Within-host evolution and virus escape of two important pandemic viruses

Oluwadamilola Aderonke Collier

UCL

This thesis is submitted for the degree of Doctor of Philosophy (PhD) in the field of Virology

May 2021

Declaration

I, Oluwadamilola Aderonke Collier confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Viruses are known to adapt to the constraints of the niche in which they find themselves. This leads to within-host adaptations, detected as nucleic acid base mutations, protein changes or epigenetic modification. These changes are driven by selection pressures which may be due to immune responses to the virus, drug treatment and requirement to exist in specific compartments within the host.

Cerebrospinal fluid (CSF) escape and central nervous system (CNS) compartmentalisation are problems in HIV-1 that are not fully understood and have implications for HIV-1 treatment and cure. The true extent of this problem in subtype C HIV-1, which is responsible for the epidemic in Southern Africa is unknown. The prevalence of these were determined in a cohort of participants co-infected with HIV-1 and cryptococcal meningitis. Following this, the phenotypic characteristics of CSF and plasma derived clones that may underly the mechanisms by which they adapt to the CNS were investigated.

CSF escape was uncommon in this cross sectional study. Four participants underwent in-depth characterisation of the genomes in both the CSF and plasma. This revealed that 2 out of 4 of them had CNS compartmentalisation according to analyses of their phylogeny. The diversity of the CSF and plasma *env* genomes were different in these individuals and are most likely reflective of their unique within-host selection pressures. All clones tested utilised the CCR5 co-receptor for viral entry with some evidence that CSF derived clones may have a higher affinity for the

CCR5 coreceptor and hence better adapted to infect CNS macrophage with low levels of CD4.

An in-depth analysis of within-host evolution of SARS-CoV-2 in a chronic infection case was undertaken. Escape mutations were identified and their functional relevance were tested in infection and neutralisation experiments. Additionally, factors related to poor immune response to mRNA BNT162b2 SARS-CoV-2 vaccine were investigated.

The findings revealed a dynamic shift in viral population *in vivo* that was likely driven by pressure from treatment with remdesivir and convalescent plasma. *In vitro* testing of emerging Spike mutants revealed that D796H mutation facilitated escape from neutralising antibodies but was accompanied by a reduction in infectivity. Meanwhile, $\triangle$H69/V70 deletions had no impact on the neutralisation activity of convalescent plasma but increased the infectivity of the pseudotyped virus.

Importantly older age i.e $\geq$80 years was a risk factor for lower Spike-specific antibody binding levels, lower neutralisation responses against Spike pseudotyped viruses and lower T cell interferon gamma (IFNγ) and IL-12 responses following the first dose of mRNA BNT162b2 vaccine. However, these poor responses were overcome by the second dose in all participants tested.

These findings will contribute to our understanding of within-host viral evolution and have implications for patients care, vaccine and cure strategies.

Impact Statement

The findings of the work presented herein provides important insights into within-host evolution of two important pandemic viruses; HIV-1 and SARS-CoV-2. It investigates proposed hypotheses of the mechanisms underlying central nervous system escape of HIV-1 from immune control in subtype C HIV-1 participants co-infected with cryptococcal meningitis. Important mutations in the Spike protein of SARS-CoV-2 that are functional adaptions, driving escape from antibody neutralisation are described. Suboptimal vaccine-elicited immune responses could also provide the optimal niche for selecting escape mutants, which may in the long-term jeopardise the vaccine program's efforts. Therefore, the findings from the study into determinants of poor immune responses to SARS-CoV-2 mRNA vaccines will potentially impact on public health strategies to bolster the vaccine program and target key groups in order to prevent vaccine resistant variants from emerging.

These findings are of interest to the academic world because they ultimately teach us something about within-host adaptation of these RNA viruses and what selection pressures are in play in the host. They have implications for vaccine development and cure strategies. They are also of interest to clinicians because they may inform the use of antiviral interventions that could help manage patients better. Finally, they have wider implications for public health as they may inform our understanding of hotspots for emergence of variants that maybe more transmissible or pathogenic and could help prioritise where investment in efforts to reduce transmission should be directed.

Acknowledgements

Firstly, I would like to thank my supervisor Professor Ravindra Gupta for his unwavering support and encouragement. For encouraging me to pursue academic medicine and undertake a PhD in his lab. For investing in my research ideas, opening doors and facilitating their execution. His mentorship has been essential to my progress as a clinician scientist and his mix of brilliance and out of the box thinking has taught me to approach scientific enquiry with alternative hypothesis rather than accept scientific dogma.

I would also like to thank Dr Laura McCoy for her mentorship, counsel and training in laboratory techniques. Without her my PhD would not have been possible. I would like to thank Professor Rob Heyderman for facilitating access to biobanked samples on which a large part of my thesis is based, again without him this PhD would not have been possible.

I would like to express my gratitude to all the members of the Gupta lab past and present, Dr Katherine Sutherland for teaching me basic lab techniques and principles when I was fresh off the wards and in the lab. Drs Petra Mlcochova, Bo Meng, Rawlings Datir and Isabella Ferreira for their unwavering enthusiasm for scientific discovery, advice and for reading drafts of this thesis. I would like to thank Dr Steven Kemp for his patience and enthusiasm in teaching me bioinformatics and helping make the best of my genomic data.

Table of Contents

List of Tables

List of Figures

PREFACE

Two viruses in recent history have managed to successfully jump species and reach pandemic proportions – human immunodeficiency virus (HIV-1) in 1981 and severe acute respiratory syndrome coronavirus- 2 (SARS-CoV-2) in 2019. The adaptation of these viruses to humans and what makes them highly transmissible and pathogenic viruses is a remarkable story. This thesis focuses on within-host adaptations that makes these viruses successful in evading the host immune system and establishing prolonged infection. It also examines the heterogeneity in vaccine responses, focusing on the impact of age on the adaptive immune response to SARS-CoV-2 vaccination with the Pfzier BNT162b2 vaccine.

Some of the research undertaken during the course of this PhD and presented in this thesis have been published in peer-reviewed journals and are listed below.

Some of the research presented was done in collaboration with other scientists and clinicians and where this is the case, it is acknowledged in the methods and figure legends. However, I have only presented work where the conception, experimental plan and interpretation of the data were my own.

List of Publications

- Collier DA, Ferreira IATM, Kotagiri P, Datir R, Lim E, Touizer E, et al. Age-related immune response heterogeneity to SARS-CoV-2 vaccine BNT162b2. Nature. 2021.

- Collier DA, De Marco A, Ferreira IATM, Meng B, Datir RP, Walls AC, et al. Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies. Nature. 2021.

- Kemp SA, Collier DA, Datir RP, Ferreira I, Gayed S, Jahun A, et al. SARS-CoV-2 evolution during treatment of chronic infection. Nature. 2021;592(7853):277-82.

- Collier DA, Assennato SM, Warne B, Sithole N, Sharrocks K, Ritchie A, et al. Point of Care Nucleic Acid Testing for SARS-CoV-2 in Hospitalized Patients: A Clinical Validation Trial and Implementation Study. Cell Rep Med. 2020;1(5):100062.

- Mlcochova P, Collier D, Ritchie A, Assennato SM, Hosmillo M, Goel N, et al. Combined Point-of-Care Nucleic Acid and Antibody Testing for SARS-CoV-2 following Emergence of D614G Spike Variant. Cell Rep Med. 2020;1(6):100099.

- Collier DA, Monit C, Gupta RK. The Impact of HIV-1 Drug Escape on the Global Treatment Landscape. Cell Host Microbe. 2019;26(1):48-60.

- Collier DA, Haddow L, Brijkumar J, Moosa MS, Benjamin L, Gupta RK. HIV Cerebrospinal Fluid Escape and Neurocognitive Pathology in the Era of Combined Antiretroviral Therapy: What Lies Beneath the Tip of the Iceberg in Sub-Saharan Africa? Brain Sci. 2018;8(10).

- Kugathasan R, Collier DA, Haddow LJ, El Bouzidi K, Edwards SG, Cartledge JD, et al. Diffuse White Matter Signal Abnormalities on Magnetic Resonance Imaging Are Associated With Human Immunodeficiency Virus Type 1 Viral Escape in the Central Nervous System Among Patients With Neurological Symptoms. Clin Infect Dis. 2017;64(8):1059-65.

CHAPTER 1: BACKGROUND AND INTRODUCTION

PART 1

1.1   The HIV-1 pandemic

The first clinical cases of Acquired Immunodeficiency Syndrome (AIDS) caused by HIV-1 were published as a case series in 1981. Affecting men who have sex with men (MSM) who presented with opportunistic pneumocystis pneumonia (PCP) caused by *Pneumocystis jirovecii* (Centers for Disease 1981). The virus was first isolated from a lymph node of a patient with pre-AIDS by Montagnier and Barre-Sinoussi in 1983 (Barre-Sinoussi et al. 1983). Archived samples tested for HIV-1 showed that the virus was present in the central Africa from as early as 1959-ZR59 and 1960- DRC60 (Worobey et al. 2008). HIV-1 originated as a zoonotic infection that became a human virus. Urine and feacal samples from chimpanzees (*Pan Troglodytes troglodytes*) in central Africa were tested in the 1980s and revealed the presence of simian immunodeficiency virus (SIVcpz) (Sharp and Hahn 2011), which is pathogenic in the natural host. Like HIV-1, it is transmitted sexually, vertically in utero and via blood contamination. Sequencing of SIVcpz and the pandemic HIV-1 subgroup M, showed a common ancestor (D'Arc et al. 2015).

The "cut hunter" hypothesis is proposed as the index case for HIV-1 whereby a hunter was exposed to the bodily fluids of a chimpanzee whilst hunting and butchering it (Pepin 2011). However, others propose more complex interactions with non-human primates that predate colonial incursions into Africa involving

27

agricultural expansion, keeping primates as pets and the extraction of natural resources from the Sangha basin forest (Rupp et al. 2016). Colonialism, population explosion and global development has seen the movement of large numbers of people from rural to urban areas and will have undoubtedly accelerated the spread of HIV-1 in Kinshasa, Democratic republic of Congo. In these urban areas, it is hypothesised HIV-1 was spread by sexual transmission and probably unhealthy medical practices such as using non-sterile equipment (Pepin 2011; Hogan et al. 2016). The global spread of HIV-1 was facilitated by the colonisation of Africa and the return of foreign nationals to their home countries. For example, french speaking doctors from Haiti arrived back to Haiti in the 1960s having been infected in DRC. Archival samples have examined the genomes from Africa and Haiti and shown that the Haitian sequence is the ancestor to the USA HIV-1 sequences (Worobey et al. 2016).

Currently, 38 million people are living with HIV/AIDS of which 26 million are receiving antiretroviral treatment (ART) (UNAIDS 2020). Since the rollout of ART the incidence of HIV-associated mortality has fallen dramatically but still, it was estimated at 690 000 in 2019 (UNAIDS 2020). HIV-1 is treated by a combination of ART with a view to suppressing the virus replication to the point where it cannot be detected by standard diagnostic RT-PCR assays. The classes of drugs available to treat HIV-1 currently in clinical use are detailed in Table 1.1.

Drug Class	Action
Nucleotide/side reverse transcription inhibitor	Competitive inhibitor of viral nucleic acid. Drugs lack a 3'-OH and act as chain terminators when they are incorporated into elongating viral DNA by RT
Non-nucleotide reverse transcription inhibitors	Allosteric inhibition of RT. The binding pocket is located ~10 Å from the polymerase active site
Protease Inhibitors	Competitive inhibitors of the protease enzyme. Bind at the catalytic site of the protease enzyme with high affinity thus blocking activity
Integrase strand transfer inhibitors	Selectively inhibit strand transfer that integrates the proviral DNA into the host's genome
Entry Inhibitors • Fusion inhibitors	Peptide mimetics of HR2. Selectively inhibit the function of gp41
• CCR5 co-receptor antagonist	Allosteric inhibitors of CCR5

Table 1.1 HIV antiretroviral drug classes and mechanism of action.

RT; reverse transcriptase, dNTP; deoxynucleotide triphosphate, CCD; catalytic core domain, CTD; C-terminal domain, NRTI; nucleotide/side reverse transcription inhibitor, NNRTI; non-nucleotide reverse transcription inhibitor, PI; protease inhibitor, INSTI; integrase strand transfer inhibitor, HR1; heptad repeat 1, HR2; heptad repeat 2, gp; glycopeptide, CCR5- β-chemokine receptor. (World Health Organisation 2018)

1.2   Human immunodeficiency Virus-1 Life Cycle

HIV-1 is a single stranded positive RNA lentivirus, which is part of the Retroviridae family. HIV-1 is 100nm in diameter. The HIV-1 genome consists of two identical single-stranded RNA molecules which encodes for the Gag (Group specific antigen), Pol (Polymerase) and Env (Envelope) structural proteins as well as non-structural proteins including Vif (Viral infectivity factor), Vpr (Viral protein R), Vpu (Viral protein U), Tat (Trans-activator of transcription), Rev (Regulator of virion expression) and Nef (Negative regulatory factor) (Figure 1.1).
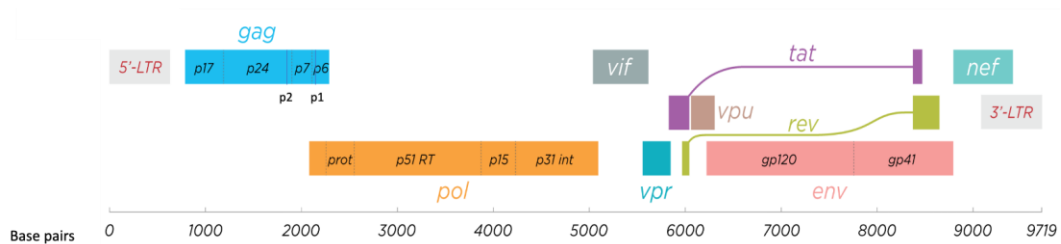
Figure 1.1 HIV-1 genome map.

Structural and non-structural proteins are encoded by the labelled genes. LTR- long terminal repeats, p17-matrix, p24 capsid, p2- spacer peptide 1, p7- nucleocapsid, p1-spacer peptide 2, p6- p6 subunit, prot- protease, P51 RT- reverse transcriptase, P15-RNaseH domain, gp- glycoprotein. Adapted from source: Thomas Splettstoesser (www.scistyle.com), CC BY-SA 3.0

1.2.1 HIV-1 attachment and Entry

The HIV-1 life cycle begins with the attachment of the HIV-1 particle to the host cell

receptor (Figure 1.2). The viral membrane is studded with Envelope (Env) proteins,

which interacts with T cell CD4 receptors. This is then followed by interaction with

a co-receptor, which maybe a β-chemokine receptor (CCR5) or an α-chemokine

receptor (CXCR4) (Figure 1.3). HIV-1 Env protein consists of trimers gp41 and

gp120 subunits. Gp120 is responsible for virion attachment to target cells via the

CD4 receptor, while gp41 mediates fusion between the virus and host cell

membranes (Merk and Subramaniam 2013).
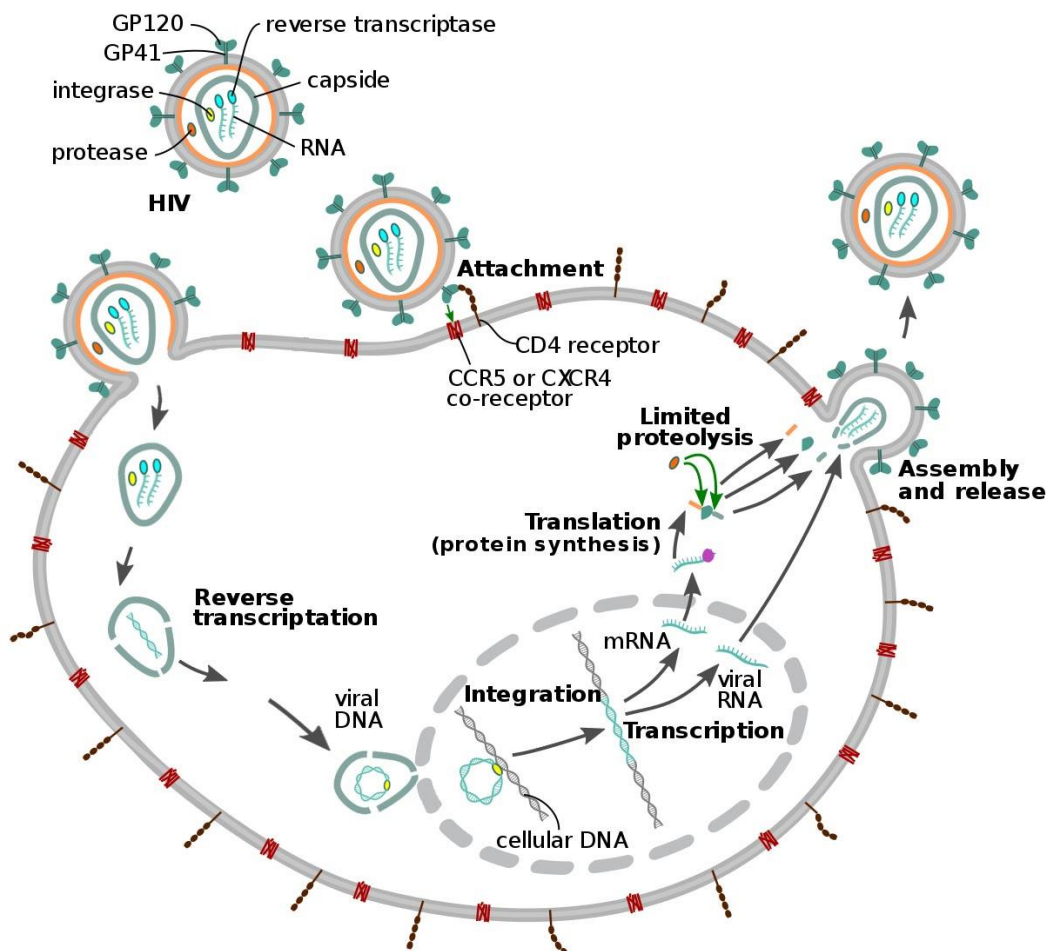
Figure 1.2 HIV-1 life cycle.

Adapted from source: Jmarchn https://commons.wikimedia.org/wiki/File:HIV-replication-cycle-en.svg CC-BY-SA 3.0 Unported.

The trimer surface unit gp120 binds to CD4 from which a phenylalanine at position 43 protrudes which is essential for binding with gp120 (Kwong et al. 1998). Binding of gp120 to CD4 leads to structural rearrangements that expose the co-receptor binding sites and secondary binding to the coreceptor occurs (Dragic et al. 1996). Next the fusion peptide (FP) in the C-Terminal domain of gp41 flips over and inserts into the cell membrane. The protein then hairpins in order to bring the membranes together.
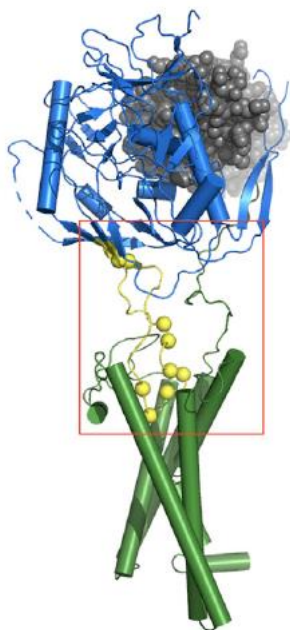
Figure 1.3. HIV-1 envelope binding. HIV-1 gp120 (blue and yellow cartoon) in complex with CD4 (gray spheres) and CCR5 (green cylinders). The V3 loop of gp120 of is shown in yellow. Determined using cryo-electron microscopy by Shaik

et al. (2019). Source: Collier et al. Brain Sci. 2018, 8, 190; doi:10.3390/brainsci8100190.

The importance of CCR5 co-receptor in HIV-1 entry is underlined by the fact that CCR5-delta 32 mutation protects against HIV-1 infection and two patients with exclusively CCR5 co-receptor using HIV-1 who received a bone marrow transplant from a donor with homozygous CCR5-delta 32 deletion following bone marrow ablative chemotherapy and were cured of HIV-1 (Gupta et al. 2019; Hütter et al. 2009).

1.2.2 Reverse Transcription

Within the viral particle is a conical shaped capsid within which reverse transcription of the viral RNA into DNA occurs. Following plasma fusion, the capsid is thought to remain intact in order to protect the viral genetic material from triggering hosts toll-like receptors (TLRs) or pattern recognition receptors (PRRs) and eliciting an innate immune response to HIV-1 (Figure 1.2) (Jacques et al. 2016). The capsid moves through the host cell cytoplasm and recruits microtubules to reach the nucleus (Sodeik, Ebersold, and Helenius 1997) where it docks at a nuclear pore (Jacques et al. 2016).

The HIV-1 particle contains its own reverse transcriptase (RT) that has two enzymatic activities; 1) DNA polymerase enzyme which reverse transcribes viral

RNA into DNA, 2) RNase H that cleaves RNA in complex with RNA or DNA. It is

a heterodimer consisting of 2 subunits; p66 and p51 (Figure 1.4). The p51 subunit is

identical to p66 except for the presence of RNaseH in p66. The fingers, palm, thumb,

connection, and RNase H subdomains of p66 together with the connection and

thumb subdomains of p51 form the nucleic-acid binding cleft or binding pocket. The

process of reverse transcription is described in Figure 1.5 and results in double

stranded DNA that will become integrated into the host DNA.
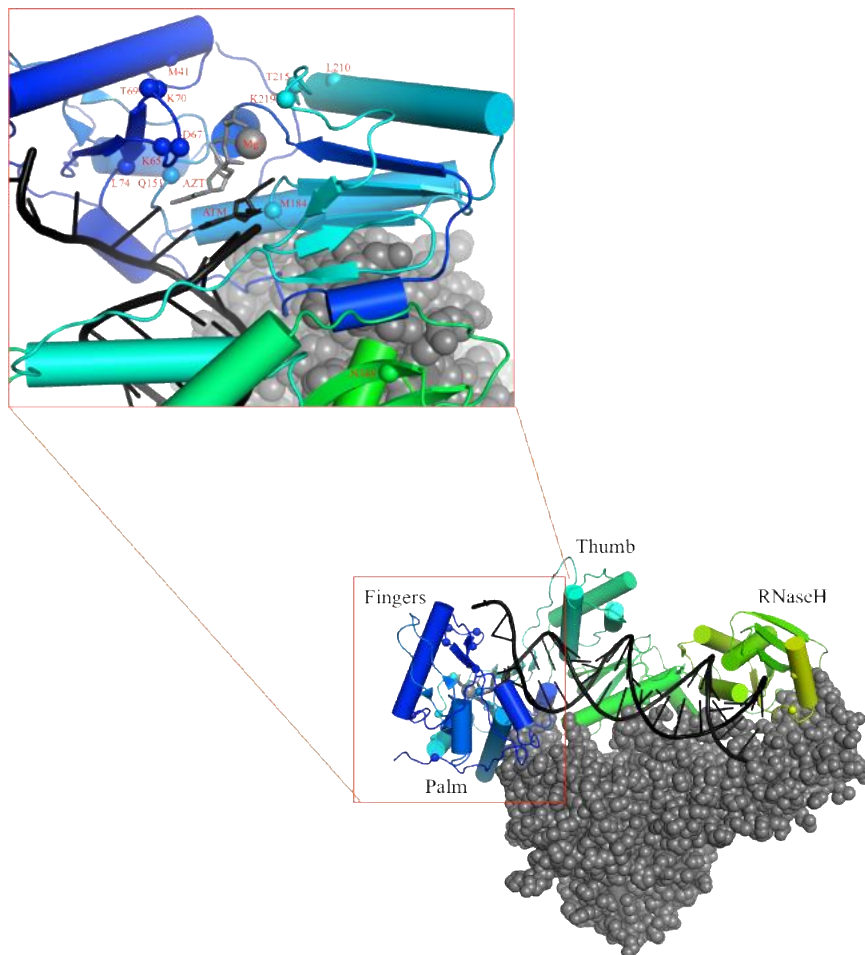
Figure 1.4. Structure of HIV-1 reverse transcriptase. The p66 subunits are

labelled and p51 subunit is shown as grey spheres. N-terminal domain (dark blue),

C-terminal domain (light green). Bound to AZT (grey) and ATM (AZT/DNA

complex; black). The nucleic-acid binding cleft (enlarged image) shows the labelled NRTI drug resistant mutation site as spheres. A divalent Mg^{2+} is shown in the polymerase active site. AZT; zidovudine, ATM; 3'-azido-3'-deoxythymidine-5'-monophosphate, NRTI- nucleoside/tide reverse transcription inhibitors. Determined by X-ray crystallography by Das et al 2012; PDC ID: 5U1C(Das et al. 2012).

Source: Collier et al. Brain Sci. 2018, 8, 190; doi:10.3390/brainsci8100190.
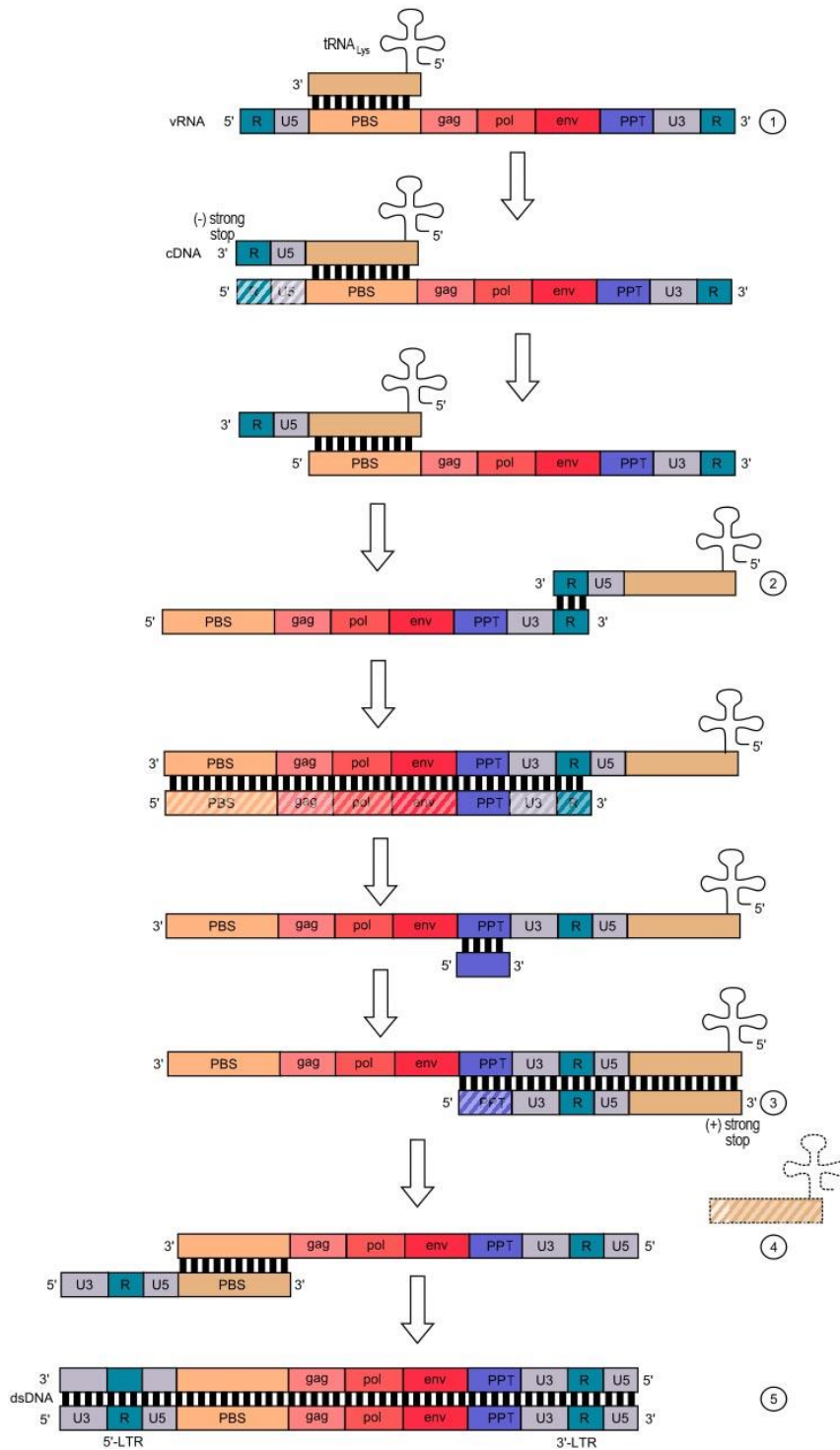
Figure 1.5. HIV-1 reverse transcription.

1) Initiation of negative (-) strand synthesis: the cellular Lysine tRNA (tRNA$_{lys}$) binds to the primer binding site (PBS) in 5` end of the LTR and polymerase

initiates transcription until the (-) strong stop is produced. The RNA template that has been transcribed is broken down by RNaseH.

2) First strand transfer: the (-) strong stop contains an R sequence which is complementary to the R on the 3` end of the RNA template and will continue to copy the RNA in a 3` to 5` direction, including the PBS. The template RNA apart from the polypurine tract (PPT) is degraded by RNAaseH as polymerisation continues.

3) Positive (+) strand synthesis: complementary DNA strand synthesis is primed at the PPT and proceeds in the 5` to 3` direction. It copies the primer site of the bound tRNA to produce the (+) strong stop DNA, complementary to the PBS on the (-) stand DNA. The tRNA and PPT are degraded by RNaseH.

4) Second strand transfer: (+) strand synthesis continues from the (+) strong stop in the 5` to 3` direction all the way to the end of the (-) strand.

5) Strand displacement synthesis: RT extends the 3` end of the (-) strand DNA to the end of the (+) stand to include U3 at the 5` end.

1.2.3 Integration

Double stranded retroviral DNA integrates into the host genome. Integrated viral
DNA is called proviral DNA. This process is catalysed by HIV-1 integrase which
has 2 enzymatic functions 1) strand transfer (ST) which integrates the proviral DNA
into the host's DNA (Pommier, Johnson, and Marchand 2005) and 2) 3′ -end
processing (3′EP) which removes two nucleotides from the 3' at both ends of HIV-1
DNA. The enzyme is encoded by the *pol* gene and is cleaved from the pol
polyprotein by HIV-1 protease enzyme (Figure 1.1). It is a tetramer and has 3
domains: 1) N-terminal domain (NTD), 2) the catalytic core domain (CCD and 3)
the C-terminal domain (CTD), which is the DNA binding site (Lodi et al. 1995)
(Figure 1.6).

Integrase binds other proteins and cofactors to facilitate the integration of the HIV-1
DNA into the host genome. Lens epithelium-derived growth factor (LEDGF) binds
integrase and chromosomal DNA at nucleosomes, which tethers the HIV-1 DNA to
sites of active transcription (Ciuffi et al. 2005). The proviral DNA hijacks the host
transcription machinery to synthesize many copies of viral mRNA by RNA
polymerase II. These mRNAs are subsequently translated into viral proteins.
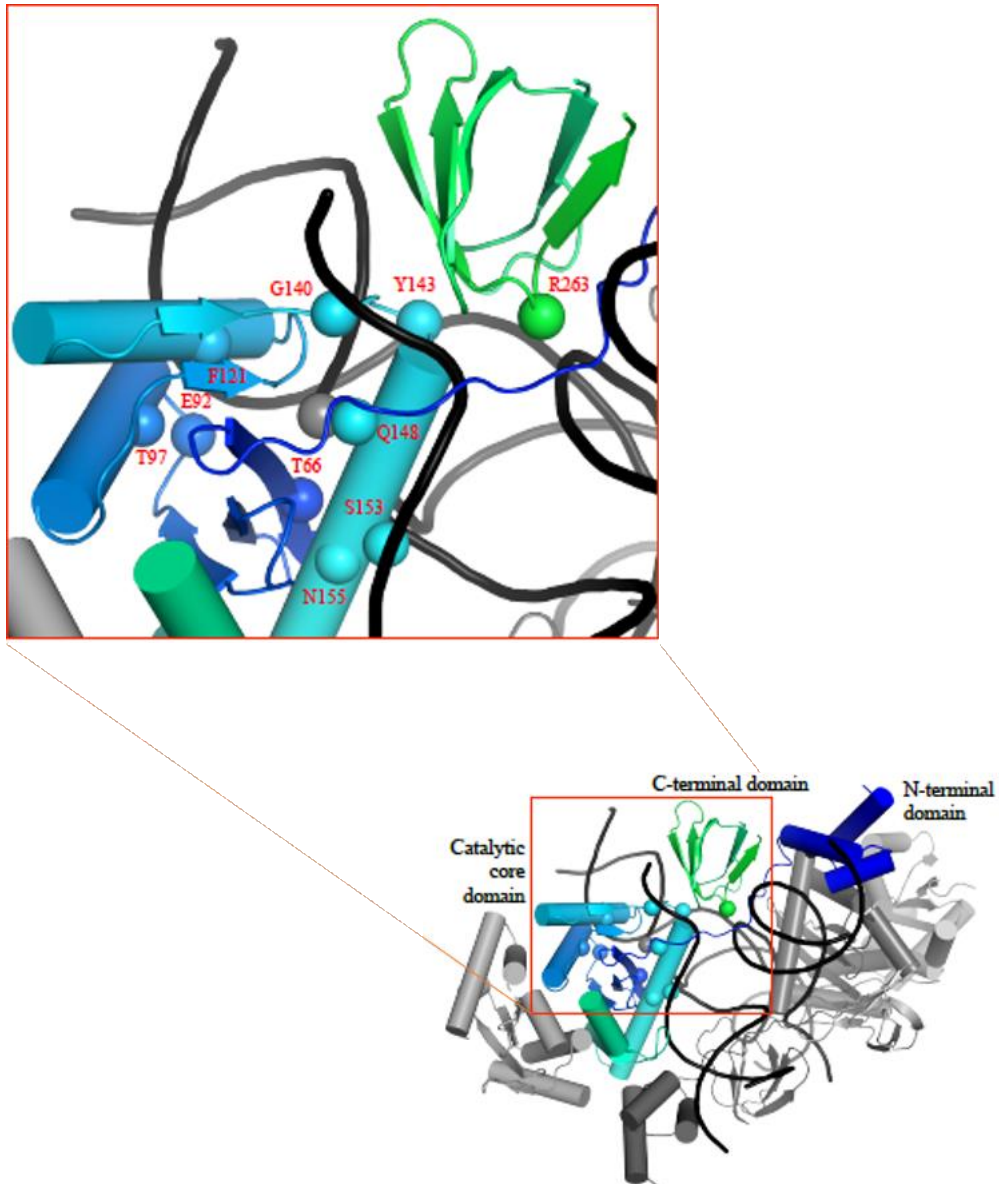
Figure 1.6. Structure of HIV-1 integrase strand transfer complex.

N-terminal domain (dark blue), C-terminal domain (light green), catalytic core domain (cyan), DNA strands (grey and black wires). Residues associated with integrase inhibitor resistance mutations are labelled (coloured spheres). Determined by cryo-electron microscopy by Passos et al 2017; PDB ID: 5U1C(Passos et al. 2017). Source Collier et al. Brain Sci. 2018, 8, 190; doi:10.3390/brainsci8100190.

1.2.4 Transcription and mRNA processing

Following integration into the host genome at transcriptionally active sites, HIV-1 gene expression is facilitated by a highly regulated process that co-opts the host's replication machinery. In the LTR of the provirus is a transcriptional control region that contains two NFκB binding sites between nucleotides -104 and -180, which are important HIV-1 transcription enhancers (Kwon et al. 1998). Also within this region are promoters essential for both basal and Tat trans-activation. It consists of a TATA box, which is a largely conserved region of DNA sequences across HIV-1 subtype that binds TATA binding proteins and is important for RNA synthesis (Montano, Nixon, and Essex 1998) and a transcription initiation site to which the trans-activation response element (TAR) is connected. TAR is a RNA stem loop structure to which the trans-activator protein Tat binds in order to increase the frequency of full length transcripts (Berkhout, Silverman, and Jeang 1989). Tat also directly interacts with the cellular cofactor positive transcription elongation factor b (P-TEFb) which results in phosphorylation of RNA polymerase II and mRNA elongation (Mancebo et al. 1997).

Transcription generates multiply spliced as well as unsliced transcripts which includes the full-length HIV-1 genomic RNA. A spliced mRNA that encodes Rev, a nuclear export protein, is exported out of the nucleus. It is translated and then re-enters the nucleus where it binds to the rev response element (RRE) at the 3` end of the viral genome (Rausch and Le Grice 2015). This complex is recognised by cellular export machinery and allows export of the whole viral genome into the

cytoplasm where two copies of RNA dimerise and form a stable RNA structure that will be packaged into virus particles (Moore and Hu 2009).

