DIVMTQSPSSLSASIGDRVTISCRASQSISNYLNWYQQKPGTAPKLLIHSAATLQSGVPSRFSGSKSGTDFTLTI
SSLQPEDFATYYCQQSYSTPRTFGRGTKLDIK

## 7.3 SARS-CoV-2 mAbs (P003, P008 and P054)

>P003_001_heavy_chain

QVQLVESGPGLVKSSQTLPLTCTVSGASINSNNYYWNWIRQPAGKGLEWIGRIHSSGNTNYSPSLRSRLTISVDT
SKNQFSLELTSVTAADTAVYYCARDDWTTALDCWGQGTLVTVSS

>P003_002_heavy_chain

QVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGRIIPIFGTANYAQKFQGRVTITADES
TSTAYMELSSLRSEDTAVYYCARDRPPIVVVTTYYYYGMDVWGQGTTVTVSS

>P003_003_heavy_chain

EVQLLESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGINKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKAFKGNYYYGMDVWGQGTTVTVSS

>P003_006_heavy_chain

QVQLVQSGAEVKNPGASVKVSCTASGYTFTSYGISWVRQAPGQGLEWMGWISAYNGNTNYAQKLQGRVTMTTDTS
TSTAYMELRSLRSDDTAVYYCALVGATDYWGQGTLVTVSS

>P003_007_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKASGYTFTGYYMHWVRQAPGQGLEWMGWINPNSGGTNYAQKFQGRVTMTRDTS
ISTAYMELSRLRSDDTAVYYCARGAASVLRFLEWLLDYWGQGALVTVSS

>P003_014_heavy_chain

QVQLVESGGGLVQPGGSLRLSCAASGFTFSSYDMHWVRQATGKGLEWVSAIGTAGDTYYPGSVKGRFTISRENAK
NSLYLQMNSLRAEDTAVYYCARAQYSSGWYLRFDYWGQGTLVTVSS

>P003_015_heavy_chain

QLQLQESGPGLVKPSGTLSLTCAVSGGFISSSNWWSWVRQPPGKGLEWIGEIYHSGSTNYNPSLKSRVTISVDKS
KNQFSLKLSSVTAADTAVYYCASRYCSGGSCGYFDYWGQGTLVTVSS

>P003_016_heavy_chain

QVQLVESGGGLVQPGGSLRLSCAASGFTFSSYWMHWVRQAPGKGLVWVSRINIYGSSTSYADSVKGRFTISRDNA
KNTLYLQMNSLRAEDTAVYYCARDLHYDSSGWDYWGQGTLVTVSS

>P003_017_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGFTVSSNYMNWVRQAPGKGLEWVSVIYSGGSTYYGDSVKGRFTISRDNSK
NTLYLQMNSLRAEDTAVYYCARDFREGAFDIWGQGTMVTVSS

>P003_022_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSTYGMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKDPTIVVVAVNPTFDYWGQGTLVTVSS

>P003_026_heavy_chain

QVQLVESGGGLVKPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWVSSISSSSSYIYYADSVKGRFTISRDNA
KNSLYLQMNSLRAEDTAVYYCARDLDIVVVPAARSYYYYGMDVWGQGTTVTVSS

>P003_027_heavy_chain

EVQLLESGGGLVQPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWVSYISSSSSTIYYADSVKGRFTISRDNA
KNSLYLQMNRLRDEDTAVYYCARDRCGDCYGPYYYGMDVWGQGTTVTVSS

>P003_030_heavy_chain

QVQLQQSGAEVKKPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGRIIPIFGTANYAQKFQGRVTITADES
TSTAYMELSSLRSEDTAVYYCASHPYYDSSGYYPNYGMDAWGQGTTVTVSS

>P003_032_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVIWYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCARDGQVGATSGIDYWGQGTLVTVSS

>P003_034_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSSYAMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKSAFGSYYYGMDVWGQGTTVTVSS

>P003_038_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTIPRDNS
KNTLYLQMNSLRAEDTAVYYCAKGGSGYRYYFDYWGQGTLVTVSS

>P003_043_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKVSGYTLTELSMHWVRQAPGKGLEWMGGFDLADGETIYAQEFQGRVTMTEDTS
TDTAYMALSSLRSEDTAVYYCATGPLTIAVAGQWFDPWGQGTLVTISS

>P003_045_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFSNYGTHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCARPDSGSYWGAFDIWGQGTMVTVSS

>P003_051_heavy_chain

EVQLVQSGAEVKKPGASVKVSCKASGYTFTSYGISWVRQAPGQGLEWMGWISAYNGNTNYAQKLQGRVTMTTDTS
TSTAYMELRSLRSDDTAVYYCARDTRGRGGHDAFDIWGQGTMVTISS

>P003_053_heavy_chain

EVQLLESGGGLVKPGGSLRLSCAASGFTFSNAWMTWVRQAPGKGLEWVGRIKSKTDGGTTDYAAPVKGRFTISRH
DSKNTLYLQMNSLKTEDTAVYYCTTASNPDYWGQGTRVTVSS

>P003_055_heavy_chain

EVQLVESGPGLVKPSETLSLTCTVSGGSISSYYWSWIRQPPGKGLEWIGYIYYSGSTNYNPSLKSRVTISVDTSK
NQFSLKLSSVTAADTAVYYCARENYDFWSGYFNGWFDPWGQGTLVTVSS

>P003_056_heavy_chain

EVQLLESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKSHSGSYFSSGDYWGQGTLVTVSS

>P003_058_heavy_chain

EVQLVQSGAEVKKPGESLKISCKGSGYSFTIYWIGWVRQMPGKGLEWMGIIYPGDSDTRYSPSFQGQVTISADKS
ISTAYLQWSSLKASDTAMYYCARRTLTTTYGMDVWGQGTTVTVSS

>P003_064_heavy_chain

EVQLLESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKGWGGYNTHFDYWGQGTLVTVSS

>P008_001_heavy_chain

EVQLVESGGGLVQPGRSLRLSCTGSGFTFGDYAMSWFRQAPGKGLEWVGFIRSEDHGGTTEYAASVKGRFTISRD
DSKSIAYLQMNSLKTEDTAVYYCSRPLTYYYDSSGYYYPYYFDYWGQGTLVTISS

>P008_002_heavy_chain

EVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSGISWNSGSIGYADSVKGRFTISRDNA
KNSLYLQMNSLRAEDTALYYCAKGNWNDVLSHYYYYYGMDVWGQGTTVTVSS

>P008_003_heavy_chain

QVQLQQWGAGLLKPSETLSLTCAAHGGSFSGYNWSWIRQPPGKGLEWIGEINHSGSTNYNPSLKSRVTISVDTSK
NQLSLKLSSVTAADTAVYYCARGGTIGTTGIYDILTGYDPFFDYWGQGTLVTVSS

>P008_004_heavy_chain

EVQLVESGGGLVHPGGSLRLSCAASGFTFSSYWMSWVRQAPGKGLEWVANIKEDGREKYYVDYVKGRFTISRDNA
KNSLYLQMNSLRAEDTAVYYCARDSVHYYYDSSGYHYSYGMDVWGQGTTVTISS

>P008_005_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGTNKYYADSVNGRFTISRDNS
KNTLYLQMDSLRAGDTAVYYCARPNSGSYSSYLDYWGQGTLVTVSS

>P008_006_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFNIYVMHWVRQAPGKGLEWVAVISYDGSNEYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKGRSGSYFNPLDYWGQGTLVTISS

>P008_007_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGFIFSNYRMNWVRQAPGKGLEWVSYISSSSSTIYYADSVKGRFTISRDNA
KNSLYLQMNSLRAEDTAVYYCARDSVPRYYYHYYGMDVWGQGTTVTVSS

>P008_010_heavy_chain

QVQLQESGPGLVKPSGTLSLTCAVSGDSISSSHWWSWVRQPPGKGLEWIGEIYHSGSTNYNPSLKSRVTISVEKS
KNQFSLKLSSVTAADTAVYYCARRRLGPSSNNSGYFYWGQGTLVTVSS

>P008_011_heavy_chain

EVQLVESGGGLVQPRGSLRLSCAASGFTFSTYWMTWVRQAPGKGLEWVANIGQDGSERYFVDSVKGRFTISRDNA
NNSLYLQMNSLRAEDSAVYYCARDYTDDYGDYGGYFDLWGRGTLVTVSS

>P008_012_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKASGYTFTIYYMHWVRQAPGQGLEWMGIINPGGGSRSYAQRFQGRVTMTRDTS
TSTVYMELSSLRSEDTAVYYCARGRYCSSSSCYIGLDYWGQGTLVTVSS

>P008_014_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASGFTFSNAWMSWVRQAPGKGLEWVGRIKSKTDGGTTDYAAPVKGRFTISRD
DSKNTLYLQMNSLKSEDTAVYYCATEEHPFAGRAKLHHLDYWGQGTLVTVSS

>P008_015_heavy_chain

EVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSGITWNSGTIGYADSVKGRFTISKDNA
KNSLYLQMNRLRAEDTAFYYCAKDLTGWGLTFGGVITNWGQGTLVTISS

>P008_016_heavy_chain

QVQLQESGPGLVKPSETLSLSCTVSGGSISPYYWSWIRQLPGKGLEWIGYIYYTGSTNYNPSLKGRVTMSADMSK
NQFSLNLSSVTAADTAMYYCARGYDFWSRGTLVTVSS

>P008_018_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGFTVSSNYMSWVRQAPGKGLEWVSFIYPGGSTYYVDSVKGRFTISRDNSK
NTLYLQMNSLRAEDTAVYYCARSYGDYYFDYWGQGTLVTVSS

>P008_022_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKMMGQYCSGGDCYSGYFDYWGQGTLVTVSS

>P008_023_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASGFTFSDYYMSWIRQAPGKGLEWVSYISSTTSYTNYADSAKGRFTISRDNA
KNSLYLQMNSLRAEDTAVYYCARDVAYYYDSGSYYYFDYWGQGTLVTVSS

>P008_024_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFSTYGMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKMVGQYCSGGNCYLGYFDYWGQGTLVTVSS

>P008_032_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFSTNAMHWVRQAPGKGLEWVALISYDGSNKYYAVSAKGRFIISRDNS
KNTLHLQMSSLRAEDTAVYYCVKGLGGNYYYFGFWGQGTLVTISS

>P008_035_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKVSGYTLTEISMHWVRQAPGKGLEWMGGFDPEEGETIYAQKFQGRVTMTEDTS
TDTAYMELSSLRSEDTAVYYCATSEVVVPGAIRHKAAYYYNYMDVWGKGTTVTISS

>P008_036_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVAVISYDGSNEYYADSVKGRFTISRDNS
KNTLYLQMNSLRVEDTAVYYCAKKGYSYGYFDYYFDYWGQGTLVTISS

>P008_038_heavy_chain

QVQLVQSGAEVKKPGSSVKVSCKASGGTFSSYTISWVRQAPGQGLEWMGRIFPILGIANYAQKFQNRVTITADRS
TSTAFMELSSLRSEDTAVYYCARDQDSGYIWWFDPWGQGTLVTVSS

>P008_039_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWVSSISSSSSYIYYADSVRGRFTISRDNA
KNSLFLQMNSLRAEDTALYYCARDPFSRWERPEGWFDPWGQGTLVTISS

>P008_040_heavy_chain

QVQLVESGGGLVQPGGSLRLSCAASGFIFDDYAIHWVRQGPGKGLEWVSGISWNSGSIGYGDSVKGRFTISRDNA
KNSLYLQMNSLRAEDTAFYYCAKGGRRAAMLLNYFDYWGQGTLVTVSA

>P008_042_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGITVSSNYMNWVRQAPGKGLEWVSVIYSGGSTFYADSVKGRFTISRDNSK
NTLYLQMNSLRAEDTAVYYCARESYGMDVWGQGTTITVSS

>P008_044_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSNYGMHWVRQAPGKGLEWVAVISYDGSNEYYADSVKGRFTISRDNS
KNTLYLQMNSLRVEDTAVYYCAKKGYSYGYFDYYFDYWGQGTLVTVSS

>P008_045_heavy_chain

EVQLVESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSGISWNSGSIAYADSVKGRFTISRDNA
KNSLYLQMNSLRAEDTALYYCAKDIKSFGIFGVVTAFDYWGQGTLVTVSS

>P008_047_heavy_chain

EVQLVETGGGLIQPGGSLRLSCATSGFTVSSNYMTWVRQAPGKGLEWVSTIYSGGSTYYADSVKGRFSISRDSSK
NTLYLQMNSLRAEDTAVYYCARGPYPHFDYWGQGTLVTVSS

>P008_051_heavy_chain

QVQLVESGPGLVKPSQTLSLTCTVSGGSISSGSYFWSWIRQPAGKGLEWIGRIFTSGSTNYNPSLKSRVTMSVDT
SKNQFSLKLSSVTAADTAVYYCARGPDIVVVPAADPRNWFDPWGQGTLVTVSS

>P008_052_heavy_chain

QVQLQESGPGLVKPSETLSLTCTVSGGSVTSNNYYWGWIRQPPGEGLEWIGSIYYTGSTFYNPSLKSRVTISVDA
SKNQFSLKLSSVTAADTAVYYCTRQLVLVRGYFDYWGQGTLVTISS

>P008_055_heavy_chain

EVQLVESGGGLIQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSGISWNGGIIDYADSVKGRFTISRDNA
KNSLYLHMRSLRAEDTAFYYCAKDIGPFEAARPGGNYYYYAMDVWGQGTTVTVSS

>P008_056_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASGFSFSSYSMNWVRQAPGKGLEWVSSISSNSNYIYYADSMKGRFTISRDNA
KNSLYLQMNSLRAEDTAVYYCASNRSPYDSSNYYFDYWGQGTRVTISS

>P008_057_heavy_chain

QVQLVETGGGLIQPGGSLRLSCAASGLTVSANYMSWVRQAPGKGLEWVSVIYSGGSTFYADSVKGRFTISRDNSK
NTLYLQMNSLRVEDTAVYYCARNIYDDAFDVWGQGTMVTVSS

>P008_060_heavy_chain

EVQLVESGGGVVQPGRSLRLTCAASGFIFSSYGMHWVRQAPGKGLEWVAVISYDGSYKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCTKADYYDFWSGYQKTYYYYMDVWGKGTTVTISS

>P008_062_heavy_chain

QLQLQQSGAEVKKPGASVKDSCKASGYTFTSYGISWVRQAPGQGLEWVGWISAYNGNTNYGLKLQGRVTMTTDTS
TSTAYMELRSLRSDDTAVYYCARAHPFYDSGGYSDYWGQGSLVTVSS

>P008_064_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWASSISSNSSCIYYADSVKGRFTISRDNA
KNSLYLQMNSLRAEDTAVYYCASGEHNYYYYYGMDVWGQGTTVTISS

>P008_065_heavy_chain

EVQLLESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGSNKYYTDSVKGRFTISRDNS
KNTLYLQMNSLRPEDRAVYYCAKDAADSRITMFGVVIISHFDYWGQGTLVTVSS

>P008_067_heavy_chain

EVQLVESGGALVQPGGSLRLSCEASEFIVSRNYMNWVRQAPGKGLEWVSVIYSGGSTFYADSVKGRFIISRDNSK
NTLYLQMNSLRAEDTAVYYCARELVGYFDYWGQGTLVTVSS

>P008_068_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGFTVSSNYMSWVRQAPGKGLEWVSVIYSGGSTYYADSVKGRFTISRDNSK
NTLYLQMNSLRAEDTAVYYCASTPRGDSYGGGAYWGQGTLVTVSS

>P008_070_heavy_chain

QVQLVESGGGVVQPGKSLRLSCAASGLTFSSYGMHWVRQAPGKGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKEGWGYSYGSYYFDYWGQGTLVTVSS

>P008_071_heavy_chain

EVQLVESGPGLVKPSETLSLTCTVSGGSISSTSYYWGWIRQPPGKGLEWIGSMYYSGNTYYNPPLQSRVTISVDT
SKNQVSLSLSSVTAADTALYYCARHPDICDFWSGYPGWFDPWGQGTLVTVSS

>P008_072_heavy_chain

QVQLVQSGAEVKKPGSSVKVSCKTSGDTFRSYSISWVRQAPGQGLEWMGRIIPILGIPNYAQKFQGRVTITADKS
TNTAYMELSSLRSEDTAVYYCARESPYCSSTTCLSDYWGQGTLVTISS

>P008_073_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKASGYSFTSYGISWVRQAPGQGLEWMGWISAFNNNANYVQKFQGRVTMTTDTS
TSTAYLELRSLRSDDAAVYYCARDDDNYDTTGYYLYWGQGTLVTISS

>P008_076_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGFIFSSYDMHWVRQATGKGLEWVSGIGNAGDTHYPGAVKGRFTISRENAK
NSLYLQMNSLRAGDTAVYYCARAHRGYYDRSGYYHNPDAFDIWGQGTTVTISS

>P008_077_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASGFMFSSYNMNWVRQAPGKGLEWVSSISSSSSYIYYEDSVKGRFTISRDNA
KNSLYLQMISLSAEDTAVYYCARMAYFDSSGYYPNAFDIWGQGTMVTVSS

>P008_079_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKVSGYTLTELSMHWVRQAPGKGLEWMGGFVAEDGETIYAQKFQGRVSMTEDTS
TDTAYMELSSLRSEDTAVYYCATDRARLDYFASGSYYGHFDYWGQGTLVTVSS

>P008_080_heavy_chain

QVQLVESGGGLVQPGGSLRLSCAASGFTFSIYWMHWVRQAPGKGLVWVSRINSDGSSTSNADSVKGRFTISRDNA
KNTLYLQMNSLRAEDTAVYYCARDHVVAATPGMDVWGQGTTITVSS

>P008_081_heavy_chain

EVQLVQSGPEVKKPGTSVKVSCKASGFTFTSSAMQWVRQARGQRLEWIGWIVVGSGNTNYAQKFQERVTITRDMS
TSTAYMELSSLRSEDTAVYYCAAPNCSRTSCQDGFDIWGQGTMVTVSS

>P008_082_heavy_chain

QVQLVQSGAEVKKPGASVRVSCKAFGYSFTSFNLNWVRQAPGQGLEWMGWISAYNGNTSYAQKFQGRVTMTTDTS
TTTVYMELRSLRSDDTAVYYCARALENYYDRNGNYYVGAFDYWGQGSLVTVSS

>P008_083_heavy_chain

QVQLVESGGGLVQPGGSLRLSCAASGITVNSNYMSWVRQAPGKGLEWVSVIYSGGSTFYADSVKGRFTISRDNSK
NTLYLQMHSLRAEDTAVYYCARDRFGRINDYWGQGTLVTVSS

>P008_085_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVAVISYDGSDKYYADSVKGRLTIFRDNS
KNTLYLQMNSLRAEDTAVYYCAREGADRSGWWGSFDYWGQGTLVTVSS

>P008_086_heavy_chain

QVQLVQSGPEVKKPGTSVKVSCKASGFTFTSSAMQWVRQARGQRLEWIGWIVVGSGNTHYAQKFQERVTITRDMS
TSTAYMELSSLRSEDTAVYYCAAPYCTTTRCHDGFDIWGQGTMVTVSS

>P008_087_heavy_chain

QVQLQESGPGLVKPSQTLSLTCTVSGGSISSYGKYWSWIRQHPGKGLEWIGYIYYSGSTQYNPSLKSRVTISVDT
SKNQFSLKLNSVTAADTAVYYCARDSAAGHFDYWGQGTLVTVSS

>P008_088_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSGYGMHWVRQAPGRGLEWVAVISYDGSNKYYADSVKGRFTISRDNS
KNTVYLQMNSLRAEDTAVYYCAKPYSGSYWSYFDYWGQGTLVTVSS

>P008_090_heavy_chain

EVQLVETGGGLIQPGGSLRLSCAASGFTVTANYMSWVRQAPGKGLEWVSTIYSGGSTFYADSVKGRFTISRDNSR
NTLYLQMNSLRAEDTAVYYCARERGAKAFDPWGQGALVTVSS

>P008_091_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFSSNAMHWVRQAPGKGLEWVAFISYDGSNQYSTDPVKGRFTFARDNS
KDTLYLQMNSLRGDDTAVYYCAKSRGGNYFDAFDMWGQGTMVTVSS

>P008_092_heavy_chain

QVQLQESGPGLVKPSETLSLTCTVSGDSISNNYAYWGWIRQPPGKGLEWIGNLYYSGSTYYNPSLKSRVTISVDT
SENQFSLRLSSVTAADTAVYYCARLPYGYDYVEAFDIWGQGTLVTISS

>P008_093_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKASGDTFTSYDVNWVRQATGQGLEWMGWMDPNSGNTGYAQNFQGRVTMTRSTS
IGTAYMELSSLRSEDTAVYYCARSKERGYYNRTGYYYPGDWFDPWGQGTLVTVSS

>P008_095_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFISYAMHWVRQAPGKGLEWVAVISYDGNNKYYSDSVKGRFTISRDNS
KNTLHLQMNSLRAEDTAVYYCAKDLSGGYSYWDYWGQGTLVTVSS

>P008_096_heavy_chain

QVQLQESGPGLVKPSETLSLTCTVSGGSTSSSFHYWGWIRQPPGKGLEWIGNIYYSGSTYYNPSLKSRVTISVDT
SKNQFSLKLSSVTAADTAVYYCARLGSGSYYTADYWGQGTRVTISS

>P008_099_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSHAMNWVRQAPGKGLEWVSGISGSGGSTYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKSMFGYDSSGYFYGEDFDYWGQGTLVTISS

>P008_100_heavy_chain

EVQLLESGGGVVQPGGSLRLSCAASGLIFSHYGMHWVRQAPGKGLEWVAFIRNDGTNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKDGGYYYESSGWFDYWGQGTLVTVSS

>P008_102_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGIHWVRQAPGKGLEWVAVISYDGSTKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYFCAKDRGTLIEGMDVWGKGTTVTVSS

>P008_103_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASEFIVSRNYMNWVRQAPGKGLEWVSLIYSGGTTYYADAVKGRFTISRDNSK
NTLYLQMNSLRAEDTAVYYCARGFGDRRLDYWGQGTLVTVSS

>P008_108_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASGFTFSDYYMSWIRQAPGKGLEWVSYITSPGSAIYYADSVKGRFTISRDNA
KNSLYLQMNSLRAEDTALYYCTRDGVIPPRFDYWGQGTMVTVSS

>P054_002_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKASGYTFTDYYMHWVRQAPGQGLEWMGWINPNSGGTDYVQKFQGRVTMTRDTS
ISTAYMELSRLRSDDTAVYYCASLSAAGPLNDVFDIWGQGTMVTVSS

>P054_003_heavy_chain

QVQLVQSGTEVKKPGASVKVSCTASGYTFTVCYIHWVRQAPGQGLEWMGWINPNSGGTNYAQKFQGRVTMTRDTS
ISTAYMELSRLRSDDTAVYYCARVARHYYDRSGNLHSADYFQHWGQGTLVTISS

>P054_004_heavy_chain

QVQLQQWGAGLLKPSETLSLTCAVYGESFSGYYWSWIRQPPGKGLEWIGEIIHSGSTNYNPPLKSRVTISVDTSK
SQFSLKLSSVTAADTAVYYCAREMSVAVVDHWGQGTLVTVSS

>P054_005_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKASGYTFTGFYIHWVRQAPGQGLEWMGWINPNSGGSNCAQKFQGRVTMTRDTS
ISTAYMEVSRLRSDDTAVYYCARYSNYYYYYGMDVWGQGTTVTVSS

>P054_006_heavy_chain

QVQLVQSGAEVKKPGESLKISCKGSGYSFTNSWIAWVRQMPGKGLEWMGIIYPGDSDTRYSPSFQGQVTISADKS
ISTAYLQWSSLKASDSAMYYCARLGSWYVYYYYYALDVWGQGPLVTVSS

>P054_009_heavy_chain

QVQLVESGGGLVQPGGSLRLSCAASGFTFSSYSMNWVRQAPGKGLEWVSYISSSGSTKYYGDSVKGRFTISRDNA
KNSLYLQLNSLRDEDTAVYYCARTIYSYDSSGYYGTERYFDYWGQGTLVTVSS

>P054_012_heavy_chain

EVQLLESGGGLVQPGRSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSVISWHSGSIDYADSVKGRFTISRDNA
KNSLYLQMNSLRAEDTALYHCAKGTGYSYGYAVDGGFDYWGQGTLVTISS

>P054_015_heavy_chain

EVQLVESGGGVVKPGGSLRLSCAASGFTFSTYSMNWVRQAPGKGLEWVSSISSSSTDIHYADSMKGRFTISRDNA
KNSLYLQMNSLRAEDTAVYYCARDFHRGWYDHSAYIFDFWGQGTLVTVSS

>P054_017_heavy_chain

QVQLVASGGGVVQPGRSLRLSCAASGFIITSYGMHWVRQAPGKGLEWVAAIWYDGSNKYYADSVKGRFTISRDNS
KNTLYLQMNSPRAEDTAVYYCARDDPTPDGDAFDIWGQGTMVTVSS

>P054_021_heavy_chain

EVQLVQSGAEVKKPGASVKVSCKVSGYTLTELSMHWVRQAPGKGLEWLGGFDPEDAETIYAQKFQGRVTMTEDTS
TDTAYMGLSSLTSEDTAVYYCATGVAVAGTQKNYSYYYGLDVWGQGTTVTVSS

>P054_022_heavy_chain

EVQLVESGSGLVKPSQTLSLTCAVSGGSIDSGGDSWSWIRQPPGKGLEWIGYIYHSGSTYYNPSLKSRVTISVDR
SNNQFSLKLSSVTAADTAVYYCARHSGYDLGGAFDIWGQGTMVTISS

>P054_026_heavy_chain

QVQLVESGGGLVQPGRSLRLTCAASGFTFDDYAMHWVRQAPGKGLEWVSGISWNSGSIGYADSVKGRFTISRDNA
KNSLYLQMNSLRPEDTALYYCAKDDSSSWYFYSRAKLGQYYYYGMDVWGQGTTVTVSS

>P054_027_heavy_chain

QVQLQESGPGLVKPSQTLSLTCTVSGGSISSGDYYWSWIRQPPGKGLTWIGHISYSGSTYYNPSLKSRLTISVDT
SKNQFSLKLSSVTAADTAVYYCARQLWLRAPFDYWGQGALVTVSS

>P054_031_heavy_chain

QVQLVQSGAEVKEPGASVKVSCKASGYTFTSYGISWVRQAPGQGLEWMGWISAYDGNTNYAQKLQGRVTMTTDTS
TSTAYMELRSLRSDDTAVYYCAREGSDYYDSSGFHDYWGQGTLVTVSS

>P054_036_heavy_chain

EVQLVQSGGGLIQPGGSLRLSCAASVFIVSSNYMSWVRQAPGKGLEWVSVIYSGGSTFYADSVKGRFTISRDDSK
NTLYLQMNSLRAEDTAVYYCARSRGGPLDYWGQGTLVTVSS

>P054_044_heavy_chain

QVQLVQSGAEMKEPGSSVKVSCKASGGTFSSYAISWVRQAPGQGLEWMGGIIPIFGTANYAQKFQDRVTITADES
TSTAYLELSSLRSEDTAVYYCARESTTIFGVVILTSYGMDVWGQGTTITISS

>P054_045_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFSSYGMHWVRQAPGKGLEWVALISYDGGNKYYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCARTRGGSYYYGMDVWGQGTTVTVSS

>P054_048_heavy_chain

QVQLQESGPGLVKPSETLSLTCTVSGGSISSSSYYWGWIRQPPGKGLEWIGTIYYSGSTYYNPSLKSRVTISVDT
SNNQFSLKLSSVTAADTAVYYCARHVQWLVLYYFDYWGQGTLVTVSS

>P054_050_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGFTFSSYDMHWVRQPTGQDLEWVSAIGTAGDTYYPDSVKGRFTISRENAK
NSSYLQMNSLRAGDTAVYYCARASFDSSGYLNYFDYWGQGTLVTVSS

>P003_001_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSVSSNYLAWYQQKPGQAPRLLIYAASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYFCQQYGSSRNTFGQGTKVEIK

>P003_002_light_chain

DIVLTQSPATLSVSPGERATLSCRASQSVSSNLAWYQQKPGQAPRLLIYGASTRATGIPARFSGSGSGTEFTLTI
SSMQSEDFAVYYCQQYNNWPPGDTFGQGTKVEIK

>P003_003_light_chain

DIQLTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQEPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYSTPFTFGPGTKLEIK

>P003_006_light_chain

EIVMTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTI
SSLEPEDFAVYYCQQRSNWPTFGQGTKVEIK

>P003_007_light_chain

AIRMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQPEDIATYYCQQYDNLPLTFGGGTKVEIK

>P003_014_light_chain

DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYSSQYTFGQGTKVEIK

>P003_015_light_chain

QSALTQPASVSGSPGQSITISCTGTSSDVGSYNLVSWYQQHPGKAPKLMIYEVSKRPSGVSNRFSGSKSGNTASL
TISGLQAEDEADYYCCSYAGSSTWVFGGGTKLTVL

>P003_016_light_chain

DIVMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYSTPPWTFGQGTKLEIK

>P003_017_light_chain

DIVMTQSPSFLSASVGDRVTITCRASQGISSYLAWYQQKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTEFTLTI
SSLQPEDFATYYCQQLNSYTLTFGGGTKVDIK

>P003_022_light_chain

QSALTQPRSVSGSPGQSVTISCTGTSSDVGGYNYVSWYQQHPGKAPKLMIYDVSKRPSGVPDRFSGSKSGNTASL
TISGLQAEDEADYYCCSYAGSYTFGYVFGTGTKVTVL

>P003_026_light_chain

SYELTQPPSVSVSPGQTARITCSGDALPKQYAYWYQQKPGQAPVLVIYKDSERPSGIPERFSGSSSGTTVTLTIS
GVQAEDEADYYCQSADSSGTYHVVFGGGTKLTVL

>P003_027_light_chain

DIQMTQSPSSLSASVGDRVTITCRASQSISTYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYSTPYTFGQGTKVEIK

>P003_030_light_chain

DIVMTQTPLSLSVTPGQPASISCKSSQSLLHSDGKTYLYWYLQKPGQPPQLLIYEVSNRFSGVPDRFSGSGSGTD
FTLKISRVEAEDVGVYYCMQSIQLARFGQGTKLEIK

>P003_032_light_chain

NFMLTQPPSVSVSPGQTARITCSGDALPKKYAYWYQQKSGQAPVLVIYEDSKRPSGIPERFSGSSSGTMATLTIS
GAQVEDEGDYYCYSTDSSGNKRVFGGGTKLTVL

>P003_034_light_chain

DIQMTQSPSSLSASVGDRVTITCRASQGISNYLAWYQQKPGKVPKLLIYAASTLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDVATYYCQKYNSAPHTFGGGTKVEIK

>P003_038_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSYTFGQGTKLEIK

>P003_043_light_chain

DIVMTQSPLSLPVTPGEPASISCRSSQSLLHSNGYNYLDWYLQKPGQSPQLLIYMGSNRASGVPDRFSGSGSGTD
FTLKISRVEAEDVGVYYCMQALQSPYNLGQGTKLEIK

>P003_045_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSFTFGPGTKVDIK

>P003_051_light_chain

SYELTQPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSNSGNTATLTIS
RVEAGDEADYYCQVWDSSSDHVVFGGGTKLTVL

>P003_053_light_chain

DIVMTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSPWTFGQGTKVDIK

>P003_055_light_chain

DIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTRESGVPDRFSGSGSGT
DFTLTISSLQAEDVAVYYCQQYYSTPRTFGQGTKVEIK

>P003_056_light_chain

NIVMTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTRESGVPDRFSGSGSGT
DFTLTISSLQAEDVAVYYCQQYYSTWTFGQGTKLEIK

>P003_058_light_chain

DVVMTQSPSSVSASVGDRVTITCRASQGISSWLAWYQQKPGKAPELPIYAASTLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQANSFPWTFGQGTKVDIK

>P003_064_light_chain

EIVLTQSPDSLAVSLGERATINCKSSQSVLYSSNNKNYLAWYQQKPGQPPKLLIYWASTRESGVPDRFSGSGSGT
DFTLTISSLQAEDVAVYYCQQYYSPPPTFGPGTKVDIK

>P008_001_light_chain

DIVMTQSPSFLSASVGDRVTITCRASQGISNYLAWYQQKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTEFTLTI
SSLQPEDFATYYCQQLNRYPLYTFGQGTKLEIK

>P008_002_light_chain

QSVLTQPPSVSVAPGQTARITCGGNNIGSKSVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSNSGNTATLTIS
RVEAGDEADYYCQVWDSSSDHWVFGGGTKLTVL

>P008_003_light_chain

QSALTQPPSVSGAPGQRVTISCTGSSSNIGADYDVHWYQQLPGTAPKLLIYGNSNRPSGVPDRFSGSKSGTSASL
AITGLQAEDEADYYCQSSDSSLSGSRVFGTGTKVTVL

>P008_004_light_chain

AIRMTQSPSTLSASVGDRVTITCRASQSFSSWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTI
SSLQPDDFATYYCQQYNSYLKTFGQGTKLEIK

>P008_005_light_chain

DIQMTQSPSTLSASVGDRVTITCRASQSFSSWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTT
SSLQPDDFATYYCQQYSTYSPLTFGGGTKVDIK

>P008_006_light_chain

AIRMTQSPSTVSASVGDRVTITCRASQSISTWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTI
SRLQPDDFATYYCQQYNSYLFGQGTKVEIK

>P008_007_light_chain

SLSQLVLTQGRLGPGQTARITCEGNNIGGKSVLWYQQKPGQAPVLVVYDDSDRPSGIPERFSASNSGNTATLTIS
RVEAGDEADYYCQVWDTSSDHAGVFGGGTKLTVL

>P008_010_light_chain

EIVMTQSPATLSVSPGERATLSCRASQSISSNLAWHQQKPGQAPRLLISDASTRATGIPARFSGSGSGTEFTLTI
DSLQSEDIAVYYCQQYNNWPPTITFGQGTRLEIK

>P008_011_light_chain

DIQMTQSPSSLSASTGDRVTITCRASQGIRTYLVWYQQKPGQAPNLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SGLQSEDFATYYCQQYYSYPLTFGGGTKVEVK

>P008_012_light_chain

DIQMTQSPSSLSASVGDRVTITCQASQDINNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTI
STLQPEDIATYYCQQYDNLLSLTFGGGTKVEIK

>P008_014_light_chain

DIQLTQSPSSLSASVGDRVTITCQASQDISNYLNWYQQKPGKAPKFLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQPEDFATYFCQQYDDLPYTFGQGTKLEIK

>P008_015_light_chain

SYELTQPPSVSVAPGQTARITCGGNNIGGKSVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSNSGNTATLTIS
RVEAGDEADYYCQVWDSSSDHVVFGGGTKLTVL

>P008_016_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSVSSIYLAWYQQKPGQAPRLLIYAASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGTSPWTFGQGTKLEIK

>P008_018_light_chain

EIVLTQSPGTLSSSPGERATLSCRASQGVSSSYLAWYQQKLGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQHSRTFGQGTKLEIK

>P008_022_light_chain

EIVMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSRTDFTFTI
SSLQPDDIATYYCQQYDKPPVTFGQGTKVDIK

>P008_023_light_chain

EIVMTQSPATLSVSPGERATLSCRASQSVSSNLAWYQQKPGQAPRLLIYGASTRATGIPARFSGSGSGTEFTLTI
SSLQSEDFAVYYCQQYNKWPRTFGQGTKLEIK

>P008_024_light_chain

DIVMTQSPSSLSASEGDRVTITCQASQDINNYLNWYQQKPGKAPKLLIYDASNLEAGVPSRFSGSGSGTDFTFTI
SSLQPEDIATYYCQQYDNLPPAFGPGTKVDIK

>P008_032_light_chain

EIVLTQSPGTLSLSPGEGATLSCRASQSVSSYLAWYQRKPGQAPRLLIYDSSNRATGIPARFSGSGSGTDFTLTI
SSLEPEDFAVYYCQERSSWPPAFGQGTRLEIK

>P008_035_light_chain

DVVMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGQAPKLLIYAASNLQSGVPSRFSGSGSGTDFTLTI
SSLQPADFATYYCQQSYSIFRTSGQGTKVEIK

>P008_036_light_chain

DIVMTQTPLSLSVTPGQPASISCKSSQSLLHSDGKTYLYWYLQRPGQPPQLLIYEVSDRFSGVPDRFSGSGSGTD
FTLKISRVEAEDVGVYYCMQSIQLPVTFGQGTRLEIK

>P008_038_light_chain

QTVVTQPPSVSGAPGQRVTISCTGSSSNIGAGYDVHWYQQLPGTAPKLLIFGNSNRPSGVPDRFSGSKSDTSASL
AITGLQAEDEADYYCQSYDSSLIGSVFGTGTKVTVL

>P008_039_light_chain

DIVMTQSPSSVSASVGDRVTITCRASQGISSGLAWYRQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQANSFPYTFGQGTKVDIK

>P008_040_light_chain

DIVMTQSPATLSVSLGERATLSCRASQSVSSNLAWYQQKPGRAPRLLIYGASTRATGIPVRFSGSGSGTEFTLTI
SSLQSEDFAVYYCQQYNNWPPSITFGQGTRLEIK

>P008_042_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSISSTYLAWYQQKPGQAPRLLIHGASSRATGISDRISGSGSGTDFILT
ISRLEPEDFAVYYCQQYGSSPGTFGQGTKVEIK

>P008_044_light_chain

DVVMTQTPLSLSVTPGQPASISCKSSQSLLHSDGKTYLYWYLQRPGQPPQLLIYEVSDRFSGVPDRFSGSGSGTD
FTLKISRVEAEDVGVYYCMQSIQLPVTFGQGTRLEIK

>P008_045_light_chain

DIVMTQTPLSLSVTPGQPASISCKSSQSLLHSDGKTYLYWYLQKPGQPPQLLIYEVSNRFSGVPDRFSGSGSGTD
FTLKISRVEAEDVGVYYCMQSIQPWRLTFGGGTKVEIK

>P008_047_light_chain

DVVMTQSPSSLSASVGDRVTITCQASQDIRNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQSEDIATYYCQQYDNLPITFGQGTRLEIK

>P008_051_light_chain

EIVLTQSPATLSLSPGEGATLSCRASQSVGRYLAWYQQKPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTI
SSLEPEDFAVYYCQQRSNWPPFTFGGGTKVDIK

>P008_052_light_chain

QAVVTQPRSVSGSPGQSVTISCTGTSSDVGAYNFVSWYQQHPGKAPKLMTYDVTKRPSGVPDRFSGSKSGNTASL
TISGLQAEDEADYYCCSYAGSFYVFGTGTKVTVL

>P008_055_light_chain

QTVVTQAPSASGTPGQRVTISCSGSSSNIGSNTVNWYQQLPGTAPKLLIYSDNQRPSGVPDRFSGSKSGTSASLA
ISGLQSEDEADYYCAAWDDSLNVVVFGGGTKLTVL

>P008_056_light_chain

AIRMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQPEDIATYYCQHHDSLPLTFGGGTKVEIK

>P008_057_light_chain

DVVMTQSPSSLSASVGDRVTITCRASQGVSSYLAWYQQKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTEFTLTI
SSLQPEDFATYYCQQLNSYPPGTFGQGTKLEIK

>P008_060_light_chain

DIQLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYGTSSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSPQITFGQGTRLEIK

>P008_062_light_chain

DIQMTQSPSSLSASVGDRVTITCRASQGITNSLAWYQQQPGKAPKLLLYAASRLASGVPSRFSGSGSGTDYTLTI
SSLQPEDFATYYCQQYYSAPPTFGQGTKVEIK

>P008_064_light_chain

DIVMTQSPLSLAVTPGEPASISCRSSQSLLHSDGYNYLDWYLQKPGQSPQLLTYLGSNRASGVPDRFSGSGSGTD
FTLKISRVEAEDVGVYYCMQALQTSITFGQGTRLEIK

>P008_065_light_chain

DIQMTQSPSTLSAFVGDRVTIICRASQSISSWLAWYQQKPGKAPKLLIYKASSLESGVPSTFSGSGSGTEFTLTI
SSLQPDDFATYFCQQYNSYPYTFGQGTKVEIK

>P008_067_light_chain

DIVMTQSPSFLSASVGDRVTITCRASQGIRNDLGWYQQKPGKAPKRLIYAASSLQSGVPSRFSGYGSGTEFTLTI
SSLQPEDFATYYCQHLNSYPRYTFGQGTKLEIK

>P008_068_light_chain

QLVLTQSPSASASLGASVKLTCTLSSGHSSNAIAWHQQQPEKGPRYLMKLNIDGSHRKGDGIPDRFSGSSSGAER
YLTISSLQSEDEADYYCQTWGTGPNWVFGGGTKLTVL

>P008_070_light_chain

DIQMTQSPSSLSASVGDRVTITCQASQDISNYLNWYHQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQPEDIATYYCQQYDNLPLTFGGGTKVDIK

>P008_071_light_chain

QAVVTQPPSVSGAPGQRVTISCTGSSSNIGAGYDIHWYQQLPGTAPKLLIYNNSNRPSGVPDRFSGSKSGTSASL
AITGLQAEDEADYYCQSYDSSLSGYVFGTGTKVTVL

>P008_072_light_chain

QSVLTQPPSVSGAPGQRVTISCTGSNSNIGAGYDVHWYQQFPGTAPKLLIYDNTNRPSGVPDRFSGSKSGTSASL
AITGLQAEDEADYYCQSYDSSRIDVVFGGGTKLTVL

>P008_073_light_chain

DIVMTQSPSTLSASVGDRVTITCRASQSISRWLAWYQQIPGKAPNLLIYQTSSLQSGVPSRFSGSGSGTEFTLTI
SSLQPDDFATYYCQQYYSYPLTFGGGTKVEIK

>P008_076_light_chain

NFMLTQPASVSGSPGQSITISCTGTSSDVGVYDYVSWYQQHPGKAPKLMIYEVSNRPSGVSNRFSGSKSGNTASL
TISGLQAEDEADYYCSSYTSSSTLVLYVFGTGTKVTVL

>P008_077_light_chain

EIVLTQSPATLSVSPGERVTLSCRASQSVSSNLAWYQQKPGQAPRLLIYGASTRATGIPARFSGSGSGTEFTLTI
SSLQSEDFAVYYCQHYYNWPPWTFGQGTKLEIK

>P008_079_light_chain

DIQLTQSPSFLSASVGDRVTITCRASQGIRSYLAWYQQKPGKAPNLLIYAASTLQSGVPSRFSGSGSGTEFTLTI
SSLQPEDFATYYCQQLNTYALTFGGGTKLEIK

>P008_080_light_chain

SYELTQPPSVSVSPEQTASITCSGDKLGDKYACWYQQKAGQSPVLVIYEDGKRPSGIPERFSGSNSGNTATLTIS
GTQTMDEADYYCQAWDRTTAVFGGGTKLTVL

>P008_081_light_chain

DIQLTQSPGTLSLSPGERATLSCRASQSVSSSHLAWYQQKSGQAPRLFIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSPPWTFGQGTKVEIK

>P008_082_light_chain

DVVMTQSPGTLSLSPGERATLSCRASHNISSTHLVWYQQKPGQAPRLLIYGASNRATGIPDRFSGGGSGTDFTLT
ISRLEPEDFAVYFCQQFGSSPQTFGQGTKVDIK

>P008_083_light_chain

DIVMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQQKPGRAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQPEDVATYYCQQYDNLPGTFGQGTKLEIK

>P008_085_light_chain

AIRMTQSPDSLAVSLGERATINCKSSQSVLHNSKNKSYLAWYQQKPGQPPNLLIYWASTRESGVPDRFSGSGSGT
DFTLTISSLQAEDVAVYYCQQYYGGRWTFGQGTKVDIK

>P008_086_light_chain

DVVMTQSPGTLSLSPGERATLSCRASQSVSRSYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSLLFTFGPGTKLEIK

>P008_087_light_chain

EIVMTQSPSSLSASVGDRVTITCQASQDINNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGAGTYFTFTI
SSLQPEDIATYYCQQYDNLPPLFTFGPGTKVEIK

>P008_088_light_chain

DIVMTQSPDSLAVSLGERATINCKSSQSVLHSSNNKNYVAWYQQKPGQPPKLLIYWASTRESGVPDRFSGSGSGT
DFTLTISSLQAEDVAVYHCQQYYSTPLTFGGGTKVEIK

>P008_090_light_chain

EIVMTQSPSSLSASVGDRVTITCQASQDIRNYLNWYQQTPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQPEDIATYYCHQYDNLPQSFGGGTKVDIK

>P008_091_light_chain

DIQLTQSPSTLSASVGDRVTITCRASQSISSWLAWYQQKPGKAPKLLIYKASSLESGVPSRFSGSGSGTEFTLTI
SSLQPDDFATYYCQQYNSYPWTFGQGTKVEIK

>P008_092_light_chain

AIRMTQSPSSLSASVGDGVTITCRASQGIRNDLGWYQQKPGKAPKRLIYAASSLQSGVPSRFSGGGSGTEFTLTI
SSLQPEDFATYYCLQHNSYPLTFGGGTKVDIK

>P008_093_light_chain

DVVMTQSPATLSLSPGERATLSCRASQSVSSSLAWYQQKPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTI
SSLGPEDFAVYYCQQRSNWPPTWTFGQGTKLEIK

>P008_095_light_chain

EIVLTQSPATLSLSPGERATLSCRASQSVSSYLAWYQQKPGQAPRLLIYDTSNRATGIPARFSGGGSGTDFTLTI
SSLEPEDFAVYYCQQRSNWRTFGQGTKLEIK

>P008_096_light_chain

DIQMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTEFTFTI
TSLQPEDIATYYCQQYDTLPPTFGPGTKVDIK

>P008_099_light_chain

DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQHKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYSIPPTFGGGTKVDIK

>P008_100_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKSGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLQPEDFAVYYCQQYGSSPMYTFGQGTKLEIK

>P008_102_light_chain

QLVLTQPPSVSEAPRQRVTISCSGSSSNIGNNPVNWYQQLPGKAPKPLIYYDDLLPSGVSDRFSGSRSATSASLA
ISGLQSEDEADYYCTAWDGSLNGYVFGSGTKVTVL

>P008_103_light_chain

EIVLTQSPSSLSASVGDRVTITCQASQDINNYLNWYQQKPGKAPKLLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQPEDIATYYCQQFGEWFTFGQGTKLEIK

>P008_108_light_chain

DIVMTQSPSSLSASVGDRVTITCQASQDISKYLNWYQRKPGTAPTLLIYDASNLETGVPSRFSGSGSGTDFTFTI
SSLQPEDIGTYYCQQYDNLPTFGGGTKVEIK

>P054_002_light_chain

QLVLTQPPTVSVSPGQTARITCSGDALPKKYAYWYQQKSGQAPVLVIYEDSKRPSGIPERISGSSSGTMATLTIS
GAQVEDGADYYCYSIDSSGNHYVFGTGTKVTVL

>P054_003_light_chain

AIRMTQSPSSLSASVGDRVTITCQASQDISNFLNWYQEKPGKAPKLLISDASNLETGVPSRFSGSGSGTDFTFTI
TSLQPEDIATYYCQQYDNLLITFGQGTRLEIK

>P054_004_light_chain

QSALTQPASVSGSPGQSITISCTGTSSDIGGYNYVSWYQQHPGKAPKLMIYDVTNRPSGVSNRFSGSKSGNTASL
TISGLQAEDEATYYCNSYTSSSTYVLFGGGTKLTVL

>P054_005_light_chain

DIQMTQTPLSLPVTPGEPASISCRSGQSLLDSDDGNTYLDWYLQKPGQSPQLLIYTLSYRASGVPDRFSGSGSGT
DFTLKISRVEAEDVGVYYCMQRIEFPWTFGQGTKLEIK

>P054_006_light_chain

EIVLTQSPATLSVSPGERATLSCRASQSVSRNLAWYQQKPGQAPRLLIYGASTRATGIPARFSGSGSGTEFTLTI
SSLQSEDFAVYYCQQYNNGRGTFGPGTKLEIK

>P054_009_light_chain

DVVMTQSPATLSLSPGERATLSCRASQSVSSYLAWYQHKPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTI
SSLEPEDFAVYYCQQRSNWPTFGQGTKLEIK

>P054_012_light_chain

QSALTQPPSASGTPGQRVTISCSGSSSNIGSYTVNWYQHLPGTAPKLLIYSNNQRPSGVPDRFSGSKSGTSASLA
ISGLQSEDEADYYCAAWDDSLNGYVFGTGTKVTVL

>P054_015_light_chain

DVVMTQSPSTLSASVGDRVTITCRASQSISSWLAWYQQKPGKAPKLLIYDASSLESGVPSRFSGSGSGTEFTLTI
SSLQPDDFATYYCQQYNSYSFTFGPGTKKDIK

>P054_017_light_chain

QLVLTQPPSVSVAPGQTARITCGGNSIGSKTVHWYQQKPGQAPVLVVYDDSDRPSGIPERFSGSKSGNTATLTIS
RVEAGDEADYYCQVWHISSDHLWVFGGGTKLTVL

>P054_021_light_chain

DIQMTQTPLSSPVTLGQPASISCRSSQSLVHSDGNTYLSWLQQRPGQPPRLLIYKISNRFSGVPDRFSGSGAGTD
FTLKISRVEAEDVGVYYCMQATQFPITFGQGTRLEIK

>P054_022_light_chain

QSALTQPPSASGSPGQSVTISCTGTSSDVGGYNYVSWYQQHPGKAPKLMISEVSKRPSGVPDRFSGSKSGNTASL
TVSGLQAEDEADYYCNSYAGSNNWVFGGGTKLTVL

>P054_026_light_chain

EIVMTQSPATLSLSPGERATLSCRASQSVNNYLAWYQQEPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTI
SSLEPEDLAVYYCQQCSNWPPSLTFGGGTKVEIK

>P054_027_light_chain

NFMLTQPHSVSESPGKTVTISCTGISGSIASNYVQWYQQRPGSAPTTVIYEDNQRPSGVPDRFSGSIDSSSNSAS
LTISGLKTEDEADYYCQSYDSSKYVVFGGGTKLTVL

>P054_031_light_chain

EIVLTQSPSTLSASVGDRVTITCRASQSISSWLAWYQQKPGKAPKVLIYKASSLESGVPSRFSGSGSGTEFTLTI
SSLQPDDFATYYCQQYNSYVTFGPGTKLEIK

>P054_036_light_chain

DVVMTQSPSSLSASVGDRVTITCRASQSIRNYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYGIPXWTFGPGTKLEIK

>P054_044_light_chain

DIQMTQSPSTLSASVGDRVTITCRASQSIRSWLAWYQQKPGKAPKVLIYDASSLESGVPSRFSGSGSGTEFTLTI
SSLQPDDFATYYCQQYNYYSVTFGQGTKVEIK

>P054_045_light_chain

EIVLTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTPTI
SSLQPEDFATYYCQQSYSTPPYTFGQGTKVEIK

>P054_048_light_chain

QSALTQPRSVSGSPGQSVTISCTGTSSDVGGYNSVSWYQHHPGKAPKLMIYDVSKRPSGVPDRFSGSKSGNTASL
TISGLQAEDEADYYCCSYAGRYTLVFGGGTKLTVL

>P054_050_light_chain

AIRMTQSPSSLSASVGDRVTITCRASQSIRNYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYGIPDWTFGPGTKVEIK

## 7.4 SARS-CoV-2 mAbs (VA014)

>VA014_01_heavy_chain

EVQLLESGGGLVQPGGSVRLSCAASGFTFTSYDMHWVRQATGKGLEWISIIGTAGDPYYADSVKGRFSVSRENAK
NSLYLQINSLRAGDTAVYYCARGGVTMLQGAIRRHYYYYMDVWGKGTTVTVSS

>VA014_02_heavy_chain

QVQLVESGGGLVQPGGSLGLSCVASGFTFSYFWMSWVRQAPGKGLEWVANINLDGSEKYYVDSVKGRFTISRDNA
KNSLFLQMNSLRVEDTAVYYCARDLRYFDWLSMGNDYWGQGTLITVSS

>VA014_03_heavy_chain

EVQLVESGGGLVKPGGSLRLSCAASGFMFSDYGMNWVRQAPGKGLEWVSYISDSGTTIHYADSVKGRFTISRDNA
KNSLYLQMTGLRDGDTAVYYCARARPPYDFWSGYHYFGHFDYWGQGTLVTVSS