1.2.5 Assembly and packaging

Translation of the HIV-1 protein precursor proteins occurs on ribosomal RNAs in the cytoplasm. Polyprotein precursors Gag and Gagpol are made in a 20:1 ratio (Shehu-Xhilaga, Crowe, and Mak 2001). The matrix subunit of Gag has a hydrophobic sequence that is myristoylated, forming a lipid residue that directs Gag to the plasma membrane. The Nucleocapsid of the Gag precursor protein interacts with Psi packaging signal that directs the two genomic RNA molecules into the forming virus particle (Sundquist and Kräusslich 2012). As the internal structures are assembled along the plasma membrane budding occurs simultaneously with the help of the cellular ESCRT (Endosomal sorting complexes required for transport) proteins by interaction between the p6 submit of Gag and proteins of the ESCRT pathway such as ALIX (Bieniasz 2009). After release from the host cell, ordered cleavage of gag results in mature infectious virus ready to infect another host cell (Fun et al. 2012).

1.3   Within-Host adaptations

1.3.1 Inherent viral diversity

HIV-1 is highly diverse. There are 4 groups – M, N, O, P. HIV-1 main group M is responsible for the global pandemic. It consists of 9 subtypes (A, B , C, D, F, G, H, J and K), 6 sub-subtypes (A1-A4 and F1 and F2) and multiple circulating recombinant forms (CRF) and unique recombinant forms (URF) (Los Alamos HIV Sequence Database).

Evolution of the virus is driven by the need to evade host immunity. This is facilitated by a high replication rate (10^{10} virions generated per day in an individual), which is error prone, resulting in one mutation for every 10^5 nucleotides copied (Mansky and Temin 1995). In addition, reverse transcription has no proof-reading mechanism and a high recombination rate serves to increase quasispecies diversity (Song et al. 2018). Consequently, HIV-1 has vast genetic diversity within and between individuals (Kearney et al. 2009). Depending on the quality of ART suppression, genetic diversity maybe broader in the presence of ART compared to drug naïve individuals (Haddad et al. 2000).

1.3.2 The host environment

HIV-1 has mechanisms to evade the host innate immune response and avoid induction of Interferon gamma (IFNγ) (Rasaiyaah et al. 2013). This has probably

facilitated its success as a pandemic virus. In addition, within-host evolution occurs. There is a constant co-evolutionary race between the virus- to escape the adaptive immune response and the host- to match the viruses antigenic evolution. The adaptive immune response includes HIV-specific CD8+ cytotoxic T-lymphocytes (CTLs) that clear infected T cells (Walker et al. 1987). This exerts an intense selection pressure on the viral population, leading to CTL escape mutants (Borrow et al. 1997). These are observed in targeted CTL Gag and Env-specific epitopes (Borrow et al. 1997; Deng et al. 2015; Walker et al. 1987).

The viral Env protein is the target of the humoral adaptive response to HIV-1 (Burton and Mascola 2015). A changing landscape of the Env protein has been described whereby the Env epitopes continually evolves to avoid neutralisation by antibodies (Wu, Wang, et al. 2012). In addition, Env residues acquire N-linked glycosylation which serve to shield the virus from antibody recognition (Wei et al. 2003).

Finally, HIV-1 may evolve in host with certain class I histocompatibility-linked leukocyte antigen (HLA) alleles such as HLA B57 and B27, which confer protection or a better prognosis once exposed to HIV-1 by promoting CTL mediated clearance of infected cells (Gao et al. 2005; Kaslow et al. 1996). This has been associated with CTL escape mutants and interestingly, these alleles are over-represented in elite controllers (Merindol et al. 2018; Migueles et al. 2000).

Some have argued that these evolutionary adaptation are "short-sighted" because although they adapt the virus to chronic infection within a host, they are a disadvantage to onward transmission (Lythgoe et al. 2017).

1.3.3 Latent reservoirs

Cellular and anatomical compartmental reservoirs of latent HIV-1 have been described. It is estimated that 1 in 10^3 resting CD4+ T cells harbours HIV-1 provirus (Bruner et al. 2016) and there is persistence of 10% of infected cells during ART (Besson et al. 2014). Different populations of HIV-1 have been identified in cellular compartments including memory, naive CD4 T cells and CD14 monocytes (Delobel et al. 2005). Long-lived cells such as neurons and macrophages in anatomical compartments such as the CNS, male genital tract and gut, may harbour replication competent HIV-1 (Collier et al. 2018; Ganor et al. 2019). Proviral DNA may lie dormant in these cells and reactivate at a later time (Sigal and Baltimore 2012). Furthermore, viruses in these anatomical reservoirs may undergo low-level replication and continue to replenish the compartment and the peripheral blood (Lorenzo-Redondo, Fryer, Bedford, Kim, Archer, Kosakovsky Pond, et al. 2016).

HIV-1 can also adapt to its cellular tropism to replicate in different cell types and anatomical compartments. This evolution may involve a change in chemokine co-receptor used i.e. CCR5 to CXCR4 shift, or a change from R5-T cell tropism to R5-macrophage tropism (Arrildt et al. 2015). Macrophages in the central nervous system (CNS) have lower levels of CD4 and are targeted by R5-Mac-tropic HIV-1 which

has adapted to infect cells with low levels of CD4 (Joseph et al. 2015). Macrophages can become highly permissive to HIV-1 by the deactivation of the restriction factor-SAMHD1 following G0 to G1 transition (Mlcochova et al. 2017).

The inherent diversity of HIV-1 and the dynamic virus and host co-evolution has implications for the persistence of HIV-1 particularly in latent reservoirs such as the CNS.

1.4   CNS replication and HIV-1

Since the rollout of ART, along with a reduction in HIV-associated mortality, the incidence of HIV-associated central nervous system (CNS) pathology has fallen (Sacktor et al. 2001). However, HIV-associated neurocognitive disorder (HAND) remains a problem, with a prevalence of up to 50% in ARV treated patients in Europe and USA (Heaton et al. 2010; Heaton et al. 2011; Simioni et al. 2010). Table 1.2 shows the 3 subclasses of HAND in order of the severity of neurocognitive impairment (NCI).

Asymptomatic Neurocognitive Impairment (ANI)	Mild Neurocognitive Disorder (MND)	HIV-Associated Dementia (HAD)
No interference with ADLs	At least mild interference with ADLs	Marked interference with ADLs
At least 1.0 SD below mean of normative population in at least two cognitive domains	At least 1.0 SD below mean of normative population in at least two cognitive domains	At least 2.0 SD below mean of normative population in at least two cognitive domains

Table 1.2. Diagnostic criteria for HIV-associated neurocognitive disorder.

According to the Frascati diagnostic criteria. ADLs; activities of daily living. SD; standard deviation. Source- Collier et al. Brain Sci. 2018, 8, 190; doi:10.3390/brainsci8100190.

It is thought that HIV-1 entry to the CNS is facilitated by loss of the integrity of the blood brain barrier (BBB) early in HIV-1 infection. HIV-1 ingresses as free particles or within activated CD4 central memory T cell or monocytes. The CNS viral reservoir is then maintained by clonal expansion of latently infected T cells or low-grade replication in macrophages (Sigal and Baltimore 2012), which may lead to an independently replicating population of HIV-1 viruses in the CNS compartment, distinct from the population in the peripheral circulating blood. This is known as CNS compartmentalisation. Due to inadequate ARV penetration into the CNS, the virus has the opportunity to sequester in the CNS and multiply (Joseph et al. 2019; Lorenzo-Redondo, Fryer, Bedford, Kim, Archer, Pond, et al. 2016). Following treatment interruption, resistant viruses may emerge in the CNS, amplifying this process (Canestri et al. 2010). The presence of replicating HIV-1 in the brain has been implicated in HIV-associated dementia (HAD); the most severe form of HAND and HIV-1 encephalitis (HIVE); a presentation of acute or subacute brain inflammation characterised by any of fever, headache, confusion, seizures or coma (Schnell et al. 2011; Filipowicz et al. 2016). Autopsy and living brain biopsy specimens from patients with HAD show positive p24 immunostaining in microglia and macrophages (Gray et al. 2013). In addition, HIV-1 isolates have been derived from brain biopsy specimen of participants with HAD (Gorry et al. 2001). Similarly in the animal models of HIVE and SIVE there is perivascular accumulation of gag

positive CD68 + and CD163+ macrophages in brain tissue (Filipowicz et al. 2016). CSF HIV-1 viral load (VL) correlates well with brain HIV-1 RNA levels (Gelman et al. 2013).

HIV-1 is suppressed by ART in both the blood and the CNS. However, this does not happen in every treated individual and the occurrence of CSF escape/discordance is now recognised (Canestri et al. 2010). It remains uncertain whether independently replicating HIV-1 in the brain parenchyma and the proxy for this—the presence of HIV-1 virus in the cerebrospinal fluid (CSF), is responsible for NCI in the ART era. CSF discordance is defined as CSF viral load (VL) greater than 0.5 or $1\log_{10}$ of the plasma VL and CSF escape is defined as any VL in the CSF above the limit of detection of the assay used (usually 40 copies/mL) when the VL in the plasma is undetectable by the same assay (Table 1.3). (Canestri et al. 2010; Peluso et al. 2012; Nightingale, Michael, et al. 2016). Although CSF HIV-1 VL does correlate well with brain HIV-1 RNA levels (Gelman et al. 2013), the causal relationship between CSF escape/discordance and HAND is unclear.

	Plasma VL	CSF VL
CSF escape	Undetectable *	Detectable *
CSF discordance	Detectable *	Greater than 0.5 or $1\log_{10}$ of the plasma VL

Table 1.3: CSF escape and CSF discordance.

* Depending on the limit of quantification of the assay used. Usually 40 copies/mL. VL; Viral Load. Source- Collier et al. Brain Sci. 2018, 8, 190; doi:10.3390/brainsci8100190.

1.5   Factors associated with HIV-1 neurotropism and neurovirulence

It has long been established that different but related strains of HIV-1 can coexist in different tissues (Gorry et al. 2001; Koyanagi et al. 1987). These strains exhibit genetic and phenotypic differences. CSF derived virus- JRCSF obtained from a patient with AIDS encephalopathy was found to infect a mixture of brain derived cells (astrocytes 70%, oligodendrocytes 30% and some fibroblasts). JRFL derived from the frontal lobe on the other hand had cellular tropism for macrophage and monocytes (Koyanagi et al. 1987). However, the factors that determine HIV-1 neurotropism are not well understand and there has been limited characterisation of this in HIV-1 subtype C viruses.

1.5.1   HIV-1 subtype and CNS disease

There is evidence to suggest that HIV-1 subtype may play a role in NCI. NCI has been associated with CRF_02AG when compared with subtype G in Nigeria (Royal et al. 2012). A greater risk of HAD was found in participants with subtype D virus compared to those with subtype A virus in Uganda (Sacktor et al. 2009). There is up to 35% variation within HIV-1 env across subtypes (Araujo and Almeida 2013). With the adaptive immune response against HIV-1 being mainly targeted at the envelope glycoprotein, inherent differences between HIV-1 env subtypes may play a role in HIV-1 neurotropism. It is therefore important to know if there are subtype dependent variations in the neurotropism and neurovirulence of HIV-1.

1.5.2   *Cellular tropism*

More recent studies characterising the phenotypic properties of CSF derived viruses have correlated macrophage tropism with neurotropism with evidence of a tropism shift from R5-T cell tropism to R5-macrophage tropism (Arrildt et al. 2015; Gorry et al. 2001; Schnell et al. 2011). Microglial cells and other macrophages in the CNS have lower levels of CD4 and are the targets for R5-Macrophage-tropic HIV-1, adapted to infect cells with low levels of CD4 (Joseph et al. 2015).

1.5.3   Coreceptor usage

The role of coreceptor tropism in viral pathogenicity is complex. The transmitted founder viruses exclusively utilise CCR5 coreceptor. Traditionally a switch from CCR5 coreceptor use to CXCR4 heralded the onset of advanced HIV/clinical AIDs. This is supported by the association of dual-tropic- CXCR4/R5 viruses with lower CD4 counts (Wilkin et al. 2007; Moyle et al. 2005). This phenomenon appears to be subtype specific with up to 50% of subtype B HIV-1 infected individuals undergoing co-receptor switching. This is approximately 20% in HIV-1 subtype C (Cilliers et al. 2003). A higher prevalence of CXCR4 variants is seen in subtype D compared with subtype A infected individuals (Huang et al. 2007). The prevalence of CXCR4 only using variants is less than 20% in most cohorts that have determined tropism using phenotypic assays (Moyle et al. 2005; Melby et al. 2006; Wilkin et al. 2007). In subtypes A and C progression of HIV-1 was accompanied by increased ability to utilise low levels of CCR5 and increased macrophage tropism. Neurotropism is also

associated with an increased ability to use low level of CD4 and or CCR5. Therefore, it would seem that the relative affinity of the virus to CD4/CCR5 receptors rather than a coreceptor switch is a better predictor of neurovirulence (Gorry et al. 2001; Gorry et al. 2002).

One study examining coreceptor usage in patients co-infected with HIV-1 subtype C and Cryptococcal meningitis (CM), found a evidence of CXCR4 usage using computer predictive algorithms (Sojane et al. 2018). This may reflect advanced HIV-1 infection or a higher prevalence of CXCR4 usage by CNS derived HIV-1 viruses.

1.5.4   Drug resistant mutations

CNS compartmentalisation has been shown (Liu et al. 2013; Schnell et al. 2011; Sturdevant et al. 2015; Stam et al. 2013), with greater genetic diversity of CSF escape viruses compared to plasma viruses according to phylogenetic analyses of pol/RT (Liu et al. 2013). Tong et al used deep sequencing to explore the range of minority resistance associated mutations in paired CSF and plasma and discovered mutations in CSF that were not identified by Sanger sequencing alone (Tong et al. 2015). The evolution of resistance mutations is also seen in CSF viruses (Canestri et al. 2010; Mukerji et al. 2017; Nightingale, Geretti, et al. 2016; Beguelin et al. 2016). Peluso et al. found CSF viral resistance in patients in whom resistance genotyping was conducted; 6/7 had NRTI mutations, 5/7 patients had PI mutations and 2/7 patients had NNRTI mutations (Peluso et al. 2012). In Mukherji's study of the pooled cohort from CHARTER, NNTC and HNRC, the CSF escape cases were

combined with all published CSF cases and showed that M184V/I mutations were detected more frequently in the CSF and plasma of patients with escape; 61% (34/56) and 30% (16/55) of samples respectively, compared to participants without escape, where it was only detected in 7% (3/43) of both CSF and plasma samples (Mukerji et al. 2018). APOBEC3F/G mediated hypermutation has been associated with compartmentalisation in CSF compared with peripheral blood mononuclear cells (PBMCs) and induced drug-resistant mutations in CSF; G73S in protease, M184V and M230I in reverse transcriptase (Fourati et al. 2014).

1.5.5   Escape from antibody neutralisation

Broadly neutralizing antibodies (bNAbs) targets 6 epitopes on the HIV-1 Env trimer-CD4 binding site, V2 apex, the V3 loop glycan patch, the membrane-proximal external region (MPER), the interface between the gp120 and gp41 subunits and 'silent face' of gp120 (Sok and Burton 2018). HIV-1 viruses in the CNS and plasma may evolve in response to differential selection pressures from both cellular immune response and neutralizing antibodies (Nabs) (Pillai et al. 2006). The glycoprotein shield associated with some of these epitopes is a target for the humoral immune response. An evolving 'glycan shield' mechanism of neutralization escape has been described by Wei et al, whereby selected changes in envelope glycosylation density and position prevented NAbs from binding but not receptor binding (Wei et al. 2003). Furthermore, over time the population of virus sensitive to neutralisation by autologous NAbs was replaced by neutralisation resistant virus. Viral escape through

mutation of the HIV-1 envelope glycoprotein (Env) maybe a means through which an independently replicating viral population can exist in the CNS.

1.5.6   CNS co-infection

Cryptococcal meningitis (CM) is an important AIDS defining illness and a leading cause of death in HIV-1 positive individuals in sub-Saharan Africa. It is uncertain to what extent opportunistic CNS infections such as CM increase the likelihood of HIV-1 CSF escape or discordance and to what extent persistent CNS HIV-1 impacts on the outcomes of CM. There is evidence of compartmentalised immune responses between CSF and blood in individuals with CNS infections such as TB meningitis (Christo et al. 2007) and CM (Chang et al. 2017), which may play a role in recruiting HIV-1 variants into the CNS, a phenomenon known as secondary CSF escape. However, it is unclear to what degree the HIV-1 variants in the CNS are an independently replicating/compartmentalised population or if they are intermixed with the plasma population as maybe the case due to BBB inflammation.

1.6   Aims

CNS compartmentalisation may lead to establishment of a long-lived viral reservoir in the CNS that may have different genotypic and phenotypic properties to the viruses in the periphery. This has implications for HIV-1 treatment, vaccine and cure strategies. The aim of the first part of thesis is therefore to investigate the factors that

determine HIV-1 subtype C CNS compartmentalisation in the context of CNS opportunistic infection/inflammation.

Specific objectives/questions are:

1   What is the prevalence of CNS compartmentalisation and what patient characteristics are associated with CSF compartmentalisation in the context of CM?

2   What are the phylogenetic characteristics of paired CNS and plasma genomes of HIV-1 subtype C in the context of co-infection with CM?

3   What are the phenotypic properties of HIV-1 clones isolated from the CNS?

2.1   Pandemic SARS-CoV-2

SARS-CoV-2 emerged in 2019 as the causative agent of a severe acute pneumonia first reported in Wuhan China (Zhou et al. 2020). This novel agent was sequenced in January 2020 from a bronchoalveolar lavage sample and was identified as a novel virus of the Coronaviridae family (Wu et al. 2020). Coronaviruses that affect humans are zoonotic and many are thought to have originated in bats (Tao et al. 2017). SARS-CoV-2 is a Betacoronavirus and one of the seven coronaviruses that infect humans. SARS-CoV and MERS-CoV cause severe disease, whilst HKU1, NL63, OC43 and 229E are associated with mild seasonal colds (Corman et al. 2018). However only SARS-CoV-2 has caused a pandemic.

The novel SARS-CoV-2 has 96% nucleotide identity with a bat coronavirus RaTG13, identified from a Rhinolophus affinis bat (Figure 2.1). However, the receptor binding domain (RBD) is where they diverge (Zhou et al. 2020). A related coronavirus found in Malayan pangolins (Manis javanica), is postulated to be the intermediary host. This is because SARS-like coronavirus found in pangolins have the 5 of the 6 amino acid changes in the RBD that are found in SARS-CoV-2 which enhances binding to their respective ACE2 (angiotensin converting enzyme-2) receptors (Andersen et al. 2020). Overall, SARS-CoV-2 is more closely related to the bat coronavirus (Figure 2.1) (Zhou et al. 2020). Also unique to SARS-CoV-2 is the insertion of a polybasic furin cleavage site between Spike subunits 1 and 2, which has been found to increase its infectivity (Xia et al. 2020). The acquisition of a

furin cleavage site in Influenza heamglutinninn also heralds an increase in

infectivity, host range and pathogenicity of Influenza (Alexander and Brown 2009;

Nao et al. 2017). The furin cleavage site is also present in human coronavirus HKU1

and MERS-CoV but absent in the sampled bat and pangolin coronaviruses

(Andersen et al. 2020).
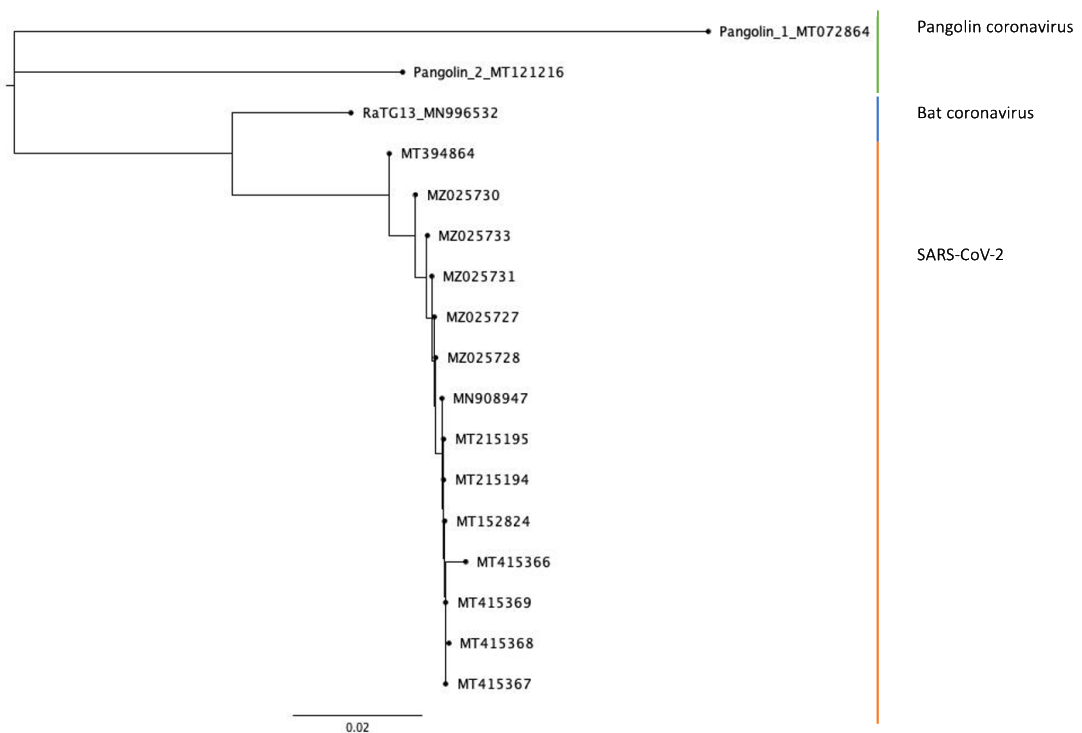
Figure 2.1 Neighbour-joining phylogenetic tree based on nucleotide sequences

of whole genome of coronaviruses. Neighbour-joining tree using HKY genetic

distance model with 1000 bootstraps. pangolin, bat and sar-cov-2. each taxon is

labelled by the genebank accession number.

Like HIV-1, the emergence of SARS-CoV-2 has been brought about by the growth

of the human population and expansion into the natural habitats of wild animals

(Jones et al. 2008). But unlike HIV-1 this emerging virus has impacted every single person on this planet. As of the 30th of April 2021 over 150 million cases of COVID-19 (coronavirus disease 2019) and 3.2 million associated deaths have been reported globally (Dong, Du, and Gardner 2020).

As yet there is no curative treatment for COVID-19. The treatments available continue to evolve during the pandemic as clinical trial evidence emerges. Many of the available drugs are directed at either mitigating or preventing the overwhelming systemic inflammation that is the hallmark of COVID-19 disease (Table 2.1). However, rapid advances have been made in the face of this global challenge and seven effective vaccines have been developed and are now being used globally in a bid to tackle this pandemic (Table 2.2). They all utilise the Spike protein antigen target. For the first time an mRNA vaccine has been approved for used to prevent a human infection (Abbasi 2020).

Drug	Drug Class	Indication	References
Dexamethasone OR Hydrocortisone/prednisolone	Corticosteroids	Requiring supplementary oxygen. 10-day course.	(The RECOVERY Collaborative Group 2020)
Remdesivir	RNA-dependent, RNA polymerase inhibitor	Hospitalised and on supplemental oxygen but not on invasive mechanical ventilation. 5-day course.	(Beigel et al. 2020; WHO Solidarity Trial Consortium 2020; Wang et al. 2020)
Tocilizumab Or Sarilumab	Interleukin-6 inhibitor	Completed a course of corticosteroids Require supplementary Oxygen and have a C-reactive protein > 75 mg/litre OR Requiring invasive mechanical ventilation.	(Horby et al. 2021; Salama et al. 2020; Gordon et al. 2021)
Baricitinib	Janus Kinase inhibitor	In combination with remdesivir For those require supplementary Oxygen OR Requiring invasive mechanical ventilation.	(Kalil et al. 2020)
Low molecular weight heparin	Anticoagulant	Likely to be hospitalised for at least 3 days and requiring supplementary oxygen	(Tang et al. 2020; Zarychanski 2021)

		14-day course or until discharge Non-invasive or invasive ventilation base decision of VTE vs bleeding risk adjust dose accordingly	
Bamlanivimab/etesevimab (LY-CoV555/ LY-CoV016 - Lilly) Casirivimab/imdevimab (REGN-CoV2 - Regeneron)	Monoclonal antibody cocktails	Mild to moderate COVID-19 in high risk individuals in order to prevent severe COVID-19 resulting in hospitalisation or death	(Gottlieb et al. 2021; Weinreich et al. 2020)
Convalescent plasma	Neutralising antibodies	High titre CCP for hospitalised patients with COVID-19 early in the course of hospitalisation	(Simonovich et al. 2020; Libster et al. 2021)

Table 2.1 Therapeutic options currently in use for managing COVID-19.

RNA- ribonucleic acid; COVID-19 (coronavirus disease 2019).

Vaccine	Platform	Reference
BNT162b2 (Pfzier BionTech)	mRNA	(Polack et al. 2020)
mRNA-1273 SARS-CoV-2 Vaccine (Moderna)	mRNA	(Baden et al. 2020)
ChAdOx1 nCoV-19 vaccine (AZD1222, AstraZeneca)	Simian adenovirus vectored (ChAdOx1)	(Voysey, Clemens, et al. 2021)
CoronaVac (Sinovac)	Inactivated vaccines	(Wu et al. 2021)
Gam-COVID-Vac (Sputnik V)	Recombinant adenovirus (rAd)-based vaccine (rAd26 and rAd5)	(Logunov et al. 2021)
NVX-CoV2373 (Novovax)	Adjuvanted recombinant protein- based nanoparticles	(Novovax 2021)
Janssen Ad26.COV2.S (Johnson &J)	Adenovirus vectored vaccines	(Sadoff et al. 2021)

Table 2.2 COVID-19 vaccines and the molecular platforms currently in use.

mRNA- messenger ribonucleic acid; COVID-19 (coronavirus disease 2019), rAd-recombinant adenovirus, ChAdOx-1- replication deficient simian adenovirus.

2.2   SARS-CoV-2 Life cycle

SARS-CoV-2 is a positive single strand RNA virus. It has a large 30kb genome (Figure 2.2). Like all viruses SARS-CoV-2 is an obligate intracellular pathogen that has to replicate by co-opting the host cellular machinery. SARS-CoV-2 gains entry into the cell via interaction between the RBD in the Spike protein S1 subunit and the ACE2 receptor (Figure 2.3). Proteolytic cleavage by cellular surface protease TMPRSS2 leads to separation of the two functional units of the Spike protein and allows the S2 subunit to mediate membrane fusion (Jaimes, Millet, and Whittaker 2020; Örd, Faustova, and Loog 2020). The virus particle may also gain entry by receptor mediated endocytosis. Acidification of endosomes leads to endosome activated cleavage by cathepsin, allowing insertion of the FP into the endosomal membrane (Jaimes, Millet, and Whittaker 2020; Örd, Faustova, and Loog 2020).
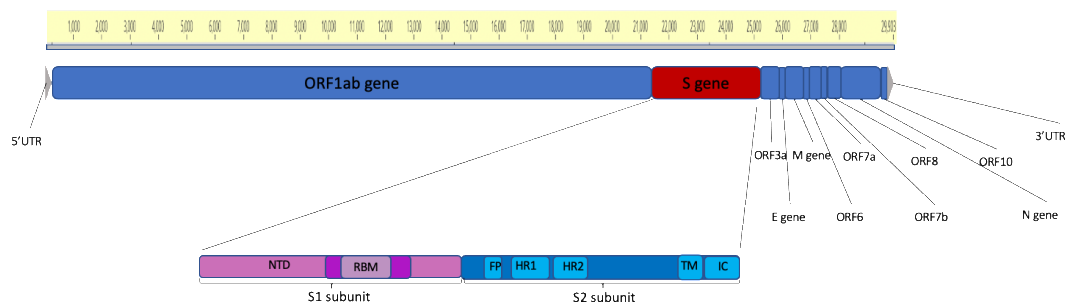
Figure 2.2 SARS-CoV-2 Genome map.

Structural and non-structural proteins are encoded by the labelled genes. The subunits and regions of the Spike gene are shown in the expanded S gene. NTD- N-terminal domain; RBM- receptor binding motif; FP- fusion peptide; HR1- heptad repeat 1; HR2- heptad repeat 2; TM- transmembrane region; IC- intracellular domain.

Translation of viral RNA into viral proteins and replication of the viral RNA occurs in the cytoplasm. ORF1a and 1b mRNAs at the 5` end of genome are translated immediately as they encode multiple non-structural proteins (nsp) involved in viral replication and transcription, including the exonuclease error reading enzyme (Shereen et al. 2020). This is then followed by replication of the viral genomic RNA for packing into viral particles and transcription of a nested set of subgenomic mRNAs. The discontinuous transcription of subgenomic mRNAs may result in recombination (V'kovski et al. 2021). The subgenomic mRNAs encode structural and accessory proteins from the 3` end of the genome and are translated in the cytoplasm. Viral assembly co-opts the host cell machinery, utilising cellular chaperones, transports systems and secretory pathways. Viral assembly occurs in the ER and transit to the ER-Golgi intermediate compartment where the genomic RNA in association with nucleocapsid are packed into the viral particles. They bud into the Golgi and are excreted from the host cell by exocytosis (Figure 2.3) (V'kovski et al. 2021).
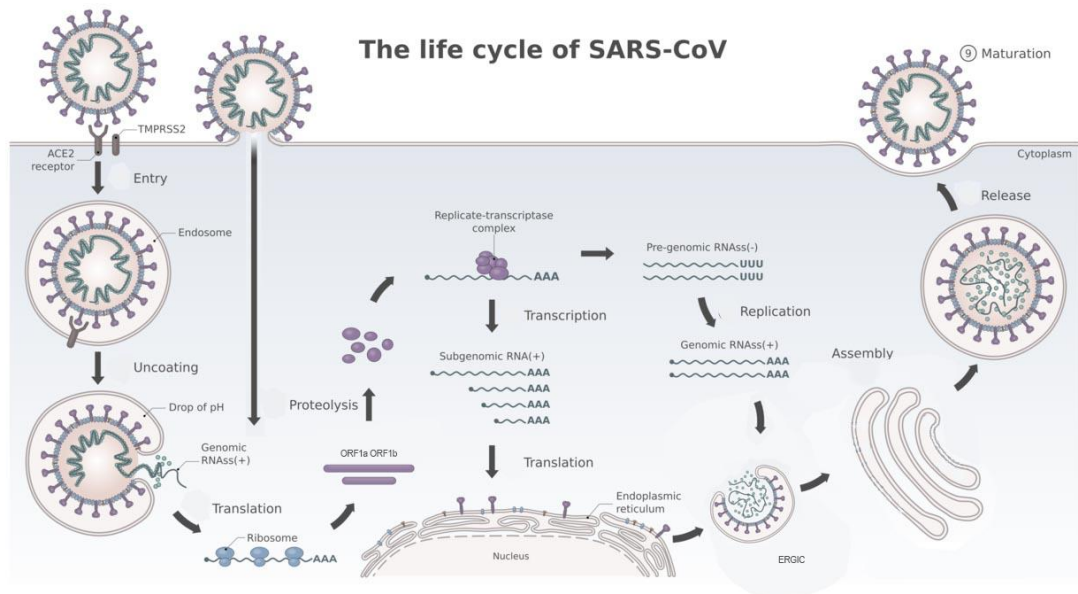
Figure 2.3 SARS-CoV-2 life cycle. ACE2- angiotensin converting enzyme- 2, ORF-open reading frame, ERGIC- endoplasmic reticulum- golgi intermediate compartment. Adapted from source: Vega Asensio - Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=88682470

2.3   Within-host adaptation

Given that SARS-CoV-2 is a novel virus, knowledge about what co-evolution might occur between the virus and human host is limited. Natural selection probably drove the acquisition of the polybasic furin cleavage site, the amino acid changes in RBD (which enhance infectivity) and the acquisition of mucin-type O-linked glycosylation of some Spike residues (which may enhance immune evasion), making SARS-CoV-2 well adapted to humans (Andersen et al. 2020). The fact that it is already so well adapted for human-to-human transmission may mean that it doesn't need to adapt

much more. However, knowledge from other human coronaviruses has highlighted that within-host adaptation is to be expected. The original SARS-CoV acquired mutations between amino acids 442- 487 of the RBD that enhanced viral interactions with both human and civet ACE2 and resulted in increased infectivity (Wu, Peng, et al. 2012). Already, it is evident that SARS-CoV-2 is mutating and has the potential to adapt, impacting transmission (Davies, Abbott, et al. 2021; Volz et al. 2021) and vaccine response (Collier, De Marco, et al. 2021; Madhi et al. 2021). However, it remains uncertain if these adaptations are impacting the severity of disease and death rates (Davies, Jarvis, et al. 2021; Frampton et al. 2021).

Compared with HIV-1 and Influenza viruses, the rate of evolution of SARS-CoV-2 is slow, with a nucleotide substitution rate of 8.4×10^{-4} substitutions per site per year (Nextstrain 2021) compared with HIV-1 which ranges from 2.7 to 6.7×10^{-3} substitutions per site per year (Leitner and Albert 1999) and Influenza A and B which ranges from 0.5 to 2.6×10^{-3} substitutions per site per year (Nobusawa and Sato 2006). Although it is slowly evolving, any adaptation to humans will be shaped by constraints of the human niche, which includes natural immunity as more people become infected and develop an adaptive immune response to SARS-CoV-2 as well as vaccine-elicited responses.

2.4   Host factors that determine immune responses to SARS-CoV-2 vaccines

Vaccines are a key tool to controlling the COVID-19 pandemic. Many vaccines have shown promise in clinical trials and real-world setting (Table 2.2). However,

duration of protection is as yet unknown. mRNA vaccine had shown good efficacy in clinical trials after two doses, separated by a three or four week gap (Baden et al. 2020; Polack et al. 2020). However, not many participants above the age of 80 years were included in the clinical trials despite being at greatest risk of severe disease and death (Docherty et al. 2020). Only twelve participants over 65 years were evaluated for neutralising responses in a BNT162b2 vaccine study (Walsh et al. 2020). It is therefore unknown what neutralising responses or efficacy will be in individual above the age of 80 years.

Furthermore, the Joint Committee on Vaccination and Immunisation (JCVI) increased the dosing interval to 12 weeks in order to vaccinate a greater proportion of the population during a period of substantial transmission (Care 2021). Due to lack of clinical trial data it is unknown if the elderly would mount adequate and sustained immune responses in the intervening period between doses to offer sufficient protection from SARS-CoV-2. There is evidence from studies in mice that immune responses in ageing mice after one dose of ChAdOx1 nCoV-19 is limited but is restored to levels comparable to young mice after the second dose (Silva-Cayetano et al. 2021).

Ageing can impact responses in many ways including but not limited to reducing the magnitude of antibody response to vaccines (Wilkinson et al. 2017), shortening the persistence of antibody response (Powers and Belshe 1993), limiting the quality of antibodies via reduced affinity and breadth, limiting induction of CD4+/8+ T cell response (Westmeier et al. 2020) and limiting induction of effector memory

(Gustafson et al. 2020). It is hypothesised that there will be age-related heterogeneity in immune response to SARS-CoV-2 vaccines.

2.5   Aims

The aims of the second part of this thesis are to:

1) Search for escape mutations in SARS-CoV-2 in a participant who had a prolonged illness with COVID-19

2) Characterise the functional importance of these escape mutations.

3) Explore age-related heterogeneity in immune responses to the mRNA BNT162b2 vaccine.

CHAPTER 2: METHODS

2.1   Section 1: HIV-1 viral escape and CNS compartmentalisation

This is a cross-sectional study design nested within a multisite, open-label, phase III randomised control trial (RCT) called the Advancing Cryptococcal Meningitis Treatment for Africa (ACTA) trial(Molloy et al. 2018). Participants in this cross-sectional study were selected from the study population recruited in Malawi. The study consists of Human immunodeficiency virus-1 (HIV-1) and Cryptococcal meningitis (CM) co-infected adults, who were older than18 years old, with bio-banked paired cerebrospinal fluid (CSF) and plasma specimens. Exclusion criteria included CSF red blood cell count > 1000 cells/mm3 in order to decrease the risk of contamination of CSF with blood that may have occurred during CSF sampling and for those on ART, a duration of treatment of less than 6 months.

2.1.1   Clinical Samples

Samples from blood and CSF were obtained in most cases simultaneously and no more that 14 days apart in a few cases. The samples were centrifuged at 1300 x g for 10 minutes and supernatant cryopreserved at $-80\,°C$ (Teunissen et al. 2011).

2.1.2 Viral load

Viral load quantification was done by the University College London Hospital Virology Laboratory. RNA extraction was done on the Qiasymphony using the DSP Virus/Pathogen kit as per manufacturer's protocols (Qiagen). Viral loads were quantified using the Panther (Hologic) and the Aptima HIV-1 viral load assay (United States). The limit of quantification is 30 copies/ml.

2.1.3 Single genome amplification (SGA)

The population of viruses in the CSF and blood compartments were sampled using SGA and sequencing of near full-length *env*. RNA extraction was done as previously described (Methods 2.1.2) cDNA was generated by reverse transcription using SuperScript IV (Thermofisher Scientific) and the *env* reverse primer- OFM19. Template RNA was degraded with RNAse H (Thermofisher Scientific). SGA involved serial dilution of cDNA template up to the point when less that 30% of wells in a 96 well plate was positive (Salazar-Gonzalez et al. 2008). This ensured that each positive well contains one amplifiable cDNA template greater than 80% of the time (Salazar-Gonzalez et al. 2008). Near full-length env was amplified using Platinum® Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA) in a nested PCR with the following primers (Table 2.1.1) (Haaland et al. 2009; Keele et al. 2008). All products derived from cDNA dilutions yielding less than 30% PCR positivity were Sanger sequenced with sequencing primers covering *env* (Table 2.1.1) (Sanger, Nicklen, and Coulson 1977).

First round primers Keele et al 2008	
Env5out (nt 5853–5877)	*TAGAGCCCTGGAAGCATCCAGGAAG*
OFM19 (nt 8913–8936)	*GCACTCAAGGCAAGCTTTATTGAGGCTTA*
Second round primers 1 Keele et al 2008	
Env5in (nt 5957–5983)	*caccTTAGGCATCTCCTATGGCAGGAAGAAG*
Env3in (nt 8904–8882)	*GTCTCGAGATACTGCTCCCACCC*
Second round primers 2 (nt 5,939 to 9,209) Haaland et al 2009	
Forward: EA1	*CCTAGGCATTTCCTATGGCAGGAAGAAGC*
Reverse: EN1	*TTGCCAATCAGGGAAGTAGCCTTGTGT*
Sequencing primers	
Forward: A_5allspl (nt 22-45)	*AAGAAGCGGAGACAGCGACGAAGA*
Reverse: B_179 (nt 601-626)	*CACATGGCTTTAGGCTTTGATCCCAT*
Forward:C_178 (nt 566-594)	*ATGGTAGAACAGATGCATGAGGATATAAT*
Reverse: D_180 (nt 1011-1041)	*TGAGTTGATACTACTGGCCTAATTCCATGTG*
Forward: E_V3Fin (nt 972-1007)	*GAACAGGACCATGTACAAATGTCAGCACAGTACAAT*
Reverse: F_mf169 (nt 1566-1586)	*TGATGGGAGGGGCATACAT*
Forward: G_181(nt 1520-1550)	*TGCAGAATAAAACAATTTATAAACATGTGGC*
Reverse: H_182 (nt 2165-2196)	*TGTATTAAGCTTGTGTAATTGTTAATTTCTCT*
Forward: J_mf159 (nt 2123-2150)	*CTGGAACAGATTTGGAATAACATGACCT*
Reverse: K_pr10 (nt 2833-2861)	*TTTTGACCACTTGCCACCCATCTTATAGC*

Table 2.1.1. PCR primer used for single genome amplification of HIV-1 *env*.

Numbering according to HXB2 reference sequence.

2.1.4   Next generation sequencing (NGS)

Short-read deep sequencing by Illumina MiSeq (Illumina, San Diego, CA, USA) was used to sequence the whole genome of viruses from CSF and plasma compartments (Gall et al. 2012). NGS and the bioinformatics quality control was performed by the diagnostic laboratory at University College London Hospital. As per the viral load testing protocol, nucleic acids were extracted from 230 µL of plasma using the Qiasymphony using the DSP Virus/Pathogen kit as per manufacturer's protocols (Qiagen, Hilden, Germany) and amplified using in-house HIV-1 primer (gag-pol codons 691-3582, pol-int gag-pol codons 2696-5527, int-env [g120] gag-pol codons 5518-7374). Library preparations were generated using the Nextera XT DNA Sample Preparation Kit (Bio Scientific, Texas, USA) and sequenced on the Illumina MiSeq platform. Prior to assembly the reads were trimmed using Trimmomatic (version 0.38). The following parameters were used; 10 bases were trimmed from the leading and training ends, a sliding window of 4:30 was used and reads with less than 50 bases were deleted. Adaptors were removed using ILLUMINACLIP 2:10:7 settings. The parameters used can be found here https://github.com/usadellab/Trimmomatic. Kraken2 was used to align reads to a database of genomes with viruses, bacteria and human. Everything that was non-viral was removed (including human). Read assembly was done using the de novo Iterative Virus Assembler (Gupta-Wright et al. 2020). Following assembly, samples were aligned to a 99ZACM9, subtype C reference sample from South Africa (GenBank: AF411967) (Papathanasopoulos et al. 2002) using the MAFFT program (version 7). A consensus sequence was generated from the mapped reads.

2.1.5　Determination of compartmentalisation

The primary outcome is central nervous system (CNS) compartmentalisation, which is defined as an independently replicating viral population in the CNS. This was determined by comparing HIV-1 genomes obtained from paired CSF and plasma samples using SGA. CNS compartmentalisation status was determined by three methods.

1. Visual inspection: Maximum likelihood (ML) phylogenetic trees were produced from the SGA sequences and limited to sequences with no gaps. Env nucleic acid sequences were aligned in Geneious version 10. RaxmlGUI was used to construct the ML tree using general time reversible model with a gamma distribution (Edler et al. 2019). The phylogenetic reconstruction was repeated 1000 times by resampling from the dataset to provide the degree of confidence in the tree branches, i.e. 1000 bootstraps. This was presented as a proportion, and a proportion greater than 70% was deemed to provide good support for that branch. The ML trees were inspected for the presence of a population structure consistent with an independently replicating population in the CSF.

2. Distance based methods: Wright's FST was used. It compares the mean pairwise genetic distance between two sequences sampled from different compartments to the mean distance between sequences sampled from the same compartment. The closer the FST is to 1 the greater the likelihood of compartmentalisation. FST <0.05 indicates very little genetic difference

between compartments. Values > 0.15 show a significant amount of genetic difference and values >0.25 show a very high degree of genetic difference. Corrections were made for unequal character frequencies, A>G, C>T and transversional bias with the Tamura-Nei model, corrections were made for distance and for rate variation using a gamma distribution. A run of 10,000 bootstrap replicates was done.

3. Tree-based methods: Slatkin-Maddison (SM) determines the minimum number of migration events between the separated populations, consistent with the structure of the reconstructed phylogenetic tree. The more migration events needed to explain the distribution of sequences the less likely that there is compartmentalisation. A p-value <0.05 is indicative of significance. Simmons association index (AI) assesses the degree of population structure in the phylogenetic tree recursively, using a bootstrap sample. A value of 1 implies that the population structures observed would be expected by chance whilst a value of 0 implies a strict genetic structuring by compartmentalisation. The tree-based methods were done using BaTS, which uses a posterior distribution of phylogenies from Bayesian MCMC analysis (Parker, Rambaut, and Pybus 2008). Given that compartmentalisation is an intra-host phenomenon, tree-based compartmentalisation analyses require at least 20 viral genomes in each compartment to consistently classify a compartment at compartmentalised (Zarate et al. 2007).

The intra-host viral diversity was further explored using multidimension scaling (MDS). This allowed mapping of the relative distances between sequences in a

visual plot. Pairwise distances between all assembled SGA sequences were calculated using the dist.dna() package, with a TN93 nucleotide-nucleotide substitution matrix (Tamura & Nei, 1993) and with pairwise deletion using the R package Ape v.5.4 (Paradis et al., 2004). Multi-dimensional scaling (MDS) was implemented using the cmdscale() function with pairwise deletion in R v4.0.4. This a data reduction method akin to principle component analysis, which is intended to simplify complex data into a more interpretable format, by reducing dimensionality of data but retaining most of the variation. In a genomics context we can use this on pairwise distance matrices, where each dimension is a sequence with data points of n-1 sequences pairwise distance, converted to a vector which can be plotted on 2-dimentional plot. This was done in collaboration with Dr Steven Kemp.

2.1.6 Drug resistance

Stanford HIV drug-resistance database was used to identify resistant mutations to non-nucleotide reverse transcriptase inhibitor (NNRTI), nucleotide(side) reverse transcriptase inhibitor (NRTI), protease inhibitor (PI) and integrase strand transfer inhibitors (INSTI) drugs (Shafer 2006). Minority variants were called at the 20% and 5% levels.

2.1.7   Prediction of coreceptor usage

Co-receptor usage was also determined genotypically by analysing the V3 loop
sequences of each amplified virus with freely available computer predictive
algorithms (CPAs), Geno2pheno (G2P) and Web position-specific scoring matrix
(WebPSSM). G2P was trained on a set of 1100 subtype B sequences with
genotypic-phenotypic correlation and set at a false positive rate (FPR) of 2.5%
(http://coreceptor.bioinf.mpi-inf.mpg.de/). PSSM uses the net charge as well as the
amino acid at positions 11 and 25, if these are either an Arginine (R) or Lysine (K),
the virus is more likely to be CXCR4 using. There are two matrices available,
WebPSSM subtype C sinsi
(https://indra.mullins.microbiol.washington.edu/webpssm) was used, which bases
predictions on known syncytium-inducing phenotypes on the MT-2 cell line (Jensen
et al. 2006). These tools predict the major coreceptor usage of *env* V3 sequences but
are not optimized to discriminate between R5X4-tropic variants from X4-tropic
variants. It is therefore necessary to use a phenotypic approach such as described
above to differentiate between dual tropic R5X4/X4 and CXCR4 viruses.

2.1.8   Phenotypic co-receptor usage

2.1.8.1. Cells

HEK293T cells and HeLa TZM-bl reporter cells were grown in Dulbecco's modified

Eagle's medium (DMEM; Gibco) supplemented with 10% fetal calf serum (FCS; Gibco) and 1% penicillin-streptomycin (10,000IU/ml, 10,000µg.ml) (Gibco). U87 cells expressing either CCR5 or CXCR4 co-receptors were grown in DMEM (Gibco) supplemented with 10% FCS (Gibco), Puromycin 1ug/ml and Neomycin 100ug/ml.

2.1.8.2. Clones

HIV-1 Gp120 *env* from distinct viruses in the blood and CSF compartments obtained from SGA were cloned into a B41 psvIII envelope expression vector (Figure 2.1.1). An additional round of PCR was done on the SGA PCR products using the primers 912 AATTGTGGGTCACAGTCTATTATGGGGTAC and 930 ATATTTGTATAATTCACTTCTCCAATTGTCCCTCATATCT (6318 to 7690) with complementarity to the Kpn I and Mfe I restriction sites respectively. The PCR product was cloned into a B41 psvIII envelope expression vector. B41 psvIII was restriction digested using restriction enzymes Kpn I and Mfe I (Neb biolabs). The digested vector was run on a 1% agarose gel and gel purified using Qiaquickgel extraction kit (Qiagen, Germany). The amplified gp120 *env* was cloned into the vector using NEBuilder® HiFi DNA Assembly Master Mix (Figure 2.1.2).
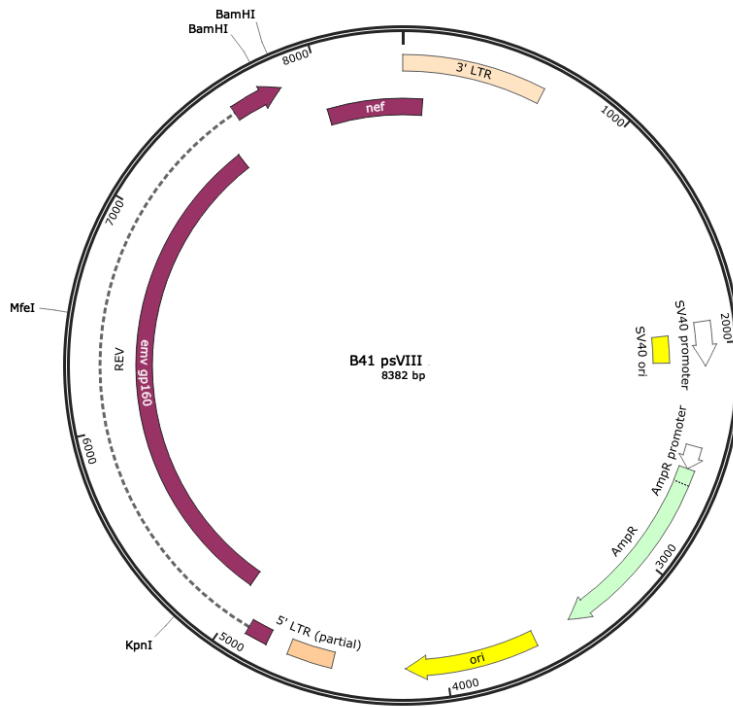
Figure 2.1.1. B41 psvIII envelope expression vector map. Donated by Dr Laura McCoy.
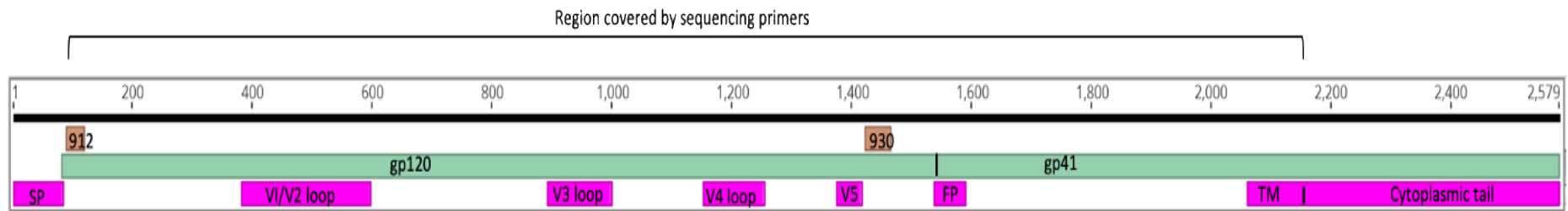
Figure 2.1.2. HIV-1 Envelope genome map. Nucleic acid numbering 1 to 2,579 base pairs. Indicating region in gp120 bound by primer 912 and 930 in brown that is cloned into the B41 psVIII envelope expression vector. SP- signal peptide; gp- glycopeptide; V-variable loop; FP-fusion peptide; TM- transmembrane domain; 912 and 930- forward and reverse cloning primers with the Kpn I and Mfe I restriction sites respectively.

2.1.8.3. Replication competent virus production

A double transfection system was used to generate pseudotyped virus using the
cloned env vector and the pNL4-3-Env Luc packaging vector, which contains a
luciferase reporter gene (Figure 2.1.3). This was done in the containment level 3
(CL3) laboratory. $1x10^6$ HEK 293T cells were plated in a 6 well plate in 2ml of
DMEM supplemented with 10% FCS, 1% 10,000 units penicillin and 1% 10 000
µg/mL Streptomycin. In a separate tube, 200ng of the participant derived Env
cloned into the envelope expression vector and 400ng of the packing vector were
added to 33µl of Optimen. 2µl of Fugene HD transfection reagent was added and left
to incubate for 20mins. The HEK293Ts were transfected with this plasmid mixture
and the produced pseudotyped viruses were harvested 48 hours later and stored at -
80ºC. Pseudotyped viruses with the Env of dual tropic WEAU-d15.410.787, CXCR4
using HXB2 and CCR5 using JRSCF were created and used as controls. The bald
packaging vectors were used as negative controls.
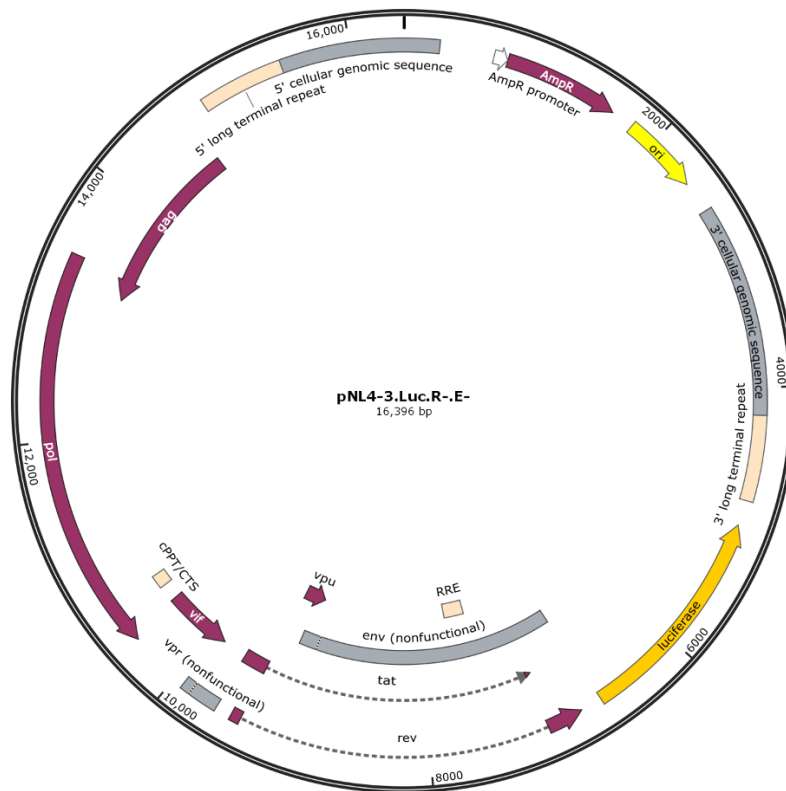
Figure 2.1.3. HIV-1 NL4-3 ΔEnv Vpr Luciferase Reporter Vector (pNL4-3.Luc.R-E-) Map. Source NIH-AIDS Reagent Program, Catalog #3418 Donated by Dr Nathaniel Landau.

2.1.8.4. Non-replication competent virus production

A triple transfection system was used to generate the pseudotyped virus using the cloned env vector, p8.91 Luc packaging vector and pCSFLW containing a luciferase reporter gene. This was done in the containment level 2 (CL2) laboratory. 1×10^6 HEK293T cells were plated in a 6 well plate in 2ml of DMEM supplemented with 10% FCS and 10,000 units penicillin and 10 mg streptomycin/mL. In a separate tube, 300ng of the participant derived Env cloned into the envelope expression vector, 300ng of the packing vector and 500ng of the luciferase reporter vector were

added to 70µl of Optimen. 6µl of Fugene HD transfection reagent was added and left to incubate for 20mins. The HEK293Ts were transfected with this plasmid mixture and the produced pseudotyped viruses were harvested 48 hours later and stored at -80ºC. Pseudotyped viruses with the Env of dual tropic WEAU-d15.410.787, CXCR4 using HXB2 and CCR5 using JRSCF were created and used as controls. The bald packaging vectors were used as negative controls.

2.1.8.5. Standardisation of virus input by tissue culture infectious dose (TCID)

The TCID was determined in a single round infectivity assay by titration of the pseudotyped viruses on target cells- U87 cells expressing either CCR5 or CXCR4 co-receptor. The relative light units (RLU) of infection was determined using the Steady-Glo Luciferase assay system (Promega). The dilution of virus that gave a RLU of 500 000 RLU was used in further experiments.

2.1.8.6. Normalisation of viral input by SYBR Green-based product-enhanced PCR assay (SG-PERT)

In addition, virus input was normalised for reverse transcriptase (RT) activity in the virus supernatant as measured by qPCR. Ten-fold dilutions of virus supernatant were lysed for 10 minutes at room temperature in a 1:1 ratio in a 2x lysis solution containing 100mM Tris-HCL buffered to pH7.4, 50mM KCl, 0.25% Triton X-100, 40% glycerol, RNase inhibitor 0.8 U/ml. 12µl of each lysate was added to 13µl of a master mix

containing SYBR Green master mix (Qiagen), 0.5µM of MS2-RNA foward and reverse primers, 3.5pmol/ml of MS2-RNA, and 0.125U/µl of Ribolock RNAse inhibitor. Thermocycling was performed in a QuantStudio. RT activity was determined from the amount of transcription of a fixed concentration of MS2 RNA bacteriophage in each sample relative to a known standard of HIV-1 RT (Merck).

2.1.8.7. Co-receptor usage phenotypic assay

Co-receptor usage was determined by infecting $1x10^4$ U87 cells expressing either CCR5 or CXCR4 co-receptors in 96 well plates with these pseudotyped viruses in serial dilution(Gupta et al. 2019). Entry efficiency was measured 48 hours later by a luciferase readout of RLUs of the lysed infected cells using Steady glo lysis buffer and substrate. Pseudotyped viruses with the Env of dual tropic WEAU-d15.410.787, X4 using HXB2 and R5 using JRSCF were created and used as controls and the bald packaging vectors were used as negative controls.

CCR5 or CXCR4 coreceptor usage was confirmed by treating target cells in the co-receptor assay with co-receptor antagonists AMD-3100 800nM to block CXCR4-mediated entry and TAK-779 400nM to block CCR5-mediated entry prior to infection. The drugs were a kind donation from Dr Clare Jolly, UCL.

2.1.9 Neutralisation

Compartmentalised cellular and immune responses in the CNS and plasma compartments may exert differential pressure on viruses in either compartment, leading to viral escape. Since epitopes on the HIV-1 envelope are the main target for humoral immune response(Sok and Burton 2018), variation in Env can occur due to selection pressure that render them better adapted to one compartment versus another. The neutralisation activity of 4 HIV-1 broadly neutralising (bnAbs) were tested against participant derived Env pseudotyped viruses. The 4 bnAbs were PGT121 and PGT128 (targeting the high mannose patch), VRC01 (CD4 binding site blocker) and PG9 (targeting the apex). Anti-dengue serotype 3 antibody was used as the negative control. This assay was done in collaboration with Dr Laura McCoy, UCL.

A single round virus neutralisation assays were performed on HeLa TZMBLs cells. HeLa TZM-bl reporter cells were grown in Dulbecco's modified Eagle's medium (DMEM; Gibco) supplemented with 10% fetal calf serum (FCS; Gibco) and penicillin-streptomycin (50 µg/ml) (Gibco). 10 µg/ml DEAE-Dextran (Sigma) was added to cell prior to infection. Participant derived Env pseudotyped viruses were produced as previously described in a CL3 laboratory. A double transfection system with packaging vector psG3-ΔEnv (Figure 2.1.4) and the participant derived Env plasmids, pseudotyped viruses were created. Envelope pseudotyped virus do not contain a luciferase gene but contain tat which will induce luciferase expression in TZM-bl cells. These pseudotyped viruses were incubated with serial dilution of bnAbs for 1h at 37°C. Virus and cell only controls were also included. Then, freshly

trypsinized Hela TZM-bl cells under tat-regulation were added to each well.

Following 48h incubation in a 5% CO_2 environment at 37°C, the luminescence was

measured using Steady-Glo Luciferase assay system (Promega).
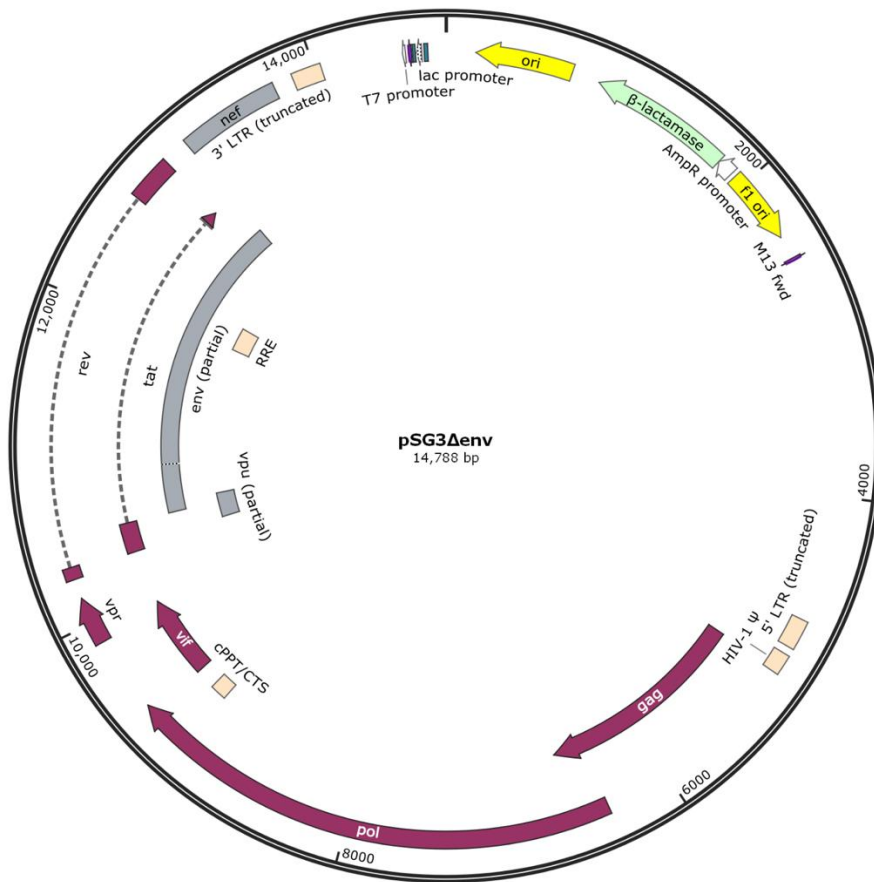
Figure 2.1.4. pSG3 ΔEnv Vector Map. Source NIH-AIDS Reagent Program,

Catalog #1105. Donated by Drs. John C. Kappes and Xiaoyun Wu.

2.1.10 Statistical Methods

Descriptive analyses of clinical and demographic data are presented as median and interquartile range (IQR) when continuous and as frequency and proportion (%) when categorical. The difference in continuous and categorical data were tested using Wilcoxon rank sum and Chi-square test respectively. A sample size of 42 is required to determine the prevalence of CNS compartmentalisation with a power of 80% and risk of type I error of 0.05.

Neutralisation activity was calculated relative to virus only controls. Dilution curves were presented as a mean neutralisation with standard error of the mean (SEM). The inhibitory concentration 50 (IC50)- the concentration of bnAbs at which 50% infection was inhibited was calculated in Prism version 9. The IC50 within groups were summarised as a geometric mean titre (GMT) and statistical comparison between groups were made with Mann-Whitney for unpaired data or Wilxocon ranked sign test for paired data. Statistical analyses were conducted using Stata version 13 and Prism version 9.

2.1.11 Ethical approval

Ethical approval for the ACTA trial was obtained from the London School of Hygiene and Tropical Medicine Research Ethics Committee- reference 8854 and the University of Malawi College of Medicine Research Ethics Committee- reference ISRCTN45035509 and all participants gave their written informed consent.

2.2   Section 2: SARS-CoV-2 Escape from Convalescent Plasma

This was an in-depth case-study of in vivo escape of SARS-CoV-2 from neutralising antibodies in an immunocompromised participant with prolonged shedding of SARS-CoV-2.

2.2.1 Clinical Sample Collection

Twenty-three samples from the upper airway- nose and throat swabs and lower airway-endotracheal aspirates were collected at various time points over 101 days of illness. Day 1 was the day on which the first SARS-CoV-2 PCR positive swab result was obtained.

2.2.2 Viral load/Cycle threshold (CT)

Viral load quantification was done by the Public Health England Clinical Microbiology. RNA extraction from 500µl of sample and was done on the easyMAG platform (Biomerieux, Marcy-l'Étoile) as per manufacturer's instructions. CT values were determined in one-step RT q-PCR assay quantified using a Rotorgene™ PCR instrument (Meredith et al. 2020). Samples with a CT values of ≤36 was considered positive.

2.2.3 SARS-CoV-2 binding antibody levels

Serology was done by the Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge, UK by multiplex particle-based flow cytometry (Luminex, Netherlands) (Xiong et al. 2020). Carboxylated beads were first activated to generate amine-reactive forms according to manufacturer's instructions (Themo Fisher Scientific). They were then covalently coupled to recombinant SARS-CoV-2 Nucleocapsid (N), Spike (S) and receptor binding domain (RBD) proteins to form a 3-plex assay. The beads were washed and then blocked for non-specific binding with a blocking buffer (containing 1% BSA, 10 mM PBS, and 0.05% NaN3). They were then incubated with the participant's sera at a dilution of 1 in 100 for 1 hour at room temperature in a 96-well plate, washed 3 times with washing buffer (containing 10 mM PBS and 0.05% Tween 20) and then incubated for 30 minutes with a PE-labeled anti–human IgG -Fc antibody (Leinco/Biotrend). The plates were read on a Luminex analyzer (Luminex / R&D Systems) using Exponent Software V31. Specific binding was reported as mean fluorescence intensity (MFI). Negative controls were unexposed healthy controls with pre-pandemic biobanked sera and positive controls were PCR positive COVID-19 cases.

2.2.4 Convalescent plasma for antibody titres

400ul sample of the 3 units of convalescent plasma (CP) administered to the participant was obtained from the NHS Blood and Transplant Service. The titres of anti- SARS-CoV-2 IgG was determined by NHS Blood and Transplant Service using

an indirect ELISA based assay- Euroimmun (Medizinische Labordiagnostika). Diluted CP was incubated with recombinant Spike 1 (S1) protein of SARS-CoV-2. The bound IgG antibody was detected by addition of an enzyme-labelled anti-human IgG antibody which catalyses a colour reaction. The ratio of optical extinction of the sample over that of a calibrator is determined. According to the manufacturer's instruction a ratio < 0.8 is negative, ratio ≥0.8 to <1.1 is borderline and ratio ≥1.1 is positive (https://www.fda.gov/media/137609/download)

2.2.5 Whole blood T cell and innate stimulation assay

This was done by the Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge, UK. Whole blood cytokine levels of IFNγ, TNFα, interleukin (IL) -1β, -2, -6, -10 and -17 were measured by multiplexed particle-based flow cytometry (Bio-Plex, Bio-Rad, UK) using a custom kit (R&D Systems, UK). T cell stimulation was done with phytohemagglutinin (PHA; 10 μg/ml; Sigma-Aldrich), or co-stimulating with anti-CD3 (MEM57, Abcam, 200 ng/ml) and IL-2 (Immunotools, 1430U/ml) or Phorbol 12-myristate 13-acetate/ionomycin, (PMA/IONO) and innate stimulation with lipopolysaccharide (LPS; 1 μg/ml, List Biochemicals). Supernatants were taken 24 hours after stimulation. Cytokines levels in pg/ml were measured by on a Luminex analyzer. The matched participant serum from the same time point was a negative control, run in parallel to account for any background cytokines released in the blood and a healthy donor samples were run in parallel as a positive control.

2.2.6 Next generation sequencing (NGS)

2.2.6.1 Long-read NGS

RNA extraction was done on the easyMAG platform (Biomerieux, Marcy-l'Étoile) as per manufacturer's instructions. Long-read NGS was done using MinION flow cells v.9.4.1 (Oxford Nanopore Technologies) by The COVID-19 Genomics UK Consortium (COG-UK; https://www.cogconsortium.uk/). The ARTICnetwork V3 protocol was used (https://dx.doi.org/10.17504/protocols.io.bbmuik6w). The bioinformatics was done using the ARTICnetwork curated reference-based assembly pipeline with minimum coverage of 20 times across the whole genome (https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html) (Loman, Rowe, and Rambaut 2020).

Phylogenetic analyses were done in collaboration with Dr Steven Kemp. All available full-genome SARS-CoV-2 sequences were downloaded from the GISAID database (http://gisaid.org/)(Shu and McCauley 2017) on the 22nd of April 2021. Low-coverage sequences were excluded. Sequences were re-aligned to the Wuhan-Hu-1 reference sequence using MAFFT v7.475 using the --keeplength and --addfragments options. Sequences were deduplicated and a random subset of 487 global sequences (from 56 countries) were obtained using seqtk (https://github.com/lh3/seqtk). To this, 239 randomly subsampled sequences were added from the UK, including England, Wales, Scotland and Ireland. Finally, the reference sequence MN908947.3, 20 sequences from the participant and 21 local sequences (to the hospital) were added to a multiple sequence alignment for a total of 768 sequences. Major SARS-CoV-2 lineage was

determined using Phylogenetic Assignment Of Named Global Outbreak Lineages (pangolin) (Rambaut et al. 2020).

A maximum-likelihood phylogeny was inferred using IQTREE v2.13 using the -fast option. (Minh et al. 2020). Evolutionary model selection for trees were inferred using ModelFinder (Kalyaanamoorthy et al. 2017) and trees were estimated using the GTR+F+I model with 1000 bootstrap replicates (Minh, Nguyen, and von Haeseler 2013). Trees were viewed with Figtree v.1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/), rooted on GenBank: MN908947.3 SARS-CoV-2 reference sequence.

2.2.6.2 Short-read NGS

In addition, 20 of the 23 samples underwent short-read NGS using Illumina MiSeq (Illumina, San Diego, CA, USA) and was done by the Department of Pathology, University of Cambridge. Library preparations were generated using the KAPA Hyper Prep kit (Roche) and sequenced on the Illumina MiSeq platform using a Miseq Nano v2 with 2x 250 paired-end sequencing. Bioinformatics processing involved removal of poor quality reads, removal of adaptors and trimming the reads with TrimGalore v0.6.6 (Martin 2011). Trimmed paired-end reads were mapped to the GenBank: MN908947.3 SARS-CoV-2 reference sequence using MiniMap2-2.17 (Li 2018). BAM files were indexed with samtools v1.11 and PCR optical duplicates removed using Picard (http://broadinstitute.github.io/picard). A consensus sequence was generated with minimum whole genome coverage of at least 20 times. BCFtools using

a 0% majority threshold was used for variant calling across the whole genome for the 20 samples. All variants were validated using custom code in the AnCovMulti package written by Dr David Pollock (https://github.com/PollockLaboratory/AnCovMulti). From this output longitudinal mutation prevalence plots across the entire genome were constructed.

Analyses of the cladal structure of the viral population in the 20 samples was done in collaboration with Dr Chris Illingworth. The SAMFIRE package (https://github.com/cjri/samfire/) was used to calculate distances between sequences using the distance metric (Lumby et al. 2020). It combines the allele frequencies across the entire genome. "Where L is the length of the genome, q(t) is 4 x L element vector describing the frequencies of each of the nucleotides A, C, G, and T at each locus in the genome sampled at time t. For any given locus i in the genome we calculate the change in allele frequencies between the times t_1 and t_2 via a generalisation of the Hamming distance, where the vertical lines indicate the absolute value of the difference.

$$d\big(q_i(t_1), q_i(t_2)\big) = \frac{1}{2} \sum_{a \in \{A,C,G,T\}} |q_i^a(t_1) - q_i^a(t_2)|$$

$$D\big(\boldsymbol{q}(t_1), \boldsymbol{q}(t_2)\big) = \sum_i d\big(q_i(t_1), q_i(t_2)\big)$$

These statistics were then combined across the genome to generate the pairwise sequence distance metric" (Kemp et al. 2021). The median pairwise distances between

samples in different clades of the phylogenetic tree were calculated and the statistically tested with the Mann-Whitney test (Kemp et al. 2021).

Long-read sequencing data have been deposited in the NCBI SRA database-BioProject PRJNA682013; accession codes SAMN16976824 - SAMN16976846 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA682013). Short reads sequence data were deposited at https://github.com/Steven-Kemp/sequence_files.

2.2.7 Single Genome Amplification (SGA)

SGA of the *spike* gene was used as an independent method to detect mutations observed in 3 samples at different time points. Viral RNA extracts were obtained as previously described Methods 2.2.2) and were reverse transcribed using SuperScript IV (Thermofisher Scientific). and an "in-house" *spike* specific primer- CoVOuter-R. Template RNA was degraded with RNAse H (Thermofisher Scientific). Terminally diluted cDNA was used to amplify partial Spike (Spike amino acids 21- 800~ 1.8 kb) by nested PCR using Platinum® Taq DNA Polymerase High Fidelity (Invitrogen, Carlsbad, CA). 'In-house' designed primers were used (Table 2.2.1) based on the participant's consensus NGS sequences. All products derived from cDNA dilutions yielding less than 30% PCR positivity were Sanger sequenced with 8 in-house sequencing primers covering *spike* by Genewiz UK (Table 2.2.1). This meant that there was at least an 80% chance that each well contained a single genome. Sequences were manually curated in DNA Dynamo software (Blue Tractor Software Ltd, UK).