>VA014_04_heavy_chain

EVQLVESGGGLVQPGESLRLSCAGSGFTFSAYDMHWVRQGTGEGLEWVSAIGTSGDTYYADSVKGRFTISRDNAK
NSLFLQMDSLTVGDTAVYYCARGDAISGVEYYFDYWGQGTLVTVSS

>VA014_05_heavy_chain

QVQLLESGGGVVQPGRSLRLSCAASGFTFSGFAINWVRQAPGKGLEWVAVISYDGINKYYADSVKGRFTISRDNS
NNTLYLQMNSLREVDTAVYFCASGRGSYRSPFDYWGPGTLVTVSS

>VA014_07_heavy_chain

QVQLVEAGPGLVKPSETLSLTCTVSGGSISNRNYDWGWIRQPPGRGLEWIGTISYSGSTSHNPSLKSRVSIFIDT
SKNQFSLKLSSVTPSDTALYFCARDGGLGWFGEQMDWYFDLWGRGTLVTVSS

>VA014_09_heavy_chain

QVQLVQSGTEVKKPGASVKVSCKASGYTFSTYGISWVRQAPGQGLEWMGWISAHNGESNYPQKFQGRVTMTTDTS
TSTAHMELRSLRSDDTAVYYCARDLGWFGELSGGFDLWGPGTMVTVSS

>VA014_10_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASRFMFNRYAMSWVRQAPGKGLEWVSGISGSGDSTQYADSVKGRFTISRDNS
KNTLYLQMNSLRAEDTAVYYCAKDLYGGSSSSTLEYWGQGTLVTVSS

>VA014_11_heavy_chain

VQLVQESGPGLVKPSETLSLTCTISGDSVSTSYWNWIRQPAGKGLQWIGRIYNSGSTNYNPSLENRVTLSVDTSK
NQFSLKLTSVTAADTAVYYCARDPGSRYSSGWYYYSFAMDVWGQGTTVTVSS

>VA014_14_heavy_chain

QVQLQESGPGLVKPSETLSLTCTVSGGSISSYYWSWLRQPPGKGLEWIGYIYYSGSSNYNPSLKSRVTISVDTSK
SQLSLKLTSVTAADTAVYYCARGFDYWGQGTLVTVSS

>VA014_15_heavy_chain

EVQLLESGPGLVKPSETLSLTCTVSGGSISFYYWSWIRQPPGKGLEWIGYFYYSGDTGSNPSLKSRLTMSLDTSK
NQFSLNLTSVTAADTAVYYCARRSLTVTPTSGRNAYHYMDVWGKGTTVTVSS

>VA014_16_heavy_chain

EVQLLESGGDLVQPGGSLRLSCAASGFTFNTYSMNWVRQAPGKGLEWVSSITSGSTTIYYADSVKGRFTISRDNA
KNSLYLQMNSLRDEDTAVYYCARDCYSSGWYTCDYLDPWGQGTLVTVSS

>VA014_18_heavy_chain

EVQLLESGGGLVKPGGSLRLSCAASGFTFSGYGMNWVRQAPGKGLEWVSSISSSSSYIYYADSVKGRFTISRDNA
KNSLYLQMNSLRAEDTAVYYCARDLTGEYNGFWSEYSTKPMDVWGKGTTVTVSS

>VA014_19_heavy_chain

EVQLVESGGGLVKPGGSLRLSCVASGFSFSDAWLSWVRQAPGKGLEWVGRIKSKLDAETTDYAKPVKGRFTISRD
DSENTLYLQMNSLKTDDTAVYYCTTGGLMWFGEEEGYWGQGTLVTISS

>VA014_21_heavy_chain

QVQLVQSGAELKKPGSSVKISCKASAGTFTTFGISWVRQAPGQGLEWMGGIMPVFQIVNYAQNLQGRITISADKS
TTTAYMELSSLRSEDTAVYYCARTQFVLRFLEWSEASMCASDIWGQGTLVTVSS

>VA014_22_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGFTSSSYSMNWVRQVPGKGLEWISYISSSSGTLYYADSVKGRFTISRDNA
KNSLYLQMNSLRDEGTAVYYCARVPAEDIVIIPPALYYYYGMDVWGQGTTVTVSS

>VA014_23_heavy_chain

EVQLVESGPGLVKPSETLSLTCTVSGGSISSTNYYWGWIRQTPGKGLEWIGSIYYSGSTYYNPSLKSRLTISVDR
SKNHFSLQLSSVTAADTAVYYCARGSFILLIFSALATGYFDFWGQGTLVTISS

>VA014_24_heavy_chain

QVQLVESGGALVKPGGSLTLSCAASGFAFNNAWMSWVRQAPGKGLEWVGRIKTKTDGGTTDYVAPVKGRFTISRD
DSKNTLYLQMNSLKTEDTAVYYCTTERDYDYIWGSYRPFQNWGQGTLVTVSS

>VA014_25_heavy_chain

QVQLVQSGAEVKKPGSSVKVSCKASGGTFSNYGITWLRQAPGQGLEWMGGIIPILGVEKYAQKFQGRVTITADKS
TGTAYMQLSSLRSEDTAVYYCARGGINMGQGFIISPYYFDYWGQGTLVTISS

>VA014_26_heavy_chain

QVQLQQWGAGLLKPSDTLSLTCGISGGPFSGYYWSWIRQPPGKGLEWLGEINHSGITTYNPSLKSRGTISVDTSK
RQISLKLASVTAADTAVYYCARFAYGDYGWDYSSDMDVWGKGTTVTVSS

>VA014_27_heavy_chain

EVQLVQSGAEVKKPGESLKISCKGSGYRFTSYWIAWVRQMPGKGLEWMGIIYPGDSDTRYSPSFQGQVTISADKS
ISTAYLQWSSLKAPDTAMYYCARIETMGVVMPAASWGGTYGMDVWGQGTTVTVSS

>VA014_30_heavy_chain

QVQLQQWGAGLLKPSETLSLTCAVSGGSFSGYYWSWIRQPPGKGLEWIGEINHIGTTNYNPSLKSRVIISVDTSK
NQFSLKLSSVTAADTAVYYCTRGRGLYTSGAYYFDFWGQGTLVTVSS

>VA014_33_heavy_chain

EVQLVESGGDLVQPGGSLRLSCGASGFTFSTYWMHWVRQGPGKGLVWVSRISGDGSSTTYADSVKGRFTISRDNA
KNTLYLQMNSLRAEDTAIYYCATSTHYDFWSGPLVHMDVWGKGTTVTVSS

>VA014_35_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKASGYTFTDYYMHWVRQAPGQGLEWMGWINPNSGGTNYAHKFQGRVTMTRDTS
ISTAYMELRRLRSDDTALYYCARGEVATIFGFYYYYGLAVWGQGTTVTISS

>VA014_36_heavy_chain

EVQLQESGPGLVKPSETLSLTCTVSGGSISNYYWSWIRQPPGKGLEWIGYIYYSGSTNYNPSLKSRVTISVDTSK
NQFSLKLTSVTAADTAVYFCAKSRGYSYGLGLGWFDPWGQGTLVTVSS

>VA014_41_heavy_chain

EVQLVESGGGVVQPGRSLRLSCVASGFTFSDYAMHWVRQAPGKGLEWVALISSDGRMDFYPDSVKGRFTISRDNS
KNALYLQMNGLRAEDTAVYYCARDKELNTPALDYWGQGSLVTISS

>VA014_43_heavy_chain

EVQLVESGGGLVQPGGSLRLSCAASGVTVSSNYMSWVRQAPGKGLEWVSVIFPGGSTFYADSVKGRFTISRHNSK
NTLYLQMDSLRAEDTAVYYCARDFFEGAFDIWGQGTMVTVSS

>VA014_44_heavy_chain

QVQLVQSGAEVKMPGSSVKVSCKASGGSFDNYGISWVRQAPGQGLEWMGGIIPILGSTKNAQKFQGRITITADKS
TNTAYMELSSLRSEDTAVYHCARGGITLGQGIMLTYYYMDVWGKGTTVTVSS

>VA014_47_heavy_chain

EVQLVESGGGVVQPGRSLRLSCGGSGFTFSSHAMHWVRQAPGKGLEWVAVISYDGSYQYYADSVKGRFTISRDNS
KNALYLQMNSLRVEDTAIYYCARDTPDLETYYFDCWGQGTLVTVSS

>VA014_48_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFSTYAIHWVRQAPGKGLEWVAVISYDGMNKYYADSVRGRFTISRDNS
QNTVFLQMNTLRAEDTALYYCARSSSGYWSAFDIWGQGTVVTVSS

>VA014_49_heavy_chain

QVQLVESGGGVVQPGRSLRLSCAASGFTFRNYGIHWVRQAPGKGLEWVAFISFDGDDTYYADSVKGRFTISRDNS
KNTLYLQMNSLRTEDTAVYYCAKQIGSYGRPDYYFDYWGQGALVTISS

>VA014_50_heavy_chain

QVQLVQSGAEVKKPGASVKVSCKASGYTFTTYGISWVRQAPGQGLEWMGWISAYNGNINYEQKFQGRVTMTTDTS
TSTAYMELRSLRSDDTAVYYCARDLGWFGELSGGFDIWGQGTMVTVSS

>VA014_52_heavy_chain

EVQLVESGGGVVQPGRSLRLSCAASGFTFTTHAMHWVRQAPGKGLDWVAVISYDGLNKYYADSVRGRFTISRDNT
KTTVYLQMNSLRPDDTAVYYCARATVAYKWFDPWGQGTLVTVSS

>VA014_57_heavy_chain

QVQLQESGPGLVKTSETLSLTCTVSGGSVTSGNYYWSWIRQSPEKRLEWIGYIYYTGGTNYNPSFRSRVAISADT
SKNQFSLKLSSVSAADTAVYYCARGINLKDFSSGYNWFDPWGQGILVTVSS

>VA014_58_heavy_chain

EVQLVESGAEVKKPGASVKVSCKASGYTFTSYDINWVRQATGQGLEWMGWMDPSTGDTGFAQKFHGRVTLTRNTS
INTAYMELSSLRSEDTAVYYCTRGSSSWTYYFDSWGRGTLVTVSS

>VA014_61_heavy_chain

EVQLVESGAEVEKPGSSVKVSCKASGGTFISNTINWVRQAPGQGLECLGGIIPIVGIANYAQKFQGRVTITADKS
TSTVYMELSSLRSEDTAVYYCAREIYPPGGFDLWGQGTMVTVSS

>VA014_63_heavy_chain

QVQLVESGGGLVQPGGSLRLSCAASGFSFSDYYMDWVRQAPGKGLEWVGRTRNKAKSYTTGYAASVKGRFIISRD
DSKNSLYLQMNSLKTEDTAVYYCARAMAEARYSSGWYYYQYYMDVWGKGTTVTVSS

>VA014_65_heavy_chain

QVQLLESGGGLVKPGGSLRLSCVASGFTFSSYSMNWVRQAPGKGLEWVSSISSSSSYIYHADSVKGRFTISRDNA
KNSLYLQVNSLRAEDMAVYYCARGSSPTGDEAWDAFDIWGQGTMLTVSS

>VA014_67_heavy_chain

QVQLVQSGAEVKKPGESLKISCKASGYIFTSYWIAWVRQMPGKGLEWMGIIYPDDSDTRYSPSFQGQVTISADKS
ITTAYLQWSSLKASDTAMYYCARLETMGVIMPAASWGGTYGLDVWGQGTTVTVSS

>VA014_68_heavy_chain

QVQLVESGPGLVKPSETLSLTCSVSGVSVTNNYWSWIRQPPGKGLEWIGYIFYYGGTNYNPSLKSRVTLSIDTSA
NQLSLKLSSVSAADTAVYYCARHAADRSVTPWFDPWGQGTLVTVSS

>VA014R_33_heavy_chain

EVQLVESGGGLVQVGGSLRLSCTTSGFIVSRNYVMWVRQAPGQGLEWVSTIYPGGSTYFADSVKGRFTISRDDSK
NRLYLQMNSLRAEDTAFYYCARYIGNINWGQGTLVTVSS

>VA014R_37_heavy_chain

QVQLVESGGGLIQPGGSLRLSCAASGLIVSSNYMSWVRQAPGKGLEWVSLIYSGGSTFYADSVKGRFTISRDNSK
NTLYLQMNSLRADDTAVYYCARELDVVGATDSWGQGTLVTVSS

>VA014R_38_heavy_chain

QVQLVESGGGVVQPGGSLRLSCAASGFTFDDYAMHWVRQAPGKGLEWVSVISGDGGSTYSADSVKGRFTISRDNR
KHSLYLQMNSLRSEDTALYYCAKDMGYVPAAIDGAFDIWGQGTMVTVSS

>VA014R_39_heavy_chain

QVQLQQWGATLLKPSETLSLTCEVYGGSFSGYYWNWIRQAPGKGLGWIAEINHSGRFNYNPSLKSRVTISVDTAK
NQFSLKLTSLTAADTAVYYCARFGYGDYGWDFDQNMDVWGKGTTVTVSS

>VA014_01_light_chain

AIRMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQRPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYSTPTWTFGQGTKVDIK

>VA014_02_light_chain

DIQMTQSPSTLSASVGDRVTITCRASQSISSWLAWYQQKPGKAPKLLIYKASNLQSGVPSRFSGSGSGTEFTLTI
SSLQPADFATYYCQQYYDYPWTFGQGTKLEIK

>VA014_03_light_chain

EIVMTQSPGTLSLSPGERATLSCRASQSVSSNYLAWYQQRPGQAPRLLIYGASRRAAGIPDRFSGGGSGTDFTLT
ISRLDPEDYAVYYCQQYGNSPGTFGPGTKVDIK

>VA014_04_light_chain

AIRMTQSPSSLSASVGDRVTLTCRASQSIRNYLNWYQLKPGKAPKLLIYAASILHSGVPTRFSGSGSGTDFTLTI
SSLQPADFATYYCQQSYIMPPWTFGPGTKLEIK

>VA014_05_light_chain

EIVMTQSPATLSVSPGERATLSCRASQSVSSNLAWYQQKFGQAPRLLIYGASTRANGIPARFSGSGSGTEFTLTI
SSLQSEDFAVYYCQQYNNWPPLTFGQGTKVEIK

>VA014_07_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYGAVSRATGIPDRFSGSGSGTDFTLT
ITRLEPEDYAVYYCQQMGTFGPGTKVDIK

>VA014_09_light_chain

EIVLTQSPATLSSSPGERATLSCRASQSVSRYLSWYQQKPGQAPRLLIYDASNRATGIPPRFSGSGSGTDFTLTI
SSLEPEDFAVYYCHQRSNPLTFGGGTKVDIK

>VA014_10_light_chain

DIQMTQSPSSLSASVGDSVTITCQASQDIGNFLNWYQQKPGKAPKLLIYDASNLQTGVPSRFSGSGSGTEFTFTI
SSLQPEDIATFYCQQYDNVHPLTFGGGTKVEIK

>VA014_11_light_chain

EIVMTQSPGTLSLSPGERATLSCKASQSVTSNYLAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGGSPQYTFGQGTKLEIK

>VA014_14_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSLSSNYLAWFQQKHGQAPRLLIYAASNRAIGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSPWTFGQGTKLEIK

>VA014_15_light_chain

DVVMTQSPLSLPVTPGEPASISCRSSQSLLHSNGYNYLDWYLQKPGQSPQLLIYLGSNRASGVPDRFSGSGSGTD
FTLKISRVEAEDVGVYYCMQALQSPFTFGPGTKVEIK

>VA014_16_light_chain

DIQMTQSPSSLSASVGDRVTITCRASQGIRSDLGWYQQKPGKAPKRLIYAASSLQSGVPSRFSGSGSGTEFTLTI
SSLQPEDFATYYCLQYNTYPALTFGGGTKVEIK

>VA014_18_light_chain

EIVMTQSPATLSVSPGERATLSCRASQSVSSNLAWYQQKAGQAPRLLIYGASTRATGIPARFSGSGSGTEFTLTI
TSLQSEDFAVYYCQHYYNWPPWTFGQGTKVEIK

>VA014_19_light_chain

DIVMTQSPSSLSASVGDRVTISCRASQSISNYLSWFQQKPGKAPKLLIYAASRLRSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYGTLPTFGGGTKLEIK

>VA014_21_light_chain

DIQLTQSPDSQAVSLGERATINCKSSQSVLFSSNNKNYLAWYQQKPGQPPKLLINWASTRESGVPDRFSGSGSGT
DFTLTISGLQAEDVAVYYCQQYYRSPPTFGQGTKLEIK

>VA014_22_light_chain

DIVMTQSPSSVSASVGDRVTITCRASQGISSWLAWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQANSSPYGFTFGPGTKVDIK

>VA014_23_light_chain

EIVLTQSPGTLSLSPGERATLSCRASQSVSSNYVAWYQQKPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDSAVYYCQQYGSSPQTFGQGTKVEIK

>VA014_24_light_chain

DIVMTQSPSSLSASVGDRVTITCRASQRINTYLNWYQQKPGKAPKLLIYAASALQSGVPSRFSGSGSGTHSTLTI
SSLQPEDFGTYYCQQSYRTSFTFGPGTKVEIK

>VA014_25_light_chain

EIVMTQSPATLSVSPGERVTLSCRASQSVSSNLAWYQQKPGQAPRLLIFGASTRATGIPARFSGSGSGTEFTLTI
SSLQSEDFAVYYCQQYNDWPIKFGQGTRLEIK

>VA014_26_light_chain

DIQLTQSPSTLSASVGDRVTITCRASQSISTWLAWHQQIPGKAPKLLIYRASTLQSGVPSRFSGSGSGTEFTLTI
SSLQPDDFATYYCQQYYSVRTFGQGTKVDIK

>VA014_27_light_chain

EIVLTQSPSSLSASVGDRVTVTCRASQSIITYLNWYQHKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYNIPETFGQGTKVEIK

>VA014_30_light_chain

EIVLTQSPATLSVSPGERATLSCRASQSVSINLAWYQQKPGQAPRLLMYGASTRVTDIPARFSGSGSGTEFTLTI
SSLQSEDFALYYCQQYNNWPRTFGQGTKVEIK

>VA014_33_light_chain

NFMLTQPHSVSESPGKTVTISCTRSSGSIASNYVQWYQQRPGSAPTSVIYEDNQRPSGVPDRFSGSIDSSSNSAS
LIISGLKTEDEADYYCQSYDSRLRVVFGGGTKLTVL

>VA014_35_light_chain

EIVLTQSPATLSVSPGERTTLSCRASQSVSSDLAWYQQKPGQAPRLLIYAASTRAPGIPARFSGSGSGTEFTLTI
SSLQSEDFAVYHCQQNNNWPWTFGQGTKLEIK

>VA014_36_light_chain

DIQMTQSPSSLSASVGDRVTITCQASEDIRYFLNWYQQKPGKAPKLLIYDASNLKTGVPSRFSGSGSGTDFTFTI
SSLQPEDIATYYCQQYDDLPPTFGPGTKVEIK

>VA014_41_light_chain

EIVMTQSPSTLSASVGDRVTITWRASQSISSYLNWYQQKPGKAPNLLIYAASNLQSGVPSRFSGRGSGTDFSLTI
TSLQPEDFATYYCQQSYSTPPLTFGGGTKVDIK

>VA014_43_light_chain

DIQLTQSPSFLSASVGDRVTITCRASQGISNYLAWYQHKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTEFTLTI
DSLQPEDFATYYCQHLNSYPRLTFGGGTKVEIK

>VA014_44_light_chain

DIVMTQSPGTLSLSPGERATLSCRASQSVGSRYLVWYQQKSGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSLSFGQGTKLEIK

>VA014_47_light_chain

EIVMTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKPGQAPRLLIYSACSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGDSPRGSFGQGTKVDIK

>VA014_48_light_chain

EIVMTQSPGTLSLSPGERATLSCRASQSVSSTYLAWHQQRPGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQHYGPSPWTFGQGTKLEIK

>VA014_49_light_chain

AIRMTQSPSSLSASVGDRVTIACQASQDISNYLSWYQQKPGKAPKVLIYDASNLEAGVPSRFSGSGSGTDFTFTI
NSLQPEDIATYYCQQYDNFPPITFGQGTRLEIK

>VA014_50_light_chain

DVVMTQSPATLSLSPGERATLSCRASQSVSRFLAWYQQRPGQAPRLLIYDASNRATGIPARFSGSGSGTDFTLTI
SSLEPEDFAVYYCHQRSNPLTFGGGTKVDIK

>VA014_52_light_chain

DIQMTQSPSFLSASVGDRVTITCRASQGINNYLAWYQQKPGKAPKLLIYAASTLQSGVPSRFSGSGSGTEFTLTI
SSLQPEDFATYYCQQLNTYPLFTFGPGTKVDIK

>VA014_57_light_chain

EIVLTQSPATLSVSPGERATLSCRASQSVSSNLAWYQQTPGQAPRLLIYGASTRATGIPARFSGSGSGTEFTLTI
SSLQSEDFAVYYCQQYNNWPRSFGQGTKVDIK

>VA014_58_light_chain

AIRMTQSPATLSLSPGERATLSCRASESLGSYLAWYQQKPGQAPRLLIFDASNRATGIPARFSGSGSGTDFTLTI
SSLEPEDFAVYFCQQRGSWPLTFGPGTKVDIK

>VA014_61_light_chain

DIVMTQSPDSLAVSLGDRATINCKSSQSVLDTSNNRNYLAWYQHKPGQPPKLLIYWASTRESGVPDRFSGSGSGA
DFTLTISSLQAEDVAVYYCQQYYSTPMYSFGQGTKVDIK

>VA014_63_light_chain

DIVMTQSPGTLSLSPGERATLSCRASQSVSSSYLAWYQQKCGQAPRLLIYGASSRATGIPDRFSGSGSGTDFTLT
ISRLEPEDFAVYYCQQYGSSPYTFGQGTKVDIK

>VA014_65_light_chain

DIVMTQSPSSLSASVGDRVTITCRASQSINIYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSSSTPRTFGQGTKVDIK

>VA014_67_light_chain

AIRMTQSPSSLSASVGDRVTITCRASQNILTFLNWYQHKPGKAPELLIYAASSLQSGVPSRFSGSGSGTDFTLTI
SSLQPEDFATYYCQQSYNIPETFGQGTKVDIK

>VA014_68_light_chain

DIVMTQSPDSLAVSLGERATINCKSSQNVFLSSNNKNYLAWYQQKPGQPPKLLIYWASIRESGVPDRFSGGGSGT
DFTLTISSLQAEDVAVYFCQQYYSTPLTFGGGTKVEIK

>VA014R_33_light_chain

DIVMTQSPSSLSASVGDRVTITCQASQDISNYLNWYQQKSGKAPKLLIFDASNLQTGVPSRFSGSGSGTDFTFTI
SSLQPEDIATYYCHQYDNLPLTFGGGTKVEIK

>VA014R_37_light_chain

EIVLTQSPGTLSLSPGESATLSCRASQSVSSSYIAWYQQKPGQSPRLLIYGASNRATGIPDRFSGSGSGTDFTLA
INRLEPEDFAVYYCQRCDTSPMYTFGQGTKLEIK

>VA014R_38_light_chain

DIQMTQSPGTLSLSPGERATLSCRASHSVSANYLAWYQHKPGQAPRLLIYAASSRATGIPDRFSGSGSGTDFTLT
VSRLEPEDFAVYYCQHYGGSPLTFGGGTKLEIK

>VA014R_39_light_chain

DIQMTQSPSTLSASVGDRVTIACRASQSIGTWLAWHQQKPGKAPNLLIYKASTLETGVPSRFSGSGSGTEFTLTI
SSLQPDDFATYYCQQYNSGRTFGQGTKVEIK

## 7.5 Extended table 1

*Extended patient information.*

| Donor number | Severity score | Hyper-inflammatory score | Comorbidities | Age range | Sex | Final outcome |
|---|---|---|---|---|---|---|
| 1 | 4 | N | asthma, HTN, obesity, T2DM, sarcoidosis | 50-59 | M | Died |
| 2 | 4 | Y | obesity | 50-59 | M | Died |
| 3 | 4 | N | HTN | 40-49 | M | Died |
| 4 | 4 | Y | AF, HTN, obesity, sarcoidosis, | 60-69 | M | Discharged |
| 5 | 4 | N | HTN, T2DM | 60-69 | M | Discharged |
| 6 | 4 | Y | T1DM | 40-49 | M | Ward |
| 7 | 4 | Y | Hypothyroidism | 70-79 | M | Discharged |
| 8 | 4 | Y | T2DM, NAFLD | 50-59 | M | Died |
| 9 | 4 | Y | HTN, IHD | 70-79 | M | Died |
| 10 | 4 | Y | T2DM | 70-79 | M | Died |
| 11 | 4 | N | HTN, prostate ca | 70-79 | M | Discharged |
| 12 | 4 | Y | none | 40-49 | M | Discharged |
| 13 | 4 | Y | none | 50-59 | M | Ward |
| 14 | 4 | N | obesity, T2DM | 50-59 | F | Discharged |
| 15 | 4 | N | COPD | 70-79 | M | Discharged |
| 16 | 2 | N | none | 50-59 | M | Discharged |
| 17 | 1 | N | obesity | 30-39 | F | Discharged |
| 18 | 1 | N | CKD, dementia, HTN, HD, obesity, VTE | 70-79 | F | Discharged |
| 19 | 2 | N | AF | 90-99 | M | Discharged |
| 20 | 2 | N | dementia, T2DM | 60-69 | M | Discharged |
| 21 | 1 | N | hepatitis C, Huntingdon's disease | 60-69 | M | Discharged |
| 22 | 1 | N | AF, dementia, HTN, stroke | 80-89 | M | Discharged |
| 23 | 1 | N | CKD, T2DM, VTE | 60-69 | M | Died |
| 24 | 4 | N | chronic pancreatitis, T2DM | 60-69 | M | Ward |
| 25 | 1 | N | AF, Marfan's syndrome | 50-59 | M | Discharged |
| 26 | 0 | N | ESRF, obesity, T1DM | 20-29 | F | Discharged |
| 28 | 4 | N | none | 40-49 | M | Discharged |
| 29 | 0 | N | ESRF, T1DM, SPK | 40-49 | M | Discharged |
| 31 | 0 | N | ESRF, obesity | 20-29 | F | Discharged |
| 32 | 4 | N | HTN, T2DM | 60-69 | M | Discharged |
| 33 | 4 | N | HTN, obesity | 60-69 | F | Discharged |
| 34 | 4 | Y | asthma, obesity | 50-59 | F | Discharged |
| 35 | 4 | N | none | 40-49 | M | Discharged |
| 36 | 4 | Y | none | 30-39 | M | Discharged |
| 37 | 4 | N | asthma, inflammatory polyarthritis, obesity | 60-69 | M | Discharged |
| 38 | 4 | N | none | 30-39 | M | Discharged |
| 39 | 4 | N | T2DM, obesity | 30-39 | M | Discharged |
| 40 | 5 | N | obesity | 30-39 | F | Discharged |
| 41 | 5 | Y | none | 60-69 | M | Discharged |
| 42 | 5 | N | asthma, obesity | 40-49 | M | Transferred to local hospital |
| 43 | 5 | Y | HTN | 50-59 | M | Transferred to local hospital |
| 44 | 5 | N | none | 50-59 | M | Died |
| 45 | 5 | N | none | 50-59 | M | Discharged |
| 46 | 5 | Y | asthma, obesity, T2DM | 50-59 | M | Died |
| 47 | 2 | N | none | 20-29 | M | Discharged |
| 48 | 3 | N | depression, obesity, T1DM | 50-59 | M | Discharged |
| 49 | 1 | N | HTN, T2DM | 60-69 | M | Discharged |
| 50 | 2 | N | none | 60-69 | M | Discharged |
| 51 | 2 | N | none | 30-39 | M | Discharged |
| 52 | 0 | N | AF, ESRF, T2DM | 60-69 | M | Ward |
| 53 | 2 | N | HTN, obesity | 50-59 | M | Discharged |
| 54 | 0 | N | AML | 20-29 | F | Discharged |
| 55 | 3 | N | bladder ca | 60-69 | M | Died |
| 56 | 1 | N | asthma | 50-59 | M | Discharged |
| 57 | 0 | N | dementia, frailty, Parkinsonism | 90-99 | M | Died |
| 58 | 1 | N | CKD, frailty | 90-99 | F | Died |
| 59 | 0 | N | CCF, IHD, PVD, stroke, T2DM | 80-89 | M | Transferred to local hospital |
| 60 | 0 | N | mycosis fungoides | 50-59 | F | Ward |
| 61 | 1 | N | COPD, bronchiectasis, renal transplant | 70-79 | M | Discharged |