First round primers	
CoVOuter-F	*AATGAAAATGGAACCATTACAGATGC*
CoVOuter-R	*ACCCGCTAACAGTGCAGAAG*
Second round primers	
CoVInner-F	*AAGTGTACGTTGAAATCCTTCACTG*
CoVInner-R	*AATGAGGTCTCTAGCAGCAA*
Sequencing primers	
Seq_F1	*TCTTGTTTTATTGCCACTAGTCTCT*
Seq_R1	*TGTTAGACTTCTCAGTGGAAGCA*
Seq _F2	*ACCACAAAAACAACAAAAGTTGGA*
Seq_R2	*CCCACATAATAAGCTGCAGCAC*
Seq_F3	*GTGTACGTTGAAATCCTTCACTGT*
Seq_R3	*TCCAAGCTATAACGCAGCCT*
Seq_F4	*CTATCAGGCCGGTAGCACAC*
Seq_R4	*ACACTGACACCACCAAAAGAAC*

Table 2.2.1. PCR primer used for single genome amplification if SARS-CoV-2

spike.

2.2.8 Generation of Spike mutants

Site directed mutagenesis (SDM) was used to introduce the following amino acid substitutions/deletions W64G, P330S, △ H69/V70 and D796H into the D614G pCDNA_SARS-CoV-2_Spike plasmid using the QuikChange Lightening Site-Directed Mutagenesis kit (Agilent Technologies, Inc., Santa Clara, CA) following the manufacturer's instructions. Primers used were designed "in-house" (Table 2.2.2).

W64G	
W64G_FWD	*TTCAGCAACGTGACCGGGTTCCACGCCATCCACG*
W64G_REV	*CGTGGATGGCGTGGAACCCGGTCACGTTGCTGAA*
△H69/V70	
H69del-V70del_FWD	*CGTGACCTGGTTCCACGCCATCTCCGGCACCAATGGCACC*
H69del-V70del_REV	*GGTGCCATTGGTGCCGGAGATGGCGTGGAACCAGGTCACG*
P330S	
P330S_FWD	*CCATCGTGCGGTTCAGCAATATCACCAATCTGTGCCCC*
P330S_REV	*GGGGCACAGATTGGTGATATTGCTGAACCGCACGATGG*
D796H	
D796H_FWD	*CCCCTCCTATCAAGCATTTCGGCGGCTTCAATTTCAGCC*
D796H_REV	*GGCTGAAATTGAAGCCGCCGAAATGCTTGATAGGAGGGG*

Table 2.2.2. Site directed mutagenesis primers.

WT- wild type Spike pseudotyped virus, mutants; D796H- aspartic acid to histidine at Spike protein amino acid position 796, △H69/V70- histidine and valine deletion at Spike protein amino acid position 69 and 70 respectively, W64G- tryptophan to

glycine at Spike protein amino acid position 64, P330S- proline to serine at Spike

protein amino acid position 330.

2.2.9 Spike pseudotyped virus preparation

Wild type (WT) Spike and mutant pseudotyped virus was generated using a

lentiviral triple transfection system. WT Spike was codon optimised, based on the

Wuhan-1 sequence but bearing the aspartic acid to glycine change at amino acid

position 614 of the Spike protein (D614G). The mutant Spikes pseudotyped viruses

were D796H- aspartic acid to histidine at Spike protein amino acid position 796,

△H69/V70- histidine and valine deletions at Spike protein amino acid position 69

and 70 respectively, W64G- tryptophan to glycine at Spike protein amino acid

position 64, P330S- proline to serine at Spike protein amino acid position 330. There

were engineered singly or in pairs consisting of D796H+△H69/V70 or W64G+

P330S. Sequences were verified by Sanger sequencing.

$1x10^6$ HEK 293T cells were seeded in a 10cm dish 24 hours prior. The cells were

transfected with a triple plasmids system consisting of 1µg of the WT or mutant

D614G pCDNA_SARS-CoV-2_Spike plasmid, 1ug of p8.91 lentiviral packaging

plasmid and 1.5µg pCSFLW expressing the firefly luciferase reporter gene using

11ul of Fugene HD transfection reagent (Promega). Supernatant was collected at 48

hours after transfection, filtered through 0.45um filter and cryopreserved at -80˚C.

The tissue culture infectious dose (TCID) of SARS-CoV-2 pseudotyped virus was

determined by titrating the viruses on HEK293T cells which had been transiently

transfected with ACE2 and TMPRSS2. Infection of the target cells were readout as relative light units (RLU) on a luminometer using Steady-Glo Luciferase assay system (Promega).

2.2.10 Normalisation of viral input by SYBR Green-based product-enhanced PCR assay (SG-PERT)

Virus input was normalised for reverse transcriptase (RT) activity in the virus supernatant as measured by qPCR using. Ten-fold dilutions of virus supernatant were lysed for 10 minutes at room temperature in a 1:1 ratio in a 2x lysis solution containing 100mM Tris-HCL buffered to pH7.4, 50mM KCl, 0.25% Triton X-100, 40% glycerol, RNase inhibitor 0.8 U/ml. 12µl of each lysate was added to 13µl of a master mix containing SYBR Green master mix (Qiagen), 0.5µM of MS2-RNA foward and reverse primers, 3.5pmol/ml of MS2-RNA, and 0.125U/µl of Ribolock RNAse inhibitor. Thermocycling was performed in a QuantStudio. RT activity was determined from the amount of transcription of a fixed concentration of MS2 RNA bacteriophage in each sample relative to a known standard of HIV-1 RT (Merck).

2.2.11 Spike expression by Western blot

Forty-eight hours after transfection of HEK 293T cells with the triple plasmid preparation, the culture supernatant was harvested and filtered through a 0.45-µm-pore-size filter. The supernatant was centrifuged at 15,000 rpm for 120 min at 4˚C to

pellet virions. This was then lysed with Laemmli reducing buffer (100% glycerol, β-mercaptoethanol, 1 M Tris-HCl [pH 6.8] SDS and bromophenol blue). The lysates were electrophoresed on SDS– 4 to 12% bis-Tris protein gels (Thermo Fisher Scientific) under reducing conditions. The protein bands were electroblotting onto polyvinylidene difluoride (PVDF) membranes, which were probed with anti-Spike (Invitrogen) and anti-p24 Gag (NIH AIDS reagents) antibodies. ChemiDoc® MP imaging system (Biorad) was used for visualisation.

2.2.12 Serum/plasma pseudotype neutralization assay

Spike pseudotype assays have been shown to have similar characteristics as neutralisation testing using fully infectious wild type SARS-CoV-2 (Schmidt et al. 2020). Serum neutralisation activity of CP and participant's sera at time points preceding and following administration CP was measured in a single-round in vitro neutralisation assay. WT, △H69/V70+ D796H, W64G+ P330 and single mutant Spike pseudotyped virus were incubated with serial dilution of heat inactivated CP or participant's serum samples in duplicate for 1h at 37˚C in a 96-well plate. 5×10^4 HEK293T cells, transiently transfected with ACE2 and TMPRSS2 were added to each well. After 48 hours incubation in a 5% CO_2 at 37°C, RLU was measured using Steady-Glo Luciferase assay system (Promega) (Mlcochova et al. 2020). Serum neutralisation curves were generated from which inhibitory dilution 50 (ID50) were determined. ID50 is the dilution of serum required to inhibit 50% of infection of a SARS-CoV-2 Spike pseudotyped virus in the assay. Human AB serum stored prior to

2019 was the negative control. Positive controls were sera from COVD-19 PCR positive cases.

2.2.13 Structural Viewing

The location of the four spike mutations of interest was mapped onto an uncleaved SARS-CoV-2 spike structure solved by cryo-electron-microscopy by Wrobel et al (PDB: 6ZGE) (Wrobel et al. 2020). The Pymol Molecular Graphics System v2.4.0 (https://github.com/schrodinger/pymol-open-source/releases) was used.

2.2.14 Ethics

Ethics approval was granted by the East of England – Cambridge Central Research Ethics Committee (17/EE/0025). Written informed consent was obtained. Additional controls with COVID-19 were enrolled to the NIHR BioResource Centre Cambridge under ethics review board (17/EE/0025).

2.3 Section 3: Age-related heterogeneity of SARS-CoV-2 mRNA vaccine-elicited responses

2.3.1 Study Design

This was a cohort study of participants receiving the mRNA BNT162b2 vaccine. Participants included members of the public receiving their vaccine at Addenbrookes Hospital, Cambridge and health care workers receiving the first dose of the vaccine. Participants had their first vaccine between the 14th of December 2020 to the 29th of January 2021. Initially the second dose was given 3 weeks after the first but during the course of the study the dosing interval was increased to 12 weeks. Participants were consecutively recruited without exclusion. Participants were seen at least 3 weeks after receiving their first does and then followed up until 3 weeks after their second dose. Blood samples were collected at each visit. The exposure of interest was age, categorised into 2 exposure levels- < 80 and ≥ 80 years. The primary outcome was inadequate vaccine-elicited serum antibody neutralisation activity at least 3 weeks after the first dose. This was measured as the dilution of serum required to inhibit infection by 50% (ID50) in an *in vitro* neutralisation assay. An ID50 of 20 or below was deemed as inadequate neutralisation. Secondary outcomes included Spike-specific binding antibody responses and IFNγ and IL-2 T cell responses.

2.3.2 Sample size calculation

The sample size was determined by assuming a risk ratio of non-neutralisation in the ≥80 years group compared with <80 years group of 5. An alpha of 0.05 and power of 90% was used. This estimated a sample size of 50 with a 1:1 ratio was required in each group.

2.3.3 Clinical samples

Blood samples collected included 10ml of serum sample, which was spun at 1000g for 30 minutes, aliquoted and cryopreserved at -80°C and 20ml of whole blood from which peripheral blood mononuclear cells (PBMCs) were prepared using Lymphoprep density gradient centrifugation and cryopreserved at -80°C.

2.3.4 Generation of pseudotyped viruses

Wild type (WT) Spike lentiviral pseudotyped virus was generated in a triple transfection system. The WT was a codon optimised Spike, based on the Wuhan-1 sequence but bearing the aspartic acid to glycine change at amino acid position 614 of the Spike protein (D614G).

1×10^6 HEK293T cells were seeded in a 10cm dish 24 hours prior transfection. The cells were transfected with a triple plasmids system consisting of 1µg of the Spike

expression plasmid- pCDNA_SARS-CoV-2_Spike plasmid, 1ug of p8.91 lentiviral packaging plasmid and 1.5µg pCSFLW expressing the firefly luciferase reporter gene. 11ul of Fugene HD transfection reagent (Promega) was used. Supernatant was collected at 48 hours after transfection, filtered through 0.45um filter and cryopreserved at -80˚C. The tissue culture infectious dose (TCID) of SARS-CoV-2 pseudotyped virus was determined by titrating the viruses on HEK 293T cells which had been transiently transfected with ACE2 and TMPRSS2. Infection of the target cells were readout as relative light units (RLU) on a luminometer using Steady-Glo Luciferase assay system (Promega).

2.3.5 Neutralisation assays

Serum neutralisation activity of vaccine-elicited sera was measured in a single-round in vitro neutralisation assay. WT Spike pseudotyped virus were incubated with serial dilution of heat inactivated participant's serum following the first and second dose of mRNA BNT162b2 vaccine in duplicate for 1h at 37˚C in a 96- well plate. 5×10^4 HEK 293T cells, transiently transfected with ACE2 and TMPRSS2 were added to each well. After 48hours incubation in a 5% CO_2 at 37°C, RLU was measured using Steady-Glo Luciferase assay system (Promega) (Mlcochova et al. 2020). Serum neutralisation curves were generated from which inhibitory dilution 50 (ID50) were determined. ID50 is the dilution of serum required to inhibit 50% of infection of a SARS-CoV-2 Spike pseudotyped virus in the assay. Human AB serum stored prior to 2019 was the negative control. Positive controls were sera from COVD-19 PCR positive cases. The cut off for 50% neutralisation was set at an ID50 of 1in 20. The ID50 within groups

were summarised as a geometric mean titre (GMT) and statistical comparison between groups were made with Mann-Whitney or Wilxocon ranked sign test for unpaired and paired analyses respectively.

2.3.6 SARS-CoV-2 binding antibody levels

Serology was done by the Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge, UK by multiplex particle-based flow cytometry (Luminex, Netherlands) (Xiong et al. 2020). Carboxylated beads were first activated to generate amine-reactive forms according to manufacturer's instructions (Themo Fisher Scientific). They were then covalently coupled to recombinant SARS-CoV-2 Nucleocapsid (N), Spike (S) and receptor binding domain (RBD) proteins to form a 3-plex assay. The beads were washed and then blocked for non-specific binding with a blocking buffer (containing 1% BSA, 10 mM PBS, and 0.05% NaN3). They were then incubated with the participant's sera at a dilution of 1 in 100 for 1 hour at room temperature in a 96-well plate, washed 3 times with washing buffer (containing 10 mM PBS and 0.05% Tween 20) and then incubated with incubated for 30 min with a PE-labeled anti–human IgG/IgA/IgM-Fc antibody (Leinco/Biotrend). The plates were read on a Luminex analyzer (Luminex / R&D Systems) using Exponent Software V31. Specific binding was reported as mean fluorescence intensity (MFI). Negative controls were unexposed healthy controls with pre-pandemic biobanked sera and positive controls were PCR positive COVID-19 cases.

2.3.7 CMV serology

Human cytomegalovirus (HCMV) IgG levels were determined using an IgG enzyme-linked immunosorbent (EIA) assay, HCMV Captia (Trinity Biotech, Didcot, UK) following manufacturer's instructions, on plasma samples. This was done in collaboration with Dr Mark Wills.

2.3.8 T cell assays- IFNγ and IL2 FLUOROSPOT

Frozen PBMCs were thawed and diluted into 10ml of TexMACS media (Miltenyi Biotech), centrifuged and resuspended in 10ml of fresh media with 10U/ml DNase (Benzonase, Merck-Millipore via Sigma-Aldrich). PBMCs were then incubated at 37°C for 1h, followed by centrifugation and resuspension in fresh media supplemented with 5% Human AB serum (Sigma Aldrich) before being counted. PBMCs were stained with 2ul of LIVE/DEAD Fixable Far Red Dead Cell Stain Kit (Thermo Fisher Scientific) and live PBMC enumerated on the BD Accuri C6 flow cytometer.

1.0 to 2.5 x 10^5 PBMCs were incubated in pre-coated FluoroSpot^{FLEX} plates (anti IFNγ and IL2 capture antibodies Mabtech AB, Nacka Strand, Sweden)) in duplicate with either peptide mixes specific for Wuhan-1(QHD43416.1) Spike SARS-CoV-2 protein (Miltenyi Biotech) or a mixture of peptides specific for Cytomegalovirus, Epstein Barr virus and Influenza virus (CEF+) (final peptide concentration 1µg/ml/peptide, Miltenyi Biotech) in addition to an unstimulated (media only) and

positive control mix (containing anti-CD3 (Mabtech AB) and Staphylococcus Enterotoxin B (SEB), (Sigma Aldrich)) at 37ºC in a humidified CO2 atmosphere for 42 hours. The cells and medium were then decanted from the plate and the assay developed following the manufacturer's instructions. Developed plates were read using an AID iSpot reader (Oxford Biosystems, Oxford, UK) and counted using AID EliSpot v7 software (Autoimmun Diagnostika GmbH, Strasberg, Germany). Peptide specific frequencies were calculated by subtracting for background cytokine specific spots (unstimulated control) and expressed as SFU/Million PBMC (Figure 2.3.1) (Collier, Ferreira, et al. 2021). This was done in collaboration with Dr Mark Wills.
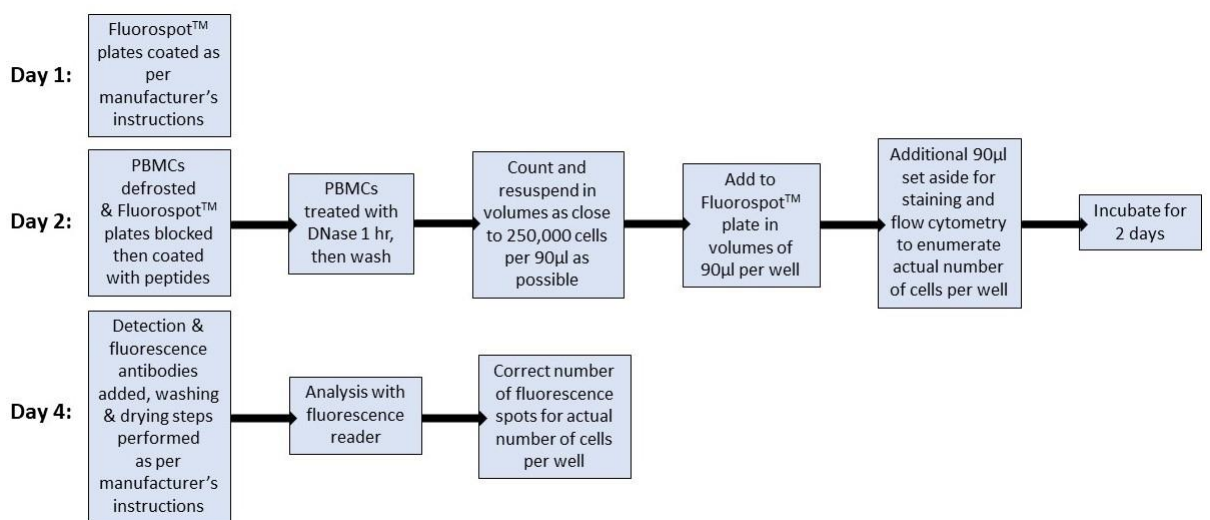
Figure 2.3.1: Schematic of the FlurospotTM assay. Courtesy of Dr Eleanor Lim.

2.3.9 Statistical Analyses

Descriptive analyses of demographic and clinical data are presented as frequency and proportion (%) when categorical and median and interquartile range (IQR) when

continuous. Test for significance included Wilcoxon rank sum for continuous variables and Chi-square test for categorical variables. Logistic regression was used to test the association between age and inadequate neutralisation by vaccine-elicited antibodies after the first dose of the BNT162b2 vaccine. The association was adjusted for the following confounders- sex and time interval from vaccination to sampling. Linear regression was also used to test the association between age and log transformed ID50, binding antibody levels, antibody subclass levels and T cell response after dose 1 and dose 2 of the BNT162b2 vaccine. Bonferroni adjustment was made for multiple comparisons. The Pearson correlation coefficient was reported. Statistical analyses were done using Stata v13, Prism v9 and R (version 3.5.1).

2.3.10 Ethical approval

Ethics approval was granted by the East of England – Cambridge Central Research Ethics Committee (17/EE/0025). PBMC from unexposed volunteers previously recruited by the NIHR BioResource Centre Cambridge through the ARIA study (2014-2016), with ethical approval from the Cambridge Human Biology Research Ethics Committee (HBREC.2014.07) and currently North of Scotland Research Ethics Committee 1 (NS/17/0110).

CHAPTER 3: RESULTS 1- HIV-1 VIRAL ESCAPE AND CNS COMPARTMENTALISATION

Key Findings:

- CSF escape is uncommon in this study population.

- Visual inspection and statistical analyses of phylogenetic trees revealed the prevalence of CNS compartmentalisation is 50%.

- No clinical or demographic factors have been identified which are associated with CNS compartmentalisation.

- Env diversity is compartmentalised in participants with CNS compartmentalisation.

- The founder virus of CNS compartmentalised viruses maybe a plasma variant.

- HIV-1 CNS compartmentalised clones characterized here utilize CCR5 co-receptors for viral entry.

- Computer predictive algorithms do not always accurately predict co-receptor usage.

- The frequency of drug resistance mutations is similar in both compartments.

- There is heterogeneity in sensitivity of CNS and plasma derived clones to broadly neutralizing antibodies.

The CNS is an immune privileged site bounded by the blood brain barrier which allows selective entry of molecules and immune cells through from the peripheral blood into the CNS (Rahimy et al. 2017). It is a recognised reservoir for HIV-1, where proviral DNA may lay dormant in long-lived cells like macrophages and neurones and be reactivated at a later date (Sigal and Baltimore 2012; Collier et al. 2018). Clonal expansion of HIV-1 or low-level replication in the CNS may lead to CNS compartmentalisation, which is the independent replication of a population of virus in an anatomical or cellular compartment, in this case the CNS (Adewumi et al. 2020; Bingham et al. 2011; Lorenzo-Redondo, Fryer, Bedford, Kim, Archer, Pond, et al. 2016; Stam et al. 2013). These compartmentalised viruses are distinct from the population in the peripheral blood. This process may also drive CSF escape (Collier et al. 2018), which for the purposes of this work is defined as a CSF viral load (VL) greater than 0.5log10 of the plasma VL.

Viruses in the CNS will be limited by the constraints in that niche and are known to have acquired some features that make them adapted to the CNS such as the switch of cellular tropism from R5 T cell to R5-Macrophage tropism (Schnell et al. 2011; Sturdevant et al. 2015), co-receptor usage with higher affinity for CCR5 co-receptor (Gorry et al. 2002) and acquisition of some drug resistant mutations (Mukerji et al. 2017). The extent to which CNS compartmentalised viruses evolve as a result of differential selection pressures from both cellular and humoral immune responses is unknown. In addition, the extent to which compartmentalised immune responses between CNS and blood in individuals with CNS co-infections drives within-host evolution in the CNS is not fully known (Adewumi et al. 2020). It is also unclear to

what degree the viruses in the CNS are an independently replicating population or if they are intermixed with the plasma population as maybe the case due to BBB inflammation. Furthermore, the phenotypic characteristics of CNS compartmentalised viruses have largely been studied in HIV-1 subtype B. This study includes participants with subtype C HIV-1.

This study includes 44 participants co-infected with HIV-1 and cryptococcal meningitis who had paired plasma and CSF samples available. They were examined for the presence of CSF escape and CNS compartmentalisation. The phylogenetic properties of paired samples were examined in detail and their phenotypic properties were characterised.

3.1 Cerebrospinal fluid (CSF) escape is uncommon in this clinical cohort

One hundred and twenty-eight participants were eligible from the Advancing Cryptococcal Meningitis Treatment for Africa (ACTA) trial (Figure 3.1). The first 44 of these were selected without exclusion. This included 23 participants who were naïve to antiretroviral treatment (ART) and 21 participants who were receiving ART for at least 6 months i.e. treatment experienced (Table 3.1). Viral loads (VL) were quantified in these 44 participants. The median $\log_{10}$ VL was higher in plasma than CSF (6.3 vs 5.3) copies/ml in treatment naïve (TN) participants (6.0 vs 5.1) copies/ml and in treatment experienced (TE) participants (Table 3.1). 4.5% (2/44) of participants had CSF escape defined as a VL in CSF 0.5 $\log_{10}$ greater that VL in the plasma. One of these was in the TN group (001) and the other in the TE group (002).
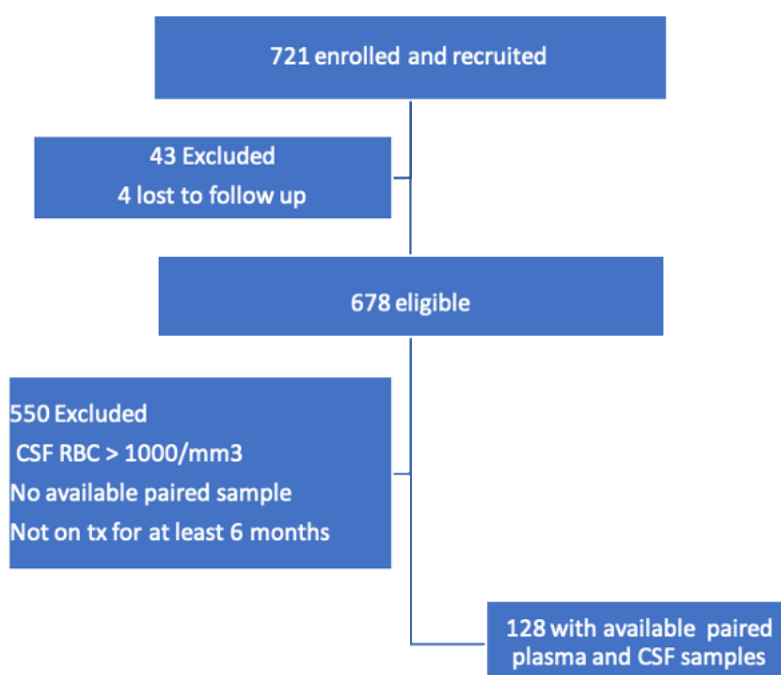
Figure 3.1. Study flow. CSF- cerebrospinal fluid; RBC- red blood count.

	Treatment naïve (23)	Treatment experienced (21)
Male gender *%*	62	53
Age median *years (IQR)*	35.0 (32.0 – 39.8)	36.0 (32.0 – 42.0)
Log$_{10}$ HIV-1 VL *copies/ml*		
CSF	5.3 (4.2 – 5.8)	5.1 (3.9 – 5.7)
Plasma	6.3 (5.8 – 6.6)	6.0 (5.4 – 6.1)
CSF escape/discordance %	4.3 (1/23)	4.8 (1/21)

Table 3.1. Clinical and demographic characteristic of study participants.

Includes 44 participants who had the VL in their paired CSF and plasma specimens

quantified. IQR- interquartile range; VL- viral load; CSF- cerebrospinal fluid.

Thirty-one participants successfully underwent next generation sequencing (NGS) of viruses in both CSF and plasma compartments by Illumina Miseq (Methods 2.1.4) and 4 of these had single genome amplification (SGA) (Methods 2.1.3) performed on paired CSF and plasma samples. Participants with successful NGS had a significantly higher median $\log_{10}$ VL in CSF and plasma compared to those without successful NGS; 5.4 (IQR 5.0-6.3) vs 3.0 (IQR 0-3.9) copies/ml (p <0.0001) in the CSF and 6.0 (5.7-6.4) vs 5.6 (0.0-6.2) copies/ml (p 0.04) in the plasma. Otherwise, these two groups had similar characteristics.

The 31 participants with paired genomic data form the basis of further analyses. The demographic and clinical data for the participants who had successful NGS done are summarised in Table 3.2. The median age was 35 years (IQR 32-40), 64.5% were male, 45.2% were on ARTs and 2 participants had CSF escape. The median mapped read depth for 31 paired specimens using the HXB2 nucleic acid positions 691-7374 was 1037 (IQR 169-3292) in the plasma and 262 (IQR 23-1879) in the CSF specimens. The lower read depth in CSF is not surprising as it is recognised that PCR amplification from CSF samples is less successful due to the presence of protein and cellular inhibitors, which can result in false negative CSF-PCR (Ratnamohan, Cunningham, and Rawlinson 1998). The read depth was variable along the length of the genome. It was at least 1000 reads in the regions spanning the gag- p17 and p24, pol- RT and env- V1-V3 genes. (Figures 3.2A, B).

	All N=31
Male % (number)	64.5 (20)
Median age years (IQR)	35 (32-40)
Receiving ARTs % (number)	45.2 (14)
Median time on ARTs months (IQR)	n=12/14 30.6 (12.5-61.1)
ART regimen % (number) **TDF+3TC/FTC+EFV** **TDF+3TC+ATV/LPV**	n=11/14 71.4 (10) 7.1 (1)
Median plasma VL $\log_{10}$ (IQR)	6.0 (5.7-6.5)
Median CSF VL $\log_{10}$ (IQR)	5.4 (4.9-6.3)
Median CD4 cells/mm^3 (IQR)	n=23/31 33 (5-65)
CSF escape % (number)	6.5 (2)
Randomisation assignment % (number) **1** **2A** **2B** **3A** **3B**	35.5 (11) 16.1 (5) 19.4 (6) 16.1 (5) 12.9 (4)
Seizures in last 72 Hours % (number)	16.1 (5)
GCS % (number) **Less than 15** **15**	25.8 (8) 74.2 (23)
10-week mortality % (number)	41.9 (13)
CSF pleocytosis % (number)	12.9 (4)

Table 3.2. Clinical and demographic data of participants with successful sequencing.

IQR- interquartile range; ART- antiretroviral treatment; TDF- tenofovir; 3TC- lamivudine; ATV- atazanavir; LPV- lopinavir; VL- viral load; CSF- cerebrospinal

fluid; GCS- Glasgow coma scale. Randomisation assignment in the ACTA study
drug regimen.

A   B
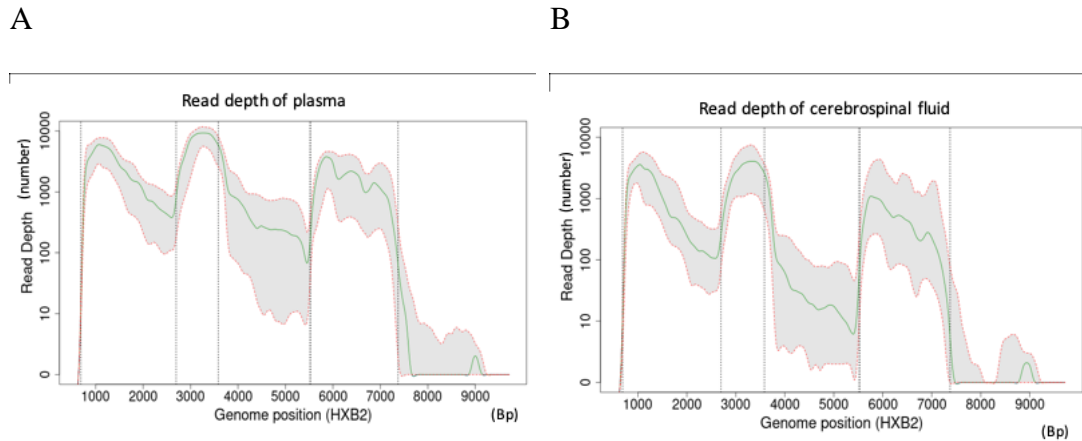
**Figures 3.2. Read depth of plasma (A) and cerebrospinal fluid (B) in Illumina
MiSeq next-generation sequencing.** Bp- base pair, HXB2- reference HIV-1
sequence. Median read depth (green line), Interquartile range (red lines).

In addition, 4 participants had SGA of the env gene performed on paired CSF and
plasma samples. Env sequences included amino acids 27 to 706 spanning the start of
gp120 to the transmembrane domain in gp41 (Figure 3.3). These participants had
between 21 and 35 viral genomes sequenced in each compartment. Table 3.3 shows
the clinical and demographic details of these 4 participants. They range between the
ages of 27 and 60 years. Two were male and 2 were on ART with a range of duration
of ART of 10.6 to 65.7 months. They universally had high HIV-1 VL in both plasma
and CSF compartments and low CD4 T cell counts. Two participants had CSF
escape; 001 and 002 and 2 participants did not; 003 and 004.
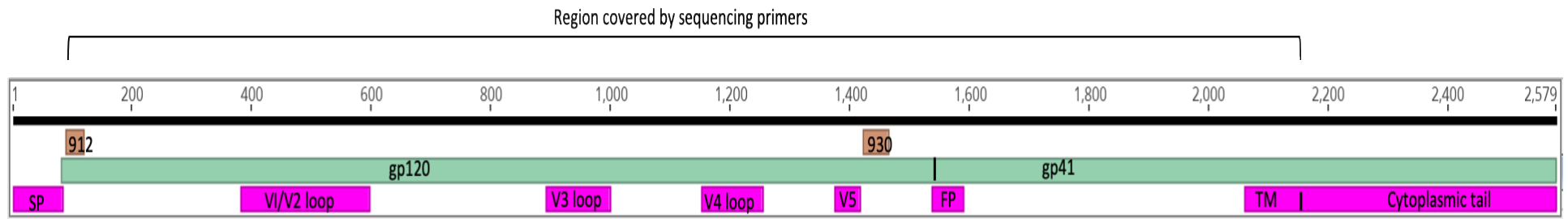
Figure 3.3. HIV-1 Envelope genome map. Nucleic acid numbering 1 to 2,579 base pairs. SP- signal peptide; gp- glycopeptide; V-variable loop; FP- fusion peptide; TM- transmembrane domain; 912 and 930- forward and reverse cloning primers with the Kpn I and Mfe I restriction sites respectively.

Variable	001	002	003	004
CNS compartmentalisation	Intermixed	Compartmentalised	Compartmentalised	Intermixed
Age	60	27	31	34
Sex	M	F	F	M
On ART	No	Yes	No	Yes
Duration of ART months	N/A	10.6	N/A	65.7
Log10 plasma VL copies/ml	5.5	5.7	6.3	6.2
Log10 CSF VL copies/ml	6.1	6.5	6.3	6.3
CD4 cells/mm3	146	98	80	missing
CSF escape	Yes	Yes	No	No
CSF pleocytosis	Yes	No	Yes	No
ACTA RCT Study Regimen	3B	3B	1	1
Seizure in last 72 hours	No	No	No	No
GCS	12	14	14	13
10-week mortality	No	Yes	No	No
DRMs	Nil	Nil	E138A in P	Nil

Table 3.3. Clinical and demographic properties of the 4 participants who successfully underwent single genome amplification.

ART- antiretroviral treatment, VL- viral load; CSF- cerebrospinal fluid; P- plasma, ACTA- Advancing Cryptococcal meningitis Treatment for Africa; RCT-randomised control trial; GCS- Glasgow coma scale, score out of 15; DRM-drug resistant mutations.

3.2 A viral population structure consistent with CNS compartmentalisation was shown in some participants

Maximum likelihood phylogenetic trees were constructed from env nucleic acid sequences and were limited to sequences with no gaps. General time reversible model with a gamma distribution and 1000 bootstraps were used in RaxML (Methods 2.1.5). Visual inspection of maximum likelihood phylogenetic trees revealed a spatial population structure that clearly segregated the taxa of the CSF derived genomes from those of the plasma in participants 002 and 003 (Figures 3.4A, B). However, genomes from the CSF and plasma from participants 001 and 004 showed no spatial population structure, therefore these plasma and CSF derived viruses are intermixed (Figures 3.4C, D). The phylogenetic trees shed some light on the sequence of evolutionary events that generated the present-day diversity of genomes. In the participant 003 with CNS compartmentalisation, the most recent common ancestor (MRCA) of a cluster of CSF genomes appeared to be a plasma variant (Figure 3.4B).
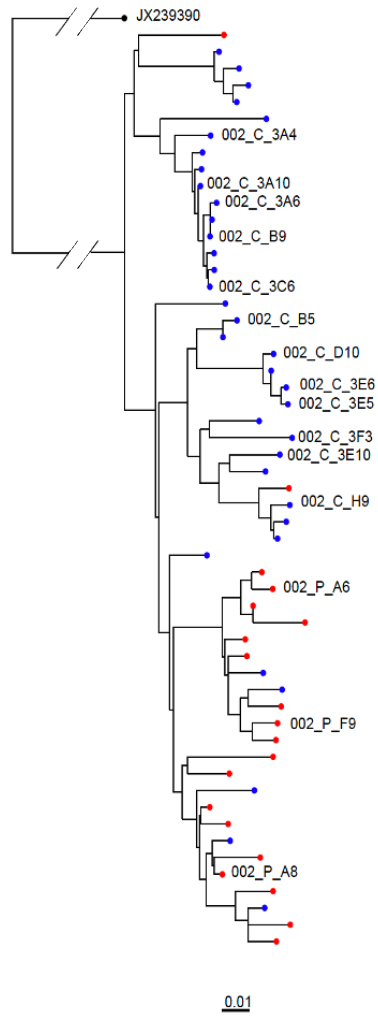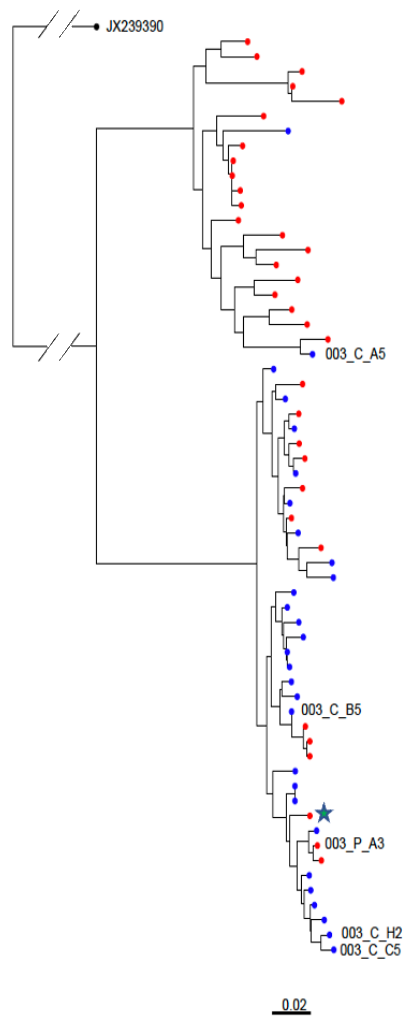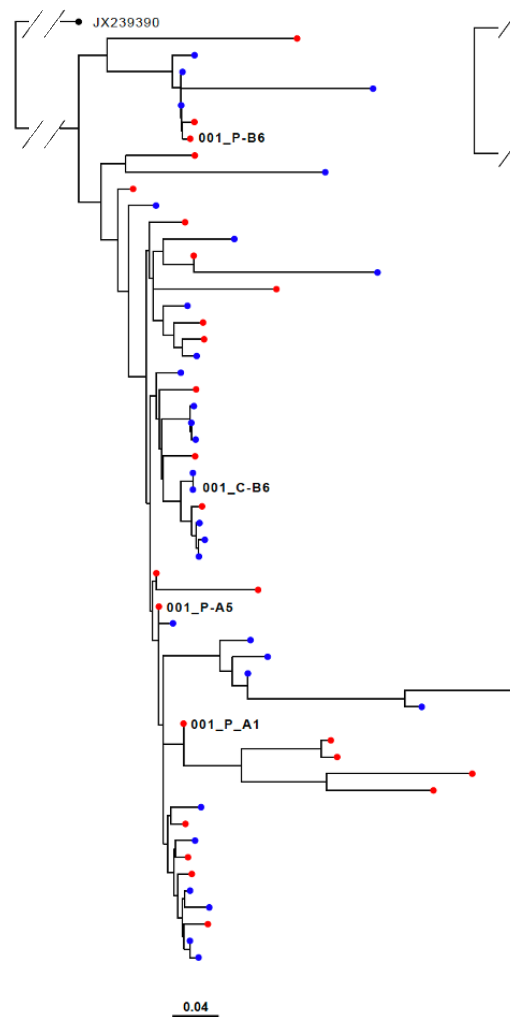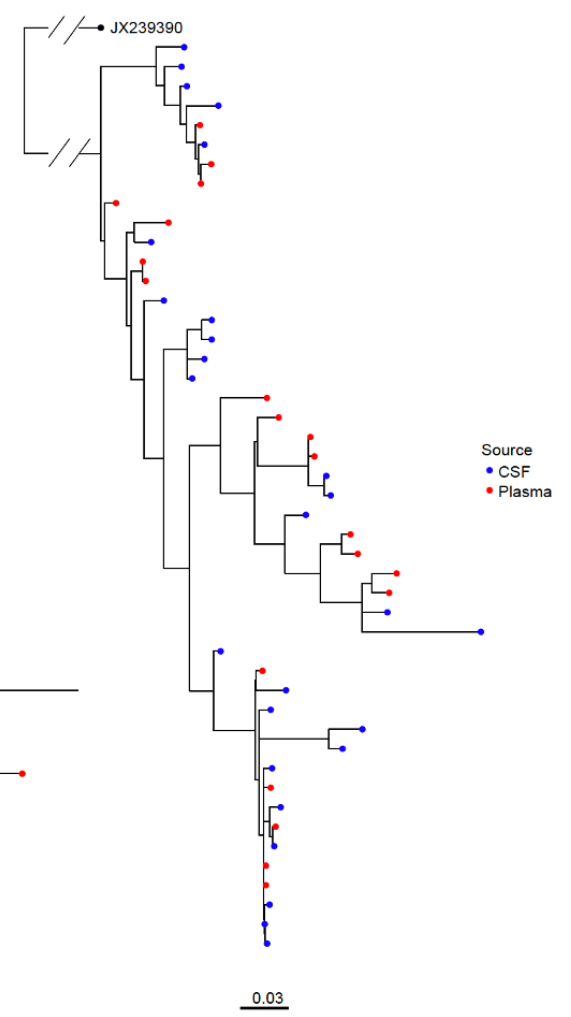
Figure 3.4. Maximum likelihood phylogenetic trees HIV-1 Env nucleic acid tree including gp120 to gp41 transmembrane domain. A) 002 (treatment experienced). B) 003 (treatment naïve) C) 001 (treatment naïve) D) 004 (treatment experienced). CSF derived viruses are shown in blue and plasma derived viruses in red. ★Represents the most recent common ancestor of a CSF cluster. CSF- cerebrospinal fluid. The genomes that are cloned and used for phenotypic characterisation in later experiments are labelled.

Distance-based and tree-based compartmentalisation analyses of the ML phylogenetic trees were done (Table 3.4). The Wright's FST analyses compared the genomes in the CSF and plasma compartments and was >0.15 in participants 002 and 003, indicating significant amount of genetic difference between the genomes on both compartments. The Wrights FST was 0.047 in participant 001, indicating very little genetic difference between compartments. The Simmons Association Index values were approaching 0 for participants 002 and 003, which supports the observations from visual inspection of the phylogenetic tree. Similarly, the Slatkin & Maddison PSS indicated that fewer migration events than expected explained the population structure observed (p<0.001) with a very small probability that this observation was due to chance. Conversely the Simmons Association Index values were higher for participants 001 and 004 and the Slatkin & Maddison PSS indicated that multiple migrations events explained the population structure. P values of 0.68 and 0.65 for participant 001 and 004 respectively was consistent with having an intermixed CNS and plasma HIV-1 viral populations.

	Visual inspection	Wright's FST	Simmons Association Index	Slatkin & Maddison PSS	
				Migration events	P value
001	Intermixed	0.047	0.339	9.9 (9.0-11.0)	0.68
002	Compartmentalised	0.158	0.079	8.5 (8.0-9.0)	<0.001
003	Compartmentalised	0.250	<0.001	3.9 (3.0- 4.0)	<0.001
004	Intermixed	Not done	0.34	15.4 (14.0- 16.0)	0.65

Table 3.4. Compartmentalisation analyses of CNS and plasma derived genomes of 4 participants.

3.3 There is increased diversity in compartmentalised participants

Env diversity was explored by calculating the average pairwise distance (APD) between *env* V2 to V4 in Mega X (Kumar et al. 2018). APD is the number of base substitutions per site from averaging over all sequence pairs within each group. Table 3.5 shows the results of analyses conducted using the Jukes-Cantor model and replicated 100 times to provide 100 bootstrap support. The rate variation among sites was modelled with a gamma distribution (shape parameter = 1).

APD amongst the participants with CNS compartmentalisation showed differences between the CSF and plasma derived viruses. In participant 002 there was more

diversity in the CSF viruses compared with the plasma viruses (5.1% vs 3.5%). In participant 003 there was more diversity in the plasma viruses compared with the CSF viruses (9.1% vs 3.7%). In participants with intermixing of CSF and plasma viruses; 001 and 004, the APD in both compartments were very similar, suggesting less genetic diversity between CSF and plasma derived viruses in participants with intermixed viral populations.

	Visual inspection	APD	SE
001, treatment naive	Intermixed		
CSF		8.3	0.007
Plasma		8.6	0.007
002, treatment experienced	Compartmentalised		
CSF		5.1	0.006
Plasma		3.5	0.005
003, treatment naïve	Compartmentalised		
CSF		3.7	0.003
Plasma		9.1	0.006
004, treatment experienced	Intermixed		
CSF		5.2	0.007
Plasma		5.5	0.006

Table 3.5: Env Variable loops 2-4 diversity in CSF and plasma compartments.
Presented as APD (%) within groups. CSF- cerebrospinal fluid; APD- average pairwise distance; SE- standard error

The intra-host viral diversity was further explored using multidimension scaling (MDS). This is a data reduction method that allows mapping of the relative distances between sequences on a plot. For participants 002 and 003 whom we have established have CNS compartmentalisation, the MDS plots show that the points representing the samples are separate by compartment, with points for CSF samples largely being separated from points for the plasma sample (Figures 3.5B and C). For participant 001 and 004, there is no clear separation between the samples (Figures 3.5A and D). This lends further weight to the cladal structure of the plasma and CSF sequences in participants 002 and 003. This was done in collaboration with Dr Steven Kemp.

A

B

001
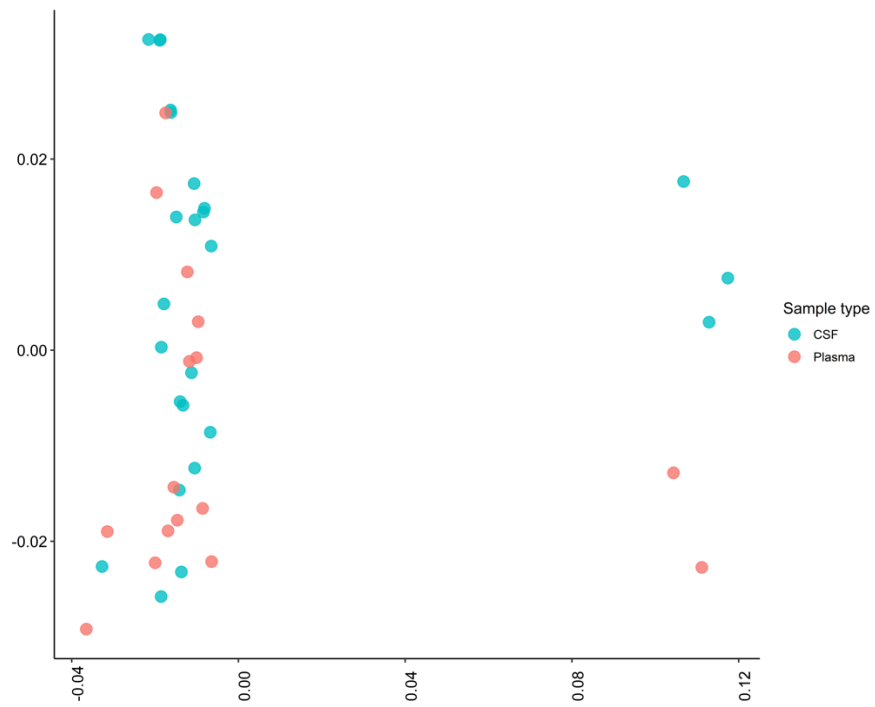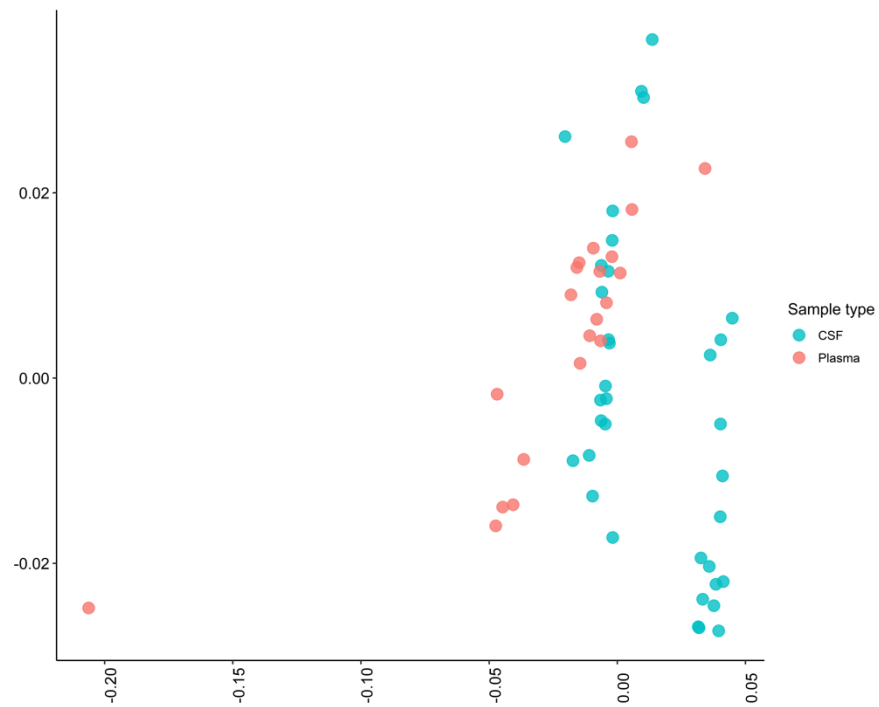
002

C
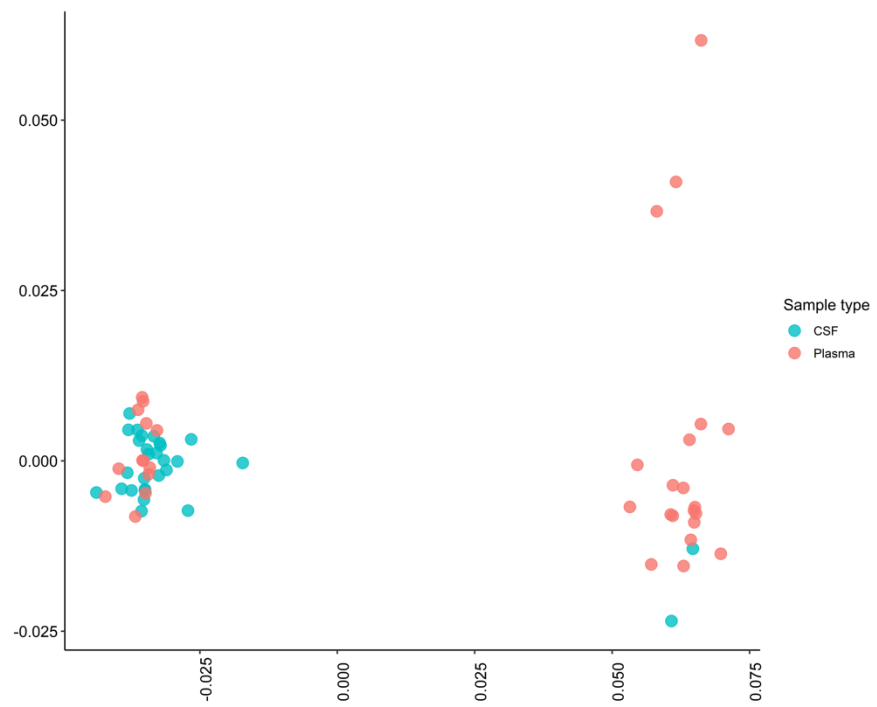
003

D
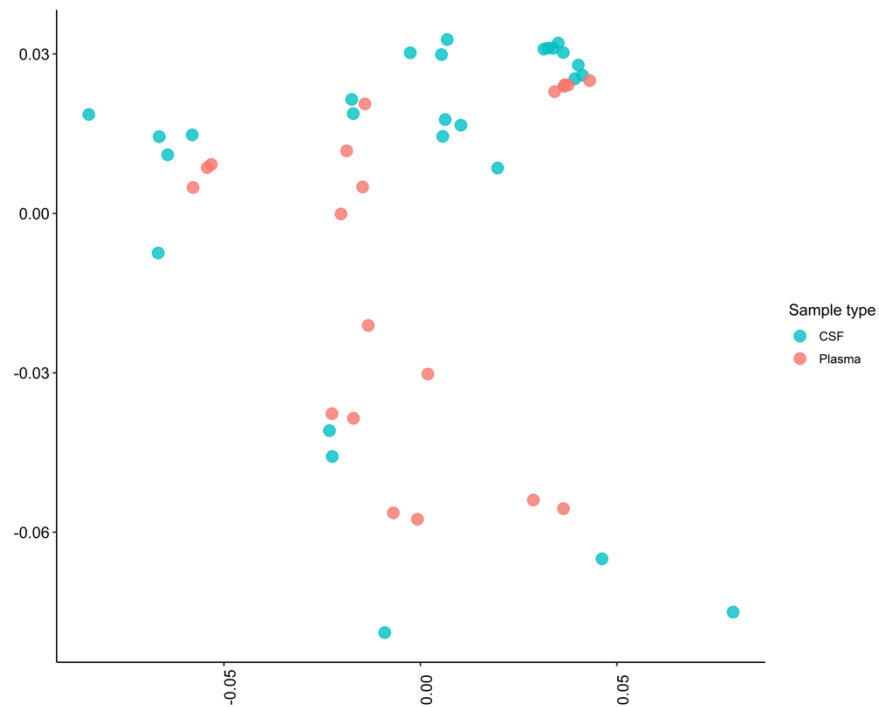
004

Figure 3.5. Multidimensional scaling plots of the average pairwise distances between sequences. A. Participant 001(treatments naive) **B.** Participant 002 (treatment experienced) **C.** Participant 003 (treatment naïve) **D.** Participant 004 (treatment experienced). Sequences are from both CSF (blue) and plasms (red) compartments for each participant and are assembled single genome amplification Env nucleotide sequences from gp120 to the TM region of gp41. TM-transmembrane domain, gp- glycopeptide. Done in collaboration with Dr Steven kemp.

3.4 No clinical or demographic factors were identified which are associated with CNS compartmentalisation

The clinical and demographic characteristics of participants with CNS compartmentalisation was compared to those without CNS compartmentalisation (Table 3.3). The age range in the CNS compartmentalised group was 27 to 31years whilst the range for the intermixed group was 34 to 60 years. The two participants in the CNS compartmentalised group were female whilst the two participants in the intermixed group were male. One participant in either group was on ART. The HIV-1 VL in CSF and plasma were similar in both groups. The range of plasma VL was 5.7 to 6.3 $\log_{10}$ copies/ml in the CNS compartmentalised group vs 5.5 to 6.2 $\log_{10}$ copies/ml in the intermixed group and the range of CSF VL was 6.3 to 6.5 $\log_{10}$ copies/ml in the CNS compartmentalised group vs 6.1 to 6.3 $\log_{10}$ copies/ml in the intermixed group. Of the two patients with CSF escape, one patient (002) had compartmentalised CNS virus and the other (001) had intermixed virus. The same

dichotomy was true in the two patients without CSF escape; 003 had CNS

compartmentalised virus whilst 004 was intermixed. There were no drug resistant

mutations identified in any compartments in 3 of the 4 participants. Participant 003

with CNS compartmentalisation had E183A which confers resistance to rilpivirine

was present in the plasma. There was a trend to a lower CD4 count; 80-98 cells/mm3

in the CNS compartmentalised group vs 146 cells/mm3 in the intermixed group was

observed. However, one CD4 count in the intermixed was missing. The Glasgow

coma scale was lower than 15 in both groups but no history of recent seizure in

either group. There was one mortality in the CNS compartmentalised group. There

does not appear to be any identified clinical or demographic factors associated with

CNS compartmentalisation. However, as there are only 4 participants in this

analysis, it is not possible to draw any firm conclusions.

3.5   Co-receptor usage by computer predicted algorithms

Effective HIV-1 entry and infection is dependent on primary binding to host CD4

receptor and then secondary binding to either CCR5 or CXCR4 co-receptors.

Although subtype dependent, co-receptor usage switch from CCR5 to CXCR4 has

also been shown to be herald the onset of advanced HIV/AIDS. CCR5 affinity is

thought to play a role in HIV-1 neurotropism. Co-receptor usage was determined

genotypically by analysing the V3 loop sequences of each amplified virus in

computer predictive algorithms (CPA)- Web position specific scoring matrix

(PSSM) and Geno2pheno (G2P). PSSM uses the net charge as well as the amino

acid at positions 11 and 25, if these are either and Arginine (R) or Lysine (K), the

virus is likely to be CXCR4 using. G2P was trained on a set of 1100 subtype B sequences with genotypic-phenotypic correlation.

Fifty-four complete V3 sequences were obtained from SGA for participant 001, 56 sequences were obtained for participant 002, 63 sequences were obtained for participant 003 and 47 sequences were obtained for participant 004. G2P consistently predicted that all genomes were CCR5 using (Figure 3.6B). In contrast PSSM predicted that co-receptor usage in participant 002 was compartmentalised, with 19 of the 35 sequences derived from the CSF predicted to utilise the CXCR4 co-receptor (Figure 3.6A). However, all the plasma sequences from participant 002 were predicted to be CCR5 using by both G2P and PSSM. The sequences derived from the CSF in participant 001 were predicted to predominantly utilise the CCR5 co-receptor-3 of 30 sequences was predicted to utilise CXCR4 by PSSM. However, all 30 CSF sequences were predicted to utilise CCR5 by G2P. All plasma sequences derived from participant 001 were predicted to be CCR5 using by both G2P and PSSM. All plasma and CSF sequences from participant 003 and 004 were predicted to utilise CCR5 co-receptor by G2P and PSSM (Figure 3.6A and B).

A   B
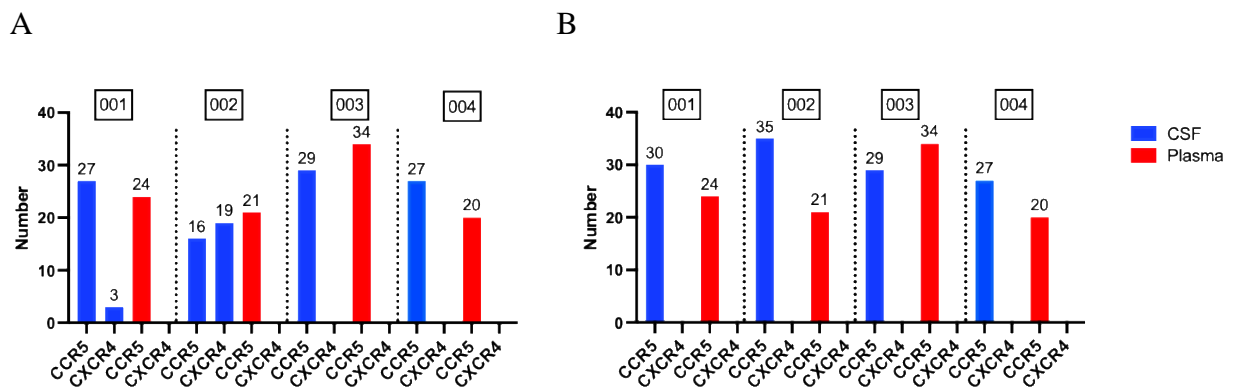
Figure 3.6: Virus co-receptor usage predicted by computer predicted algorithms.

A) By position-specific scoring matrix (PSSM) and B) Genotopheno (G2P) and using V3 loop sequences of CSF and plasma derived sequences. CSF derived sequences are in blue and plasma derived sequences are in red. CSF- cerebrospinal fluid.

3.6 Infectivity of pseudotyped viruses

Pseudotyped viruses were generated by dual transfection of the participant derived env clone in B41 psvIII env expression plasmid with pNL4-3-Env Luc packaging vectors (Methods 2.1.8.3). The infectivity of the pseudotyped viruses generated from clones from participant 002 was determine by two methods. Firstly, the pseudotyped viruses were titrated on the target cells; U87 bearing CCR5 coreceptors to determine the TCID and expressed as RLU/µl of virus input (Methods 2.1.8.5) (Figure 3.7A and 3.8A). In order to demonstrate equivalence between the TCID and SG-PERT determined infectivity, one single round infection experiment for each participant derived Env pseudotyped virus done using TCID standardised input was back corrected for SG-PERT determined RT-activity. This showed equivalence between both approaches (Methods 2.1.8.6) (Figure 3.7A, B). The TCID normalised input at the dilution of virus that gave a RLU of 500 000 RLU was used in further experiments.

3.7 Co-receptor usage in a phenotypic assay

The predicted co-receptor usage by CSF and plasma clones was tested in a phenotypic co-receptor usage experiment. A subset of clones from both CSF and plasma compartments from participants 001, 002 and 003 were tested as indicated on the phylogenetic trees (Figure 3.5A, B, C). Clones from participant 001 (Figures 3.8A, B); P_A1, P_C5 and P_B6 from plasma and C_B6 from CSF utilised CCR5 co-receptor. They were able to infect U87 CCR5+ cells at least as well as the positive controls WEAU and JRCSF (Figure 3.8A). However, they were not able to infect U87 CXCR4 cells compared to the positive control WEAU (Figure 3.8B). This corroborated the predicted co-receptor usage by the CPAs.

Clones from participant 002 (3.9A, B); P_A6, P_A8 and P_F9 from plasma and C_B5, C_B9, C_H9, C_D10, C_3F3, C_3A4, C_3A6, C_3E10, C_3A10, C_3E5, C_3C6 and C_3E6 from CSF utilised CCR5 co-receptor and were able to infect U87 CCR5+ cells as well as the positive controls WEAU and JRCSF but in contrast to the negative control HXB2 (Figure 3.9A).However these clones could not infect U87 CXCR4+ cells when compared to the WEAU positive control (Figure 3.9B). Of interest this was inconsistent with the co-receptor usage prediction by PSSM which predicted that majority of CSF clones tested were CXCR4 co-receptor utilising.

Clones from participant 003 (figures 3.10A B); P_A3 from plasma and C_A5, C_C5, C_B5 and C_H2 from CSF utilised CCR5 co-receptor and were able to infect U87

CCR5 positive cells but not U87 CXCR4 cells. This corroborated the predicted co-receptor usage by the CPAs.
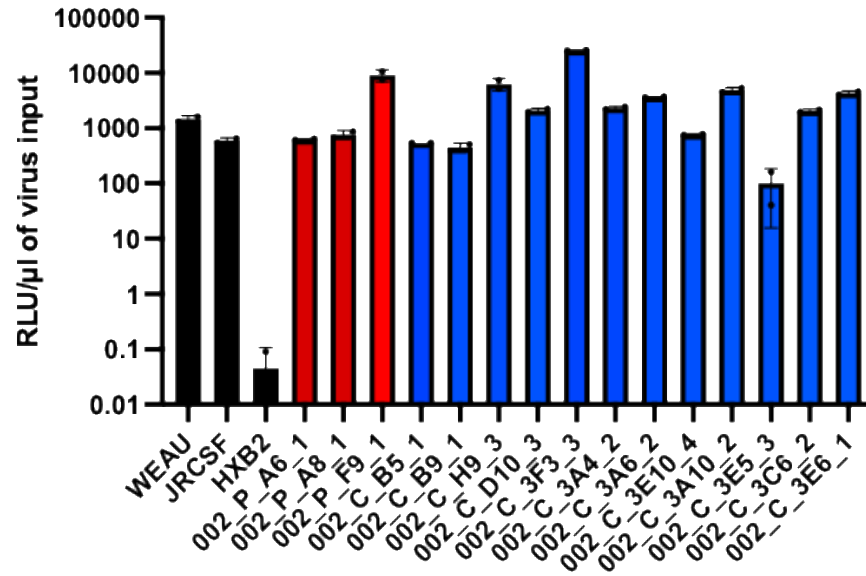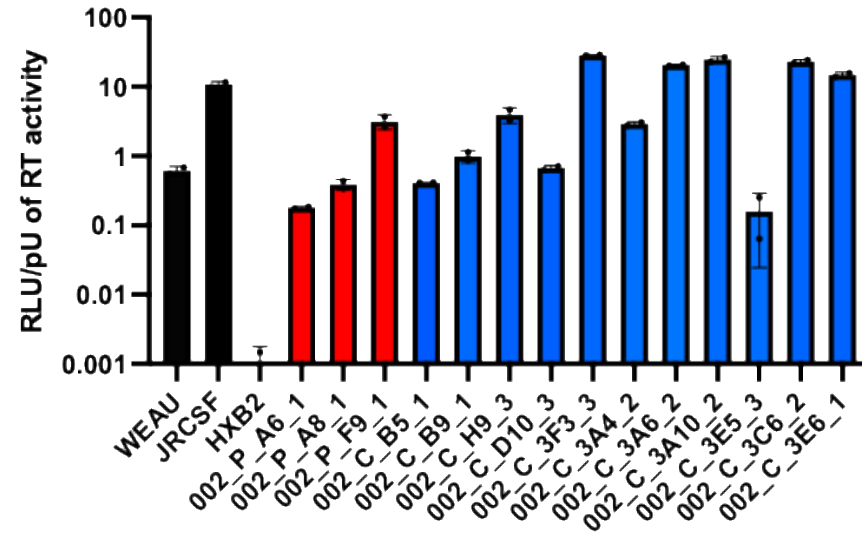
Figure 3.7 A and B. Infectivity and co-receptor usage by pseudotyped viral clones from participant 002. Infection of U87 CCR5+ target cells. **A.** TCID; Titration on U87 CCR5+ cells showing infectivity per μl of virus input. **B.** Pseudotyped virus input normalised for reverse transcriptase (RT) activity. Target cells are U87 CCR5+. Plasma clones (red), CSF clones (blue), controls (black). The clones used are on the X-axis, WEAU- WEAU-d15.410.787 used both CXCR4 and CCR5 co-receptors, JRCSF- CCR5 using variant, HXB2- CXCR4 using variant. Data are representative of two experimental repeats and they are the mean and s.d of two technical replicates.

A
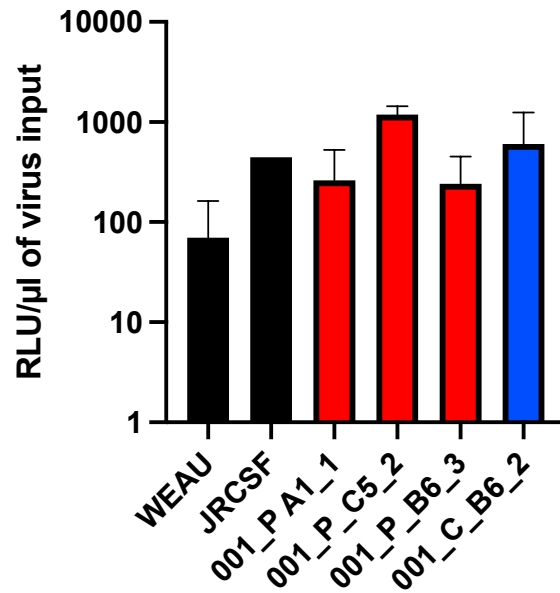
B
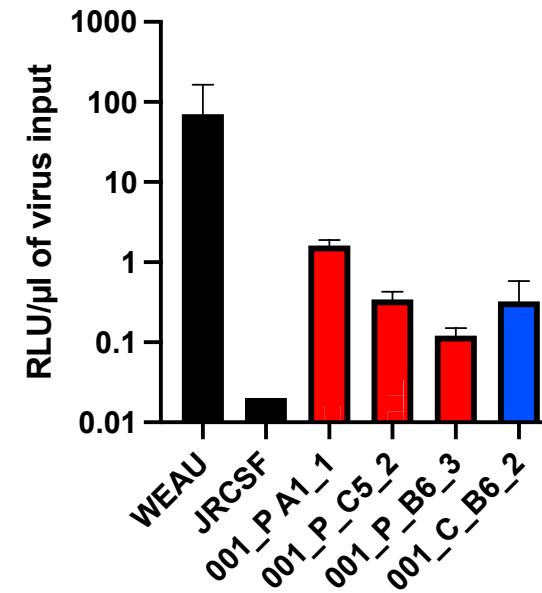
Figure 3.8 A and B. Coreceptor usage by pseudotyped viral clones from participant 001. A. Infection of U87 CCR5+ target cells. **B.** Infection of U87 CXCR4+ target cells. Plasma clones (red), CSF clones (blue), controls (black). The clones used are on the X-axis, WEAU-WEAU-d15.410.787 used both CXCR4 and CCR5 co-receptors, JRCSF- CCR5 using variant. Data are representative of two independent experiments and they are the mean and s.d of two technical replicates.
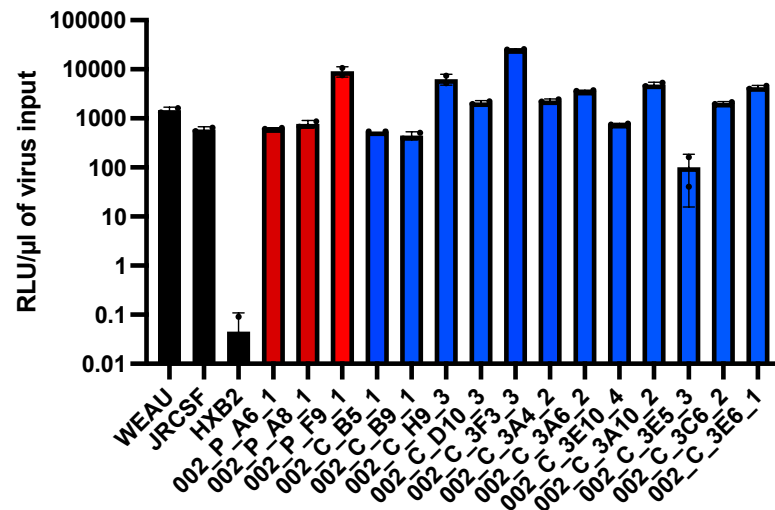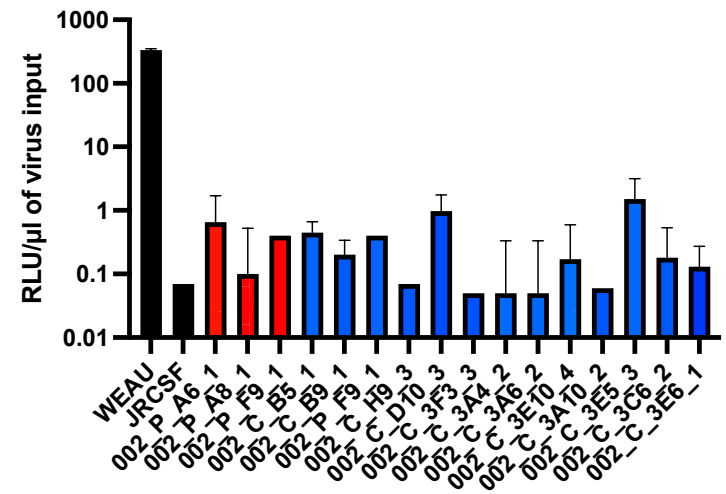
Figure 3.9 A and B. Coreceptor usage by pseudotyped viral clones from participant 002. A. Infection of U87 CCR5+ target cells. **B.** Infection of U87 CXCR4+ target cells. Plasma clones (red), CSF clones (blue), controls (black). The clones used are on the X-axis, WEAU-WEAU-d15.410.787 used both CXCR4 and CCR5 co-receptors, JRCSF- CCR5 using variant, HXB2- CXCR4 using variant. Data are representative of two independent experiments and they are the mean and s.d of two technical replicates.
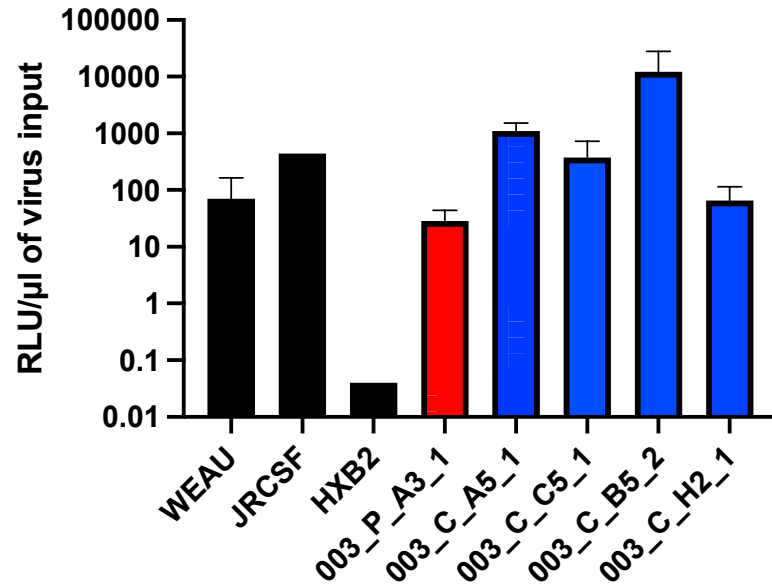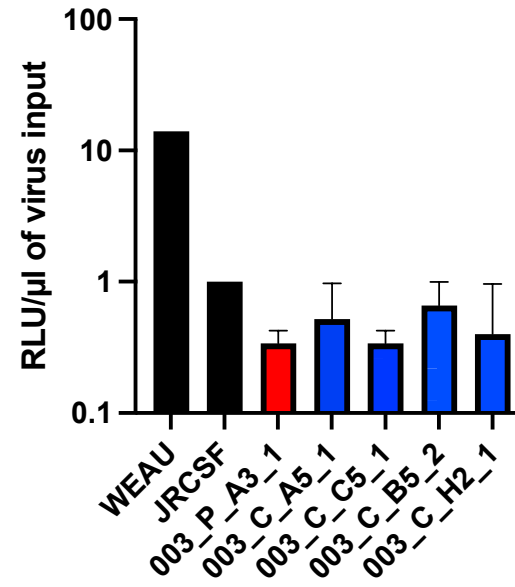
Figure 3.10 A and B. Coreceptor usage by pseudotyped viral clones from participant 003. A. Infection of U87 CCR5+ target cells **B.** Infection of U87 CXCR4+ target cells. Infection of U87 CCR5+ target cells. Plasma clones (red), CSF clones (blue), controls (black). The clones used are on the X-axis, WEAU- WEAU-d15.410.787 used both CXCR4 and CCR5 co-receptors, JRCSF- CCR5 using variant, HXB2- CXCR4 using variant. Data are representative of two independent experiment and they are the mean and s.d of two technical replicates.