## 7.6 Extended table 2

*Summary table of the characteristics of mAbs isolated from SARS-CoV-2 infected donors.*

| mAb name | Specificty | Competition group | S EC50 | RBD EC50 | NTD EC50 | S2 ELISA | SARS-CoV-2 WT FL IC50 | SARS-CoV-2 WT PV IC50 | SARS-CoV PV IC50 | % ACE2 comp | VH | VL |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| P003_014 | RBD | 1 | 0.0076 | 0.0086 | n.d. | n.d. | >100 | >100 | >50 | 86.3 | IGHV3-13 | IGKV1-39 |
| P054_004 | RBD | 1 | 0.0360 | 0.0511 | n.d. | n.d. | 42.2300 | 0.2000 | 5.9585 | 95.2 | IGHV4-34 | IGLV2-14 |
| P054_027 | RBD | 1 | 0.0077 | 0.0127 | n.d. | n.d. | >100 | 0.7000 | >50 | 92.9 | IGHV4-30-4 | IGLV6-57 |
| P054_003 | RBD | 2 | 0.1756 | 0.2001 | n.d. | n.d. | 1.7900 | 0.0050 | >100 | 98.5 | IGHV1-2 | IGKV1-33 |
| P003_017 | RBD | 3 | 0.0057 | 0.0068 | n.d. | n.d. | 1.0800 | 0.1627 | >100 | 96.5 | IGHV3-66 | IGKV1-9 |
| P008_003 | RBD | 3 | 0.0756 | 0.0825 | n.d. | n.d. | 0.0170 | 0.0764 | >50 | 99.7 | IGHV4-34 | IGLV1-40 |
| P008_018 | RBD | 3 | 0.0714 | 0.0824 | n.d. | n.d. | 0.0550 | 0.0008 | >50 | 99.9 | IGHV3-66 | IGKV3-20 |
| P008_042 | RBD | 3 | 0.1277 | 0.1732 | n.d. | n.d. | 0.0240 | 0.0044 | >50 | 99.9 | IGHV3-66 | IGKV3-20 |
| P008_047 | RBD | 3 | 0.1731 | 0.2327 | n.d. | n.d. | 0.1893 | 0.0017 | >50 | 99.7 | IGHV3-53 | IGKV1-33 |
| P008_067 | RBD | 3 | 0.1638 | 0.1859 | >20 | n.d. | 0.1700 | 0.0004 | >50 | 99.1 | IGHV3-66 | IGKV1-17 |
| P008_081 | RBD | 3 | 0.0384 | 0.0389 | n.d. | n.d. | 0.1159 | 0.0219 | >50 | 99.5 | IGHV1-58 | IGKV3-20 |
| P008_083 | RBD | 3 | 0.0203 | 0.1193 | n.d. | n.d. | 0.0260 | 0.0029 | >50 | 99.9 | IGHV3-66 | IGKV1-33 |
| P008_090 | RBD | 3 | 0.0182 | 0.0259 | n.d. | n.d. | 0.0709 | 0.0042 | >50 | 99.9 | IGHV3-53 | IGKV1-33 |
| P008_108 | RBD | 3 | 0.0092 | 0.0138 | n.d. | n.d. | 0.0023 | 0.0031 | >50 | 100.0 | IGHV3-11 | IGKV1-33 |
| P054_022 | RBD | 3 | 0.1741 | 0.1944 | n.d. | n.d. | 2.4400 | 0.5941 | >50 | 98.5 | IGHV4-30-2 | IGLV2-8 |
| P008_015 | RBD | 4 | 0.0337 | 0.0362 | n.d. | n.d. | 0.0553 | 0.0001 | >50 | 80.2 | IGHV3-9 | IGLV3-21 |
| P008_038 | RBD | 4 | 0.0578 | 0.0926 | >10 | n.d. | 0.1100 | 0.0003 | >50 | 42.3 | IGHV1-69 | IGLV1-40 |
| P008_087 | RBD | 4 | 0.0212 | 0.0560 | n.d. | n.d. | 4.9600 | 0.0100 | 10.4600 | 48.5 | IGHV4-31 | IGKV1-33 |
| P008_096 | RBD | 4 | 0.0299 | 0.0549 | n.d. | n.d. | 9.0920 | 0.4106 | >50 | 50.6 | IGHV4-39 | IGKV1-33 |
| P008_039 | NTD | 5 | 0.1014 | >10 | 0.0240 | n.d. | 25.2800 | 0.0075 | >50 | 91.2 | IGHV3-21 | IGKV1-12 |
| P008_051 | NTD | 5 | 0.0914 | >10 | 0.0561 | n.d. | 48.8200 | 0.3000 | >50 | 74.5 | IGHV4-61 | IGKV3-11 |
| P008_052 | NTD | 5 | 0.0853 | >10 | 0.0625 | n.d. | 48.6500 | 0.1250 | 34.6400 | 80.2 | IGHV4-39 | IGLV2-11 |
| P003_027 | NTD | 6 | 0.0039 | >10 | 0.0065 | n.d. | 1.6400 | 12.9150 | >50 | 47.4 | IGHV3-48 | IGKV1-39 |
| P008_007 | NTD | 6 | 0.0700 | >10 | 0.1710 | n.d. | 0.0549 | 0.0006 | >100 | 66.3 | IGHV3-48 | IGLV3-21 |
| P008_056 | NTD | 6 | 0.0602 | >10 | 0.0488 | n.d. | 0.0136 | >100 | >50 | 36.8 | IGHV3-21 | IGKV1-33 |
| P054_021 | NTD | 6 | 0.0263 | >10 | 0.0643 | n.d. | 0.0421 | 0.0004 | >100 | 51.0 | IGHV1-24 | IGKV2-24 |
| P008_060 | NON-S1 | 7 | 0.0311 | >10 | >10 | n.d. | 32.9000 | 0.0950 | 0.6434 | 50.7 | IGHV3-30 | IGKV3-20 |
| P008_057 | RBD | n.d. | 0.0324 | 0.0597 | >20 | n.d. | 0.0154 | 0.0063 | >100 | 99.8 | IGHV3-53 | IGKV1-9 |
| P008_070 | RBD | n.d. | 0.0342 | 0.1075 | n.d. | n.d. | 34.9500 | 4.0147 | >100 | 47.3 | IGHV3-30 | IGKV1-33 |
| P008_076 | RBD | n.d. | 0.0354 | 0.0384 | n.d. | n.d. | 0.8466 | 0.0889 | 0.4648 | 99.0 | IGHV3-13 | IGLV2-14 |
| P008_086 | RBD | n.d. | 0.0338 | 0.0326 | >20 | n.d. | 0.0052 | 0.0012 | >100 | 99.3 | IGHV1-58 | IGKV3-20 |
| P008_100 | RBD | n.d. | 0.0327 | 0.0364 | >10 | n.d. | 11.3000 | 5.9766 | >100 | 38.4 | IGHV3-30 | IGKV3-20 |
| P008_103 | RBD | n.d. | 0.0663 | 0.1014 | n.d. | n.d. | 0.0215 | 0.0055 | >100 | 99.8 | IGHV3-66 | IGKV1-33 |
| P054_036 | RBD | n.d. | 0.0130 | 0.0293 | >10 | n.d. | 0.0549 | 0.0100 | >100 | 99.5 | IGHV3-53 | IGKV1-39 |
| P054_050 | RBD | n.d. | 0.0268 | 0.0473 | n.d. | n.d. | >100 | 66.6033 | >100 | 60.1 | IGHV3-13 | IGKV1-39 |
| P003_016 | NTD | n.d. | 0.1304 | >20 | 2.0000 | n.d. | 0.1530 | 0.4494 | >100 | 25.9 | IGHV3-74 | IGKV1-39 |
| P003_055 | NTD | n.d. | 0.0600 | >10 | 0.0546 | n.d. | >100 | >100 | >100 | n.d. | IGHV4-59 | IGKV4-1 |
| P008_001 | NTD | n.d. | 0.8157 | >20 | 1.2000 | n.d. | >100 | >100 | >100 | 37.3 | IGHV3-49 | IGKV1-9 |
| P008_002 | NTD | n.d. | 0.0084 | >20 | 0.0146 | n.d. | >100 | >100 | >100 | 24.2 | IGHV3-9 | IGLV3-21 |
| P008_014 | NTD | n.d. | 0.0662 | >20 | 0.0215 | n.d. | 1.8120 | >100 | >100 | 43.0 | IGHV3-15 | IGKV1-33 |
| P008_099 | NTD | n.d. | 0.0222 | >20 | 0.1500 | n.d. | >100 | >100 | >100 | 46.8 | IGHV3-23 | IGKV1-39 |
| P054_044 | NTD | n.d. | 0.0308 | >20 | 4.0000 | n.d. | 0.9428 | 0.7403 | >100 | 46.7 | IGHV1-69 | IGKV1-5 |
| P008_004 | NON-S1 | n.d. | 0.0450 | >10 | >10 | + | >100 | >100 | >100 | 61.4 | IGHV3-7 | IGKV1-5 |
| P008_005 | NON-S1 | n.d. | 0.2000 | >10 | >10 | + | >100 | >100 | >100 | 22.8 | IGHV3-30 | IGKV1-5 |
| P008_006 | NON-S1 | n.d. | 0.1000 | >10 | >10 | + | >100 | >100 | >100 | 16.0 | IGHV3-30 | IGKV1-5 |
| P008_016 | NON-S1 | n.d. | 0.1075 | >10 | >10 | + | >100 | >100 | >100 | 39.9 | IGHV4-59 | IGKV3-20 |
| P008_023 | NON-S1 | n.d. | 0.4000 | >10 | >10 | + | >100 | >100 | >100 | 22.7 | IGHV3-11 | IGKV3-15 |
| P008_032 | NON-S1 | n.d. | 0.0600 | >10 | >10 | + | >100 | >100 | >100 | 23.3 | IGHV3-30 | IGKV3-11 |

# References


1. Pinho, S. & Frenette, P. S. Haematopoietic stem cell activity and interactions with the niche. *Nat. Rev. Mol. Cell Biol.* **20**, 303–320 (2019).

2. Siminovitch, L., Mcculloch, E. A. & Till, J. E. The distribution of colony-forming cells among spleen colonies. *J. Cell. Physiol.* **62**, 327–36 (1963).

3. Andrews, R. G., Singer, J. W. & Bernstein, I. D. Precursors of colony-forming cells in humans can be distinguished from colony-forming cells by expression of the CD33 and CD34 antigens and light scatter properties. *J. Exp. Med.* **169**, 1721–31 (1989).

4. Yang, L. *et al.* Identification of Lin-Sca1+kit+CD34 +Flt3- short-term hematopoietic stem cells capable of rapidly reconstituting and rescuing myeloablated transplant recipients. *Blood* **105**, 2717–2723 (2005).

5. Zhu, J. & Emerson, S. G. Hematopoietic cytokines, transcription factors and lineage commitment. *Oncogene* **21**, 3295–3313 (2002).

6. Seita, J. & Weissman, I. L. Hematopoietic stem cell: Self-renewal versus differentiation. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2**, 640–653 (2010).

7. Cheng, H., Zheng, Z. & Cheng, T. New paradigms on hematopoietic stem cell differentiation. *Protein Cell* **11**, 34–44 (2020).

8. Marshall, J. S., Warrington, R., Watson, W. & Kim, H. L. An introduction to immunology and immunopathology. *Allergy, Asthma Clin. Immunol.* **14**, 1–10 (2018).

9. von Behring, E. & Kitasato, S. Ueber das zustandekommen der diphtherie-immunität und der tetanus-immunität bei thieren. *Dtsch. Medizinsche Wochenschrift* **16**, 1113–1114 (1890).

10. Coombs, R. R., Feinstein, A. & Wilson, A. B. Immunoglobulin determinants on the surface of human lymphocytes. *Lancet* **2**, 1157–1160 (1969).

11. Fagraeus, A. The plasma cellular reaction and its relation to the formation of antibodies in vitro. *J. Immunol.* **58**, 1–13 (1948).

12. Porter, R. R. The hydrolysis of rabbit γ-globulin and antibodies with crystalline papain. *Biochem. J.* **73**, 119–126 (1959).

13. Arnold, J. N., Wormald, M. R., Sim, R. B., Rudd, P. M. & Dwek, R. A. The impact of glycosylation on the biological function and structure of human immunoglobulins. *Annu. Rev. Immunol.* **25**, 21–50 (2007).

14. Valentine, R. C. & Green, N. M. Electron microscopy of an antibody-hapten complex. *J. Mol. Biol.* **27**, 615–617 (1967).

15. Fleischman, J. B., Porter, R. R. & Press, E. M. The arrangement of the peptide chains in γ-globulin. *Biochem. J.* **88**, 220–228 (1963).

16. Molé, C. M., Béné, M. C., Montagne, P. M., Seilles, E. & Faure, G. C. Light chains of immunoglobulins in human secretions. *Clin. Chim. Acta* **224**, 191–197 (1994).

17. Spira, G., Yuan, R., Paizi, M. & Weisendal, M. Simultaneous expression of kappa and lambda light chains in a murine IgG3 anti-Cryptococcus neoformans hybridoma cell line. *Hybridoma* **13**, 531–535 (1994).

18. Edelman, G. M. *et al.* The covalent structure of an entire γG immunoglobulin molecule. *Proc. Natl. Acad. Sci. U. S. A.* **63**, 78–85 (1969).

19. Yguerabide, J., Epstein, H. F. & Stryer, L. Segmental flexibility in an antibody molecule. *J. Mol. Biol.* **51**, 573–590 (1970).

20. Elhanati, Y. *et al.* Inferring processes underlying B-cell repertoire diversity. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, (2015).

21. Lefranc, M.-P. Nomenclature of the Human Immunoglobulin Heavy (IGH) Genes. *Exp. Clin. Immunogenet.* **18**, 100–116 (2001).

22. Lefranc, M.-P. Nomenclature of the Human Immunoglobulin Kappa (IGK) Genes. *Exp. Clin. Immunogenet.* **18**, 161–174 (2001).

23. Lefranc, M.-P. Nomenclature of the Human Immunoglobulin Lambda (IGL) Genes. *Exp. Clin. Immunogenet.* **18**, 242–254 (2001).

24. Boyd, S. D. *et al.* Individual Variation in the Germline Ig Gene Repertoire Inferred from Variable Region Gene Rearrangements. *J. Immunol.* **184**, 6986–6992 (2010).

25. Schatz, D. G. & Ji, Y. Recombination centres and the orchestration of V(D)J recombination. *Nat. Rev. Immunol.* **11**, 251–263 (2011).

26. Ramsden, D. A., Baetz, K. & Wu, G. E. Conservation of sequence in recombination signal sequence spacers. *Nucleic Acids Res.* **22**, 1785–1796 (1994).

27. Van Gent, D. C., Ramsden, D. A. & Gellert, M. The RAG1 and RAG2 proteins establish the 12/23 rule in V(D)J recombination. *Cell* **85**, 107–113 (1996).

28. Jung, D. & Alt, F. W. Unraveling V(D)J Recombination: Insights into Gene Regulation. *Cell* **116**, 299–311 (2004).

29. Van Gent, D. C., Mizuuchi, K. & Gellert, M. Similarities between initiation of V(D)J recombination and retroviral integration. *Science (80-. ).* **271**, 1592–1594 (1996).

30. West, R. B., Yaneva, M. & Lieber, M. R. Productive and Nonproductive Complexes of Ku and DNA-Dependent Protein Kinase at DNA Termini. *Mol. Cell. Biol.* **18**, 5908–5920 (1998).

31. Ma, Y., Pannicke, U., Schwarz, K. & Lieber, M. R. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell* **108**, 781–794 (2002).

32. Motea, E. A. & Berdis, A. J. Terminal deoxynucleotidyl transferase: The story of a misguided DNA polymerase. *Biochim. Biophys. Acta - Proteins Proteomics* **1804**, 1151–1166 (2010).

33. Roth, D. B. V(D)J recombination: Mechanism, errors, and fidelity. *Mob. DNA III* **2**, 313–324 (2015).

34. Feeney, A. J., Tang, A. & Ogwaro, K. M. B-cell repertoire formation: Role of the recombination signal sequence in non-random V segment utilization. *Immunol. Rev.* **175**, 59–69 (2000).

35. Derdeyn, C. A., Moore, P. L. & Morris, L. Development of broadly neutralizing antibodies from autologous neutralizing antibody responses in HIV infection. *Curr. Opin. HIV AIDS* **9**, 210–216 (2014).

36. Mu, Z., Haynes, B. F. & Cain, D. W. Strategies for eliciting multiple lineages of broadly neutralizing antibodies to HIV by vaccination. *Curr. Opin. Virol.* **51**, 172–178 (2021).

37. Shi, B. *et al.* Comparative analysis of human and mouse immunoglobulin variable heavy regions from IMGT/LIGM-DB with IMGT/HighV-QUEST. *Theor. Biol. Med. Model.* **11**, 1–11 (2014).

38. Collins, A. M. & Watson, C. T. Immunoglobulin light chain gene rearrangements, receptor editing and the development of a self-tolerant antibody repertoire. *Front. Immunol.* **9**, 1–12 (2018).

39. Chiu, M. L., Goulet, D. R., Teplyakov, A. & Gilliland, G. L. Antibody Structure and Function: The Basis for Engineering Therapeutics. *Antibodies* **8**, 55 (2019).

40. Wu, T. T. & Kabat, E. A. An analysis of the sequences of the variable regions of bence jones proteins and myeloma light chains and their implications for antibody complementarity. *J. Exp. Med.* **132**, 211–250 (1970).

41. Henry, K. A. *et al.* Role of the non-hypervariable FR3 D-E loop in single-domain antibody recognition of haptens and carbohydrates. *J. Mol. Recognit.* **32**, 1–11 (2019).

42. Liu, Q. *et al.* Improvement of antibody functionality by structure-guided paratope engraftment. *Nat. Commun.* **10**, 1–13 (2019).

43. Larimore, K., McCormick, M. W., Robins, H. S. & Greenberg, P. D. Shaping of Human Germline IgH Repertoires Revealed by Deep Sequencing. *J. Immunol.* **189**, 3221–3230 (2012).

44. Xu, J. L. & Davis, M. M. Diversity in the CDR3 region of VH is sufficient for most antibody specificities. *Immunity* **13**, 37–45 (2000).

45. Persson, H. *et al.* CDR-H3 diversity is not required for antigen recognition by synthetic antibodies. *J. Mol. Biol.* **425**, 803–811 (2013).

46. Simonich, C. A. *et al.* HIV-1 Neutralizing Antibodies with Limited Hypermutation from an Infant. *Cell* **166**, 77–87 (2016).

47. Odegard, V. H. & Schatz, D. G. Targeting of somatic hypermutation. *Nat. Rev. Immunol.* **6**, 573–583 (2006).

48. Bransteitter, R., Pham, P., Scharfft, M. D. & Goodman, M. F. Activation-induced cytidine deaminase deaminates deoxycytidine on single-stranded DNA but requires the action of RNase. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 4102–4107 (2003).

49. Rogozin, I. B. & Diaz, M. Cutting Edge: DGYW/WRCH Is a Better Predictor of Mutability at G:C Bases in Ig Hypermutation Than the Widely Accepted RGYW/WRCY Motif and Probably Reflects a Two-Step Activation-Induced Cytidine Deaminase-Triggered Process. *J. Immunol.* **172**, 3382–3384 (2004).

50. Pilzecker, B. & Jacobs, H. Mutating for good: DNA damage responses during somatic hypermutation. *Front. Immunol.* **10**, 1–13 (2019).

51. Robertson, A. B., Klungland, A., Rognes, T. & Leiros, I. Base excision repair: The long and short of it. *Cell. Mol. Life Sci.* **66**, 981–993 (2009).

52. Jiricny, J. The multifaceted mismatch-repair system. *Nat. Rev. Mol. Cell Biol.* **7**, 335–346 (2006).

53. Haracska, L. *et al.* Roles of yeast DNA polymerases δ and ζ of Rev 1 in the bypass of abasic sites. *Genes Dev.* **15**, 945–954 (2001).

54. Hwang, J. K., Alt, F. W. & Yeap, L.-S. Related Mechanisms of Antibody Somatic Hypermutation and Class Switch Recombination. *Microbiol. Spectr.* **3**, 1–35 (2015).

55. Kepler, T. B. *et al.* Immunoglobulin gene insertions and deletions in the affinity maturation of HIV-1 broadly reactive neutralizing antibodies. *Cell Host Microbe* **16**, 304–313 (2014).

56. Doores, K. J. *et al.* Two Classes of Broadly Neutralizing Antibodies within a Single Lineage Directed to the High-Mannose Patch of HIV Envelope. *J. Virol.* **89**, 1105–1118 (2015).

57. Di Noia, J. M. & Neuberger, M. S. Molecular mechanisms of antibody somatic hypermutation. *Annu. Rev. Biochem.* **76**, 1–22 (2007).

58. Jolly, C. J. *et al.* The targeting of somatic hypermutation. *Semin. Immunol.* **8**, 159–168 (1996).

59. Liu, M. *et al.* Two levels of protection for the B cell genome during somatic hypermutation. *Nature* **451**, 841–845 (2008).

60. Stavnezer, J. & Schrader, C. E. Ig heavy chain class switch recombination: mechanism and regulation. *J. Immunol.* **193**, 5370–5378 (2014).

61. Schroeder, H. W. & Cavacini, L. Structure and function of immunoglobulins. *J. Allergy Clin. Immunol.* **125**, S41–S52 (2010).

62. Xu, Z., Zan, H., Pone, E. J., Mai, T. & Casali, P. Immunoglobulin class-switch DNA recombination: Induction, targeting and beyond. *Nat. Rev. Immunol.* **12**, 517–531 (2012).

63. Matthews, A. J., Zheng, S., DiMenna, L. J. & Chaudhuri, J. Regulation of Immunoglobulin Class-Switch Recombination. *Adv. Immunol.* **122**, 1–57 (2014).

64. Marchalot, A. *et al.* Uncoupling Splicing From Transcription Using Antisense Oligonucleotides Reveals a Dual Role for I Exon Donor Splice Sites in Antibody Class Switching. *Front. Immunol.* **11**, 1–13 (2020).

65. Yu, K., Chedin, F., Hsieh, C. L., Wilson, T. E. & Lieber, M. R. R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* **4**, 442–451 (2003).

66. Kitaura, K. *et al.* Different somatic hypermutation levels among antibody subclasses disclosed by a new next-generation sequencing-based antibody repertoire analysis. *Front. Immunol.* **8**, 1–11 (2017).

67. Lu, L. L., Suscovich, T. J., Fortune, S. M. & Alter, G. Beyond binding: Antibody effector functions in infectious diseases. *Nat. Rev. Immunol.* **18**, 46–61 (2018).

68. Grönwall, C. & Silverman, G. J. Natural IgM: Beneficial autoantibodies for the control of inflammatory and autoimmune disease. *J. Clin. Immunol.* **34**, 12–21 (2014).

69. Díaz-Zaragoza, M., Hernández-Ávila, R., Viedma-Rodríguez, R., Arenas-Aranda, D. & Ostoa-Saloma, P. Natural and adaptive IgM antibodies in the recognition of tumor-associated antigens of breast cancer (Review). *Oncol. Rep.* **34**, 1106–1114 (2015).

70. Chen, K. *et al.* Immunoglobulin D enhances immune surveillance by activating antimicrobial, proinflammatory and B cell-stimulating programs in basophils. *Nat. Immunol.* **10**, 889–898 (2009).

71. Liu, H. & May, K. Disulfide bond structures of IgG molecules: Structural variations, chemical modifications and possible impacts to stability and biological function. *MAbs* **4**, 17–23 (2012).

72. Chan, K. R., Ong, E. Z., Mok, D. Z. & Ooi, E. E. Fc receptors and their influence on efficacy of therapeutic antibodies for treatment of viral diseases. *Expert Rev. Anti. Infect. Ther.* **13**, 1351–1360 (2015).

73. Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: From structure to effector functions. *Front. Immunol.* **5**, 1–17 (2014).