3.8 Co-receptor inhibitors confirm co-receptor usage in phenotypic assay

CCR5 or CXCR4 coreceptor usage was confirmed by treating target cells in the co-receptor assay with co-receptor antagonists 800nM AMD-3100 to block CXCR4-mediated entry and 400nM TAK-779 to block CCR5-mediated entry prior to infection. Due to the global COVID-19 pandemic and interruption to access to the CL3 laboratory as SARS-CoV-2 work was prioritised only a limited number of experiments could be done on the replication competent pseudotyped viruses. I optimised an assay that could be performed in a CL2 lab that utilised a triple plasmid transfection system to generate pseudotyped virus with the participant derived clones (Methods 2.1.8.4). This utilised p8.91 packaging vector containing gag-pol, tat and rev but lacking the other accessory proteins present in pNL43△env. The infectivity of a subset of viruses generated by both systems was compared and this showed that the triple plasmid system with p8.91 gave higher virus titre compared with pnl43 (Figure 3.11).
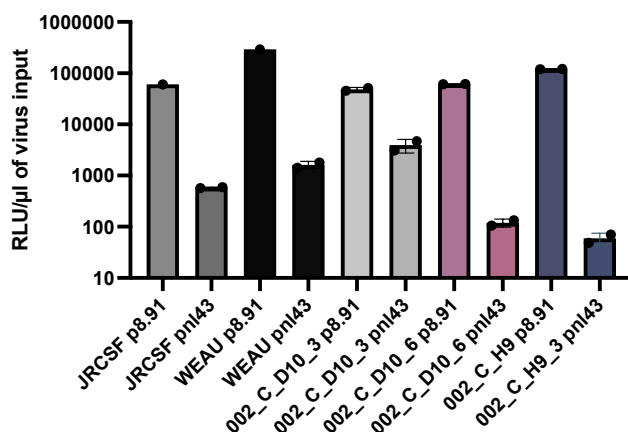
Figure 3.11. Virus infectivity comparing 5 pseudotyped viruses generated by a double or triple plasmid transfection system. The clones used are on the X-axis,

WEAU- WEAU-d15.410.787 used both CXCR4 and CCR5 co-receptors, JRCSF-CCR5 using variant. p8.91- lentiviral packing vector containing gag-pol, tat and rev. pnl43△env is a packing vector containing gag-pol, and all HIV-1 accessory proteins.

Drug inhibition by the CCR5 inhibitor TAK-779 showed a drug dose response on the JRCSF; the CCR5 utilising control (Figure 3.12A and 3.13A). There was a similar dose response by all the plasma and CSF clones tested for participant 002. The CSF clones were inhibited to a higher degree than the plasma clones and were and there was complete inhibition of infection to at least the level of the bald p8.91 control at the highest dose used at 400nM. The plasma clones also showed a dose-response to TAK-779 with a 2-log reduction in infection at the highest dose used but overall was not as sensitive to inhibition by the CCR5 blocker as the CSF clones (Figure 3.12A).

Inhibition by the CXCR4 inhibitor AMD 3100 had the expected effect on HXB2- the CXCR4 utilising clone, with complete inhibition of infection at the highest dose used 800nM (Figure 3.12B and 3.13B). CXCR4 blocking however had no effect on infection by the plasma or CSF clones from participant 002, as these were all CCR5 co-receptor using (Figure 3.12B). The plasma clones from participant 003 similarly showed a dose response to inhibition by TAK-779 (Figure 3.13A) but no effect of AMD-3100 (Figure 3.13B) as again they were CCR5 co-receptor using.

A
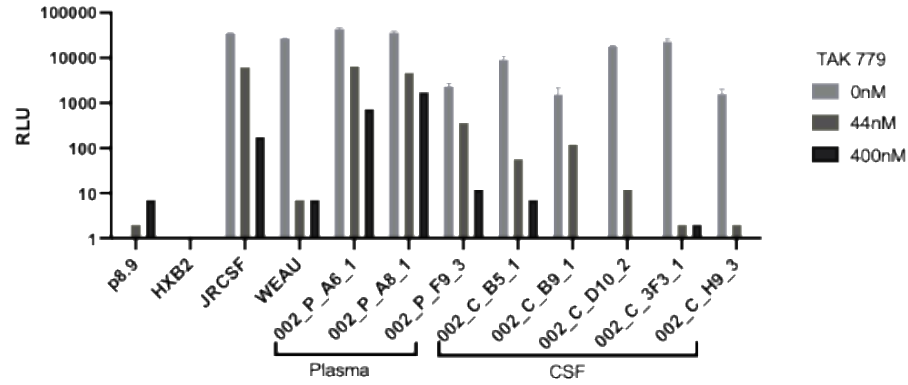
B
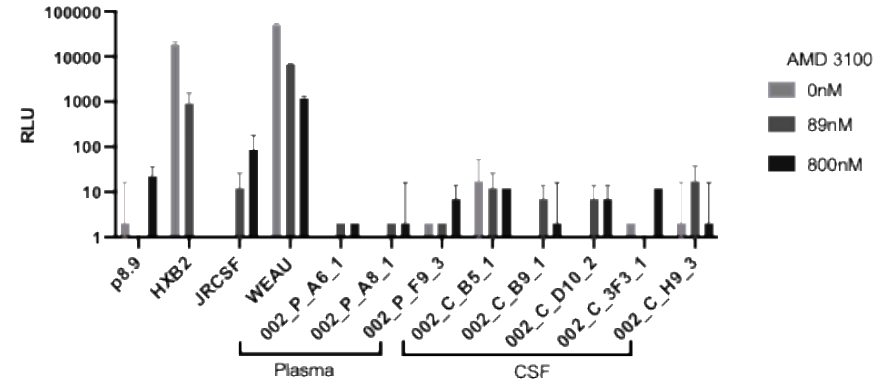
Figure 3.12. Dose response of coreceptor inhibitors on infection by pseudotyped viral clones from participant 002. A. Infection of U87 CCR5+ target cells. **B.** Infection of U87 CXCR4+ target cells. TAK 779- CCR5 inhibitor, AMD 3100- CXCR4 inhibitor, RLU- relative light units. Data are representative of two independent experiment and they are the mean and s.d of two technical replicates.
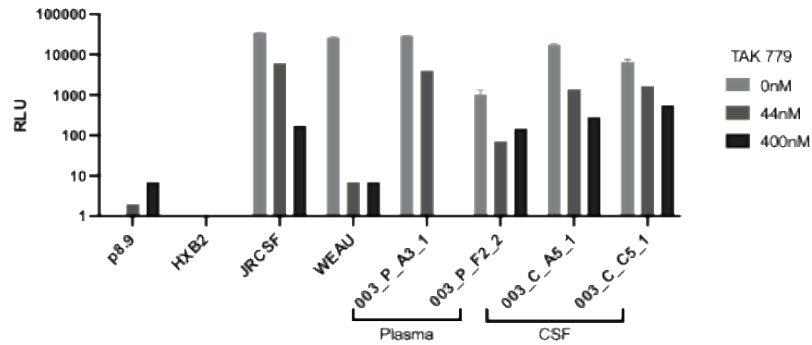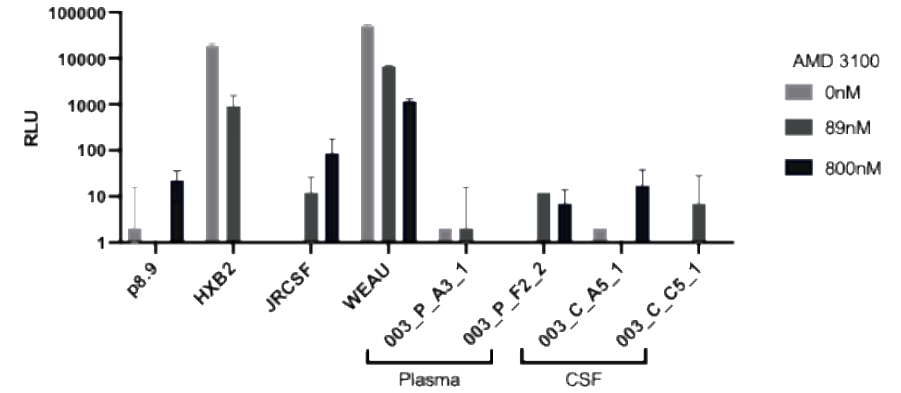
Figure 3.13. Dose response of coreceptor inhibitors on infection by pseudotyped viral clones from participant 003. A. Infection of U87 CCR5+ target cells. **B.** Infection of U87 CXCR4+ target cells. TAK 779- CCR5 inhibitor, AMD 3100- CXCR4 inhibitor, RLU- relative light units. Data are representative of two independent experiment and they are the mean and s.d of two technical replicates.

3.9 The frequency of drug resistance mutations is similar in both compartments

Consensus sequences of HIV-1 from paired plasma and CSF from the 31 participants with successful next generation sequencing were generated and analysed for drug resistance mutations (DRMs). The pol gene was examined for DRMs using the Stanford HIV drug resistance database (https://hivdb.stanford.edu/hivdb/by-sequences/report/). The frequency of NRTI and NNRTI DRMs encountered in CSF and plasma compartments were largely similar, with some small differences (figures 3.14A-D). Previously reported DRMs associated with CNS compartmentalisation - M184V and M230L were found in equal frequencies in plasma and paired CSF samples. Thymidine analogue mutations (TAMs) were more frequent in the plasma compared with the CSF specimens (11 in plasma vs 8 in CSF). NRTI mutations E44D (1 in CSF vs 0 in plasma) and K70E/R (5 in CSF vs 3 in plasma) were more frequent in CSF compared with plasma samples. However, in 3 of the 4 participants analysed for CNS compartmentalisation, there were no drug resistant mutations found in either compartment. In one participant-003 with CNS compartmentalisation, E138A conferring resistance to rilpirivine was found in the plasma (Table 3.3).
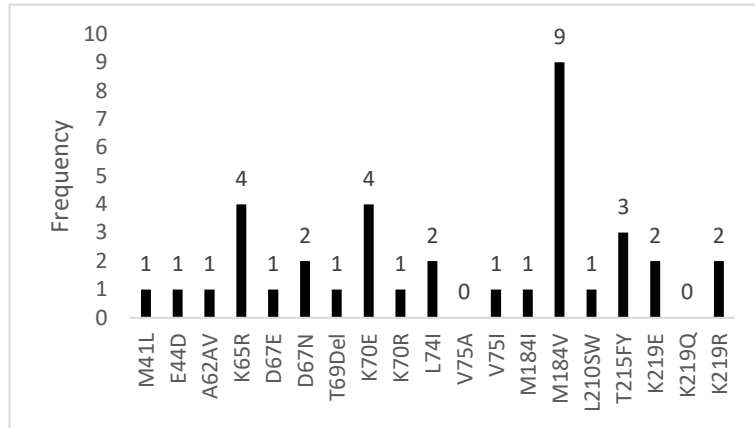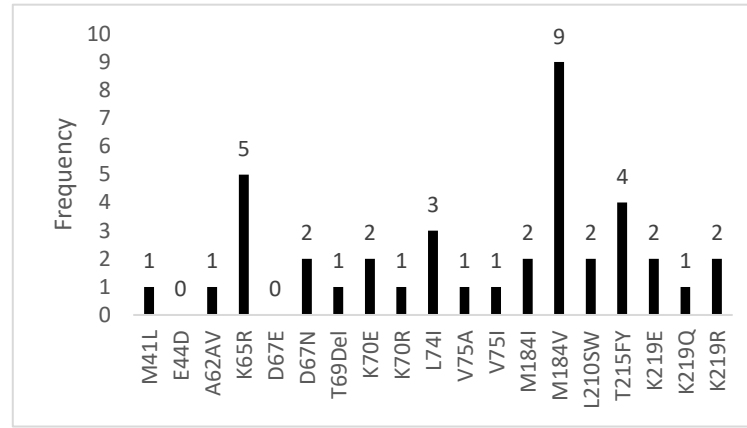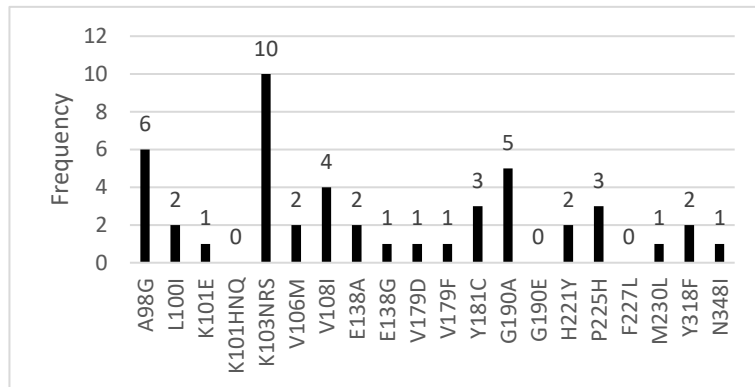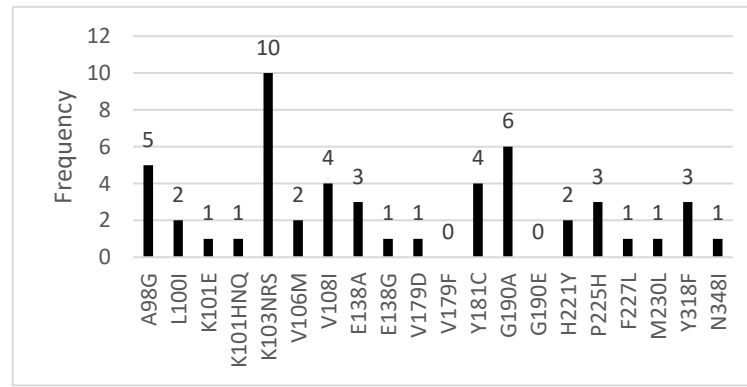
Figure 3.14. Drug resistance mutation in consensus sequences of CSF and plasma derived HIV-1.

A. NRTI DRMs in CSF compartment B. NRTI DRMs in plasma compartment C. NNRTI DRMs in CSF compartment D. NNRTI DRMs in plasma compartment. DRMs- drug resistant mutations; NRTI- nucleotide reverse transcriptase inhibitors; NNRTI- non-nucleotide reverse transcriptase inhibitors; CSF-cerebrospinal fluid.

3.10 Heterogeneity of CNS and plasma derived Env clones to bnAbs

The neutralising activity of 4 broadly neutralising antibodies (bnAbs) against participant derived Env pseudotyped viruses was measured in a single-round *in vitro* neutralisation assay (Methods 2.1.9). Serum neutralisation curves were generated from which inhibitory concentration 50 (IC50) were determined. IC50 is the concentration of bnAbs required to inhibit 50% of infection of a Env pseudotyped virus in the assay. The 4 bnAbs were PGT121 and PGT128 (targeting the high mannose patch), VRC01 (CD4 binding site blocker) and PG9 (targeting the apex). Anti-dengue serotype 3 antibody was used as the negative control. This was done for the 2 participants 002 and 003 who had evidence of CNS compartmentalisation.

Clones from participant 002 were universally resistant to neutralisation by PGT121 and PGT128 (Figure 3.15, Table 3.5). Clone 002_C_B9 form the CSF was sensitive to neutralisation by VRCO1(Figure 3.15A) and clone 002_P_F9 from the plasma was sensitive to neutralisation by PG9 (Figure 3.15B). The CSF derived clones from participant 003 were between 3 to 30- fold times more sensitive to neutralisation by VRCO1 compared with the plasma clone (Figure 3.16A, Table 3.15). None of the 003 clones were neutralised by PG9 (Figure 3.16B). All the 003 clones were sensitive to neutralisation by PG121 (figure 3.16C). The CSF clone 003_C_A5 was neutralised by PGT128 and the plasma clone 003_P_A3 only just reached 50% neutralisation at the highest concentration of PGT128 (Figure 3.16D).
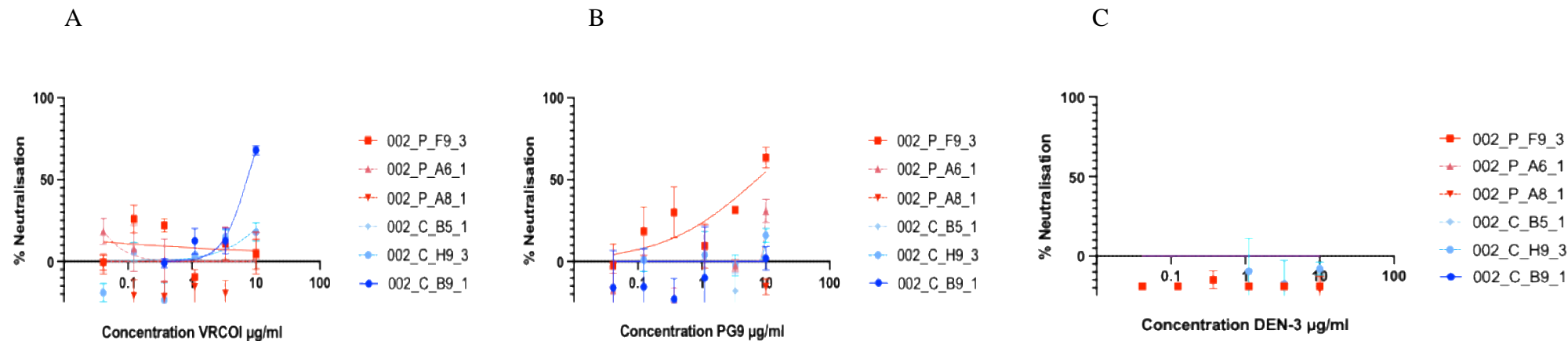
Figure 3.15. Neutralisation sensitivity of Env pseudotyped lentiviruses from participant 002 to bnAbs. A. Representative neutralisation curves for Env clones against VRCO1. B. Representative neutralisation curves for Env clones against PG9. C. Representative neutralisation curves for Env clones against DEN-3 (negative control). bnAbs- broadly neutralising antibodies. Data are mean and s.d of two technical repeats and are based on one experiment.
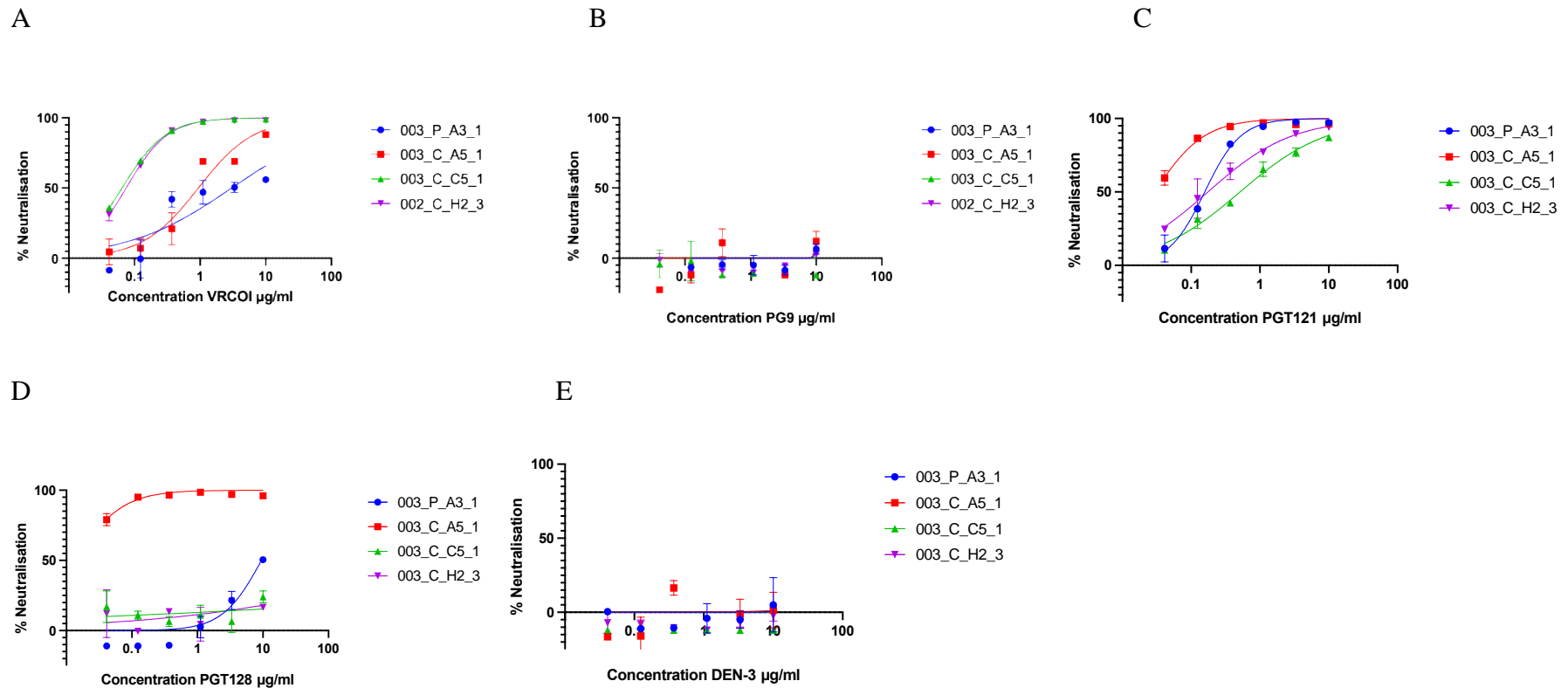
Figure 3.16. Neutralisation sensitivity of Env pseudotyped lentiviruses from participant 003 to bnAbs. A. Representative neutralisation curves for Env clones against VRCO1. **B.** Representative neutralisation curves for Env clones against PG9. **C.** Representative neutralisation curves for Env clones against PGT121. **D.** Representative neutralisation curves for Env clones against PGT128. E. Representative neutralisation curves for Env clones against DEN-3 (negative control). bnAbs- broadly neutralising antibodies. Data are mean and s.d of two technical repeats and are based on one experiment.

| bnAb | IC50 µg/ml | | | | | | | | | |
| | Participant 002 | | | | | | Participant 003 | | | |
	002_P_A6_1	002_P_A8_1	002_C_B5_1	002_C_B9_1	002_P_F9_3	002_C_H9_3	003_P_A3_1	003_C_A5_1	003_C_C5_1	002_C_H2_3
PGT121	-	-	-	-	-	-	0.2	0.03	0.5	0.2
PGT128	-	-	-	-	-	-	9.5	0.01	-	-
VRC01	-	-	-	0.9	-	-	3.0	0.9	0.1	0.1
PG9	-	-	-	-	7.2	-	-	-	-	-
DEN-3	-	-	-	-	-	-	-	-	-	-

Table 3.6. Inhibitory concentration 50 of Env clones from participants 002 and 003.

IC50- inhibitory concentration 50, bnAb- broadly neutralising antibody. Good neutralisation (green), moderate neutralization (orange), poor neutralisation (pink), no neutralisation (red).

3.11 Discussion

In this cross-sectional study CSF escape was found to be uncommon at a prevalence of 4%. The prevalence of CSF escape is heavily influenced by the limit of detection of the assay used to quantify viral load. On the other hand, there was evidence of CNS compartmentalisation in 2 out of the 4 participants who underwent SGA and sequencing. Multiple approaches including visual inspection, distance and tree-based methods were used to determine CNS compartmentalisation, providing confidence in the assignation of this status.

Pairwise distance analyses of sequences revealed compartment-related heterogeneity in Env diversity amongst participants with CNS compartmentalisation. It has been shown that there is a linear correlation between pairwise distance and recency of HIV-1 infection (Moyo et al. 2017). It was hypothesised that if CNS compartmentalised viruses gained entry into the CNS early in acute HIV-1 infection and established latency in long-lived cells, then the diversity of CNS compartmentalised viruses will be lower than that of the plasma. This was the case in the treatment naïve participant but not the treatment experienced. It is certainly plausible the ART may have exerted a selective sweep on the viruses in the plasma, limiting its diversity (Feder et al. 2016b) compared to the CNS where there is poor drug penetration (Letendre 2011). Analyses of recency of HIV-1 infection was limited by the lack of longitudinal sampling and also it was not possible to make inference about the evolutionary rate of the viruses in either compartment.

The phenotypic co-receptor inhibition assay revealed all clones were CCR5 using, and infection was blocked by CCR5 inhibitors. If as hypothesised, that CSF clones had adapted their cellular tropism for macrophages in the CNS with low-level CD4, it would be compelling to have tested for greater affinity for the CCR5 co-receptor by the CSF clones. In future work the co-receptor usage assay will be repeated with the inclusion of Affinophile target cells which have inducible levels of CD4 and CCR5 using doxycycline and Ponasterone A respectively (Joseph, Lee, and Swanstrom 2014).

An interesting observation was that the CSF clones from participant 003 were 3 to 30 -fold more sensitive to neutralisation activity of VRCO1- CD4 binding site antibody but more resistant to CCR5 inhibition. This may suggest that these CSF clones have become adapted to binding CCR5 with higher affinity. The antibody neutralisation data will be more meaningful if I was able to clone autologous antibodies from each compartment in the participants in order to test neutralisation activity of autologous antibodies and cross neutralisation against their own virus from the plasma. This would give more insight into the antibody selection pressure that drove evolution of the virus in that compartment.

The *env* coverage from Illumina MiSeq was insufficient to be validated against the SGA data, which precluded analyses of the NGS data for compartmentalisation. In addition, analyses of clinical and demographic factors associated with CNS compartmentalisation was limited by the small sample size.

This chapter highted the importance of longitudinal sampling and hence time-solved samples in studying within-host adaptation. The following chapter explores within-host adaptation of SARS-CoV-2 in a case of chronic infection that had sampling from multiple timepoints over the course of 101 days.

CHAPTER 4: RESULTS 2- SARS-COV-2 ESCAPE FROM POLYCLONAL CONVALESCENT SERA

Key Findings:

- Description of *in vivo* evolution of SARS-CoV-2 in an immunocompromised host with limited T cell and innate responses and absent B cell responses.

- Dynamic shifts in viral population can occurs in a host driven by immune pressure from convalescent plasma.

- D796H mutation enabled the virus escape serum neutralisation, co-occurred with △H69/V70 which compensated for the loss of infectivity.

- These mutations had no impact on Spike incorporation and are not within the Spike RBD but may impact on antibody binding by allosteric mechanisms.

The Spike protein SARS-CoV-2 engages with of the human ACE2 receptor and is a major antibody target. This is because it is the means by which the virus attaches and gains entry into host cells (V'kovski et al. 2021). Natural infection by SARS-CoV-2 elicits potently neutralizing monoclonal antibodies that recognize the RBD and other

epitopes of the Spike protein (Chi et al. 2020). Using mutagenesis in yeast display vector in combination with Deep Mutational Scanning, the Bloom lab have been able to map mutations in RBD that escape antibody binding (Greaney et al. 2021). They provided a picture of the functional consequences of single mutations within the RBD, most of which are deleterious to protein folding, protein expression and ACE2 binding (Starr et al. 2020). Outside of the RBD immunodominant epitopes to polyclonal sera have been found at the fusion peptide and linker region upstream of the heptad repeat region 2 (Garrett et al. 2020; Lu et al. 2021; Li et al. 2020).

Given the rapid spread and amplification of SARS-CoV-2, variants that can evade the immune response are likely to arise. Already the virus has adapted with various Variants of Concern emerging since the start of the pandemic. It is known that some of these changes increase binding to the human ACE2 receptor such as N501Y (Collier, De Marco, et al. 2021; Starr et al. 2020), other changes for example the B.1.1.7 variant has have been associated with increased transmission and higher viral loads (Frampton et al. 2021), some changes have been associated with vaccine-elicited antibody escape for example E484K (Collier, De Marco, et al. 2021).

This was an in-depth case- study of an immunocompromised man with chronic infection with SARS-CoV-2. The aim of this study was to characterise the dynamics, genotypic and phenotypic properties of the virus population in this individual's respiratory samples over 101 days of illness, with a view to providing insights into *in vivo* escape of SARS-CoV-2 from neutralising antibodies.

4.1 Clinical Case History

4.1.1 COVID-19 clinical case in an immunocompromised host

A man in his 70s presenting to the Emergency Room with acute shortness of breath
and dry cough. He was diagnosed with COVID-19 after testing positive for SARS-
CoV-2 RT-PCR test on a nasopharyngeal swab with a CT value of 25, 35 days prior.
The day of his first PCR positive SARS-CoV-2 test was designated day 1 (Figure
4.1). His past medical history was significant for Marginal B Cell Lymphoma
diagnosed 8 years prior and was treated with vincristine, prednisolone,
cyclophosphamide and anti-CD20 B cell depletion with rituximab, last given one
year prior. Clinical examination revealed an elevated respiratory rate of 42/min and
oxygen saturation of 92% on 6 L of oxygen. Laboratory tests on admission were
significant for a C-reactive protein (CRP) of 245 mg/L, lymphopaenia and IgG <1.4
g/l. Imaging studies on admission including a Chest x ray showing widespread
interstitial opacification and CT chest done showed widespread, basal and dependent
predominant ground glass and peri broncho vascular consolidation. He was
transferred to the intensive care unit (ITU) where he received non-invasive
ventilation.
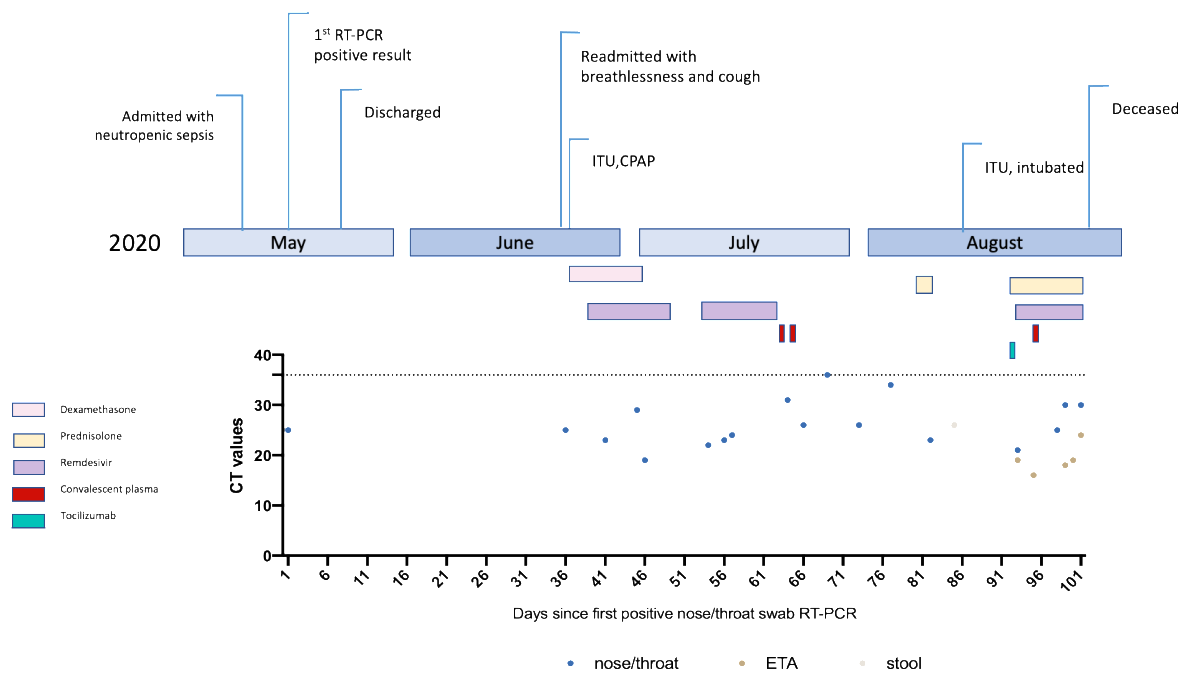
Figure 4.1. Time course of clinical events with day of sampling and CT values.

CT, cycle threshold; CPAP, continuous positive airway pressure; ITU, intensive therapy unit; ETA, endotracheal aspirate; RT-PCR- real time- polymerase chain reaction. Source: Kemp and Collier et al, SARS-CoV-2 evolution during treatment of chronic infection, Nature 592, 277–282 (2021). https://doi.org/10.1038/s41586-021-03291-y.

Treatment included antibiotics, dexamethasone, and two ten-day courses of remdesivir- an inhibitor of the viral RNA-dependent, RNA polymerase, with a five-day gap in between (Figure 4.1). He subsequently received convalescent plasma (CP) on days 63 and 65. His condition remained stable on the ward despite having persistently positive SARS-CoV-2 nose/throat swabs for RNA (Figure 4.1). On day 77, he developed fever and fast atrial fibrillation with persistently elevated CRP of 303 mg/L. Chest CT on day 86 showed new dense left upper lobe consolidation,

148

worsening inflammatory changes, chronic organizing pneumonia and ongoing

changes compatible with persistent SARS-CoV2 infection (Figure 4.2).
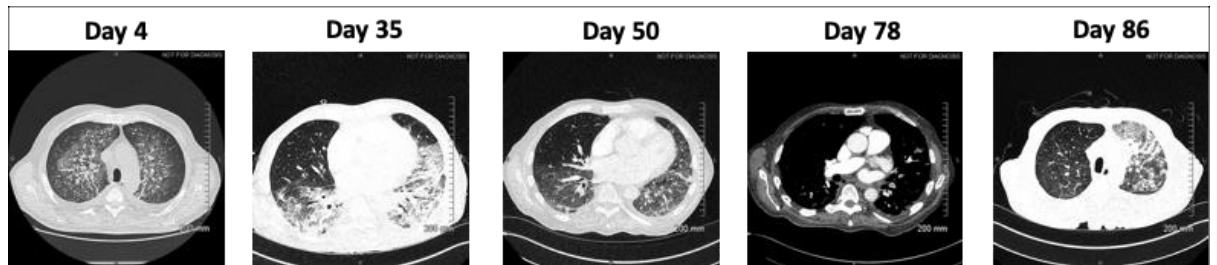
**Figure 4.2. Time course of CT chest radiological changes following PCR

detection of SARS-CoV-2.** CT- computer tomography. Day 4- Initial presentation

with ground glass and peribroncho-vascular consolidation, intralobular septal

thickening and honeycombing. Day 35- Improvement with evidence of resolving

pneumonitis. Day 78- Bilateral pulmonary embolism. Day 86- Worsening

inflammatory changes, chronic organizing pneumonia and ongoing changes

compatible with persistent SARS-CoV2 infection.

Clinical deterioration ensued with persistent fevers, low blood pressure, tachycardia,

and hypoxia requiring high-flow nasal oxygen. Broad spectrum antibacterial,

antifungal and steroid treatments were given. All cultures and fungal antigen tests

were negative. Due to increasing oxygen requirements, hypotension and renal failure

he was transferred to ITU on day 93 where he was intubated and ventilated. He

received tocilizumab- the Interleukin-6 (IL-6) monoclonal antibody and

prednisolone on day 93 followed by remdesivir and 1 further unit of CP on day 95.

His inflammatory markers began to decline following tocilizumab (Figure 4.3). CT

values on tracheal aspirate samples also showed signs of increasing following the combination of treatments but unfortunately the individual died on day 101 (Figure 4.1).
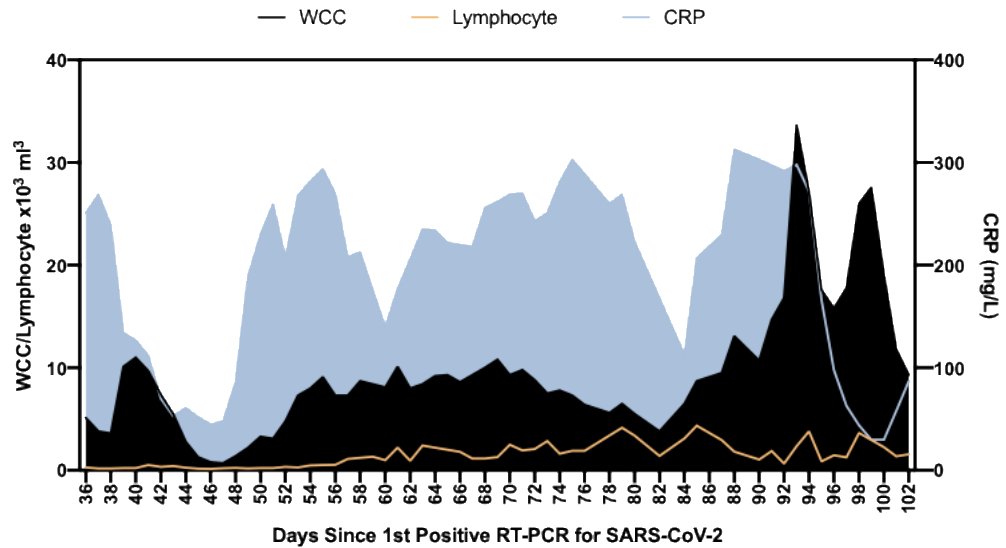
Figure 4.3. Time course of inflammatory markers in this clinical case.

WCC- white cell count, CRP- C reactive protein.