74. Irani, V. *et al.* Molecular properties of human IgG subclasses and their implications for

designing therapeutic monoclonal antibodies against infectious diseases. *Mol. Immunol.* **67**, 171–182 (2015).

75. Barrett, D. J. & Ayoub, E. M. IgG2 subclass restriction of antibody to pneumococcal polysaccharides. *Clin. Exp. Immunol.* **63**, 127–34 (1986).

76. Jefferis, R. Isotype and glycoform selection for antibody therapeutics. *Arch. Biochem. Biophys.* **526**, 159–166 (2012).

77. van de Veen, W. & Akdis, M. Role of IgG4 in IgE-mediated allergic responses. *J. Allergy Clin. Immunol.* **138**, 1434–1435 (2016).

78. Woof, J. M. & Kerr, M. A. IgA function - Variations on a theme. *Immunology* **113**, 175–177 (2004).

79. Roos, A. *et al.* Human IgA Activates the Complement System Via the Mannan-Binding Lectin Pathway. *J. Immunol.* **167**, 2861–2868 (2001).

80. Bakema, J. E. & Van Egmond, M. Immunoglobulin A: A next generation of therapeutic antibodies? *MAbs* **3**, 352–361 (2011).

81. Gould, H. J. *et al.* The biology of IgE and the basis of allergic disease. *Annu. Rev. Immunol.* **21**, 579–628 (2003).

82. Gould, H. J. & Sutton, B. J. IgE in allergy and asthma today. *Nat. Rev. Immunol.* **8**, 205–217 (2008).

83. Clark, M. R., Mandal, M., Ochiai, K. & Singh, H. Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat. Rev. Immunol.* **14**, 69–80 (2014).

84. Kwang, D., Tjin, G. & Purton, L. E. Regulation of murine B lymphopoiesis by stromal cells. *Immunol. Rev.* **302**, 47–67 (2021).

85. Alt, F. W. *et al.* Ordered rearrangement of immunoglobulin heavy chain variable region segments. *EMBO J.* **3**, 1209–1219 (1984).

86. Stamatopoulos, K., Kosmas, C., Stavroyianni, N., Belessi, C. & Papadaki, T. Selection of immunoglobulin diversity gene reading frames in B cell lymphoproliferative disorders. *Leukemia* **13**, 601–604 (1999).

87. Vettermann, C. & Schlissel, M. S. Allelic exclusion of immunoglobulin genes: Models and mechanisms. *Immunol. Rev.* **237**, 22–42 (2010).

88. Nemazee, D. Receptor editing in lymphocyte development and central tolerance. *Nat. Rev. Immunol.* **6**, 728–740 (2006).

89. Pillai, S., Mattoo, H. & Cariappa, A. B cells and autoimmunity. *Curr. Opin. Immunol.* **23**, 721–731 (2011).

90. Phan, T. G., Gray, E. E. & Cyster, J. G. The microanatomy of B cell activation. *Curr. Opin. Immunol.* **21**, 258–265 (2009).

91. Okada, T. & Cyster, J. G. B cell migration and interactions in the early phase of antibody responses. *Curr. Opin. Immunol.* **18**, 278–285 (2006).

92. Akkaya, M., Kwak, K. & Pierce, S. K. B cell memory: building two walls of protection against pathogens. *Nat. Rev. Immunol.* **20**, 229–238 (2020).

93. Parker, D. C. T cell-dependent B cell activation. *Annu. Rev. Immunol.* **11**, 331–360 (1993).

94. Bortnick, A. & Allman, D. What Is and What Should Always Have Been: Long-Lived Plasma Cells Induced by T Cell–Independent Antigens. *J. Immunol.* **190**, 5913–5918 (2013).

95. Taylor, J. J., Pape, K. A. & Jenkins, M. K. A germinal center-independent pathway generates unswitched memory B cells early in the primary response. *J. Exp. Med.* **209**, 597–606 (2012).

96. Victora, G. D. & Nussenzweig, M. C. Germinal centers. *Annu. Rev. Immunol.* **30**, 429–457 (2012).

97. Liu, Y. J. *et al.* Within germinal centers, isotype switching of immunoglobulin genes occurs after the onset of somatic mutation. *Immunity* **4**, 241–250 (1996).

98. Manz, R. A., Löhning, M., Cassese, G., Thiel, A. & Radbruch, A. Survival of long-lived plasma cells is independent of antigen. *Int. Immunol.* **10**, 1703–1711 (1998).

99. Krueger, C. C., Thoms, F., Keller, E., Vogel, M. & Bachmann, M. F. Virus-Specific Secondary Plasma Cells Produce Elevated Levels of High-Avidity Antibodies but Are Functionally Short Lived. *Front. Immunol.* **10**, 1831–1843 (2019).

100. Nicolai, O. *et al.* Antibody production in murine Polymicrobial Sepsis—Kinetics and Key Players. *Front. Immunol.* **11**, 828–839 (2020).

101. Fehr, T., Skrastina, D., Pumpens, P. & Zinkernagel, R. M. T cell-independent type I antibody response against B cell epitopes expressed repetitively on recombinant virus particles. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 9477–9481 (1998).

102. Vos, Q., Lees, A., Wu, Z. Q., Snapper, C. M. & Mond, J. J. B-cell activation by T-cell-independent type 2 antigens as an integral part of the humoral immune response to pathogenic microorganisms. *Immunol. Rev.* **176**, 154–170 (2000).

103. Obukhanych, T. V. & Nussenzweig, M. C. T-independent type II immune responses generate memory B cells. *J. Exp. Med.* **203**, 305–310 (2006).

104. Toellner, K. M. *et al.* Low-level hypermutation in T cell-independent germinal centers compared with high mutation rates associated with T cell-dependent germinal centers. *J. Exp. Med.* **195**, 383–389 (2002).

105. Guo, W. *et al.* Somatic hypermutation as a generator of antinuclear antibodies in a murine model of systemic autoimmunity. *J. Exp. Med.* **207**, 2225–2237 (2010).

106. Packham, G. *et al.* The outcome of B-cell receptor signaling in chronic lymphocytic leukemia: Proliferation or anergy. *Haematologica* **99**, 1138–1148 (2014).

107. Eldershaw, S. A., Sansom, D. M. & Narendran, P. Expression and function of the autoimmune regulator (Aire) gene in non-thymic tissue. *Clin. Exp. Immunol.* **163**, 296–308 (2011).

108. Sakaguchi, S., Yamaguchi, T., Nomura, T. & Ono, M. Regulatory T Cells and Immune Tolerance. *Cell* **133**, 775–787 (2008).

109. Tussiwand, R., Bosco, N., Ceredig, R. & Rolink, A. G. Tolerance checkpoints in B-cell development: Johnny B good. *Eur. J. Immunol.* **39**, 2317–2324 (2009).

110. Iikuni, N., Lourenço, E. V., Hahn, B. H. & La Cava, A. Cutting Edge: Regulatory T Cells Directly Suppress B Cells in Systemic Lupus Erythematosus. *J. Immunol.* **183**, 1518–1522 (2009).

111. Crampton, S. P., Voynova, E. & Bolland, S. Innate pathways to B-cell activation and tolerance. *Ann. N. Y. Acad. Sci.* **1183**, 58–68 (2010).

112. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **124**, 783–801 (2006).

113. Berkeredjian-Ding, I. B. *et al.* Plasmacytoid dendritic cells control TLR7 sensitivity of naive B cells via type I IFN. *J. Immunol.* **174**, 5884b – 5884 (2005).

114. Diebold, S. S., Kaisho, T., Hemmi, H., Akira, S. & Reis E Sousa, C. Innate Antiviral Responses by Means of TLR7-Mediated Recognition of Single-Stranded RNA. *Science (80-. ).* **303**, 1529–1531 (2004).

115. Douagi, I. *et al.* Human B Cell Responses to TLR Ligands Are Differentially Modulated by Myeloid and Plasmacytoid Dendritic Cells. *J. Immunol.* **182**, 1991–2001 (2009).

116. Kilmon, M. A., Rutan, J. A., Clarke, S. H. & Vilen, B. J. Cutting Edge: Low-Affinity, Smith Antigen-Specific B Cells Are Tolerized by Dendritic Cells and Macrophages. *J. Immunol.* **175**, 37–41 (2005).

117. Kilmon, M. A. *et al.* Macrophages prevent the differentiation of autoreactive B cells by secreting CD40 ligand and interleukin-6. *Blood* **110**, 1595–1602 (2007).

118. Lefranc, M. P. *et al.* IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Dev. Comp. Immunol.* **27**, 55–77 (2003).

119. Gottlieb, M. S. *et al.* Pneumocystis carinii Pneumonia and Mucosal Candidiasis in Previously Healthy Homosexual Men. *N. Engl. J. Med.* **305**, 1425–1431 (1981).

120. Sharp, P. M. & Hahn, B. H. Origins of HIV and the AIDS Pandemic. *Cold Spring Harb. Perspect. Med.* **1**, 6841–6856 (2011).

121. Greene, W. C. A history of AIDS: Looking back to see ahead. *Eur. J. Immunol.* **37**, 94–102 (2007).

122. Levy, J. A. *et al.* Isolation of Lymphocytopathic Retroviruses from San Francisco Patients with AIDS. *Science (80-. ).* **225**, 840–842 (1984).

123. Gallo, R. C. *et al.* Frequent Detection and Isolation of Cytopathic Retroviruses (HTLV-III) from Patients with AIDS and at Risk for AIDS. *Science (80-. ).* **224**, 500–503 (1984).

124. Barré-Sinoussi, F. *et al.* Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science (80-. ).* **220**, 868–871 (1983).

125. Coffin, J. *et al.* What to call the AIDS virus? *Nature* **321**, 10–21 (1986).

126. Steinbrook, R. One Step Forward, Two Steps Back — Will There Ever Be an AIDS Vaccine? *N. Engl. J. Med.* **357**, 2653–2655 (2007).

127. Peate, I. World AIDS Day 2021: a comparison of HIV and COVID-19. *Br. J. Healthc. Assist.* **15**, 506–509 (2021).

128. Peeters, M. *et al.* Risk to human health from a plethora of Simian immunodeficiency viruses in primate bushmeat. *Emerg. Infect. Dis.* **8**, 451–457 (2002).

129. Klatt, N. R., Silvestri, G. & Hirsch, V. Nonpathogenic Simian Immunodeficiency Virus Infections. *Cold Spring Harb. Perspect. Med.* **2**, 7153–7164 (2012).

130. Keele, B. F. *et al.* Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. *Nature* **460**, 515–519 (2009).

131. Hemelaar, J. *et al.* Global and regional molecular epidemiology of HIV-1, 1990–2015: a

systematic review, global survey, and trend analysis. *Lancet Infect. Dis.* **19**, 143–155 (2019).

132. Mourez, T., Simon, F. & Plantiera, J. C. Non-M variants of human immunodeficiency virus type. *Clin. Microbiol. Rev.* **26**, 448–461 (2013).

133. D'Arc, M. *et al.* Origin of the HIV-1 group O epidemic in western lowland gorillas. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 1343–1352 (2015).

134. Gottlieb, G. S., Raugi, D. N. & Smith, R. A. 90-90-90 for HIV-2? Ending the HIV-2 epidemic by enhancing care and clinical management of patients infected with HIV-2. *Lancet HIV* **5**, 390–399 (2018).

135. Seitz, R. Human Immunodeficiency Virus (HIV). *Transfus. Med. Hemotherapy* **43**, 203–222 (2016).

136. Bolinger, C. & Boris-Lawrie, K. Mechanisms employed by retroviruses to exploit host factors for translational control of a complicated proteome. *Retrovirology* **6**, 1–20 (2009).

137. Watts, J. M. *et al.* Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**, 711–716 (2009).

138. Zhu, P. *et al.* Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* **441**, 847–852 (2006).

139. Jolly, C., Kashefi, K., Hollinshead, M. & Sattentau, Q. J. HIV-1 Cell to Cell Transfer across an Env-induced, Actin-dependent Synapse. *J. Exp. Med.* **199**, 283–293 (2004).

140. Kandathil, A. J., Sugawara, S. & Balagopal, A. Are T cells the only HIV-1 reservoir? *Retrovirology* **13**, 86 (2016).

141. Berger, E. A. *et al.* A new classification for HIV-1. *Nature* **391**, 240–240 (1998).

142. Lobritz, M. A., Ratcliff, A. N. & Arts, E. J. HIV-1 entry, inhibitors, and resistance. *Viruses* **2**, 1069–1105 (2010).

143. Campbell, E. M. & Hope, T. J. HIV-1 capsid: The multifaceted key player in HIV-1 infection. *Nat. Rev. Microbiol.* **13**, 471–483 (2015).

144. Hu, W. S. & Hughes, S. H. HIV-1 reverse transcription. *Cold Spring Harb. Perspect. Med.* **2**, 1–22 (2012).

145. Abram, M. E., Ferris, A. L., Shao, W., Alvord, W. G. & Hughes, S. H. Nature, Position, and Frequency of Mutations Made in a Single Cycle of HIV-1 Replication. *J. Virol.* **84**, 9864–9878 (2010).

146. Hu, W. S. & Temin, H. M. Genetic consequences of packaging two RNA genomes in one retroviral particle: Pseudodiploidy and high rate of genetic recombination. *Proc. Natl. Acad. Sci. U. S. A.* **87**, 1556–1560 (1990).

147. Lusic, M. & Siliciano, R. F. Nuclear landscape of HIV-1 infection and integration. *Nat. Rev. Microbiol.* **15**, 69–82 (2017).

148. Craigie, R. & Bushman, F. D. HIV DNA integration. *Cold Spring Harb. Perspect. Med.* **2**, 1–18 (2012).

149. Roebuck, K. A. & Saifuddin, M. Regulation o f HIV-1 Transcription. *Gene Expr.* **8**, 67–84 (1999).

150. Hidalgo, L. & Swanson, C. M. Regulation of human immunodeficiency virus type 1 (HIV-1) mRNA translation. *Biochem. Soc. Trans.* **45**, 353–364 (2017).

151. Fischer, U., Huber, J., Boelens, W. C., Mattajt, L. W. & Lührmann, R. The HIV-1 Rev Activation Domain is a nuclear export signal that accesses an export pathway used by specific cellular RNAs. *Cell* **82**, 475–483 (1995).

152. Karn, J. & Stoltzfus, C. M. Transcriptional and posttranscriptional regulation of HIV-1 gene expression. *Cold Spring Harb. Perspect. Med.* **2**, 1–17 (2012).

153. Mousseau, G. & Valente, S. T. Role of Host Factors on the Regulation of Tat-Mediated HIV-1 Transcription. *Curr. Pharm. Des.* **23**, 4079–4090 (2017).

154. Ho, J. S. Y., Zhu, Z. & Marazzi, I. Unconventional viral gene expression mechanisms as therapeutic targets. *Nature* **593**, 362–371 (2021).

155. Jolly, C., Mitar, I. & Sattentau, Q. J. Requirement for an Intact T-Cell Actin and Tubulin Cytoskeleton for Efficient Assembly and Spread of Human Immunodeficiency Virus Type 1. *J. Virol.* **81**, 5547–5560 (2007).

156. Ono, A., Ablan, S. D., Lockett, S. J., Nagashima, K. & Freed, E. O. Phosphatidylinositol (4,5) bisphosphate regulates HIV-1 Gag targeting to the plasma membrane. *Proc. Natl. Acad. Sci.* **101**, 14889–14894 (2004).

157. Kleiman, L. tRNALys3: The primer tRNA for reverse transcription in HIV-1. *IUBMB Life* **53**, 107–114 (2002).

158. Dussupt, V. *et al.* Basic Residues in the Nucleocapsid Domain of Gag Are Critical for Late Events of HIV-1 Budding. *J. Virol.* **85**, 2304–2315 (2011).

159. Votteler, J. & Sundquist, W. I. Virus Budding and the ESCRT Pathway. *Cell Host Microbe* **14**, 232–241 (2013).

160. Wanaguru, M. & Bishop, K. N. HIV-1 Gag Recruits Oligomeric Vpr via Two Binding Sites in p6, but Both Mature p6 and Vpr Are Rapidly Lost upon Target Cell Entry. *J. Virol.* **95**, 1–15 (2021).

161. Wang, W. *et al.* A systematic study of the N-glycosylation sites of HIV-1 envelope protein on infectivity and antibody-mediated neutralization. *Retrovirology* **10**, 1–14 (2013).

162. Kornfeld, R. & Kornfeld, S. Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* **54**, 631–664 (1985).

163. Ware, F. E. *et al.* The molecular chaperone calnexin binds Glc1Man9GlcNAc2 oligosaccharide as an initial step in recognizing unfolded glycoproteins. *J. Biol. Chem.* **270**, 4697–4704 (1995).

164. Checkley, M. A., Luttge, B. G. & Freed, E. O. HIV-1 envelope glycoprotein biosynthesis, trafficking, and incorporation. *J. Mol. Biol.* **410**, 582–608 (2011).

165. Helenius, A. & Aebi, M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu. Rev. Biochem.* **73**, 1019–1049 (2004).

166. Doores, K. J. *et al.* Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13800–13805 (2010).

167. Freed, E. O. HIV-1 assembly, release and maturation. *Nat. Rev. Microbiol.* **13**, 484–496 (2015).

168. Joseph, S. B., Swanstrom, R., Kashuba, A. D. M. & Cohen, M. S. Bottlenecks in HIV-1 transmission: insights from the study of founder viruses. *Nat. Rev. Microbiol.* **13**, 414–425 (2015).

169. Weber, J. The pathogenesis of HIV-1 infection. *Br. Med. Bull.* **58**, 61–72 (2001).

170. Hu, J., Gardner, M. B. & Miller, C. J. Simian Immunodeficiency Virus Rapidly Penetrates the Cervicovaginal Mucosa after Intravaginal Inoculation and Infects Intraepithelial Dendritic Cells. *J. Virol.* **74**, 6087–6095 (2000).

171. Stieh, D. J. *et al.* Th17 Cells Are Preferentially Infected Very Early after Vaginal Transmission of SIV in Macaques. *Cell Host Microbe* **19**, 529–540 (2016).

172. Lackner, A. A., Lederman, M. M. & Rodriguez, B. HIV pathogenesis: The host. *Cold Spring Harb. Perspect. Med.* **2**, 1–23 (2012).

173. Coffin, J. & Swanstrom, R. HIV pathogenesis: Dynamics and genetics of viral populations and infected cells. *Cold Spring Harb. Perspect. Med.* **3**, 12526–12542 (2013).

174. Sterling, T. R. *et al.* Initial Plasma HIV-1 RNA Levels and Progression to AIDS in Women and Men. *N. Engl. J. Med.* **344**, 720–725 (2001).

175. Kumar, P. Long term non-progressor (LTNP) HIV infection. *Indian J. Med. Res.* **138**, 291–293 (2013).

176. Wibmer, C. K. *et al.* Common helical V1V2 conformations of HIV-1 Envelope expose the α4β7 binding site on intact virions. *Nat. Commun.* **9**, 4489–4501 (2018).

177. van Anken, E. *et al.* Only Five of 10 Strictly Conserved Disulfide Bonds Are Essential for Folding and Eight for Function of the HIV-1 Envelope Glycoprotein. *Mol. Biol. Cell* **19**, 4298–4309 (2008).

178. Karray, S. & Zouali, M. Identification of the B cell superantigen-binding site of HIV-1 gp120. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 1356–1360 (1997).

179. Curlin, M. E. *et al.* HIV-1 Envelope Subregion Length Variation during Disease Progression. *PLoS Pathog.* **6**, 1001228–1001239 (2010).

180. Hartley, O., Klasse, P. J., Sattentau, Q. J. & Moore, J. P. V3: HIV's Switch-Hitter. *AIDS Res. Hum. Retroviruses* **21**, 171–189 (2005).

181. Peeters, M. & Sharp, P. M. Genetic diversity of HIV-1: the moving target. *AIDS* **14**, 129–140 (2000).

182. Klasse, P. J., Ozorowski, G., Sanders, R. W. & Moore, J. P. Env Exceptionalism: Why Are HIV-1 Env Glycoproteins Atypical Immunogens? *Cell Host Microbe* **27**, 507–518 (2020).

183. Burton, D. R. & Hangartner, L. Broadly Neutralizing Antibodies to HIV and Their Role in Vaccine Design. *Annu. Rev. Immunol.* **34**, 635–659 (2016).

184. Cory, T. J., Schacker, T. W., Stevenson, M. & Fletcher, C. V. Overcoming pharmacologic sanctuaries. *Curr. Opin. HIV AIDS* **8**, 190–195 (2013).

185. Moore, P. L. *et al.* Nature of Nonfunctional Envelope Proteins on the Surface of Human Immunodeficiency Virus Type 1. *J. Virol.* **80**, 2515–2528 (2006).

186. Lee, W. S. & Kent, S. J. Anti-HIV-1 antibody-dependent cellular cytotoxicity: Is there more to antibodies than neutralization? *Curr. Opin. HIV AIDS* **13**, 160–166 (2018).

187. Moore, P. L., Gray, E. S. & Morris, L. Specificity of the autologous neutralizing antibody response. *Curr. Opin. HIV AIDS* **4**, 358–363 (2009).

188. Allen, T. M. & Altfeld, M. HIV-1 superinfection. *J. Allergy Clin. Immunol.* **112**, 829–835 (2003).

189. Wei, X. *et al.* Antibody neutralization and escape by HIV-1. *Nature* **422**, 307–312 (2003).

190. Mascola, J. R. & Montefiori, D. C. HIV-1: nature's master of disguise. *Nat. Med.* **9**, 393–394 (2003).

191. Korber, B. *et al.* Evolutionary and immunological implications of contemporary HIV-1 variation. *Br. Med. Bull.* **58**, 19–42 (2001).

192. Sather, D. N. *et al.* Broadly Neutralizing Antibodies Developed by an HIV-Positive Elite Neutralizer Exact a Replication Fitness Cost on the Contemporaneous Virus. *J. Virol.* **86**, 12676–12685 (2012).

193. Bunnik, E. M., Quakkelaar, E. D., van Nuenen, A. C., Boeser-Nunnink, B. & Schuitemaker, H. Increased Neutralization Sensitivity of Recently Emerged CXCR4-Using Human Immunodeficiency Virus Type 1 Strains Compared to Coexisting CCR5-Using Variants from the Same Patient. *J. Virol.* **81**, 525–531 (2007).

194. Mikell, I. *et al.* Characteristics of the Earliest Cross-Neutralizing Antibody Response to HIV-1. *PLoS Pathog.* **7**, 1001251–1001262 (2011).

195. Hraber, P. *et al.* Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection. *AIDS* **28**, 163–169 (2014).

196. Griffith, S. A. & McCoy, L. E. To bnAb or Not to bnAb: Defining Broadly Neutralising Antibodies Against HIV-1. *Front. Immunol.* **12**, 1–16 (2021).

197. Doria-Rose, N. A. *et al.* Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* **509**, 55–62 (2014).

198. Haynes, B. F. *et al.* Cardiolipin Polyspecific Autoreactivity in Two Broadly Neutralizing HIV-1 Antibodies. *Science (80-. ).* **308**, 1906–1908 (2005).

199. Landais, E. *et al.* HIV Envelope Glycoform Heterogeneity and Localized Diversity Govern the Initiation and Maturation of a V2 Apex Broadly Neutralizing Antibody Lineage. *Immunity* **47**, 990-1003.e9 (2017).

200. Sok, D. & Burton, D. R. Recent progress in broadly neutralizing antibodies to HIV. *Nat. Immunol.* **19**, 1179–1188 (2018).

201. Lynch, R. M. *et al.* HIV-1 Fitness Cost Associated with Escape from the VRC01 Class of CD4 Binding Site Neutralizing Antibodies. *J. Virol.* **89**, 4201–4213 (2015).

202. Conti, S. *et al.* Design of immunogens to elicit broadly neutralizing antibodies against HIV targeting the CD4 binding site. *Proc. Natl. Acad. Sci.* **118**, 56–75 (2021).

203. Zhou, T. *et al.* Multidonor Analysis Reveals Structural Elements, Genetic Determinants, and Maturation Pathway for HIV-1 Neutralization by VRC01-Class Antibodies. *Immunity* **39**, 245–258 (2013).

204. Bonsignori, M. *et al.* Maturation Pathway from Germline to Broad HIV-1 Neutralizer of a CD4-Mimic Antibody. *Cell* **165**, 449–463 (2016).

205. Gao, F. *et al.* Cooperation of B Cell Lineages in Induction of HIV-1-Broadly Neutralizing Antibodies. *Cell* **158**, 481–491 (2014).

206. Liao, H.-X. *et al.* Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* **496**, 469–476 (2013).

207. Umotoy, J. *et al.* Rapid and Focused Maturation of a VRC01-Class HIV Broadly Neutralizing Antibody Lineage Involves Both Binding and Accommodation of the N276-Glycan. *Immunity*

**51**, 141-154.e6 (2019).

208. Doria-Rose, N. A. & Landais, E. Coevolution of HIV-1 and broadly neutralizing antibodies. *Curr. Opin. HIV AIDS* **14**, 286–293 (2019).

209. Bhiman, J. N. *et al.* Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. *Nat. Med.* **21**, 1332–1336 (2015).

210. Sheward, D. J. *et al.* HIV Superinfection Drives De Novo Antibody Responses and Not Neutralization Breadth. *Cell Host Microbe* **24**, 593–599 (2018).

211. Rantalainen, K. *et al.* Co-evolution of HIV Envelope and Apex-Targeting Neutralizing Antibody Lineage Provides Benchmarks for Vaccine Design. *Cell Rep.* **23**, 3249–3261 (2018).

212. Landais, E. *et al.* Broadly Neutralizing Antibody Responses in a Large Longitudinal Sub-Saharan HIV Primary Infection Cohort. *PLoS Pathog.* **12**, 1–22 (2016).

213. Coss, K. P. *et al.* HIV-1 Glycan Density Drives the Persistence of the Mannose Patch within an Infected Individual. *J. Virol.* **90**, 11132–11144 (2016).

214. Moyo, T., Kitchin, D. & Moore, P. L. Targeting the N332-supersite of the HIV-1 envelope for vaccine design. *Expert Opin. Ther. Targets* **24**, 499–509 (2020).

215. Krumm, S. A. *et al.* Mechanisms of escape from the PGT128 family of anti-HIV broadly neutralizing antibodies. *Retrovirology* **13**, 1–15 (2016).