4.1.2 B cell depletions and deficient SARS-CoV-2 binding antibodies

The participant had received repeated cycles of rituximab anti-CD20 B cell depletion therapy and consequently he was hypogammaglobulinaemic with reduction in total serum IgG and IgA. The serum IgM was within the normal reference range which is either due to his underlying lymphoproliferative state or ongoing infection (Table 4.1). SARS-CoV-2 specific Spike (S), Receptor binding domain (RBD) and Nucleocapsid (N) IgG antibodies were measured in the participant's serum at

multiple timepoints using multiplex particle-based flow cytometry (Luminex Corp, Texas, USA) and an immunoassay (Siemens) in the hospital diagnostic laboratory. This was compared with antibody levels in unexposed healthy controls with pre-pandemic biobanked sera and PCR positive COVID-19 cases (Xiong et al. 2020). Levels of SARS-CoV-2 specific antibodies were also measured in the 3 units of CP received by the participant. SARS-CoV-2 serology was negative by both Luminex or Siemens on Days 39, 44 and 50, which preceded the first doses of CP. After the first dose of CP on day 63, SARS-CoV-2 serology from the participants sera was positive at multiple timepoints from day 64 up till day 101 (Figure 4.4, Table 4.1) and as expected, following the second dose of CP on day 65 SARS-CoV-2 specific antibody titres fell over time before increasing after the third unit of CP on day 93 (Figure 4.4).

	Reference range	Day 44	Day 49	Day 50	Day 68	Day 89	Day 90	Day 101
IgG	6.00 – 16.00 g/l		<1.40			1.9		
IgA	0.80 – 4.00 g/l		<0.30			<0.30		
IgM	0.50 – 2.00 g/l		1.27			1.92		
Total S	(Siemens)	Neg		Neg	Pos	-	Pos	Pos

Table 4.1. Serum immunoglobulin levels pre- and post- receipt of convalescent plasma (CP).

Following receipt of the 1st unit of CP on day 63, the presence of SARS-CoV-2 Spike antibodies is evident. Ig- Immunoglobulin, S- Spike protein, Neg- negative, Pos-positive.
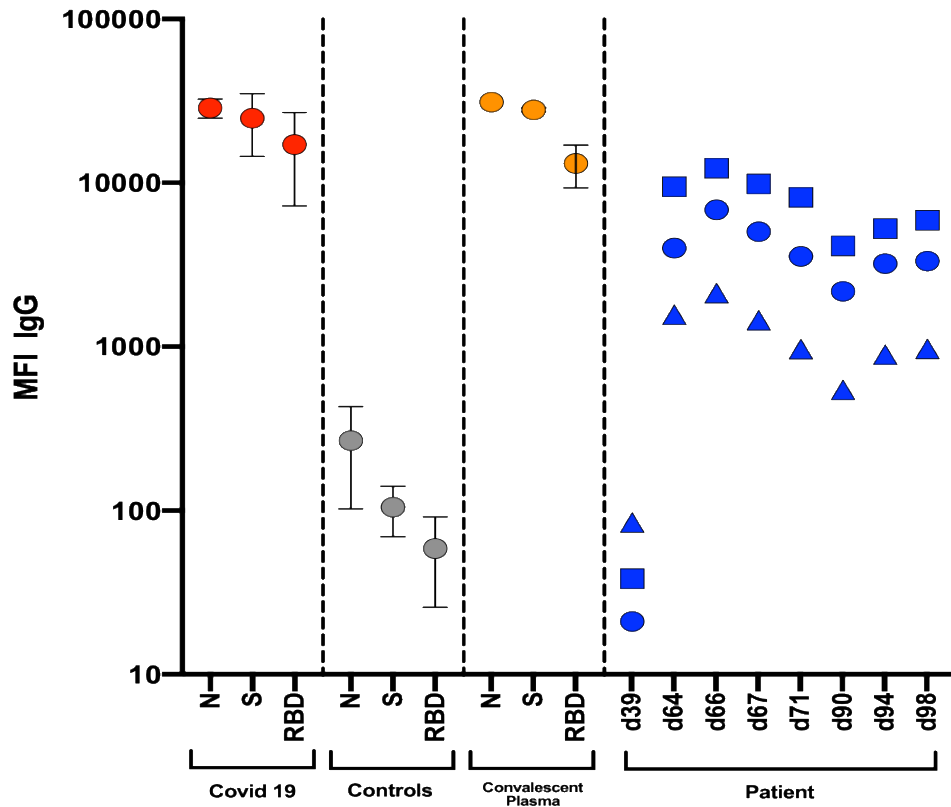
Figure 4.4. Serum antibody levels against SARS-CoV-2. Anti-SARS-CoV-2 IgG antibodies against the nucleocapsid (N) protein, trimeric spike (S) protein and RBD of SARS-CoV-2 measured by multiplexed particle-based flow cytometry (Luminex). PCR positive COVID-19 cases (red, n = 20), unexposed healthy controls (grey, n = 20), Convalescent plasma (gold, n=3). Participant sera from day 39 to 101 (blue). Anti-SARS-CoV-2 IgG to nucleocapsid (squares), Spike (circles) and RBD (triangles). MFI- mean fluorescent intensity (MFI). Ig- immunoglobulin. Data are MFI ± s.d. Serology was done by the Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge, UK. Source: Kemp and Collier et al, SARS-CoV-2 evolution during treatment of chronic infection, Nature 592, 277–282 (2021). https://doi.org/10.1038/s41586-021-03291-y.

Samples of the 3 units of CP administered to the participant were obtained from the NHS Blood and Transplant Clearance Registry. This had been tested at source using the Euroimmun assay (Medizinische Labordiagnostika) and had a level of 6 and above, indicating high levels of binding antibodies (Figure 4.5). Multiplexed particle-based flow cytometry was also used to measure the levels of binding antibodies, which was comparable to PCR-positive COVID-19 cases (Figure 4.4). After these potent CPs were administered, SARS-CoV-2 specific antibodies were detected in the participants sera. Serology was done by the Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge, UK.
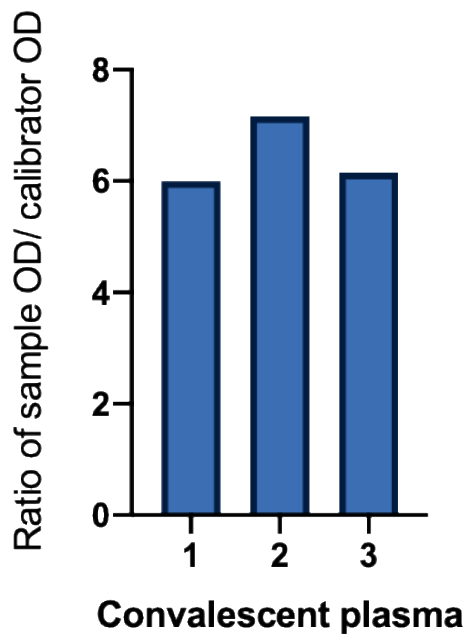
Figure 4.5. SARS-CoV-2 antibody titres in convalescent plasma. Measurement of SARS-CoV-2 specific IgG antibody titres in three units of convalescent plasma (CP) by Euroimmun assay. OD, optical density. Serology was done by the NHS Blood and Transplant Service.

4.1.3 Innate and T cell response despite absence of B cell responses

Whole blood cytokine responses were measured at days 71 and 98. IFNγ, IL-17, IL-2, TNF, IL-6, IL-1b and IL-10 levels were measured by multiplex particle-based flow cytometry after stimulation for 24 hours either after T cell stimulation with phytohaemagglutinin (PHA), anti-CD3/IL2 antibodies or Phorbol 12-myristate 13-acetate/ionomycin, (PMA/IONO) or innate stimulation with lipopolysaccharide (LPS). This was compared to responses from unexposed health controls (UHCs). T cell stimulation of the participant's samples by PHA and α CD3/IL2 showed reduced IFNγ, IL-17 and IL-2 responses compared with UHCs, particularly at the later day 98 timepoint. The day 71 response was lower than UHCs whilst the day 98 IFNγ response to either stimulant was completely absent (Figure 4.6). This timepoint just preceded the participant's death. A similar pattern was seen for IL-17 responses to PHA and α CD3/IL2 and IL-2 responses to PHA. IL-2 levels in response to PMA/IONO stimulation of T cells were higher than PHA stimulation but still lower than responses in UHCs (Figure 4.6). Innate responses were measured following stimulation with LPS. This is resulted in release of TNFα, IL-1β, LI-6 and IL-10 proinflammatory cytokines, although the levels were lower than responses from UHCs, with the day 98 responses being lower than the day 71 responses.
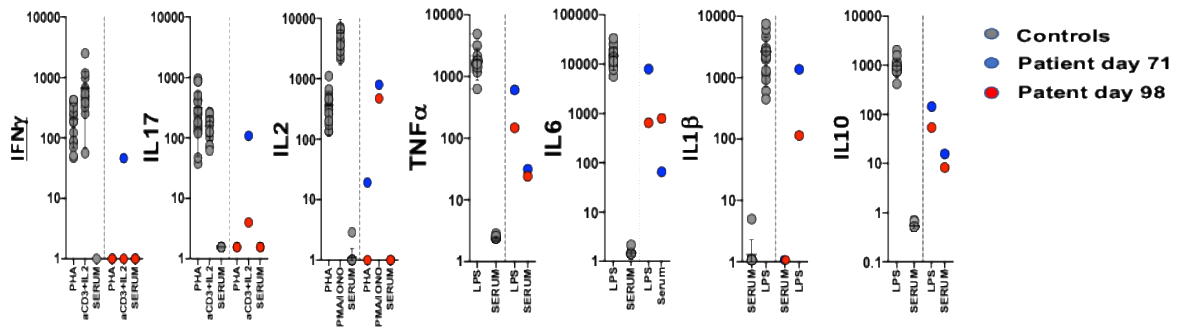
Figure 4.6. Assessment of T cell and innate function. IFNγ, IL-17, IL-2, TNF, IL-6, IL-1b and IL-10 levels in a whole blood cytokine assay measured in pg ml^{-1}. Data from unexposed healthy controls (grey, n= 15), participant samples at days 71 (blue) and 98 (red). Data are mean ± s.d. PHA- phytohaemagglutinin, αCD3+IL-2- anti-CD3/Interleukin- 2 antibodies, serum- matched serum from the same time point to account for background signal, PMA/IONO- Phorbol 12-myristate 13-acetate/ionomycin, LPS- lipopolysaccharide, IFNγ- interferon gamma, IL-interleukin, TNFα- tissue necrosis factor alpha. Experiment done in collaboration with Professor Rainer Doffinger.

4.2 Deep sequencing of participants samples revealed diversity in the SAR-CoV-2 viral population

Twenty-three respiratory samples were collected over 101 days. The cycle threshold (CT) values ranged from 16 to 34. All 23 successfully underwent long-read deep sequencing using MinION flow cells v.9.4.1 (Oxford Nanopore Technologies) by The COVID-19 Genomics UK Consortium (COG-UK; https://www.cogconsortium.uk/). Twenty of these, additionally underwent short-read deep sequencing using Illumina MiSeq, done by the Department of Pathology, University of Cambridge.

Independently, single genome amplification (SGA) and sequencing of the *spike* gene was used to detect mutations in 3 samples at different time points (Table 4.2).

Timepoint (Day)	Sequence Identifier	CT value	Sample type	Nanopore	Illumina	SGA
1	NB16_CAMB-1B19D5	25	NT	Y	Y	Y
37	NB23_CAMB-1B4FB2	-	NT	Y	Y	Y
41	NB01_CAMB-1B5124	23	NT	Y	Y	-
45	NB08_CAMB-1B529A	29	NT	Y	Y	-
50	NB13_CAMB-1B54A3	-	NT	Y	Y	-
54	NB10_CAMB-1B5643	22	NT	Y	-	-
55	NB08_CAMB-1B5467	-	NT	Y	Y	-
56	NB11_CAMB-1B5616	23	NT	Y	-	-
57	NB14_CAMB-1B5607	24	NT	Y	-	-
66	NB05_CAMB-1B55FB	26	NT	Y	Y	-
82	NB06_CAMB-1B55CE	23	NT	Y	Y	-
86	NB07_CAMB-1B5A29	-	NT	Y	-	-
89	NB01_CAMB-1B5ADE	-	NT	Y	Y	-
93	NB16_CAMB-1B5BAE	21	NT	Y	Y	-
93	NB17_CAMB-1B5B71	19	ETA	Y	Y	-
95	NB02_CAMB-1B5BCC	16	NT	Y	Y	-
98	NB01_CAMB-1B5BDB	25	NT	Y	Y	Y
99	NB09_CAMB-1B5C50	30	NT	Y	Y	-
99	NB10_CAMB-1B5C6F	18	ETA	Y	Y	-
100	NB13_CAMB-1B5CE7	-	NT	Y	Y	-
100	NB14_CAMB-1B5CBA	19	ETA	Y	Y	-
101	NB15_CAMB-1B5CC9	24	ETA	Y	Y	-
101	NB16_CAMB-1B5C9C	30	NT	Y	Y	-

Table 4.2. Sampling timepoint, type and sequencing methods.

Sampling time point indicates the day since 1st positive PCR for SARS-CoV-2 that the sample was taken. CT values are reported where available. NT- nose and throat, ETA- endotracheal aspirate, Y- Yes, - Not done, NA not available

1,133,032 complete sequences globally and a length of more than 29,000 bp were downloaded from GISAID on 22nd April 2021. Sequences were deduplicated and randomly subsampled using seqtk (https://github.com/lh3/seqtk) to select a random subset of 768 sequences. The maximum-likelihood phylogeny was inferred using IQTREE v2.13 using the -fast option. All of the participant's sequences clustered together, distinct from the majority of global sequences. The most recent common ancestors to the participant appeared to be from Italy and England (Figure 4.7). Lineages were assigned to each sequence using Phylogenetic Assignment Of Named Global Outbreak Lineages (PANGOLIN). The participant belonged to lineage B.1.1.1.
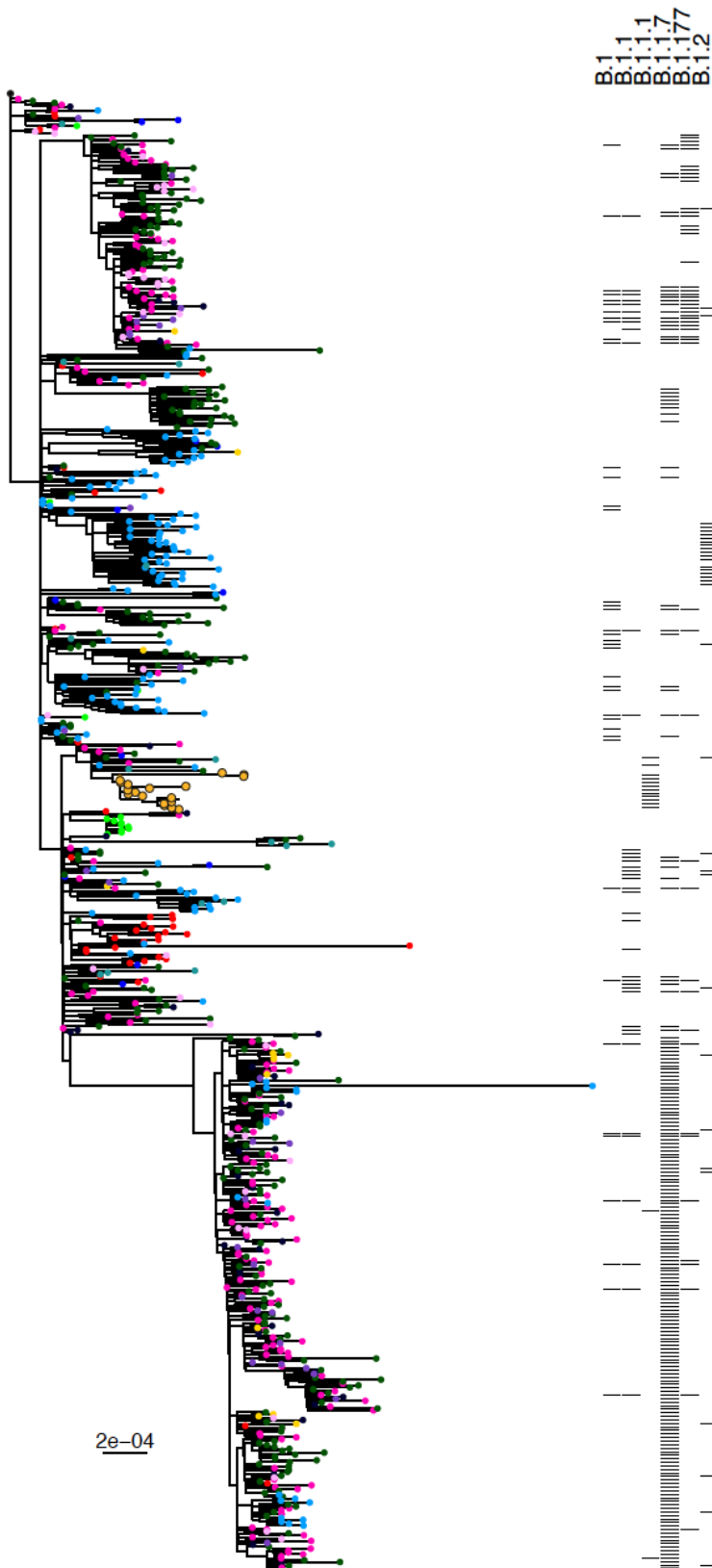
Figure 4.7. Analysis of whole-genome sequences of SARS-CoV-2 of the participant in the context of local, UK and globally derived sequences. The maximum-likelihood phylogenetic tree is rooted on the Wuhan-Hu-1 reference sequence, showing a subset of total of 768 sequences of SARS-CoV-2 from GISAID.

The Illumina Miseq short read sequences were used for further analyses. Visual inspection of the maximum likelihood phylogenetic tree of the participant's sequences revealed the existence of two distinct viral populations. The samples collected at days 93 and 95 bearing the W64G and P330S Spike mutations were rooted within, but significantly divergent from the original population (Figure 4.8). The relationship of the divergent samples to those at earlier time points argues against superinfection. All samples were consistent with having arisen from a single underlying viral population.
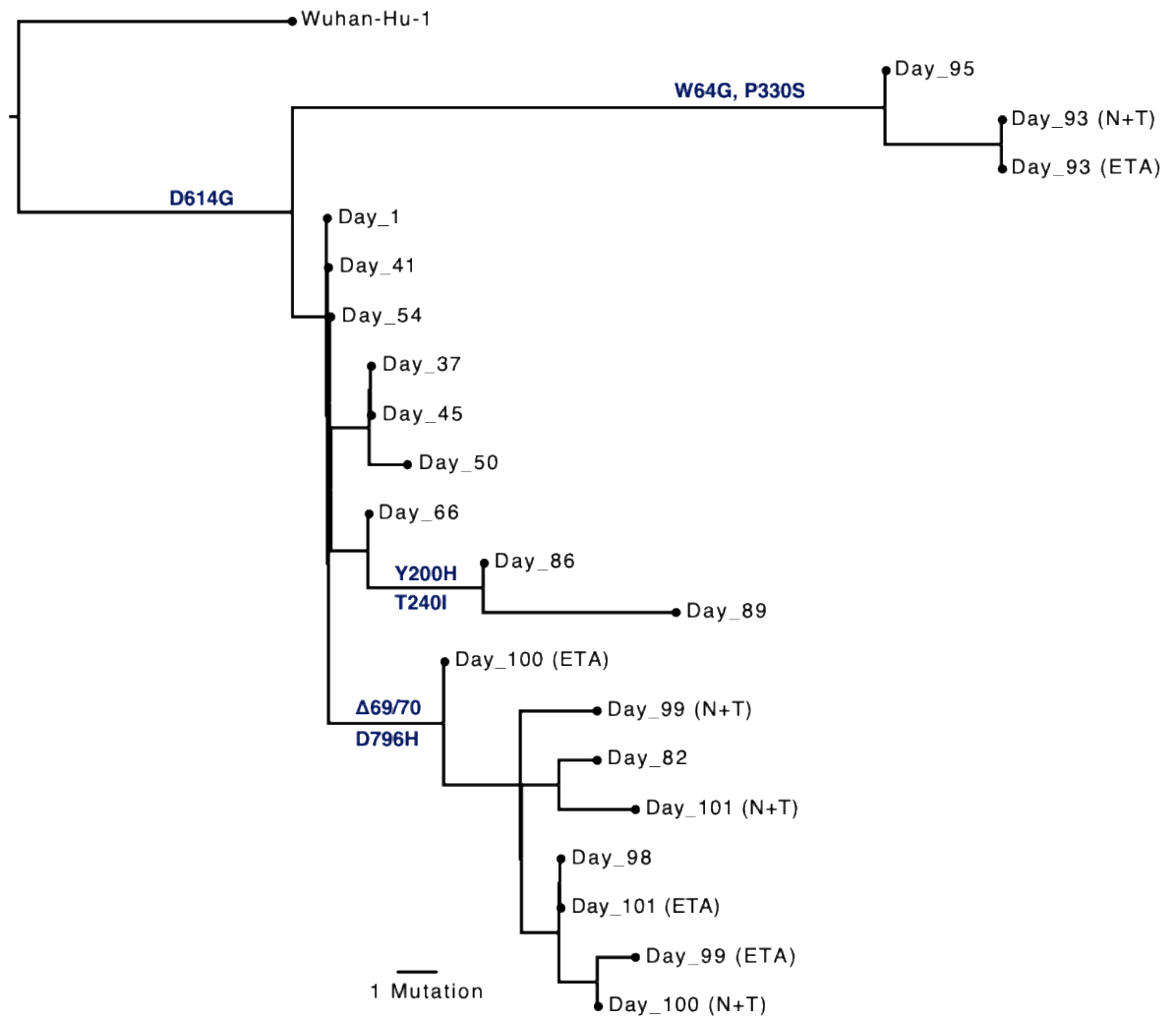
Figure 4.8. Maximum likelihood phylogenetic tree of Spike amino acid of the participant. Day of sampling is indicated. Spike mutations defining each clade are shown on the ancestral branch on which they emerged. Sample type is also indicated as endotracheal aspirate (ETA) and Nose + throat swabs (N+T). Source: Kemp and Collier et al, SARS-CoV-2 evolution during treatment of chronic infection, Nature 592, 277–282 (2021). https://doi.org/10.1038/s41586-021-03291-y.

Another way to explore the within-host population structure is to investigate the relationship between the sequence distance and the time distance i.e the sum of allele frequency distances across the whole genome, calculated between pairs of samples,

plotted against the difference in times at which the samples were collected (Figure 4.9A). This was done in collaboration with Dr Chris Illingworth using the distance metric (Lumby et al. 2020). The median pairwise distance between samples in the main clade and those collected on days 93 and 95, bearing W64G, P330S Spike mutations was significantly different (Figure 4.9B, $p<0.001$). However, when the sample collected on days 93 and 95 were removed the median pairwise distance between samples in the main clade and those collected on days 86 and 89, bearing the Y200H and T240I Spike mutation was not significantly different (Figure 4.9C, p 0.066). This provided further support for the cladal structure of the viral population.
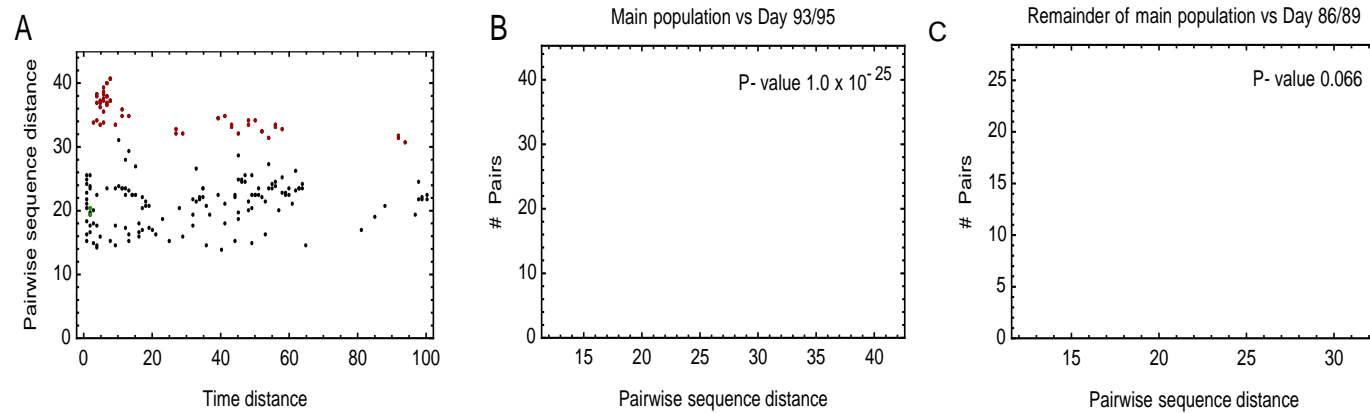
Figure 4.9. Pairwise distances between samples to explore within-host cladal structure. A. Distances between sequences from all sampling time points and days 93 and 95 (red). Internal distances between samples from samples days 93 and 95 (green). All others (black). **B.** Pairwise distances between samples in the larger clade (black) and between these samples and those collected on days 93 and 95 (red). **C.** Pairwise distances between samples in the main clade (black) and those collected on days 86 and 89 (red). P value for Mann-Whitney test. Analyses done in collaboration with Dr Chris Illingworth.

4.3 SARS-CoV-2 evolves in vivo in response to pressure from remdesivir and convalescent plasma

These analyses were done in collaboration with Dr Steven Kemp using custom code for variant calling in an AnCoVMulti package written by Professor David Pollock and is available at https://github.com/PollockLaboratory/AnCovMulti. The frequency of Single Nucelotide Polymorphisms (SNPs) were identified for each timepoint from the Illumina MiSeq reads and plotted to show dynamic population changes which were tracked longitudinally during the entire period (Figure 4.10). Shifts in viral population were noted after treatment with the first course of remdesivir at day 41, T39I (C27509T) mutation in *ORF7a* reaching 79% on day 45 and thereafter diminished (Figure 4.10A, pink, Table 4.3). Following the second dose of remdesivir I513T in NSP2 (T2343C) and V157L (G13936T) in RNA dependent, RNA polymerase (RdRp) rose to 100% frequency (Figure 4.10A, red and green dashed lines). This was probably driven by the RdRp inhibitor.

Following the first two units of CP, a variant bearing D796H in Spike S2 and **Δ**H69/**Δ**V70 in the Spike S1 N-terminal domain (NTD) becoming the dominant population at day 82. This was diminished down to less than 5% by day 86, consistent with washout of the CP. This was overtaken by Spike mutations Y200H and T240I on day 86 and 89, accompanied at a mutant pair that had been seen previously- I513T in NSP2 and V157L in RdRp, lending support that this new lineage emerged out of a previously existing population

(Figure 4.10B, 4.11, red and green lines). At day 93, Spike mutations P330S at the edge of the RBD and W64G in S1 NTD arose to close to 100% prevalence, with the Y200H and T240I paired mutations reduced to <2% (Figure 4.10B, 4.11, black lines). These shifts in the viral population suggests competition between them. Following the 3rd unit of CP the paired mutant D796H and **Δ**H69/V70 re-emerged at high frequency (Figure 4.10B) strongly supporting the hypothesis that the combination conferred selective advantage.

A

B
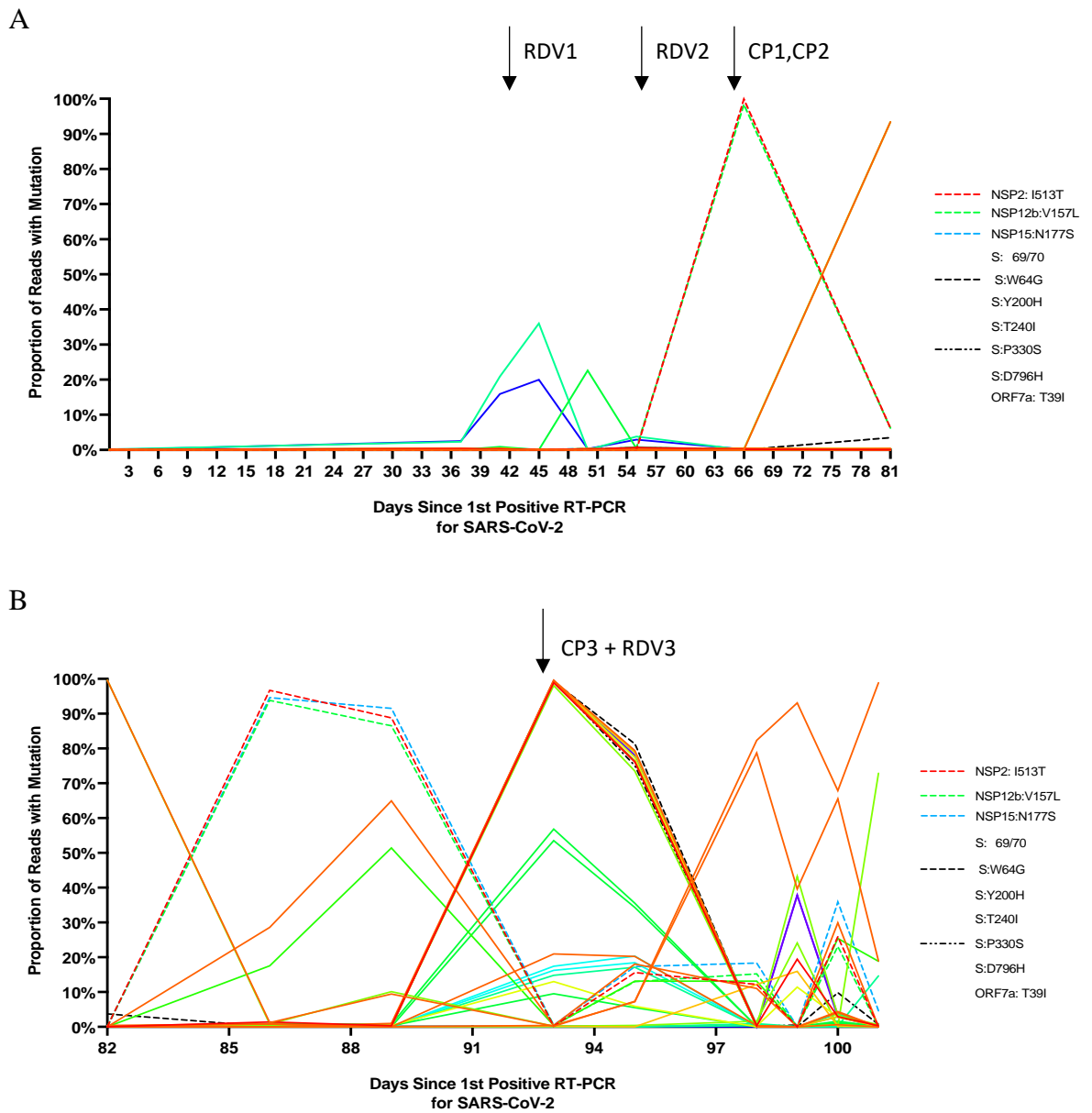
Figure 4.10. Whole-genome variant frequency plots. A. Variants detected in the participant's samples on days 1–82 after the first positive RT–PCR test for SARS-CoV-2. Spike(D796H*) (light blue) has the same frequency as NSP3(K902N) (orange) and is hidden beneath the orange line. **B.** Variants detected in the participant's samples on days 82–101. Showing amino acid changes and the timing of treatment with remdesivir (RDV) and convalescent plasm (CP). Variants shown reached a frequency of at least 10% in at least two samples. Analyses done in collaboration with Dr Steven Kemp and Professor David Pollock.
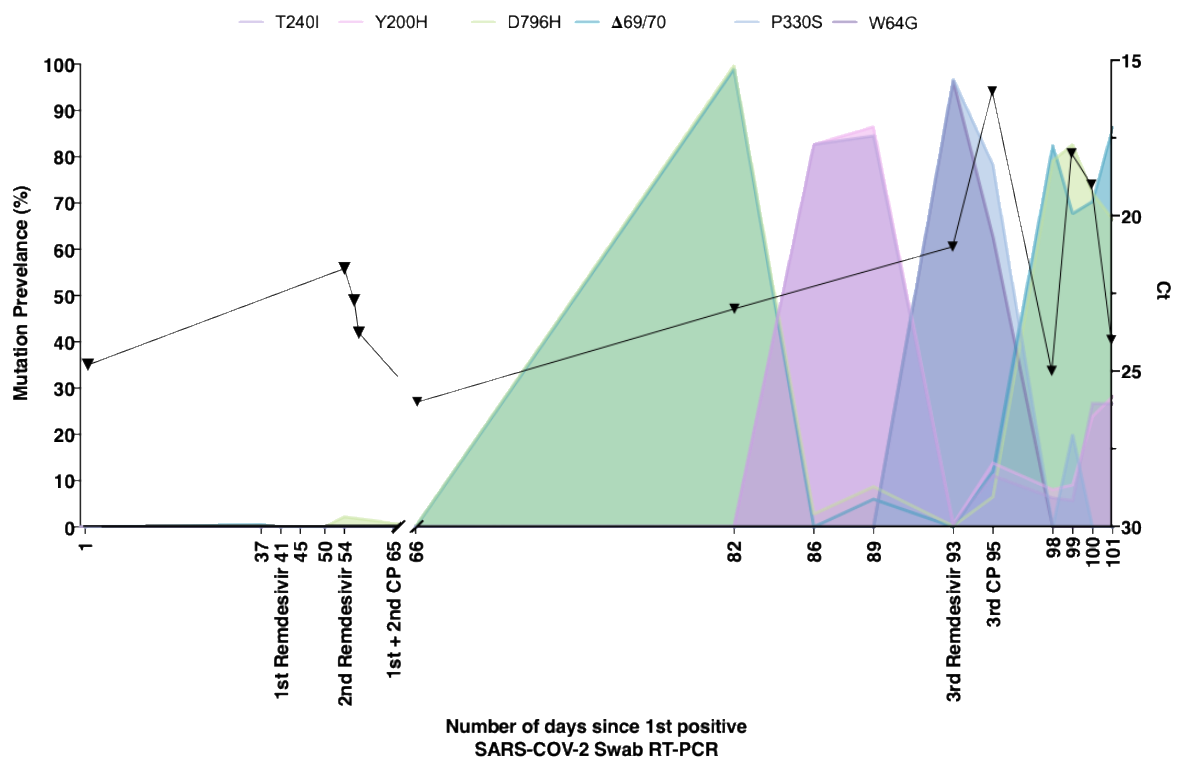
Figure 4.11. Longitudinal mutation frequency plot of 6 key mutations in SARS-CoV-2 Spike protein. Frequencies and dynamic phylogenetic relationships for virus populations bearing **Δ**69/70 + D796H, W64G+ P330 and Y200H + T240I. D796H-

aspartic acid to histidine at Spike protein amino acid position 796, △H69/V70-

histidine and valine deletion at Spike protein amino acid position 69 and 70

respectively, W64G- tryptophan to glycine at Spike protein amino acid position 64,

P330S- Proline to serine at Spike protein amino acid position 330, T200I- threonine

to isoleucine at Spike protein amino acid position 200, Y240H- tyrosine to histidine

at Spike protein amino acid position 240. Analyses done in collaboration with Dr

Steven Kemp and Professor David Pollock. Source: Kemp and Collier et al, SARS-

CoV-2 evolution during treatment of chronic infection, Nature 592, 277–282 (2021).

https://doi.org/10.1038/s41586-021-03291-y.
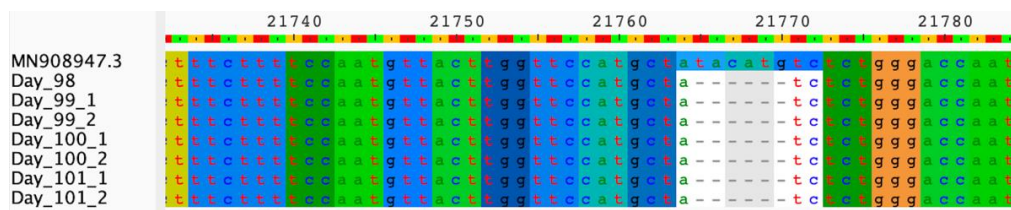
Figure 4.12. SARS-CoV-2 Spike Δ69/70 deletion. Multiple sequence alignment of

participant's SARS-CoV-2 Genomes containing the **Δ**69/70 deletion. In all cases there

is an out of frame 6 nucleotide deletion. Wuhan-1 GenBank reference MN908947.3 is

the reference variant at the top.