216. Scanlan, C. N. *et al.* The Broadly Neutralizing Anti-Human Immunodeficiency Virus Type 1 Antibody 2G12 Recognizes a Cluster of α1→2 Mannose Residues on the Outer Face of gp120. *J. Virol.* **76**, 7306–7321 (2002).

217. Caskey, M. *et al.* Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. *Nat. Med.* **23**, 185–191 (2017).

218. Walker, L. M. *et al.* Broad neutralization coverage of HIV by multiple highly potent antibodies. *Nature* **477**, 466–470 (2011).

219. MacLeod, D. T. *et al.* Early Antibody Lineage Diversification and Independent Limb Maturation Lead to Broad HIV-1 Neutralization Targeting the Env High-Mannose Patch. *Immunity* **44**, 1215–1226 (2016).

220. Landais, E. & Moore, P. L. Development of broadly neutralizing antibodies in HIV-1 infected elite neutralizers. *Retrovirology* **15**, 1–14 (2018).

221. Bonsignori, M. *et al.* Staged induction of HIV-1 glycan–dependent broadly neutralizing antibodies. *Sci. Transl. Med.* **9**, 1–26 (2017).

222. Wibmer, C. K. *et al.* Structure and Recognition of a Novel HIV-1 gp120-gp41 Interface Antibody that Caused MPER Exposure through Viral Escape. *PLoS Pathog.* **13**, 1–30 (2017).

223. Derking, R. *et al.* Comprehensive Antigenic Map of a Cleaved Soluble HIV-1 Envelope Trimer. *PLoS Pathog.* **11**, 1–22 (2015).

224. Schoofs, T. *et al.* Broad and Potent Neutralizing Antibodies Recognize the Silent Face of the HIV Envelope. *Immunity* **50**, 1513-1529.e9 (2019).

225. Molinos-Albert, L. M., Clotet, B., Blanco, J. & Carrillo, J. Immunologic insights on the membrane proximal external region: A major human immunodeficiency virus type-1 vaccine target. *Front. Immunol.* **8**, 1–12 (2017).

226. Fu, Q. *et al.* Structure of the membrane proximal external region of HIV-1 envelope glycoprotein. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 8892–8899 (2018).

227. Carravilla, P. *et al.* Molecular recognition of the native HIV-1 MPER revealed by STED microscopy of single virions. *Nat. Commun.* **10**, 78 (2019).

228. Ruprecht, C. R. *et al.* MPER-specific antibodies induce gp120 shedding and irreversibly neutralize HIV-1. *J. Exp. Med.* **208**, 439–454 (2011).

229. Krebs, S. J. *et al.* Longitudinal Analysis Reveals Early Development of Three MPER-Directed Neutralizing Antibody Lineages from an HIV-1-Infected Individual. *Immunity* **50**, 677–693 (2019).

230. Arts, E. J. & Hazuda, D. J. HIV-1 Antiretroviral Drug Therapy. *Cold Spring Harb. Perspect. Med.* **2**, 7161–7178 (2012).

231. Lv, Z., Chu, Y. & Wang, Y. HIV protease inhibitors: A review of molecular selectivity and toxicity. *HIV/AIDS - Res. Palliat. Care* **7**, 95–104 (2015).

232. Hardy, H. & Skolnik, P. R. Enfuvirtide, a New Fusion Inhibitor for Therapy of Human Immunodeficiency Virus Infection. *Pharmacotherapy* **24**, 198–211 (2004).

233. Gulick, R. M. *et al.* Maraviroc for Previously Treated Patients with R5 HIV-1 Infection. *N. Engl. J. Med.* **359**, 1429–1441 (2008).

234. Kuritzkes, D. R. HIV-1 entry inhibitors: An overview. *Curr. Opin. HIV AIDS* **4**, 82–87 (2009).

235. Scheid, J. F. *et al.* HIV-1 antibody 3BNC117 suppresses viral rebound in humans during treatment interruption. *Nature* **535**, 556–560 (2016).

236. Lynch, R. M. *et al.* Virologic effects of broadly neutralizing antibody VRC01 administration during chronic HIV-1 infection. *Sci. Transl. Med.* **7**, 1–15 (2015).

237. Mendoza, P. *et al.* Combination therapy with anti-HIV-1 antibodies maintains viral suppression. *Nature* **561**, 479–484 (2018).

238. Deeks, S. G., Overbaugh, J., Phillips, A. & Buchbinder, S. HIV infection. *Nat. Rev. Dis. Prim.* **1**, 15035–12044 (2015).

239. Hütter, G. *et al.* Long-Term Control of HIV by CCR5 Delta32/Delta32 Stem-Cell Transplantation. *N. Engl. J. Med.* **360**, 692–698 (2009).

240. Lederman, M. M. & Pike, E. Ten Years HIV Free: An Interview with “The Berlin Patient,” Timothy Ray Brown. *Pathog. Immun.* **2**, 422–433 (2017).

241. Gupta, R. K. *et al.* HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation. *Nature* **568**, 244–248 (2019).

242. Rerks-Ngarm, S. *et al.* Vaccination with ALVAC and AIDSVAX to Prevent HIV-1 Infection in Thailand. *N. Engl. J. Med.* **361**, 2209–2220 (2009).

243. Corey, L. *et al.* Immune correlates of vaccine protection against HIV-1 acquisition. *Sci. Transl. Med.* **7**, 100–106 (2015).

244. Laher, F. *et al.* Safety and immune responses after a 12-month booster in healthy HIV-uninfected adults in HVTN 100 in South Africa: A randomized double-blind placebo-controlled trial of ALVAC-HIV (vCP2438) and bivalent subtype C gp120/MF59 vaccines. *PLoS Med.* **17**, e1003038 (2020).

245. Seabright, G. E., Doores, K. J., Burton, D. R. & Crispin, M. Protein and Glycan Mimicry in HIV Vaccine Design. *J. Mol. Biol.* **431**, 2223–2247 (2019).

246. Wu, X. *et al.* Neutralization Escape Variants of Human Immunodeficiency Virus Type 1 Are Transmitted from Mother to Infant. *J. Virol.* **80**, 835–844 (2006).

247. Sanders, R. W. *et al.* A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies. *PLoS Pathog.* **9**, 1003618–1003633 (2013).

248. Sharma, S. K. *et al.* Cleavage-Independent HIV-1 Env Trimers Engineered as Soluble Native Spike Mimetics for Vaccine Design. *Cell Rep.* **11**, 539–550 (2015).

249. Kong, L. *et al.* Uncleaved prefusion-optimized gp140 trimers derived from analysis of HIV-1 envelope metastability. *Nat. Commun.* **7**, 1–15 (2016).

250. Sanders, R. W. *et al.* HIV-1 neutralizing antibodies induced by native-like envelope trimers. *Science (80-. ).* **349**, 52–66 (2015).

251. McCoy, L. E. *et al.* Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies. *Cell Rep.* **16**, 2327–2338 (2016).

252. Sliepen, K. *et al.* Structure and immunogenicity of a stabilized HIV-1 envelope trimer based on a group-M consensus sequence. *Nat. Commun.* **10**, 1–16 (2019).

253. Barouch, D. H. *et al.* Evaluation of a mosaic HIV-1 vaccine in a multicentre, randomised, double-blind, placebo-controlled, phase 1/2a clinical trial (APPROACH) and in rhesus monkeys (NHP 13-19). *Lancet* **392**, 232–243 (2018).

254. del Moral-Sánchez, I. & Sliepen, K. Strategies for inducing effective neutralizing antibody responses against HIV-1. *Expert Rev. Vaccines* **18**, 1127–1143 (2019).

255. Bianchi, M. *et al.* Electron-Microscopy-Based Epitope Mapping Defines Specificities of Polyclonal Antibodies Elicited during HIV-1 BG505 Envelope Trimer Immunization. *Immunity* **49**, 288-300.e8 (2018).

256. Brouwer, P. J. M. & Sanders, R. W. Presentation of HIV-1 envelope glycoprotein trimers on diverse nanoparticle platforms. *Curr. Opin. HIV AIDS* **14**, 302–308 (2019).

257. Schiller, J. & Chackerian, B. Why HIV Virions Have Low Numbers of Envelope Spikes: Implications for Vaccine Development. *PLoS Pathog.* **10**, 8–11 (2014).

258. Kelly, H. G., Kent, S. J. & Wheatley, A. K. Immunological basis for enhanced immunity of nanoparticle vaccines. *Expert Rev. Vaccines* **18**, 269–280 (2019).

259. The SPARTAC Trial Investigators. Short-Course Antiretroviral Therapy in Primary HIV Infection. *N. Engl. J. Med.* **368**, 207–217 (2013).

260. Spencer, D. A., Shapiro, M. B., Haigwood, N. L. & Hessell, A. J. Advancing HIV Broadly Neutralizing Antibodies: From Discovery to the Clinic. *Front. Public Heal.* **9**, 1–17 (2021).

261. Shultz, L. D., Ishikawa, F. & Greiner, D. L. Humanized mice in translational biomedical research. *Nat. Rev. Immunol.* **7**, 118–130 (2007).

262. Klein, F. *et al.* HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. *Nature* **492**, 118–122 (2012).

263. Deruaz, M. *et al.* Protection of Humanized Mice From Repeated Intravaginal HIV Challenge by Passive Immunization: A Model for Studying the Efficacy of Neutralizing Antibodies In Vivo. *J.*

*Infect. Dis.* **214**, 612–616 (2016).

264. Balazs, A. B. *et al.* Vectored immunoprophylaxis protects humanized mice from mucosal HIV transmission. *Nat. Med.* **20**, 296–300 (2014).

265. Hu, S.-L. Non-Human Primate Models for AIDS Vaccine Research. *Curr. Drug Target -Infectious Disord.* **5**, 193–201 (2005).

266. Bauer, A. M. & Bar, K. J. Advances in simian--human immunodeficiency viruses for nonhuman primate studies of HIV prevention and cure. *Curr. Opin. HIV AIDS* **15**, 275–281 (2020).

267. Humes, D., Emery, S., Laws, E. & Overbaugh, J. A Species-Specific Amino Acid Difference in the Macaque CD4 Receptor Restricts Replication by Global Circulating HIV-1 Variants Representing Viruses from Recent Infection. *J. Virol.* **86**, 12472–12483 (2012).

268. Rosenberg, Y. J. *et al.* Protection against SHIV Challenge by Subcutaneous Administration of the Plant-Derived PGT121 Broadly Neutralizing Antibody in Macaques. *PLoS One* **11**, 152760–152772 (2016).

269. Virgin, H. W., Wherry, E. J. & Ahmed, R. Redefining Chronic Viral Infection. *Cell* **138**, 30–50 (2009).

270. Ortega-Prieto, A. & Dorner, M. Immune Evasion Strategies during Chronic Hepatitis B and C Virus Infection. *Vaccines* **5**, 24 (2017).

271. Lin, M. V., King, L. Y. & Chung, R. T. Hepatitis C Virus–Associated Cancer. *Annu. Rev. Pathol. Mech. Dis.* **10**, 345–370 (2015).

272. Yano, Y. Variations and mutations in the hepatitis B virus genome and their associations with clinical characteristics. *World J. Hepatol.* **7**, 583 (2015).

273. Tisthammer, K. H. *et al.* Assessing in vivo mutation frequencies and creating a high-resolution genome-wide map of fitness costs of Hepatitis C virus. *PLOS Genet.* **18**, e1010179 (2022).

274. Patiño-Galindo, J. Á. & González-Candelas, F. The substitution rate of HIV-1 subtypes: a genomic approach. *Virus Evol.* **3**, 1–7 (2017).

275. Hehle, V. *et al.* Potent human broadly neutralizing antibodies to hepatitis B virus from natural controllers. *J. Exp. Med.* **217**, (2020).

276. Ganem, D. Persistent Infection of Humans with Hepatitis B Virus: Mechanisms and Consequences. *Clin. Infect. Dis.* **4**, 1026–1047 (1982).

277. Ciupe, S. M., Ribeiro, R. M. & Perelson, A. S. Antibody Responses during Hepatitis B Viral Infection. *PLoS Comput. Biol.* **10**, e1003730 (2014).

278. Trépo, C., Chan, H. L. Y. & Lok, A. Hepatitis B virus infection. *Lancet* **384**, 2053–2063 (2014).

279. Giang, E. *et al.* Human broadly neutralizing antibodies to the envelope glycoprotein complex of hepatitis C virus. *Proc. Natl. Acad. Sci.* **109**, 6205–6210 (2012).

280. Corman, V. M., Muth, D., Niemeyer, D. & Drosten, C. Hosts and Sources of Endemic Human Coronaviruses. *Adv. Virus Res.* **100**, 163–188 (2018).

281. Woo, P. C. Y. *et al.* Characterization and Complete Genome Sequence of a Novel Coronavirus, Coronavirus HKU1, from Patients with Pneumonia. *J. Virol.* **79**, 884–895 (2005).

282. Van Der Hoek, L. *et al.* Identification of a new human coronavirus. *Nat. Med.* **10**, 368–373 (2004).

283. McIntosh, K., Becker, W. B. & Chanock, R. M. Growth in suckling-mouse brain of 'IBV-like' viruses from patients with upper respiratory tract disease. *Proc. Natl. Acad. Sci. U. S. A.* **58**, 2268–2273 (1967).

284. Hamre, D. & Procknow, J. J. A New Virus Isolated from the Human Respiratory Tract. *Proc. Soc. Exp. Biol. Med.* **121**, 190–193 (1966).

285. Oosterhof, L., Christensen, C. B. & Sengeløv, H. Fatal lower respiratory tract disease with human corona virus NL63 in an adult haematopoietic cell transplant recipient. *Bone Marrow Transplant.* **45**, 1115–1116 (2010).

286. Konca, C. *et al.* The first infant death associated with human coronavirus NL63 infection. *Pediatr. Infect. Dis. J.* **36**, 231–233 (2017).

287. Ksiazek, T. G. *et al.* A Novel Coronavirus Associated with Severe Acute Respiratory Syndrome. *N. Engl. J. Med.* **348**, 1953–1966 (2003).

288. Zaki, A. M., van Boheemen, S., Bestebroer, T. M., Osterhaus, A. D. M. E. & Fouchier, R. A. M. Isolation of a Novel Coronavirus from a Man with Pneumonia in Saudi Arabia. *N. Engl. J. Med.* **367**, 1814–1820 (2012).

289. Pavli, A., Tsiodras, S. & Maltezou, H. C. Middle East respiratory syndrome coronavirus (MERS-CoV): Prevention in travelers. *Travel Med. Infect. Dis.* **12**, 602–608 (2014).

290. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).

291. do Vale, B. *et al.* Bats, pangolins, minks and other animals - villains or victims of SARS-CoV-2? *Vet. Res. Commun.* **45**, 1–19 (2021).

292. Hu, B., Guo, H., Zhou, P. & Shi, Z. L. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **19**, 141–154 (2021).

293. Shang, Y., Li, H. & Zhang, R. Effects of Pandemic Outbreak on Economies: Evidence From Business History Context. *Front. Public Heal.* **9**, 1–12 (2021).

294. Gao, Z. *et al.* A systematic review of asymptomatic infections with COVID-19. *J. Microbiol. Immunol. Infect.* **54**, 12–16 (2021).

295. National Health Commission of People's Republic of China. Diagnosis and treatment plan of Corona Virus Disease 2019 (tentative sixth edition). *Glob. Heal. J.* **4**, 1–5 (2020).

296. Ahmed-Hassan, H. *et al.* Innate Immune Responses to Highly Pathogenic Coronaviruses and Other Significant Respiratory Viral Infections. *Front. Immunol.* **11**, 1979–1992 (2020).

297. Stokes, E. K. *et al.* Coronavirus Disease 2019 Case Surveillance — United States, January 22–May 30, 2020. *MMWR. Morb. Mortal. Wkly. Rep.* **69**, 759–765 (2020).

298. Jin, J. M. *et al.* Gender Differences in Patients With COVID-19: Focus on Severity and Mortality. *Front. Public Heal.* **8**, 1–6 (2020).

299. Gebhard, C., Regitz-Zagrosek, V., Neuhauser, H. K., Morgan, R. & Klein, S. L. Impact of sex and gender on COVID-19 outcomes in Europe. *Biol. Sex Differ.* **11**, 1–13 (2020).

300. Sze, S. *et al.* Ethnicity and clinical outcomes in COVID-19: A systematic review and meta-analysis. *EClinicalMedicine* **29**, 100630–100642 (2020).

301. Khafaie, M. A. & Rahim, F. Osong Public Health and Research Perspectives Cross-Country Comparison of Case Fatality Rates of. *Osong Public Heal. Res Perspect.* **11**, 74–80 (2020).

302. Ahn, D. G. *et al.* Current status of epidemiology, diagnosis, therapeutics, and vaccines for novel coronavirus disease 2019 (COVID-19). *J. Microbiol. Biotechnol.* **30**, 313–324 (2020).

303. Long, Q. X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* **26**, 1200–1204 (2020).

304. Drain, P. K. Rapid Diagnostic Testing for SARS-CoV-2. *N. Engl. J. Med.* **386**, 264–272 (2022).

305. Ayouni, I. *et al.* Effective public health measures to mitigate the spread of COVID-19: a systematic review. *BMC Public Health* **21**, 1–14 (2021).

306. Choi, E. P. H., Hui, B. P. H. & Wan, E. Y. F. Depression and anxiety in Hong Kong during covid-19. *Int. J. Environ. Res. Public Health* **17**, 3740–3755 (2020).

307. García-Montero, C. *et al.* An updated review of sars-cov-2 vaccines and the importance of effective vaccination programs in pandemic times. *Vaccines* **9**, 1–22 (2021).

308. Pai, S. M. *et al.* Science and Evidence-Based Review and Approval of COVID-19 Vaccines: A Statement of Support for the US FDA. *J. Clin. Pharmacol.* **61**, 277–279 (2021).

309. Hasan, T., Beardsley, J., Marais, B. J., Nguyen, T. A. & Fox, G. J. The implementation of mass-vaccination against SARS-CoV-2: A systematic review of existing strategies and guidelines. *Vaccines* **9**, 1–15 (2021).

310. Català, M., Li, X., Prats, C. & Prieto-Alhambra, D. The impact of prioritisation and dosing intervals on the effects of COVID-19 vaccination in Europe: an agent-based cohort model. *Sci. Rep.* **11**, 1–10 (2021).

311. Tao, K. *et al.* The biological and clinical significance of emerging SARS-CoV-2 variants. *Nat. Rev. Genet.* **22**, 757–773 (2021).

312. Liu, Q., Qin, C., Liu, M. & Liu, J. Effectiveness and safety of SARS-CoV-2 vaccine in real-world studies: a systematic review and meta-analysis. *Infect. Dis. Poverty* **10**, 1–15 (2021).

313. Tao, K. *et al.* SARS-CoV-2 Antiviral Therapy. *Clin. Microbiol. Rev.* **34**, 10921–10932 (2021).

314. Coopersmith, C. M. *et al.* The Surviving Sepsis Campaign: Research Priorities for Coronavirus Disease 2019 in Critical Illness. *Crit. Care Med.* **49**, 598–622 (2021).

315. Hotez, P. J. *et al.* Global public health security and justice for vaccines and therapeutics in the COVID-19 pandemic. *EClinicalMedicine* **39**, 101053–101063 (2021).

316. Cui, J., Li, F. & Shi, Z. L. Origin and evolution of pathogenic coronaviruses. *Nat. Rev. Microbiol.* **17**, 181–192 (2019).

317. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).

318. Zhang, T., Wu, Q. & Zhang, Z. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. *Curr. Biol.* **30**, 1346–1351 (2020).

319. Holmes, E. C. *et al.* The origins of SARS-CoV-2: A critical review. *Cell* **184**, 4848–4856 (2021).

320. Lauring, A. S. & Hodcroft, E. B. Genetic Variants of SARS-CoV-2—What Do They Mean? *JAMA* **325**, 529–539 (2021).

321. Harvey, W. T. *et al.* SARS-CoV-2 variants, spike mutations and immune escape. *Nat. Rev. Microbiol.* **19**, 409–424 (2021).

322. Corey, L. *et al.* SARS-CoV-2 Variants in Patients with Immunosuppression. *N. Engl. J. Med.*

**385**, 562–566 (2021).

323. Haddad, D. *et al.* SARS-CoV-2: Possible recombination and emergence of potentially more virulent strains. *PLoS One* **16**, 1–20 (2021).

324. Koelle, K., Martin, M. A., Antia, R., Lopman, B. & Dean, N. E. The changing epidemiology of SARS-CoV-2. *Science (80-. ).* **375**, 1116–1121 (2022).

325. Rastogi, M., Pandey, N., Shukla, A. & Singh, S. K. SARS coronavirus 2: from genome to infectome. *Respir. Res.* **21**, 1–15 (2020).

326. Bai, C., Zhong, Q. & Gao, G. F. Overview of SARS-CoV-2 genome-encoded proteins. *Sci. China Life Sci.* **65**, 280–294 (2022).

327. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H. & Thiel, V. Coronavirus biology and replication: implications for SARS-CoV-2. *Nat. Rev. Microbiol.* **19**, 155–170 (2021).

328. Ke, Z. *et al.* Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature* **588**, 498–502 (2020).

329. Yao, H. *et al.* Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **183**, 730–738 (2020).

330. Zeng, C. *et al.* SARS-CoV-2 spreads through cell-to-cell transmission. *Proc. Natl. Acad. Sci. U. S. A.* **119**, 1–12 (2022).

331. Jackson, C. B., Farzan, M., Chen, B. & Choe, H. Mechanisms of SARS-CoV-2 entry into cells. *Nat. Rev. Mol. Cell Biol.* **23**, 3–20 (2022).

332. Song, W., Gui, M., Wang, X. & Xiang, Y. Cryo-EM structure of the SARS coronavirus spike glycoprotein in complex with its host cell receptor ACE2. *PLOS Pathog.* **14**, 1–19 (2018).

333. Wang, M. Y. *et al.* SARS-CoV-2: Structure, Biology, and Structure-Based Therapeutics Development. *Front. Cell. Infect. Microbiol.* **10**, 1–17 (2020).

334. Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S. & Crispin, M. Site-specific glycan analysis of the SARS-CoV-2 spike. *Science (80-. ).* **369**, 330–333 (2020).

335. Shang, J. *et al.* Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci.* **117**, 11727–11734 (2020).

336. Hoffmann, M. *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271–280 (2020).

337. Ou, X. *et al.* Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity with SARS-CoV. *Nat. Commun.* **11**, 1620–1632 (2020).

338. Xia, S. *et al.* Inhibition of SARS-CoV-2 (previously 2019-nCoV) infection by a highly potent pan-coronavirus fusion inhibitor targeting its spike protein that harbors a high capacity to mediate membrane fusion. *Cell Res.* **30**, 343–355 (2020).

339. Romano, M., Ruggiero, A., Squeglia, F., Maga, G. & Berisio, R. A Structural View of SARS-CoV-2 RNA Replication Machinery: RNA Synthesis, Proofreading and Final Capping. *Cells* **9**, 1267–1278 (2020).

340. Thoms, M. *et al.* Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science (80-. ).* **369**, 1249–1256 (2020).

341. Yan, L. *et al.* Architecture of a SARS-CoV-2 mini replication and transcription complex. *Nat. Commun.* **11**, 3–8 (2020).

342. Sawicki, S. G. & Sawicki, D. L. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv. Exp. Med. Biol.* **380**, 499–506 (1995).

343. Klein, S. *et al.* SARS-CoV-2 structure and replication characterized by in situ cryo-electron tomography. *Nat. Commun.* **11**, 1–10 (2020).

344. Wolff, G. *et al.* A molecular pore spans the double membrane of the coronavirus replication organelle. *Science (80-. ).* **369**, 1395–1398 (2020).

345. Bracquemond, D. & Muriaux, D. Betacoronavirus Assembly: Clues and Perspectives for Elucidating SARS-CoV-2 Particle Formation and Egress. *MBio* **12**, 2371–2385 (2021).

346. Mendonça, L. *et al.* Correlative multi-scale cryo-imaging unveils SARS-CoV-2 assembly and egress. *Nat. Commun.* **12**, 1–10 (2021).

347. Khan, M. *et al.* Visualizing in deceased COVID-19 patients how SARS-CoV-2 attacks the respiratory and olfactory mucosae but spares the olfactory bulb. *Cell* **184**, 5932–5949 (2021).

348. Ahn, J. H. *et al.* Nasal ciliated cells are primary targets for SARS-CoV-2 replication in the early stage of COVID-19. *J. Clin. Invest.* **131**, 1–14 (2021).

349. Lamers, M. M. & Haagmans, B. L. SARS-CoV-2 pathogenesis. *Nat. Rev. Microbiol.* **2**, 1–15 (2022).

350. Robinot, R. *et al.* SARS-CoV-2 infection induces the dedifferentiation of multiciliated cells and impairs mucociliary clearance. *Nat. Commun.* **12**, 1–16 (2021).

351. Nova, Z., Skovierova, H. & Calkovska, A. Alveolar-capillary membrane-related pulmonary cells as a target in endotoxin-induced acute lung injury. *Int. J. Mol. Sci.* **20**, 831–843 (2019).

352. Hou, Y. J. *et al.* SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell* **182**, 429–446 (2020).

353. Melms, J. C. *et al.* A molecular single-cell lung atlas of lethal COVID-19. *Nature* **595**, 114–119 (2021).

354. Georg, P. *et al.* Complement activation induces excessive T cell cytotoxicity in severe COVID-19. *Cell* **185**, 493–512 (2022).

355. Ouwendijk, W. J. D. *et al.* High Levels of Neutrophil Extracellular Traps Persist in the Lower Respiratory Tract of Critically Ill Patients With Coronavirus Disease 2019. *J. Infect. Dis.* **223**, 1512–1521 (2021).

356. Zuo, Y. *et al.* Plasma tissue plasminogen activator and plasminogen activator inhibitor-1 in hospitalized COVID-19 patients. *Sci. Rep.* **11**, 1–9 (2021).

357. Gupta, A. *et al.* Extrapulmonary manifestations of COVID-19. *Nat. Med.* **26**, 1017–1032 (2020).

358. Rodriguez-Guerra, M., Jadhav, P. & Vittorio, T. J. Current treatment in COVID-19 disease: a rapid review. *Drugs Context* **10**, 1–8 (2021).

359. Gandhi, R. T., Lynch, J. B. & del Rio, C. Mild or Moderate Covid-19. *N. Engl. J. Med.* **383**, 1757–1766 (2020).

360. Kokic, G. *et al.* Mechanism of SARS-CoV-2 polymerase stalling by remdesivir. *Nat. Commun.* **12**, 1–7 (2021).

361. Beigel, J. H. *et al.* Remdesivir for the Treatment of Covid-19 — Final Report. *N. Engl. J. Med.*

**383**, 1813–1826 (2020).

362. Spinner, C. D. *et al.* Effect of Remdesivir vs Standard Care on Clinical Status at 11 Days in Patients with Moderate COVID-19: A Randomized Clinical Trial. *JAMA - J. Am. Med. Assoc.* **324**, 1048–1057 (2020).

363. WHO Solidarity Trial Consortium. Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results. *N. Engl. J. Med.* **384**, 497–511 (2021).

364. Wahl, A. *et al.* SARS-CoV-2 infection is effectively treated and prevented by EIDD-2801. *Nature* **591**, 451–457 (2021).

365. Kabinger, F. *et al.* Mechanism of molnupiravir-induced SARS-CoV-2 mutagenesis. *Nat. Struct. Mol. Biol.* **28**, 740–746 (2021).

366. Jayk Bernal, A. *et al.* Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. *N. Engl. J. Med.* **386**, 509–520 (2022).

367. Lamb, Y. N. Nirmatrelvir Plus Ritonavir: First Approval. *Drugs* **82**, 585–591 (2022).

368. Hammond, J. *et al.* Oral Nirmatrelvir for High-Risk, Nonhospitalized Adults with Covid-19. *N. Engl. J. Med.* **386**, 1397–1408 (2022).

369. Whitley, R. Molnupiravir — A Step toward Orally Bioavailable Therapies for Covid-19. *N. Engl. J. Med.* **386**, 592–593 (2022).

370. Ramamoorthy, S. & Cidlowski, J. A. Corticosteroids-Mechanisms of Action in Health and Disease. *Rheum. Dis. Clin. North Am.* **42**, 15–31 (2016).

371. The RECOVERY Collaborative Group. Dexamethasone in Hospitalized Patients with Covid-19. *N. Engl. J. Med.* **384**, 693–704 (2021).

372. Shankar-Hari, M. *et al.* Association between Administration of IL-6 Antagonists and Mortality among Patients Hospitalized for COVID-19: A Meta-analysis. *JAMA - J. Am. Med. Assoc.* **326**, 499–518 (2021).

373. Grant, A. H. *et al.* The Many Faces of JAKs and STATs Within the COVID-19 Storm. *Front. Immunol.* **12**, 1–12 (2021).

374. Kalil, A. C. *et al.* Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19. *N. Engl. J. Med.* **384**, 795–807 (2021).

375. Kamel, A. M., Sobhy, M., Magdy, N., Sabry, N. & Farid, S. Anticoagulation outcomes in hospitalized Covid-19 patients: A systematic review and meta-analysis of case-control and cohort studies. *Rev. Med. Virol.* **31**, 1–23 (2021).

376. Wooding, D. J. & Bach, H. Treatment of COVID-19 with convalescent plasma: lessons from past coronavirus outbreaks. *Clin. Microbiol. Infect.* **26**, 1436–1446 (2020).

377. Axfors, C. *et al.* Association between convalescent plasma treatment and mortality in COVID-19: a collaborative systematic review and meta-analysis of randomized clinical trials. *BMC Infect. Dis.* **21**, 1–23 (2021).

378. Rijnders, B. J. A., Huygens, S. & Mitjà, O. Evidence-based dosing of convalescent plasma for COVID-19 in future trials. *Clin. Microbiol. Infect.* **1**, 1–19 (2022).

379. Taylor, P. C. *et al.* Neutralizing monoclonal antibodies for treatment of COVID-19. *Nat. Rev. Immunol.* **21**, 382–393 (2021).

380. Baum, A. *et al.* REGN-COV2 antibodies prevent and treat SARS-CoV-2 infection in rhesus macaques and hamsters. *Science (80-. ).* **370**, 1110–1115 (2020).

381. Weinreich, D. M. *et al.* REGN-COV2, a Neutralizing Antibody Cocktail, in Outpatients with Covid-19. *N. Engl. J. Med.* **384**, 238–251 (2021).

382. Weinreich, D. M. *et al.* REGEN-COV Antibody Combination and Outcomes in Outpatients with Covid-19. *N. Engl. J. Med.* **385**, 81–99 (2021).

383. Laurini, E., Marson, D., Aulic, S., Fermeglia, A. & Pricl, S. Molecular rationale for SARS-CoV-2 spike circulating mutations able to escape bamlanivimab and etesevimab monoclonal antibodies. *Sci. Rep.* **11**, 1–20 (2021).

384. Gottlieb, R. L. *et al.* Effect of bamlanivimab as monotherapy or in combination with etesevimab on viral load in patients with mild to moderate COVID-19: A randomized clinical trial. *JAMA - J. Am. Med. Assoc.* **325**, 632–644 (2021).

385. Dougan, M. *et al.* Bamlanivimab plus Etesevimab in Mild or Moderate Covid-19. *N. Engl. J. Med.* **385**, 1382–1392 (2021).

386. VanBlargan, L. A. *et al.* An infectious SARS-CoV-2 B.1.1.529 Omicron virus escapes neutralization by therapeutic monoclonal antibodies. *Nat. Med.* **28**, 490–495 (2022).

387. Cameroni, E. *et al.* Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* **602**, 664–670 (2022).

388. Pinto, D. *et al.* Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295 (2020).

389. Gupta, A. *et al.* Early Treatment for Covid-19 with SARS-CoV-2 Neutralizing Antibody Sotrovimab. *N. Engl. J. Med.* **385**, 1941–1950 (2021).

390. Krammer, F. SARS-CoV-2 vaccines in development. *Nature* **586**, 516–527 (2020).

391. Malone, R. W., Felgner, P. L. & Verma, I. M. Cationic liposome-mediated RNA transfection. *Proc. Natl. Acad. Sci. U. S. A.* **86**, 6077–6081 (1989).

392. Chaudhary, N., Weissman, D. & Whitehead, K. A. mRNA vaccines for infectious diseases: principles, delivery and clinical translation. *Nat. Rev. Drug Discov.* **20**, 817–838 (2021).

393. Xia, X. Domains and functions of spike protein in sars-cov-2 in the context of vaccine design. *Viruses* **13**, 1–16 (2021).

394. Hou, X., Zaks, T., Langer, R. & Dong, Y. Lipid nanoparticles for mRNA delivery. *Nat. Rev. Mater.* **6**, 1078–1094 (2021).

395. Baden, L. R. *et al.* Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. *N. Engl. J. Med.* **384**, 403–416 (2021).

396. Polack, F. P. *et al.* Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).

397. Shimabukuro, T. T., Cole, M. & Su, J. R. Reports of Anaphylaxis after Receipt of mRNA COVID-19 Vaccines in the US-December 14, 2020-January 18, 2021. *JAMA - J. Am. Med. Assoc.* **325**, 1101–1102 (2021).

398. Sellaturay, P., Nasser, S., Islam, S., Gurugama, P. & Ewan, P. W. Polyethylene glycol (PEG) is a cause of anaphylaxis to the Pfizer/BioNTech mRNA COVID-19 vaccine. *Clin. Exp. Allergy* **51**, 861–863 (2021).

399. Mendonça, S. A., Lorincz, R., Boucher, P. & Curiel, D. T. Adenoviral vector vaccine platforms in the SARS-CoV-2 pandemic. *npj Vaccines* **6**, 1–18 (2021).

400. Rowe, W. P., Huebner, R. J., Gilmore, L. K., Parrott, R. H. & Ward, T. G. Isolation of a Cytopathogenic Agent from Human Adenoids Undergoing Spontaneous Degeneration in Tissue Culture. *Exp. Biol. Med.* **84**, 570–573 (1953).

401. Vanaparthy, R., Mohan, G., Vasireddy, D. & Atluri, P. Review of covid-19 viral vector-based vaccines and covid-19 variants. *Infez. Med.* **29**, 328–338 (2021).

402. Bos, R. *et al.* Ad26 vector-based COVID-19 vaccine encoding a prefusion-stabilized SARS-CoV-2 Spike immunogen induces potent humoral and cellular immune responses. *npj Vaccines* **5**, 1–11 (2020).

403. Afkhami, S., Yao, Y. & Xing, Z. Methods and clinical development of adenovirus-vectored vaccines against mucosal pathogens. *Mol. Ther. - Methods Clin. Dev.* **3**, 16030–16045 (2016).

404. Joe, C. C. D. *et al.* Manufacturing a chimpanzee adenovirus-vectored SARS-CoV-2 vaccine to meet global needs. *Biotechnol. Bioeng.* **119**, 48–58 (2022).

405. Sadoff, J. *et al.* Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. *N. Engl. J. Med.* **384**, 2187–2201 (2021).

406. Voysey, M. *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. *Lancet* **397**, 99–111 (2021).

407. Greinacher, A. *et al.* Thrombotic Thrombocytopenia after ChAdOx1 nCov-19 Vaccination. *N. Engl. J. Med.* **384**, 2092–2101 (2021).

408. Muir, K.-L., Kallam, A., Koepsell, S. A. & Gundabolu, K. Thrombotic Thrombocytopenia after Ad26.COV2.S Vaccination. *N. Engl. J. Med.* **384**, 1964–1965 (2021).

409. Haynes, B. F. SARS-CoV-2 and HIV-1 — a tale of two vaccines. *Nat. Rev. Immunol.* **21**, 543–544 (2021).

410. Zhang, P. *et al.* A multiclade env–gag VLP mRNA vaccine elicits tier-2 HIV-1-neutralizing antibodies and reduces the risk of heterologous SHIV infection in macaques. *Nat. Med.* **27**, 2234–2245 (2021).

411. Rubens, M., Ramamoorthy, V., Saxena, A., Shehadeh, N. & Appunni, S. HIV Vaccine: Recent Advances, Current Roadblocks, and Future Directions. *J. Immunol. Res.* **2015**, 1–9 (2015).

412. Bloom, K., van den Berg, F. & Arbuthnot, P. Self-amplifying RNA vaccines for infectious diseases. *Gene Ther.* **28**, 117–129 (2021).

413. Abdelrahman, Z., Li, M. & Wang, X. Comparative Review of SARS-CoV-2, SARS-CoV, MERS-CoV, and Influenza A Respiratory Viruses. *Front. Immunol.* **11**, 2002–2012 (2020).

414. Glezen, W. P. Emerging Infections: Pandemic Influenza. *Epidemiol. Rev.* **18**, 64–76 (1996).

415. Viboud, C., Grais, R. F., Lafont, B. A. P., Miller, M. A. & Simonsen, L. Multinational Impact of the 1968 Hong Kong Influenza Pandemic: Evidence for a Smoldering Pandemic. *J. Infect. Dis.* **192**, 233–248 (2005).

416. Garten, R. J. *et al.* Antigenic and Genetic Characteristics of Swine-Origin 2009 A(H1N1) Influenza Viruses Circulating in Humans. *Science (80-. ).* **325**, 197–201 (2009).

417. Shieh, W.-J. *et al.* 2009 Pandemic Influenza A (H1N1). *Am. J. Pathol.* **177**, 166–175 (2010).

418. Krammer, F. The human antibody response to influenza A virus infection and vaccination. *Nat. Rev. Immunol.* **19**, 383–397 (2019).

419. Doud, M. & Bloom, J. Accurate Measurement of the Effects of All Amino-Acid Mutations on Influenza Hemagglutinin. *Viruses* **8**, 155–168 (2016).

420. Nobusawa, E. & Sato, K. Comparison of the Mutation Rates of Human Influenza A and B Viruses. *J. Virol.* **80**, 3675–3678 (2006).

421. Gerdil, C. The annual production cycle for influenza vaccine. *Vaccine* **21**, 1776–1779 (2003).

422. Wang, Q. *et al.* SARS-CoV-2 neutralising antibodies after bivalent versus monovalent booster. *Lancet Infect. Dis.* **23**, 527–528 (2023).

423. Ekiert, D. C. & Wilson, I. A. Broadly neutralizing antibodies against influenza virus and prospects for universal therapies. *Curr. Opin. Virol.* **2**, 134–141 (2012).

424. Wang, T. T. *et al.* Broadly Protective Monoclonal Antibodies against H3 Influenza Viruses following Sequential Immunization with Different Hemagglutinins. *PLoS Pathog.* **6**, e1000796 (2010).

425. Demoliou, C., Papaneophytou, C. & Nicolaidou, V. SARS-CoV-2 and HIV-1: So Different yet so Alike . Immune Response at the Cellular and Molecular Level. *Int. J. Med. Sci.* **19**, 1787–1795 (2022).

426. Engelman, A. N. & Singh, P. K. Cellular and molecular mechanisms of HIV-1 integration targeting. *Cell. Mol. Life Sci.* **75**, 2491–2507 (2018).

427. Smits, N. *et al.* No evidence of human genome integration of SARS-CoV-2 found by long-read DNA sequencing. *Cell Rep.* **36**, 109530–109546 (2021).

428. Cohn, L. B. & Deeks, S. G. The immune response fails to control HIV early in initial virus spread. *J. Clin. Invest.* **130**, 2803–2805 (2020).

429. Barouch, D. H. Challenges in the development of an HIV-1 vaccine. *Nature* **455**, 613–619 (2008).

430. Lauring, A. S. *et al.* Clinical severity of, and effectiveness of mRNA vaccines against, covid-19 from omicron, delta, and alpha SARS-CoV-2 variants in the United States: prospective observational study. *BMJ* **376**, 69761–69780 (2022).

431. Davis, M. W. & Jorgensen, E. M. ApE, A Plasmid Editor: A Freely Available DNA Manipulation and Visualization Program. *Front. Bioinforma.* **2**, 1–15 (2022).

432. Von Boehmer, L. *et al.* Sequencing and cloning of antigen-specific antibodies from mouse memory B cells. *Nat. Protoc.* **11**, 1908–1923 (2016).

433. Brochet, X., Lefranc, M. P. & Giudicelli, V. IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis. *Nucleic Acids Res.* **36**, 503–508 (2008).

434. Sehnal, D. *et al.* Mol∗Viewer: Modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Res.* **49**, 431–437 (2021).

435. Burton, D. R. *et al.* A Blueprint for HIV Vaccine Discovery. *Cell Host Microbe* **12**, 396–407 (2012).

436. Moldt, B. *et al.* Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **109**,

18921–18925 (2012).

437. Hessell, A. J. *et al.* Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. *Nat. Med.* **15**, 951–954 (2009).

438. Freund, N. T. *et al.* Coexistence of potent HIV-1 broadly neutralizing antibodies and antibody-sensitive viruses in a viremic controller. *Sci. Transl. Med.* **9**, 139–148 (2017).

439. Bonsignori, M. *et al.* Antibody-virus co-evolution in HIV infection: paths for HIV vaccine development. *Immunol. Rev.* **275**, 145–160 (2017).

440. Van Schooten, J. & Van Gils, M. J. HIV-1 immunogens and strategies to drive antibody responses towards neutralization breadth. *Retrovirology* **15**, 1–11 (2018).

441. Dubrovskaya, V. *et al.* Vaccination with Glycan-Modified HIV NFL Envelope Trimer-Liposomes Elicits Broadly Neutralizing Antibodies to Multiple Sites of Vulnerability. *Immunity* **51**, 915–929 (2019).

442. The INSIGHT START Study Group. Initiation of Antiretroviral Therapy in Early Asymptomatic HIV Infection. *N. Engl. J. Med.* **373**, 795–807 (2015).

443. Anglemyer, A. *et al.* Early initiation of antiretroviral therapy in HIV-infected adults and adolescents. *AIDS* **28**, 105–118 (2014).

444. Granger, L. A. *et al.* Broadly neutralizing antibody responses in the longitudinal primary HIV-1 infection Short Pulse Anti-Retroviral Therapy at Seroconversion cohort. *AIDS* **35**, 2073–2084 (2021).

445. Anthony, C. *et al.* Cooperation between Strain-Specific and Broadly Neutralizing Responses Limited Viral Escape and Prolonged the Exposure of the Broadly Neutralizing Epitope. *J. Virol.* **91**, 1–16 (2017).

446. Simek, M. D. *et al.* Human Immunodeficiency Virus Type 1 Elite Neutralizers: Individuals with Broad and Potent Neutralizing Activity Identified by Using a High-Throughput Neutralization Assay together with an Analytical Selection Algorithm. *J. Virol.* **83**, 7337–7348 (2009).

447. deCamp, A. *et al.* Global Panel of HIV-1 Env Reference Strains for Standardized Assessments of Vaccine-Elicited Neutralizing Antibodies. *J. Virol.* **88**, 2489–2507 (2014).

448. Montefiori, D. C. Evaluating Neutralizing Antibodies Against HIV, SIV, and SHIV in Luciferase Reporter Gene Assays. *Curr. Protoc. Immunol.* **64**, 1–17 (2004).

449. Sarzotti-Kelsoe, M. *et al.* Optimization and validation of a neutralizing antibody assay for HIV-1 in A3R5 cells. *J. Immunol. Methods* **409**, 147–160 (2014).

450. Deng, W. *et al.* DIVEIN: A web server to analyze phylogenies, sequence divergence, diversity, and informative sites. *Biotechniques* **48**, 405–408 (2010).

451. Kong, L. *et al.* Supersite of immune vulnerability on the glycosylated face of HIV-1 envelope glycoprotein gp120. *Nat. Struct. Mol. Biol.* **20**, 796–803 (2013).

452. Julien, J. P. *et al.* Crystal structure of a soluble cleaved HIV-1 envelope trimer. *Science (80-. ).* **342**, 1477–1483 (2013).

453. Pejchal, R. *et al.* A Potent and Broad Neutralizing Antibody Recognizes and Penetrates the HIV Glycan Shield. *Science (80-. ).* **334**, 1097–1103 (2011).

454. Pancera, M. *et al.* Crystal Structure of PG16 and Chimeric Dissection with Somatically Related PG9: Structure-Function Analysis of Two Quaternary-Specific Antibodies That Effectively

Neutralize HIV-1. *J. Virol.* **84**, 8098–8110 (2010).

455. Silver, Z. A., Dickinson, G. M., Seaman, M. S. & Desrosiers, R. C. A Highly Unusual V1 Region of Env in an Elite Controller of HIV Infection. *J. Virol.* **93**, 1–13 (2019).

456. Kerkhof, T. L. G. M. *et al.* HIV-1 escapes from N332-directed antibody neutralization in an elite neutralizer by envelope glycoprotein elongation and introduction of unusual disulfide bonds. *Retrovirology* **13**, 1–19 (2016).

457. Deshpande, S. *et al.* HIV-1 clade C escapes broadly neutralizing autologous antibodies with N332 glycan specificity by distinct mechanisms. *Retrovirology* **13**, 1–9 (2016).

458. Williams, K. L. *et al.* Superinfection Drives HIV Neutralizing Antibody Responses from Several B Cell Lineages that Contribute to a Polyclonal Repertoire. *Cell Rep.* **23**, 682–691 (2018).

459. Piantadosi, A. *et al.* Breadth of Neutralizing Antibody Response to Human Immunodeficiency Virus Type 1 Is Affected by Factors Early in Infection but Does Not Influence Disease Progression. *J. Virol.* **83**, 10269–10274 (2009).

460. Sather, D. N. *et al.* Factors Associated with the Development of Cross-Reactive Neutralizing Antibodies during Human Immunodeficiency Virus Type 1 Infection. *J. Virol.* **83**, 757–769 (2009).

461. Moore, P. L. *et al.* Multiple Pathways of Escape from HIV Broadly Cross-Neutralizing V2-Dependent Antibodies. *J. Virol.* **87**, 4882–4894 (2013).

462. Moore, P. L. *et al.* Potent and Broad Neutralization of HIV-1 Subtype C by Plasma Antibodies Targeting a Quaternary Epitope Including Residues in the V2 Loop. *J. Virol.* **85**, 3128–3141 (2011).

463. Moore, P. L. *et al.* Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. *Nat. Med.* **18**, 1688–1692 (2012).

464. Sok, D. *et al.* Promiscuous Glycan Site Recognition by Antibodies to the High-Mannose Patch of gp120 Broadens Neutralization of HIV. *Sci. Transl. Med.* **6**, 1–7 (2014).

465. Köhler, G., Milstein, C. Cultivos continuos de células fusionadas que secretan anticuerpos de especificidad predefinida. *Nature* **256**, 495–497 (1975).

466. Pedrioli, A. & Oxenius, A. Single B cell technologies for monoclonal antibody discovery. *Trends Immunol.* **42**, 1143–1158 (2021).

467. Smith, G. P. Filamentous fusion phage: Novel expression vectors that display cloned antigens on the virion surface. *Science (80-. ).* **228**, 1315–1317 (1985).

468. Bazan, J., Całkosiński, I. & Gamian, A. Phage display—A powerful technique for immunotherapy. *Hum. Vaccin. Immunother.* **8**, 1817–1828 (2012).

469. Liu, Y., Cao, W., Sun, M. & Li, T. Broadly neutralizing antibodies for HIV-1: efficacies, challenges and opportunities. *Emerg. Microbes Infect.* **9**, 194–206 (2020).

470. Mascola, J. R. *et al.* Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. *Nat. Med.* **6**, 207–210 (2000).

471. Klein, F. *et al.* Antibodies in HIV-1 Vaccine Development and Therapy. *Science (80-. ).* **341**, 1199–1204 (2013).

472. Sok, D. *et al.* The Effects of Somatic Hypermutation on Neutralization and Binding in the

PGT121 Family of Broadly Neutralizing HIV Antibodies. *PLoS Pathog.* **9**, 1003754–1003767 (2013).

473. Wang, M. *et al.* Quantifying CD4 receptor protein in two human CD4+ lymphocyte preparations for quantitative flow cytometry. *Clin. Proteomics* **11**, 1–10 (2014).

474. Wang, Q. *et al.* Stabilized diverse HIV-1 envelope trimers for vaccine design. *Emerg. Microbes Infect.* **9**, 775–786 (2020).

475. Moore, D. K. *et al.* Isolation of B-cells using Miltenyi MACS bead isolation kits. *PLoS One* **14**, 1–15 (2019).

476. Laoharawee, K. *et al.* Genome Engineering of Primary Human B Cells Using CRISPR/Cas9. *J. Vis. Exp.* **176**, 139–148 (2020).

477. Yang, Q., Li, C., Wei, Y., Huang, W. & Wang, L.-X. Expression, Glycoform Characterization, and Antibody-Binding of HIV-1 V3 Glycopeptide Domain Fused with Human IgG1-Fc. *Bioconjug. Chem.* **21**, 875–883 (2010).

478. Chen, F., Tzarum, N., Wilson, I. A. & Law, M. VH1-69 antiviral broadly neutralizing antibodies: genetics, structures, and relevance to rational vaccine design. *Curr. Opin. Virol.* **34**, 149–159 (2019).

479. Huang, C. C. *et al.* Structural basis of tyrosine sulfation and VH-gene usage in antibodies that recognize the HIV type 1 coreceptor-binding site on gp120. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 2706–2711 (2004).

480. Williams, W. B. *et al.* Diversion of HIV-1 vaccine–induced immunity by gp41-microbiota cross-reactive antibodies. *Science (80-. ).* **349**, 139–148 (2015).

481. Haynes, B. F. *et al.* Strategies for HIV-1 vaccines that induce broadly neutralizing antibodies. *Nat. Rev. Immunol.* **1**, 1–17 (2022).

482. Zhang, L. *et al.* An MPER antibody neutralizes HIV-1 using germline features shared among donors. *Nat. Commun.* **10**, 5389 (2019).

483. Smith, S. A. *et al.* VH1-69 utilizing antibodies are capable of mediating non-neutralizing Fc-mediated effector functions against the transmitted/founder gp120. *Front. Immunol.* **10**, 1–19 (2019).

484. Zheng, J. SARS-CoV-2: An emerging coronavirus that causes a global threat. *Int. J. Biol. Sci.* **16**, 1678–1685 (2020).

485. Pickering, S. *et al.* Comparative assessment of multiple COVID-19 serological technologies supports continued evaluation of point-of-care lateral flow assays in hospital and community healthcare settings. *PLoS Pathog.* **16**, 1008817–1008825 (2020).

486. Zou, L. *et al.* SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *N. Engl. J. Med.* **382**, 1177–1179 (2020).

487. To, K. K.-W. *et al.* Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect. Dis.* **20**, 565–574 (2020).

488. Amanat, F. *et al.* A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036 (2020).

489. Long, Q. X. *et al.* Antibody responses to SARS-CoV-2 in patients with COVID-19. *Nat. Med.* **26**,

845–848 (2020).

490. Brouwer, P. J. M. *et al.* Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability. *Science (80-. ).* **369**, 643–650 (2020).

491. Cao, Y. *et al.* Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* **182**, 73–84 (2020).

492. Rogers, T. F. *et al.* Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. *Science (80-. ).* **369**, 956–963 (2020).

493. Luchsinger, L. L. *et al.* Serological assays estimate highly variable SARS-CoV-2 neutralizing antibody activity in recovered COVID-19 patients. *J. Clin. Microbiol.* **58**, 2005–2020 (2020).

494. Okba, N. M. A. *et al.* Severe Acute Respiratory Syndrome Coronavirus 2-Specific Antibody Responses in Coronavirus Disease Patients. *Emerg. Infect. Dis.* **26**, 1478–1488 (2020).

495. Crawford, K. H. D. *et al.* Dynamics of Neutralizing Antibody Titers in the Months after Severe Acute Respiratory Syndrome Coronavirus 2 Infection. *J. Infect. Dis.* **223**, 197–205 (2021).

496. Wajnberg, A. *et al.* Robust neutralizing antibodies to SARS-CoV-2 infection persist for months. *Science (80-. ).* **370**, 1227–1230 (2020).

497. Callow, K. A., Parry, H. F., Sergeant, M. & Tyrrell, D. A. J. The time course of the immune response to experimental coronavirus infection of man. *Epidemiol. Infect.* **105**, 435–446 (1990).

498. Mo, H. *et al.* Longitudinal profile of antibodies against SARS-coronavirus in SARS patients and their clinical significance. *Respirology* **11**, 49–53 (2006).

499. Kellam, P. & Barclay, W. The dynamics of humoral immune responses following SARS-CoV-2 infection and the potential for reinfection. *J. Gen. Virol.* **101**, 791–797 (2020).

500. Cao, W.-C., Liu, W., Zhang, P.-H., Zhang, F. & Richardus, J. H. Disappearance of Antibodies to SARS-Associated Coronavirus after Recovery. *N. Engl. J. Med.* **357**, 1162–1163 (2007).

501. Edridge, A. W. D. *et al.* Seasonal coronavirus protective immunity is short-lasting. *Nat. Med.* **26**, 1691–1693 (2020).

502. Wang, M. *et al.* Antibody dynamics of 2009 influenza A (H1N1) virus in infected patients and vaccinated people in China. *PLoS One* **6**, 16809–16815 (2011).

503. Grehan, K., Ferrara, F. & Temperton, N. An optimised method for the production of MERS-CoV spike expressing viral pseudotypes. *MethodsX* **2**, 379–384 (2015).

504. Sterlin, D. *et al.* IgA dominates the early neutralizing antibody response to SARS-CoV-2. *Sci. Transl. Med.* **13**, 968–970 (2021).

505. Schmidt, F. *et al.* Measuring SARS-CoV-2 neutralizing antibody activity using pseudotyped and chimeric viruses. *J. Exp. Med.* **217**, 1–15 (2020).

506. Ju, B. *et al.* Human neutralizing antibodies elicited by SARS-CoV-2 infection. *Nature* **584**, 115–119 (2020).

507. Laing, A. G. *et al.* A dynamic COVID-19 immune signature includes associations with poor prognosis. *Nat. Med.* **26**, 1623–1635 (2020).

508. Wu, F. *et al.* Evaluating the Association of Clinical Characteristics with Neutralizing Antibody Levels in Patients Who Have Recovered from Mild COVID-19 in Shanghai, China. *JAMA Intern.*

*Med.* **180**, 1356–1362 (2020).

509. Chen, W. *et al.* SARS-CoV-2 neutralizing antibody levels are correlated with severity of COVID-19 pneumonia. *Biomed. Pharmacother.* **130**, 110629–110641 (2020).

510. Legros, V. *et al.* A longitudinal study of SARS-CoV-2-infected patients reveals a high correlation between neutralizing antibodies and COVID-19 severity. *Cell. Mol. Immunol.* **18**, 318–327 (2021).

511. Pang, N. Y. L., Pang, A. S. R., Chow, V. T. & Wang, D. Y. Understanding neutralising antibodies against SARS-CoV-2 and their implications in clinical practice. *Mil. Med. Res.* **8**, 1–17 (2021).

512. Batra, M. *et al.* Role of IgG against N-protein of SARS-CoV2 in COVID19 clinical outcomes. *Sci. Rep.* **11**, 1–9 (2021).

513. Muecksch, F. *et al.* Longitudinal serological analysis and neutralizing antibody levels in coronavirus disease 2019 convalescent patients. *J. Infect. Dis.* **223**, 389–398 (2021).

514. Wang, K. *et al.* Longitudinal Dynamics of the Neutralizing Antibody Response to Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection. *Clin. Infect. Dis.* **73**, 531–539 (2021).

515. Isho, B. *et al.* Persistence of serum and saliva antibody responses to SARS-CoV-2 spike antigens in COVID-19 patients. *Sci. Immunol.* **5**, 1–20 (2020).

516. Goel, R. R. *et al.* Distinct antibody and memory B cell responses in SARSCoV-2 naïve and recovered individuals following mRNA vaccination. *Sci. Immunol.* **6**, 1–19 (2021).

517. Dan, J. M. *et al.* Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science (80-. ).* **371**, 4063–4078 (2021).

518. Lechmere, T. *et al.* Broad Neutralization of SARS-CoV-2 Variants, Including Omicron, following Breakthrough Infection with Delta in COVID-19-Vaccinated Individuals. *MBio* **13**, 3798–3821 (2022).

519. Sekine, T. *et al.* Robust T Cell Immunity in Convalescent Individuals with Asymptomatic or Mild COVID-19. *Cell* **183**, 158-168.e14 (2020).

520. Dupont, L. *et al.* Neutralizing antibody activity in convalescent sera from infection in humans with SARS-CoV-2 and variants of concern. *Nat. Microbiol.* **6**, 1433–1442 (2021).

521. Chia, W. N. *et al.* Dynamics of SARS-CoV-2 neutralising antibody responses and duration of immunity: a longitudinal study. *The Lancet Microbe* **2**, 240–249 (2021).

522. Anand, S. P. *et al.* Longitudinal analysis of humoral immunity against SARS-CoV-2 Spike in convalescent individuals up to 8 months post-symptom onset. *Cell Reports Med.* **2**, 100290–100311 (2021).

523. Cohen, K. W. *et al.* Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Reports Med.* **2**, 100354–100378 (2021).

524. Cromer, D. *et al.* Prospects for durable immune control of SARS-CoV-2 and prevention of reinfection. *Nat. Rev. Immunol.* **21**, 395–404 (2021).

525. Gaebler, C. *et al.* Evolution of antibody immunity to SARS-CoV-2. *Nature* **591**, 639–644 (2021).

526. Petersen, E. *et al.* Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics. *Lancet Infect. Dis.* **20**, 238–244 (2020).

527. Gorse, G. J., Donovan, M. M. & Patel, G. B. Antibodies to coronaviruses are higher in older compared with younger adults and binding antibodies are more sensitive than neutralizing antibodies in identifying coronavirus-associated illnesses. *J. Med. Virol.* **92**, 512–517 (2020).

528. Mao, Y., Wang, W., Ma, J., Wu, S. & Sun, F. Reinfection rates among patients previously infected by SARS-CoV-2: Systematic review and meta-analysis. *Chin. Med. J. (Engl).* **135**, 145–152 (2022).

529. Shi, R. *et al.* A human neutralizing antibody targets the receptor-binding site of SARS-CoV-2. *Nature* **584**, 120–124 (2020).

530. Chandrashekar, A. *et al.* SARS-CoV-2 infection protects against rechallenge in rhesus macaques. *Science (80-. ).* **369**, 812–817 (2020).

531. Deng, W. *et al.* Primary exposure to SARS-CoV-2 protects against reinfection in rhesus macaques. *Science (80-. ).* **369**, 818–823 (2020).

532. Corbett, K. S. *et al.* Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. *N. Engl. J. Med.* **383**, 1544–1555 (2020).

533. Vogel, A. B. *et al.* BNT162b vaccines protect rhesus macaques from SARS-CoV-2. *Nature* **592**, 283–289 (2021).

534. Earle, K. A. *et al.* Evidence for antibody as a protective correlate for COVID-19 vaccines. *Vaccine* **39**, 4423–4428 (2021).

535. Khoury, D. S. *et al.* Neutralizing antibody levels are highly predictive of immune protection from symptomatic SARS-CoV-2 infection. *Nat. Med.* **27**, 1205–1211 (2021).

536. Seow, J. *et al.* Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans. *Nat. Microbiol.* **5**, 1598–1607 (2020).

537. Piccoli, L. *et al.* Mapping Neutralizing and Immunodominant Sites on the SARS-CoV-2 Spike Receptor-Binding Domain by Structure-Guided High-Resolution Serology. *Cell* **183**, 1024-1042.e21 (2020).

538. Greaney, A. J. *et al.* Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **29**, 463–476 (2021).

539. Tortorici, M. A. *et al.* Ultrapotent human antibodies protect against SARS-CoV-2 challenge via multiple mechanisms. *Science (80-. ).* **370**, 950–957 (2020).

540. Barnes, C. O. *et al.* SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **588**, 682–687 (2020).

541. Liu, L. *et al.* Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450–456 (2020).

542. Chi, X. *et al.* A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2. *Science (80-. ).* **369**, 650–655 (2020).

543. Korber, B. *et al.* Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell* **182**, 812–827 (2020).

544. Li, Q. *et al.* The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* **182**, 1284–1294 (2020).

545. Weissman, D. *et al.* D614G Spike Mutation Increases SARS CoV-2 Susceptibility to Neutralization. *Cell Host Microbe* **29**, 23–31 (2021).

546. Wang, P. *et al.* Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130–135 (2021).

547. Frampton, D. *et al.* Genomic characteristics and clinical effect of the emergent SARS-CoV-2 B.1.1.7 lineage in London, UK: a whole-genome sequencing and hospital-based cohort study. *Lancet Infect. Dis.* **21**, 1246–1256 (2021).

548. Meng, B. *et al.* Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. *Cell Rep.* **35**, 109292–109313 (2021).

549. Kemp, S. A. *et al.* SARS-CoV-2 evolution during treatment of chronic infection. *Nature* **592**, 277–282 (2021).

550. McCarthy, K. R. *et al.* Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science (80-. ).* **371**, 1139–1142 (2021).

551. Avanzato, V. A. *et al.* Case Study: Prolonged Infectious SARS-CoV-2 Shedding from an Asymptomatic Immunocompromised Individual with Cancer. *Cell* **183**, 1901–1912 (2020).

552. Collier, D. A. *et al.* Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. *Nature* **593**, 136–141 (2021).

553. Edara, V. V., Hudson, W. H., Xie, X., Ahmed, R. & Suthar, M. S. Neutralizing Antibodies Against SARS-CoV-2 Variants After Infection and Vaccination. *JAMA* **325**, 1896 (2021).

554. Rees-Spear, C. *et al.* The effect of spike mutations on SARS-CoV-2 neutralization. *Cell Rep.* **34**, 108890–108911 (2021).

555. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science (80-. ).* **367**, 1260–1263 (2020).

556. Briney, B., Inderbitzin, A., Joyce, C. & Burton, D. R. Commonality despite exceptional diversity in the baseline human antibody repertoire. *Nature* **566**, 393–397 (2019).

557. Robbiani, D. F. *et al.* Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437–442 (2020).

558. Kim, S. Il *et al.* Stereotypic neutralizing VHantibodies against SARS-CoV-2 spike protein receptor binding domain in patients with COVID-19 and healthy individuals. *Sci. Transl. Med.* **13**, 1–15 (2021).

559. Andreano, E. *et al.* Extremely potent human monoclonal antibodies from COVID-19 convalescent patients. *Cell* **184**, 1821–1835 (2021).

560. McCallum, M. *et al.* N-terminal domain antigenic mapping reveals a site of vulnerability for SARS-CoV-2. *Cell* **184**, 2332–2347 (2021).

561. Ter Meulen, J. *et al.* Human monoclonal antibody combination against SARS coronavirus: Synergy and coverage of escape mutants. *PLoS Med.* **3**, 1071–1079 (2006).

562. Yuan, M., Liu, H., Wu, N. C. & Wilson, I. A. Recognition of the SARS-CoV-2 receptor binding domain by neutralizing antibodies. *Biochem. Biophys. Res. Commun.* **538**, 192–203 (2021).

563. Rosa, A. *et al.* SARS-CoV-2 can recruit a heme metabolite to evade antibody immunity. *Sci. Adv.* **7**, 1–15 (2021).

564. Seow, J. *et al.* ChAdOx1 nCoV-19 (AZD1222) vaccine elicits monoclonal antibodies with potent cross-neutralizing activity against SARS-CoV-2 viral variants. *Cell Rep.* **39**, 110757–110780 (2022).

565. Doores, K. J. & Burton, D. R. Variable Loop Glycan Dependency of the Broad and Potent HIV-1-Neutralizing Antibodies PG9 and PG16. *J. Virol.* **84**, 10510–10521 (2010).

566. Yang, Q. *et al.* Inhibition of SARS-CoV-2 viral entry upon blocking N-and O-glycan elaboration. *Elife* **9**, 1–44 (2020).

567. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**, 281–292 (2020).

568. Yuan, M. *et al.* A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science (80-. ).* **368**, 630–633 (2020).

569. Li, W. *et al.* Rapid identification of a human antibody with high prophylactic and therapeutic efficacy in three animal models of SARS-CoV-2 infection. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 29832–29838 (2020).

570. Yurkovetskiy, L. *et al.* Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **183**, 739–751 (2020).

571. Wibmer, C. K. *et al.* SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nat. Med.* **27**, 622–625 (2021).

572. Shah, P., Canziani, G. A., Carter, E. P. & Chaiken, I. The Case for S2: The Potential Benefits of the S2 Subunit of the SARS-CoV-2 Spike Protein as an Immunogen in Fighting the COVID-19 Pandemic. *Front. Immunol.* **12**, 1–7 (2021).

573. Fielding, C. A. *et al.* SARS-CoV-2 host-shutoff impacts innate NK cell functions, but antibody-dependent NK activity is strongly activated through non-spike antibodies. *Elife* **11**, 74489–74499 (2022).

574. Li, C. *et al.* Neutralizing Monoclonal Antibodies Inhibit SARS-CoV-2 Infection through Blocking Membrane Fusion. *Microbiol. Spectr.* **10**, 1814–1821 (2022).

575. Chen, R. E. *et al.* Resistance of SARS-CoV-2 variants to neutralization by monoclonal and serum-derived polyclonal antibodies. *Nat. Med.* **27**, 717–726 (2021).

576. Weisblum, Y. *et al.* Escape from neutralizing antibodies 1 by SARS-CoV-2 spike protein variants. *Elife* **9**, 61312–61326 (2020).

577. Starr, T. N. *et al.* Prospective mapping of viral mutations that escape antibodies used to treat COVID-19. *Science (80-. ).* **371**, 850–854 (2021).

578. Supasa, P. *et al.* Reduced neutralization of SARS-CoV-2 B.1.1.7 variant by convalescent and vaccine sera. *Cell* **184**, 2201–2211 (2021).

579. Schmidt, F. *et al.* Plasma Neutralization of the SARS-CoV-2 Omicron Variant. *N. Engl. J. Med.* **386**, 599–601 (2022).

580. Graham, C. *et al.* Neutralization potency of monoclonal antibodies recognizing dominant and subdominant epitopes on SARS-CoV-2 Spike is impacted by the B.1.1.7 variant. *Immunity* **54**, 1276–1289 (2021).

581. Du, L., Yang, Y. & Zhang, X. Neutralizing antibodies for the prevention and treatment of COVID-19. *Cell. Mol. Immunol.* **18**, 2293–2306 (2021).

582. Cerutti, G. *et al.* Potent SARS-CoV-2 neutralizing antibodies directed against spike N-terminal domain target a single supersite. *Cell Host Microbe* **29**, 819–833 (2021).

583. Anderson, E. M. *et al.* Seasonal human coronavirus antibodies are boosted upon SARS-CoV-2 infection but not associated with protection. *Cell* **184**, 1858–1864 (2021).

584. Beaudoin-Bussières, G. *et al.* A Fc-enhanced NTD-binding non-neutralizing antibody delays virus spread and synergizes with a nAb to protect mice from lethal SARS-CoV-2 infection. *Cell Rep.* **38**, 110368–110377 (2022).

585. Corti, D., Purcell, L. A., Snell, G. & Veesler, D. Tackling COVID-19 with neutralizing monoclonal antibodies. *Cell* **184**, 3086–3108 (2021).

586. Dejnirattisai, W. *et al.* The antigenic anatomy of SARS-CoV-2 receptor binding domain. *Cell* **184**, 2183–2200 (2021).

587. Alter, G. *et al.* Immunogenicity of Ad26.COV2.S vaccine against SARS-CoV-2 variants in humans. *Nature* **596**, 268–272 (2021).

588. Monin, L. *et al.* Safety and immunogenicity of one versus two doses of the COVID-19 vaccine BNT162b2 for patients with cancer: interim analysis of a prospective observational study. *Lancet Oncol.* **22**, 765–778 (2021).

589. Emary, K. R. W. *et al.* Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B.1.1.7): an exploratory analysis of a randomised controlled trial. *Lancet* **397**, 1351–1362 (2021).

590. Lopez Bernal, J. *et al.* Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *N. Engl. J. Med.* **385**, 585–594 (2021).

591. Madhi, S. A. *et al.* Efficacy of the ChAdOx1 nCoV-19 Covid-19 Vaccine against the B.1.351 Variant. *N. Engl. J. Med.* **384**, 1885–1898 (2021).

592. Payne, R. P. *et al.* Immunogenicity of standard and extended dosing intervals of BNT162b2 mRNA vaccine. *Cell* **184**, 5699–5714 (2021).

593. Barros-Martins, J. *et al.* Immune responses against SARS-CoV-2 variants after heterologous and homologous ChAdOx1 nCoV-19/BNT162b2 vaccination. *Nat. Med.* **27**, 1525–1529 (2021).

594. Amanat, F. *et al.* SARS-CoV-2 mRNA vaccination induces functionally diverse antibodies to NTD, RBD, and S2. *Cell* **184**, 3936–3948 (2021).

595. Cho, A. *et al.* Anti-SARS-CoV-2 receptor-binding domain antibody evolution after mRNA vaccination. *Nature* **600**, 517–522 (2021).

596. Wang, Z. *et al.* mRNA vaccine-elicited antibodies to SARS-CoV-2 and circulating variants. *Nature* **592**, 616–622 (2021).

597. Raybould, M. I. J., Kovaltsuk, A., Marks, C. & Deane, C. M. CoV-AbDab: The coronavirus antibody database. *Bioinformatics* **37**, 734–735 (2021).

598. Siu, J. H. Y. *et al.* Two subsets of human marginal zone B cells resolved by global analysis of lymphoid tissues and blood. *Sci. Immunol.* **7**, 9060–9082 (2022).

599. Cele, S. *et al.* Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature* **602**, 654–656 (2022).

600. Garcia-Beltran, W. F. *et al.* mRNA-based COVID-19 vaccine boosters induce neutralizing immunity against SARS-CoV-2 Omicron variant. *Cell* **185**, 457–466 (2022).

601. Cao, Y. *et al.* Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* **602**, 657–663 (2022).

602. Clark, S. A. *et al.* SARS-CoV-2 evolution in an immunocompromised host reveals shared neutralization escape mechanisms. *Cell* **184**, 2605–2617 (2021).

603. Yuan, M. *et al.* Structural basis of a shared antibody response to SARS-CoV-2. *Science (80-. ).* **369**, 1119–1123 (2020).

604. Wall, E. C. *et al.* AZD1222-induced neutralising antibody activity against SARS-CoV-2 Delta VOC. *Lancet* **398**, 207–209 (2021).

605. Zhou, D. *et al.* Evidence of escape of SARS-CoV-2 variant B.1.351 from natural and vaccine-induced sera. *Cell* **184**, 2348–2361 (2021).

606. Liu, C. *et al.* Reduced neutralization of SARS-CoV-2 B.1.617 by vaccine and convalescent serum. *Cell* **184**, 4220–4236 (2021).

607. Wu, M. *et al.* Three-dose vaccination elicits neutralising antibodies against omicron. *Lancet* **399**, 715–717 (2022).

608. Sheward, D. J. *et al.* Neutralisation sensitivity of the SARS-CoV-2 omicron (B.1.1.529) variant: a cross-sectional study. *Lancet Infect. Dis.* **22**, 813–820 (2022).

609. Planas, D. *et al.* Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **602**, 671–675 (2022).

610. Hurlburt, N. K. *et al.* Structural basis for potent neutralization of SARS-CoV-2 and role of antibody affinity maturation. *Nat. Commun.* **11**, 1–7 (2020).

611. De Mattos Barbosa, M. G. *et al.* IgV somatic mutation of human anti-SARSCoV- 2 monoclonal antibodies governs neutralization and breadth of reactivity. *JCI Insight* **6**, 1–13 (2021).

612. Muecksch, F. *et al.* Affinity maturation of SARS-CoV-2 neutralizing antibodies confers potency, breadth, and resilience to viral escape mutations. *Immunity* **54**, 1853–1868 (2021).

613. Andreano, E. *et al.* Hybrid immunity improves B cells and antibodies against SARS-CoV-2 variants. *Nature* **600**, 530–535 (2021).

614. Kaku, C. I. *et al.* Broad anti–SARS-CoV-2 antibody immunity induced by heterologous ChAdOx1/mRNA-1273 vaccination. *Science (80-. ).* **375**, 1041–1047 (2022).

615. Seydoux, E. *et al.* Analysis of a SARS-CoV-2-Infected Individual Reveals Development of Potent Neutralizing Antibodies with Limited Somatic Mutation. *Immunity* **53**, 98–105 (2020).

616. Shrotri, M. *et al.* Spike-antibody waning after second dose of BNT162b2 or ChAdOx1. *Lancet* **398**, 385–387 (2021).

617. Lim, S. Y. *et al.* Breakthrough infections and waning immune responses with ChAdOx1 nCoV-19 or mRNA vaccine in healthcare workers. *Clin. Transl. Med.* **12**, 1–7 (2022).

618. Goel, R. R. *et al.* Efficient recall of Omicron-reactive B cell memory after a third dose of SARS-CoV-2 mRNA vaccine. *Cell* **185**, 1875–1887 (2022).

619. Gruell, H. *et al.* mRNA booster immunization elicits potent neutralizing serum activity against the SARS-CoV-2 Omicron variant. *Nat. Med.* **28**, 477–480 (2022).

620. Saag, M. S. *et al.* Antiretroviral Drugs for Treatment and Prevention of HIV Infection in Adults: 2020 Recommendations of the International Antiviral Society-USA Panel. *JAMA - J. Am. Med.*

*Assoc.* **324**, 1651–1669 (2020).

621. Cortez, V., Odem-Davis, K., McClelland, R. S., Jaoko, W. & Overbaugh, J. HIV-1 superinfection in women broadens and strengthens the neutralizing antibody response. *PLoS Pathog.* **8**, 1–11 (2012).

622. Powell, R. L. R., Kinge, T. & Nyambi, P. N. Infection by Discordant Strains of HIV-1 Markedly Enhances the Neutralizing Antibody Response against Heterologous Virus. *J. Virol.* **84**, 9415–9426 (2010).

623. Dawa, N. & Narain, J. COVID-19 & HIV/AIDS pandemics: Parallels & lessons. *Indian J. Med. Res.* **153**, 214 (2021).

624. Kristiansen, P. A. *et al.* WHO International Standard for anti-SARS-CoV-2 immunoglobulin. *Lancet* **397**, 1347–1348 (2021).

625. Feng, S. *et al.* Correlates of protection against symptomatic and asymptomatic SARS-CoV-2 infection. *Nat. Med.* **27**, 2032–2040 (2021).

626. Rijkers, G. T. & van Overveld, F. J. The "original antigenic sin" and its relevance for SARS-CoV-2 (COVID-19) vaccination. *Clin. Immunol. Commun.* **1**, 13–16 (2021).

627. Moyano, J. & Aguirre, L. Opioids in the immune system: from experimental studies to clinical practice. *Rev. Assoc. Med. Bras.* **65**, 262–269 (2019).

628. Murphy, K., Travers, P., Walport, M. & Janeway, C. *Janeway's immunobiology*. (Garland science, 2012).

629. Harris, L. J., Skaletsky, E. & McPherson, A. Crystallographic structure of an intact IgG1 monoclonal antibody. *J. Mol. Biol.* **275**, 861–872 (1998).

630. Dunn-Walters, D., Townsend, C., Sinclair, E. & Stewart, A. Immunoglobulin gene analysis as a tool for investigating human immune responses. *Immunol. Rev.* **284**, 132–147 (2018).

631. Perkins, S. J., Nealis, A. S., Sutton, B. J. & Feinstein, A. Solution structure of human and mouse immunoglobulin M by synchrotron X-ray scattering and molecular graphics modelling. A possible mechanism for complement activation. *J. Mol. Biol.* **221**, 1345–1366 (1991).

632. Sun, Z. *et al.* Semi-extended solution structure of human myeloma immunoglobulin D determined by constrained X-ray scattering. *J. Mol. Biol.* **353**, 155–173 (2005).

633. Chen, J. B. *et al.* Structural basis for selective inhibition of immunoglobulin E-receptor interactions by an anti-IgE antibody. *Sci. Rep.* **8**, 1–13 (2018).

634. Niemi, M. *et al.* Molecular Interactions between a Recombinant IgE Antibody and the β-Lactoglobulin Allergen. *Structure* **15**, 1413–1421 (2007).

635. Bonner, A., Almogren, A., Furtado, P. B., Kerr, M. A. & Perkins, S. J. Location of secretory component on the Fc edge of dimeric IgA1 reveals insight into the role of secretory IgA1 in mucosal immunity. *Mucosal Immunol.* **2**, 74–84 (2009).

636. Castro-Nallar, E., Crandall, K. A. & Pérez-Losada, M. Genetic diversity and molecular epidemiology of HIV transmission. *Future Virol.* **7**, 239–252 (2012).

637. Munier, M. L. & Kelleher, A. D. Acutely dysregulated, chronically disabled by the enemy within: T-cell responses to HIV-1 infection. *Immunol. Cell Biol.* **85**, 6–15 (2007).

638. Piai, A. *et al.* Structural basis of transmembrane coupling of the HIV-1 envelope glycoprotein. *Nat. Commun.* **11**, 1–12 (2020).

639. McLellan, J. S. *et al.* Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480**, 336–343 (2011).

640. Zhou, T. *et al.* Structural Basis for Broad and Potent Neutralization of HIV-1 by Antibody VRC01. *Science (80-. ).* **329**, 811–817 (2010).

641. Scharf, L. *et al.* Broadly Neutralizing Antibody 8ANC195 Recognizes Closed and Open States of HIV-1 Env. *Cell* **162**, 1379–1390 (2015).

642. Pancera, M. *et al.* Structure and immune recognition of trimeric pre-fusion HIV-1 Env. *Nature* **514**, 455–461 (2014).

643. Varki, A. *et al. Essentials of glycobiology*. (Cold Spring Harbor Laboratory Press, 2017).

644. Casalino, L. *et al.* Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein. *ACS Cent. Sci.* **6**, 1722–1734 (2020).

645. Wu, N. C. *et al.* An Alternative Binding Mode of IGHV3-53 Antibodies to the SARS-CoV-2 Receptor Binding Domain. *Cell Rep.* **33**, 108274 (2020).

646. Liu, H. *et al.* Cross-Neutralization of a SARS-CoV-2 Antibody to a Functionally Conserved Site Is Mediated by Avidity. *Immunity* **53**, 1272-1280.e5 (2020).

647. Leem, J., Dunbar, J., Georges, G., Shi, J. & Deane, C. M. ABodyBuilder: Automated antibody structure prediction with data–driven accuracy estimation. *MAbs* **8**, 1259–1268 (2016).

648. Irimia, A. *et al.* Lipid interactions and angle of approach to the HIV-1 viral membrane of broadly neutralizing antibody 10E8: Insights for vaccine and therapeutic design. *PLoS Pathog.* **13**, 1–20 (2017).