					Days since 1st Positive RT-PCR for SARS-CoV-2																			
Locus	From	To	Gene	Protein	1	37	41	45	50	55	66	82	86	89	93	93 (ETA)	95	98	99	99 (ETA)	100 (ETA)	100	101 (ETA)	101
401	C	T	ORF1ab	NSP1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.1	0.0	0.0	0.6	10.1
541	C	T	ORF1ab	NSP1	0.0	0.5	0.4	27.1	0.0	0.0	0.3	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1
596	G	T	ORF1ab	NSP1	0.0	0.0	0.0	0.0	0.0	0.4	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	14.9	0.1	0.0	0.0	0.6	0.0
635	C	T	ORF1ab	NSP1	0.0	3.5	0.3	18.2	0.4	0.7	1.1	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.1	0.0	0.2	0.0	0.0
1115	A	G	ORF1ab	NSP2	0.0	0.2	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17.7	0.4	0.0	0.0
1721	G	C	ORF1ab	NSP2	0.0	0.3	0.2	0.0	0.0	0.0	0.3	0.1	0.4	0.0	0.1	0.0	0.2	0.2	10.2	0.1	1.3	0.3	0.2	0.0
2236	T	C	ORF1ab	NSP2	0.0	0.3	0.3	0.0	0.2	0.7	0.2	0.0	1.4	0.3	98.9	96.8	76.1	0.3	19.4	5.5	0.3	2.9	6.8	0.3
2343	**T**	**C**	**ORF1ab**	**NSP2**	0.0	0.0	0.0	0.0	0.0	0.0	99.9	0.0	96.7	88.8	0.2	2.2	15.6	12.2	0.0	31.9	29.5	25.9	28.7	0.0
2570	C	T	ORF1ab	NSP2	0.0	0.9	16.8	6.3	0.0	1.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3030	G	T	ORF1ab	NSP3	0.0	0.2	1.3	0.2	13.3	0.0	1.0	0.7	2.0	0.3	0.2	0.5	1.7	2.0	0.3	0.4	0.3	0.3	0.0	0.0
3057	A	G	ORF1ab	NSP3	0.0	0.0	0.1	0.2	0.0	0.4	0.5	0.3	1.2	0.6	0.2	0.0	2.1	1.0	10.6	1.4	8.0	0.0	0.3	0.0
3303	A	T	ORF1ab	NSP3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	52.3
3426	C	T	ORF1ab	NSP3	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.0	0.0	0.3	16.5	0.0	0.0	0.2	0.0	0.0
3918	G	A	ORF1ab	NSP3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	99.5	98.0	76.3	0.0	0.0	0.0	0.0	4.4	7.1	0.0
4975	G	T	ORF1ab	NSP3	0.0	0.2	0.0	0.1	0.0	0.4	0.3	0.1	0.0	0.0	0.3	0.0	0.2	0.1	20.2	0.1	0.0	0.1	0.6	0.0
5011	A	T	ORF1ab	NSP3	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17.2	0.0	0.0	0.0	0.0	0.0

					Days since 1st Positive RT-PCR for SARS-CoV-2																	
Locus	From	To	Gene	Protein	1	37	41	45	50	55	66	82	86	89	93	93 (ETA)	95	98	99	99 (ETA)	100 (ETA)	100	101 (ETA)	101
5012	G	T	ORF1ab	NSP3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17.2	0.0	0.0	0.0	0.0	0.0
5339	C	T	ORF1ab	NSP3	0.0	0.0	0.0	0.0	0.0	0.5	23.0	0.1	0.0	0.5	0.2	0.5	0.1	0.0	0.0	0.1	0.0	0.1	0.0	1.4
5393	T	C	ORF1ab	NSP3	0.0	0.5	0.0	0.0	0.3	0.0	0.0	0.4	0.8	0.0	0.4	0.0	7.2	78.7	39.6	64.3	28.4	65.5	65.4	18.9
5406	T	G	ORF1ab	NSP3	0.0	0.4	0.0	0.0	0.3	0.0	0.4	0.0	28.6	64.9	0.4	2.1	18.0	11.0	0.0	30.9	16.3	29.9	28.4	0.0
5425	G	T	ORF1ab	NSP3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	99.6	1.2	9.4	0.2	0.8	7.3	82.3	93.1	65.6	53.9	67.9	65.8	98.9
5590	T	C	ORF1ab	NSP3	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.2	0.8	0.0	99.6	95.9	79.4	0.3	0.0	4.4	0.0	4.1	7.4	0.0
6070	C	A	ORF1ab	NSP3	0.2	0.6	6.3	19.2	0.0	0.0	0.4	0.4	0.3	0.6	0.0	0.0	0.3	0.0	0.0	0.1	0.0	0.0	0.3	0.0
6568	C	T	ORF1ab	NSP3	0.0	0.0	1.9	0.0	31.4	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.4	0.0	0.5	0.0
7667	G	A	ORF1ab	NSP3	0.3	0.2	0.0	0.0	0.0	0.0	0.3	0.0	0.4	0.0	0.0	0.2	0.0	0.0	0.3	0.3	0.0	0.0	0.0	13.6
7729	T	G	ORF1ab	NSP3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.3
7805	G	A	ORF1ab	NSP3	0.0	0.2	0.3	0.0	0.0	0.6	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	11.1
8389	C	T	ORF1ab	NSP3	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	20.9	18.6	20.2	0.3	0.0	1.1	0.0	0.6	1.1	0.0
8496	G	C	ORF1ab	NSP3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.0	0.0	12.1	0.0	0.0	0.0	0.0	0.0
9130	G	T	ORF1ab	NSP4	0.0	0.2	0.0	0.0	0.2	0.0	0.2	0.0	0.0	0.0	0.1	0.0	0.0	12.2	15.9	0.0	19.6	0.2	0.0	0.0
9438	C	T	ORF1ab	NSP4	0.0	0.0	0.3	0.0	0.3	0.0	0.3	0.3	0.0	0.5	99.2	97.6	78.9	0.0	0.0	5.3	0.0	2.9	9.2	0.0
9773	G	A	ORF1ab	NSP4	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	17.4	0.0	0.0	0.0	0.0	0.0
10642	G	C	ORF1ab	NSP5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	27.0

Locus	From	To	Gene	Protein	— Days since 1st Positive RT-PCR for SARS-CoV-2																			
					1	37	41	45	50	55	66	82	86	89	93	93 (ETA)	95	98	99	99 (ETA)	100 (ETA)	100	101 (ETA)	101
10700	G	A	ORF1ab	NSP5	0.0	0.0	0.5	0.0	0.0	0.2	0.4	0.3	0.0	0.3	99.3	96.4	77.4	0.2	11.4	4.6	0.2	2.8	6.6	0.2
10906	T	C	ORF1ab	NSP5	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.4	13.0	19.6	5.9	0.0	0.0	0.5	0.0	0.6	0.8	0.0
11306	G	A	ORF1ab	NSP6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	22.8	0.0	0.0	0.0	0.0	0.0
11620	C	T	ORF1ab	NSP6	0.2	0.3	0.2	0.1	0.0	0.4	0.2	99.6	0.8	10.1	0.2	0.0	0.4	1.6	24.1	0.0	8.6	0.2	0.8	72.9
11770	A	G	ORF1ab	NSP6	0.0	0.0	0.2	0.0	0.2	0.0	0.2	0.2	0.4	0.3	98.1	91.0	73.5	0.2	43.2	5.9	0.6	3.2	5.5	0.6
12043	C	T	ORF1ab	NSP7	0.0	0.1	0.0	0.0	0.3	0.0	0.0	0.0	17.5	51.4	0.4	4.1	13.1	13.1	0.0	27.5	12.0	25.4	27.7	18.8
12344	A	G	ORF1ab	NSP8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.2	12.7	0.0	0.0	0.2	0.0	0.0
12459	C	T	ORF1ab	NSP8	0.0	0.6	10.1	0.2	0.0	4.4	0.2	0.1	0.0	0.0	0.2	0.0	0.2	0.5	0.0	0.0	0.1	0.2	0.2	0.2
13527	T	C	ORF1ab	NSP12b	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.2	0.9	0.7	56.8	50.2	35.5	0.0	0.0	1.1	0.0	1.1	3.4	0.0
13936	**G**	**T**	**ORF1ab**	**NSP12b**	0.0	0.2	0.0	0.0	0.1	0.6	98.4	0.1	93.8	86.5	0.4	1.1	13.0	15.2	0.0	25.9	40.5	23.1	26.2	0.0
14068	C	T	ORF1ab	NSP12b	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.2	11.0
14776	G	T	ORF1ab	NSP12b	0.0	0.0	0.2	0.2	0.2	0.0	0.1	0.1	0.2	0.0	53.5	51.0	34.3	0.1	0.0	1.8	0.3	1.6	1.9	0.1
14913	C	T	ORF1ab	NSP12b	0.0	0.0	0.9	0.0	22.6	0.4	0.0	0.0	0.0	0.3	9.5	16.9	5.5	0.0	0.0	0.5	0.0	1.1	1.3	0.0
15814	G	A	ORF1ab	NSP12b	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.8	0.3	99.2	97.1	76.4	0.0	0.0	4.5	0.0	2.9	6.7	0.0
16733	C	T	ORF1ab	NSP13	0.2	2.3	20.9	36.0	0.0	3.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0
16901	G	T	ORF1ab	NSP13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.2	0.2	25.7	0.1	0.0	0.2	0.0	0.0
17207	A	G	ORF1ab	NSP13	0.2	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.3	0.2	99.3	97.4	78.8	0.6	0.0	5.2	0.2	3.7	9.9	0.1

Locus	From	To	Gene	Protein	Days since 1st Positive RT-PCR for SARS-CoV-2																				
					1	37	41	45	50	55	66	82	86	89	93	93 (ETA)	95	98	99	99 (ETA)	100 (ETA)	100	101 (ETA)	101
17304	C	T	ORF1ab	NSP13	0.1	41.6	0.4	0.3	0.1	1.1	0.0	0.0	0.1	0.1	0.0	0.0	0.2	0.1	0.0	0.2	2.2	0.2	0.4	0.1
17320	G	T	ORF1ab	NSP13	0.2	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.9	0.1	0.4	19.4	0.2	0.2	14.6
17358	T	A	ORF1ab	NSP13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.1	11.2	0.0	0.0	0.2	0.0	0.0
17436	C	T	ORF1ab	NSP13	0.0	0.2	0.0	0.2	0.0	0.0	21.1	0.0	0.0	0.2	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
17550	C	T	ORF1ab	NSP13	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	99.2	0.1	0.0	0.1	0.0	0.0
17703	C	T	ORF1ab	NSP13	0.0	0.2	0.0	0.1	0.0	0.8	0.0	0.0	0.0	0.0	14.8	7.8	17.1	0.0	0.0	0.8	0.0	0.7	0.9	0.1
18488	T	C	ORF1ab	NSP14	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	16.2	8.9	18.4	0.2	0.0	1.4	0.0	0.2	0.7	0.0
19388	G	A	ORF1ab	NSP14	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	17.4	10.9	20.3	0.0	0.0	0.9	0.0	1.1	1.5	0.0
20150	**A**	**G**	**ORF1ab**	**NSP15**	0.0	0.0	0.0	0.0	0.4	0.0	0.0	0.0	94.6	91.5	0.5	2.0	17.3	18.3	0.0	31.2	42.3	35.9	27.1	4.6
21600	**G**	**T**	**S**	**Spike**	0.2	0.0	0.2	0.0	0.3	0.0	0.0	0.0	0.0	0.9	99.3	96.8	76.4	0.7	15.2	22.2	0.0	66.7	6.5	0.0
21635	C	A	S	Spike	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	1.0	26.6	15.0	23.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0
21752	**T**	**G**	**S**	**Spike**	0.0	0.0	0.0	0.0	0.0	0.2	0.0	3.7	0.0	0.3	99.2	95.8	81.3	0.0	0.0	10.6	0.0	9.8	25.3	0.7
21765-2170	**T**	**-**	**S**	**Spike**	0.0	0.0	0.0	0.0	0.0	0.0	0.0	83.0	0.0	6.0	0.0	0.0	12	72.9	75.8	61.4	49.7	58.3	60.7	78.7
21855	C	T	S	Spike	0.0	0.0	0.4	0.0	0.0	11.9	0.0	0.0	0.0	0.0	0.2	0.0	0.2	0.0	0.0	0.1	0.0	0.2	0.1	0.1
22088	C	T	S	Spike	0.2	0.5	0.2	0.0	0.0	0.6	0.0	0.2	2.4	0.0	16.9	6.2	17.3	0.0	0.0	0.0	0.0	0.0	0.6	1.8
22160	**T**	**C**	**S**	**Spike**	0.2	0.0	0.0	0.0	0.0	0.0	0.2	0.2	82.0	85.8	0.5	0.6	13.9	8.5	21.2	0.0	23.2	28.6	27.6	1.7
22281	**C**	**T**	**S**	**Spike**	0.2	0.0	0.0	0.0	0.0	0.8	0.0	0.2	86.6	84.2	0.5	1.9	12.9	6.8	17.7	0.0	28.6	28.6	25.8	1.7

Locus	From	To	Gene	Protein	1	37	41	45	50	55	66	82	86	89	93	93 (ETA)	95	98	99	99 (ETA)	100 (ETA)	100	101 (ETA)	101
22550	**C**	**T**	**S**	**Spike**	0.0	0.0	0.0	0.0	0.1	0.7	0.0	0.0	0.0	0.3	99.4	96.9	75.0	0.2	0.5	6.3	0.5	0.0	5.8	0.0
22679	T	A	S	Spike	0.0	0.0	0.0	0.2	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	12.5
22959	G	T	S	Spike	0.0	0.0	1.5	0.0	0.2	0.2	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	13.4	0.0	0.0	0.1	0.0	0.0
23063	A	T	S	Spike	0.0	0.0	0.0	0.0	0.0	32.7	0.1	0.0	1.7	0.1	0.0	0.0	0.1	1.2	0.0	0.1	0.1	0.1	0.0	0.0
23398	T	C	S	Spike	0.0	0.0	0.2	0.0	0.0	0.2	0.1	0.2	24.9	65.2	0.4	2.5	16.1	12.4	0.1	36.3	34.8	31.1	26.8	0.2
23948	**G**	**C**	**S**	**Spike**	0.0	0.3	0.0	0.3	0.1	2.2	0.1	99.5	2.8	9.5	0.2	0.2	7.6	79.3	82.4	76.0	70.1	82.8	65.4	97.7
23996	C	T	S	Spike	0.2	0.0	0.0	0.4	0.0	0.3	0.0	0.2	0.0	0.0	0.0	0.2	0.1	0.2	0.0	0.0	0.0	0.0	0.0	95.4
24257	G	T	S	Spike	0.0	0.0	2.1	0.1	27.0	2.4	0.0	0.0	0.2	0.1	0.2	0.0	0.3	0.0	0.0	0.2	10.2	0.0	0.3	0.1
25033	G	A	S	Spike	0.0	0.1	0.2	1.1	0.0	0.3	0.0	0.0	0.0	0.2	99.3	96.3	77.4	0.3	1.1	4.9	0.1	4.0	6.9	0.2
25334	G	A	S	Spike	0.0	0.0	0.2	0.1	0.0	0.2	0.0	0.3	0.0	0.1	0.0	0.0	0.0	0.0	20.1	0.0	0.0	0.0	0.0	0.0
25537	G	A	ORF3a	ORF3a Protein	0.0	2.5	15.9	20.0	0.3	2.9	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
26299	C	A	E	Env	0.0	0.0	0.0	0.0	0.0	0.2	0.1	0.0	0.1	0.3	99.3	97.3	77.9	0.0	37.9	1.3	0.0	2.8	7.0	0.0
26333	C	T	E	Env	0.2	0.0	0.2	0.0	0.0	0.2	0.0	0.0	0.0	0.3	99.5	97.4	78.0	0.0	37.7	1.3	0.0	3.4	7.1	0.0
26373	C	T	E	Env	0.0	0.0	0.0	0.2	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.8	0.0	0.0	0.0	0.0	0.0
26529	G	A	M	Membrane	0.1	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.2	0.0	0.0	0.0	14.7	0.0	0.0	0.0	0.0	0.0
26634	G	T	M	Membrane	0.2	0.0	0.2	0.1	0.0	0.0	0.1	0.0	0.2	0.0	43.2	32.4	10.7	0.0	1.8	0.3	0.0	0.0	0.7	0.0

Locus	From	To	Gene	Protein	Days since 1st Positive RT-PCR for SARS-CoV-2																			
					1	37	41	45	50	55	66	82	86	89	93	93 (ETA)	95	98	99	99 (ETA)	100 (ETA)	100	101 (ETA)	101
26647	G	A	M	Membrane	0.2	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	22.6	26.4	6.6	0.2	0.0	0.3	3.1	0.0	0.0	0.1
26895	C	T	M	Membrane	0.3	0.0	0.0	0.0	0.3	0.0	0.0	0.3	0.0	0.0	99.6	97.6	76.3	0.3	43.6	5.0	0.0	4.5	7.7	0.0
27087	G	A	M	Membrane	0.0	0.6	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	6.2	10.5	22.0	0.2	10.5	0.2	0.0	0.2	1.1	0.0
27408	T	C	ORF7a	ORF7a Protein	0.0	0.1	0.0	0.1	0.0	0.0	0.1	0.0	28.1	65.0	0.3	2.7	14.5	10.5	6.2	29.6	17.5	32.6	26.9	0.0
27459	G	T	ORF7a	ORF7a Protein	0.3	0.1	0.2	0.2	0.3	0.0	0.1	0.4	0.0	3.2	0.2	0.0	7.5	81.8	76.8	65.0	51.8	61.8	65.2	81.6
27509	**C**	**T**	**ORF7a**	**ORF7a Protein**	0.0	65.5	44.9	79.2	67.5	26.6	0.0	0.1	0.0	0.0	0.0	0.7	0.1	0.2	0.1	0.1	1.3	0.2	1.7	0.0
27618	T	C	ORF7a	ORF7a Protein	0.3	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	1.6	99.3	98.0	85.4	0.9	29.6	55.6	0.0	2.4	10.6	0.0
27634	T	C	ORF7a	ORF7a Protein	0.0	0.0	3.1	0.0	26.3	1.2	0.0	0.3	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	3.9	0.0	0.0	0.0
28068	G	T	ORF7b	ORF7B Protein	0.0	0.7	10.6	0.1	0.3	1.1	0.0	0.2	0.0	0.0	0.0	0.5	0.0	0.2	0.1	0.1	0.0	0.0	0.0	0.4
28121	C	T	ORF8	ORF8 Protein	0.0	0.3	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.2	0.2	0.0	0.0	0.0	0.0	0.0	15.9

Locus	From	To	Gene	Protein	Days since 1st Positive RT-PCR for SARS-CoV-2
					1	37	41	45	50	55	66	82	86	89	93	93 (ETA)	95	98	99	99 (ETA)	100 (ETA)	100	101 (ETA)	101
28209	G	C	ORF9	ORF8 Protein	0.0	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.7	0.2	1.5	11.9	3.3	0.0	24.9	1.3	28.2	27.7	0.2
28313	C	T	N	Nucleocapsid	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.2	0.2	0.0	0.2	0.0	0.1	0.0	0.0	10.6
28356	A	G	N	Nucleocapsid	0.0	0.3	2.7	0.2	48.3	0.2	0.2	99.9	2.3	7.1	0.3	0.4	6.5	79.9	88.2	59.7	46.7	62.4	61.7	89.0
28748	C	T	N	Nucleocapsid	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.2	1.7	0.0	0.0	1.9	14.0	3.8	0.0	26.7	0.1	27.9	27.5	0.0
29198	C	T	N	Nucleocapsid	0.0	0.0	0.3	0.2	0.3	0.0	0.0	0.1	0.0	0.0	0.1	0.0	0.2	0.1	21.9	0.2	0.1	0.0	0.2	0.1
29259	C	T	N	Nucleocapsid	0.0	0.2	0.0	0.0	0.0	0.2	0.1	0.0	0.0	0.0	0.1	0.0	0.2	0.1	0.1	0.1	0.1	0.3	0.0	18.8
29825	G	C	N	Nucleocapsid	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.8	0.7	0.0	8.7	0.0	8.1	12.0	0.0
29825	G	T	N	Nucleocapsid	0.0	0.2	0.0	0.0	0.4	0.0	0.0	99.3	3.3	9.2	0.2	1.1	6.1	80.4	99.6	55.1	40.5	54.2	54.8	99.6

Table 4.3. Prevalence of all nucleotide variants at sequential time points measured by short-read (Illumina MiSeq).

Mutations in Figure 4.10 are shown here in bold. N = Nucleocapsid, E = envelope, S = Spike; Locus 2343 = NSP2: I513T, 13936 = NSP12b: V157L, 20150 = NSP15:N177S, 21600 = S:S13I, 21752 = S:W64G, 21765 = S:Δ69, 22160 = S:Y200H, 22281 = S:T240I, 22550 = S:P330S, 23948 = S:D796H, 27509= ORF7a: T39I. Analyses done in collaboration with Dr Steven Kemp and Professor David Pollock.

The presence of the mixed population of viruses in this participant's samples was independently confirmed using SGA on samples from three time points - days 1 and 37, which preceded administration of CP, and day 98, which was after the third dose of CP. The absence at days 1 and 37 and the presence at day 98 of a mixed population of viruses bearing the different Spike mutations, provides further evidence of dynamic population shifts in response to convalescent plasma (Table 4.4).

Timepoint	N	Prevalence of mutations (%)					
		W64G	P330S	ΔH69/V70	D796H	T200I	Y240H
Day 1	7	0	0	0	0	0	0
Day 37	38	0	0	0	0	0	0
Day 98	21	4.8 (1)	4.8 (1)	81.0 (17)	68.4 (13*)	14.3 (3)	14.3 (3)

Table 4.4: Single genome amplification data showing the prevalence of each mutation in respiratory samples at indicated time points.

N- indicates the number of single genomes obtained at each time point. * The denominator is 19 as the sequence reads for 2 genomes covering amino acid position 796 were poor quality. WT- wild type Spike pseudotyped virus, mutants; D796H-aspartic acid to histidine at Spike protein amino acid position 796, △H69/V70-histidine and valine deletion at Spike protein amino acid position 69 and 70 respectively, W64G- tryptophan to glycine at Spike protein amino acid position 64, P330S- proline to serine at Spike protein amino acid position 330, T200I- threonine to isoleucine at Spike protein amino acid position 200, Y240H- tyrosine to histidine at Spike protein amino acid position 240.

4.4 Spike Mutants D796H causes a decrease in infectivity which is compensated for by acquisition of Δ69/70 deletion

The phenotypic properties of Spike mutant virus bearing four of the observed mutants W64G + P330S and **Δ**H69/V70 + D796H were investigated using lentivirus pseudotyping infectivity and neutralisation assays. These mutations were engineered into codon optimised Spike expression pCDNA_SARS-CoV-2_S plasmid bearing D614G by site direct mutatgenesis (SDM), either as single or double mutation as indicated (Methods 2.2.8). Pseudotyped viruses were generated by transfection of HEK293T cells with p8.91 lentivirus packaging plasmid, pCSFLW luciferase expression plasmid and pCDNA_SARS-CoV-2_S Spike expression plasmid. Pseudotyped viruses were harvested 48 hours later (Methods 2.2.9). The wild type (WT) was the Spike pseudotyped virus bearing D614G and the pseudotyped virus bearing W64G + P330S and **Δ**H69/V70 + D796H either individually or in pairs were designated single or double mutants prefixed by the amino acid mutation and position in Spike protein.

Firstly, the infectivity of these pseudotyped viruses were determined in a single round infectivity assay (Methods 2.2.9) by titration on target cells- HEK293T cells which had transiently been transfected with ACE2 and TMPRSS2 or HeLA expressing ACE2 receptors. The relative light units (RLU) per µl of infectious dose was determined using the Steady-Glo Luciferase assay system (Promega). (Figures 4.13A, B). In another experiment, infectivity was measured by normalising virus input for reverse RT activity in the virus supernatant as measured by qPCR (Methods 2.2.10) (Figures 4.13 C, D). The normalised input was used to infect either HeLa

cells expressing ACE2 receptors or HEK293T cells which had transiently been transfected with ACE2 and TMPRSS2. Both approaches were consistent in that over a single round of infection compared to WT, the acquisition of the single D796H mutation led to a 3.5-fold decrease in infectivity, the **Δ**H69/V70 deletions appeared to have two-fold higher infectivity compared to WT and the acquisition of the double mutant **Δ**H69/V70 + D796H restored infectivity to levels similar to the WT (Figures 4.13A, B, C, D). Acquisition of W64G+P3300S also led to a two-fold decrease in infectivity compared with WT (Figure 4.13A).

A
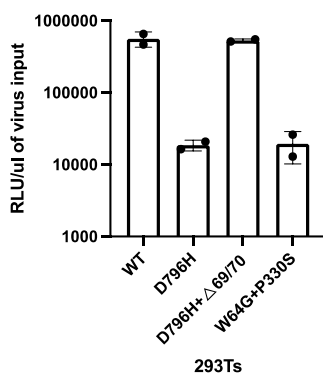
B
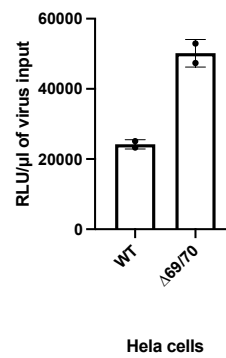
C
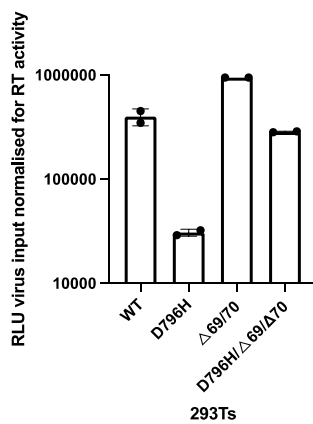
D
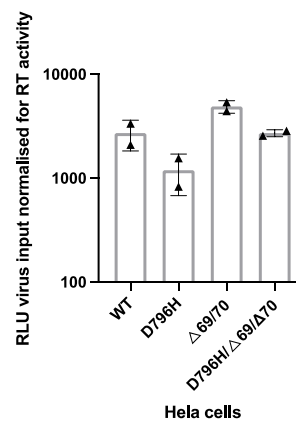
Figure 4.13. SARS-CoV-2 Spike pseudotyped virus infectivity. A, B. Infectivity determined by titration of the pseudotyped viruses on the indicated target cell line. Data points are RLU/µl of input virus, are mean ± s.d. of two technical replicates and represent two independent experiments. **C, D.** Infectivity determined by quantification of reverse transcriptase activity by qPCR. Data points are RLU normalised for RT-activity are mean ± s.d. of two technical replicates and represent two independent experiments. WT- wild type Spike pseudotyped virus, mutants; D796H- aspartic acid to histidine at Spike protein amino acid position 796, △H69/V70- histidine and valine deletion at Spike protein amino acid position 69 and 70 respectively, W64G- tryptophan to glycine at Spike protein amino acid position 64, P330S- proline to serine at Spike protein amino acid position 330.

4.5 Spike Mutants D796H and Δ69/70 individually or in combination does not impact Spike incorporation

Spike protein from each pseudotyped virus was detected by Western blot in pelleted virions obtained by centrifugation of the supernatant from transfected cells (Methods 2.2.11). Equal amounts of full-length Spike protein was detected in pelleted virions in the WT and each mutant (Figure 4.14). HIV-1 p24 antibody probing was used to monitor levels of lentiviral particle production.
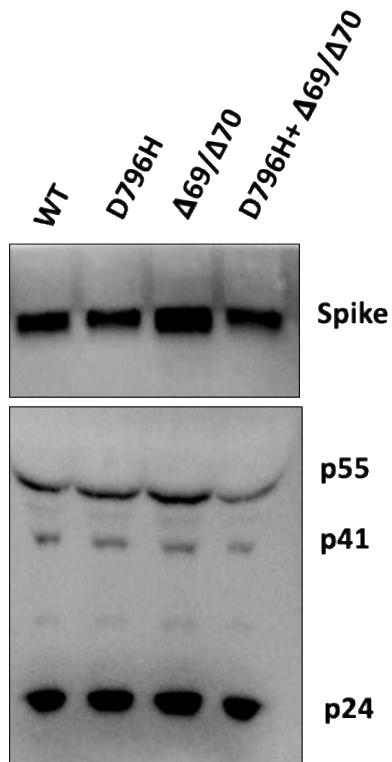
Figure 4.14. Spike expression of WT and mutant pseudotyped viruses. Western blot of pseudotyped virus pellets from supernatants from cells transfected with the triple plasmids to produce the WT and mutant pseudotyped viruses that includes the Spike protein. Blots are representative of two independent transfections. WT- wild type Spike pseudotyped virus, mutants; D796H- aspartic acid to histidine at Spike protein amino acid position 796, $\triangle$H69/V70- histidine and valine deletion at Spike protein amino acid position 69 and 70 respectively. Done in collaboration with Dr Rawlings Datir.

4.6 The spike D796H evolved to escape neutralisation from polyclonal sera

Serum neutralisation was measured in a single-round *in vitro* neutralisation assay (Methods 2.2.12). Using a lentiviral pseudotyping method, WT, $\triangle$H69/V70 +

D796H, W64G+ P330 and single mutant Spike proteins in enveloped virions were generated in order to measure neutralisation activity of CP and participant's sera at time points preceding and following administration CP against these viruses. Serum neutralisation curves were generated from which inhibitory dilution 50 (ID_{50}) were determined. ID_{50} is the dilution of serum required to inhibit 50% of infection of a SARS-CoV-2 Spike pseudotyped virus in the assay. Human AB serum stored prior to 2019 was the negative control. Positive controls were sera from COVID-19 PCR positive cases.

Examples of neutralisation curves from participant's sera against WT are compared with neutralisation by a pre-pandemic negative control, a SARs-CoV-2 PCR positive control and 3 samples of neutralising convalescent sera from independent recovered patients (Figure 4.15). The participant sera from days 38 and 52 had no neutralisation activity against the WT virus. However, following administration of the first 2 units of CP, participants' sera from days 66 - 80 had neutralisation activity against WT (Figure 4.15).
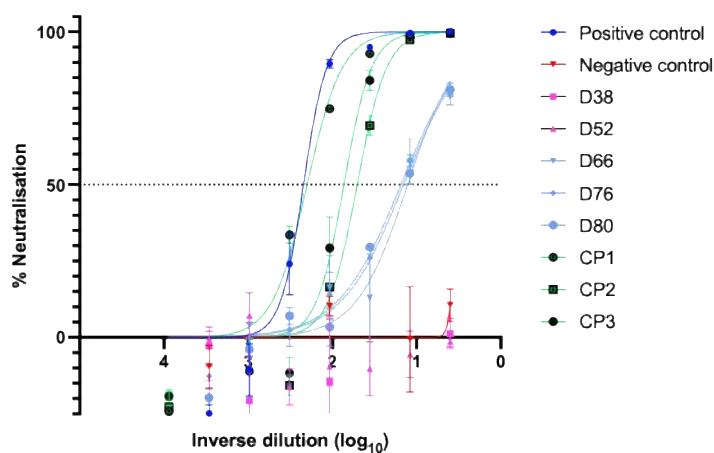
Figure 4.15. Neutralisation curves against the WT pseudotyped virus. Days 38 and 52 of the participant's sera with no neutralisation activity against WT (pink). Days 66 to 80 of the participant's sera with neutralisation activity against WT (light blue). CP- Convalescent plasma, units 1 to 3 with activity against the WT (green). Negative control- pre-pandemic human AB serum (red). Positive control- serum from a SARS-CoV-2 PCR positive case (dark blue). D- day sample was taken, day 1 being first PCR positive SARS-CoV-2 swab. WT- wild type Spike pseudotyped virus.

D796H alone and the D796H + △H69/V70 double mutant were less sensitive to neutralisation by all CP samples. The neutralisation curves show a shift to the right for both these mutants (Figure 4.16A). The was at least a 2-fold decrease in ID50 against the D796H + △H69/V70 mutants by CP samples 2 and 3 and a 1.3-fold decrease by CP 1 (Table 4.5). These mutations emerged after administration of CP units 1 and 2, diminished and then rebounded after CP unit 3, which may explain why the fold decrease by CP 1 was lower than that of CP 2 and 3. △H69/V70 deletion on its own had no impact in neutralisation activity by any of the CPs (Figure 4.16A). Patient derived serum from days 64 and 66, one day after CP1 and CP2 respectively, similarly showed lower potency against the D796H + △H69/V70 mutants, with 1.7 to 2.6-fold decrease in ID50s (Figure 4.16B, Table 4.5).

W64G + P330S as a double mutant was not detected in the participants' samples until day 93 when it reached a prevalence 99%. It was already diminishing on day 95 when the 3rd unit of CP was given, probably in response to remdesivir given on day

93. It completely disappeared and rebounded again to 25% (W64G) and 6% (P330S) by day 101. This double mutant was less sensitive to neutralisation by CP unit 3, with a 3.4-fold decrease in ID_{50} (Fig 4.17, Table 4.6). CP unit 2 had a 1.5-fold decrease in ID_{50} against the mutant whilst CP1 has equal neutralisation activity against the mutant and the WT. (Fig 4.17, Table 4.6).
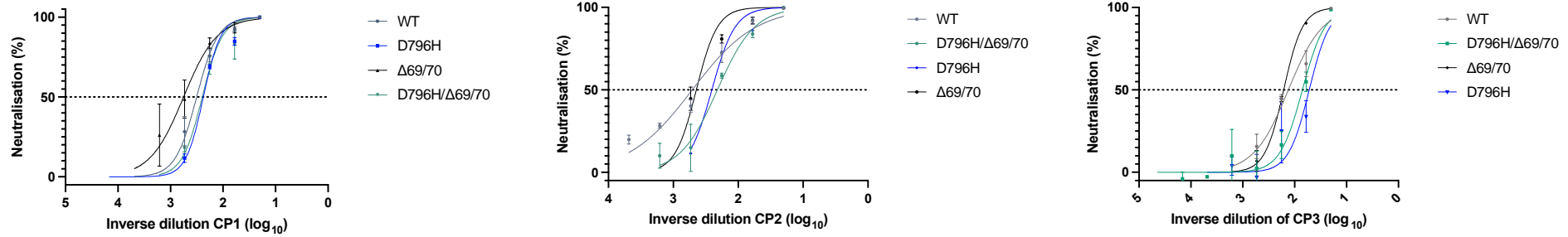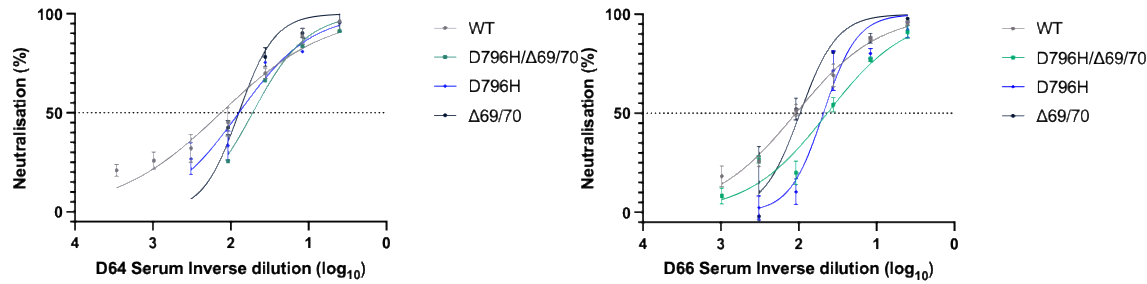
Fig 4.16. Neutralisation sensitivity of Spike WT, D796H, △H69/V70 and D796H +△H69/V70 pseudotyped lentiviruses. A. Representative neutralisation curves for Spike variants against convalescent plasma units 1-3. **B.** Representative neutralisation curves for Spike variants against participant's sera from days 64 and 66. CP- convalescent plasma, D- day sample was taken, day 1 being first PCR positive SARS-CoV-2 swab.

WT- wild type Spike pseudotyped virus, mutants; D796H- aspartic acid to histidine at Spike protein amino acid position 796, $\triangle$H69/V70- histidine and valine deletion at Spike protein amino acid position 69 and 70 respectively. Data points represent mean neutralisation of technical replicates and error bars represent standard error of the mean. Data are representative of two independent experiments.

Sera	ID50				Fold decrease		
	WT	D796H	△H69/V70	D796H+△ H69/V70	D796H	△H69/V70	D796H+△ H69/V70
CPI	317.5	236.8	562.9	248.3	1.3	0.6	1.3
CP2	511.5	255.1	442.3	209.6	2.0	1.2	2.4
CP3	132.3	52.4	162.8	71.0	2.5	0.8	1.9
D64	133.5	78.48	78.70	52.25	1.7	1.7	2.6
D66	109.2	48.23	96.13	43.04	2.3	1.1	2.5

Table 4.5. Inhibitory dilution 50 and fold changes for D796H, △H69/V70 and D796H +△H69/V70 spike mutations.

ID50- inhibitory dilution 50, CP- convalescent plasma, D- day sample was taken, day 1 being first PCR positive SARS-CoV-2 swab. WT- wild type Spike pseudotyped virus, mutants; D796H aspartic acid to histidine at Spike protein amino acid position 796, △H69/V70- histidine and valine deletion at Spike protein amino acid position 69 and 70 respectively. No fold decrease (green), moderate fold decrease (orange), significant fold decrease (red).
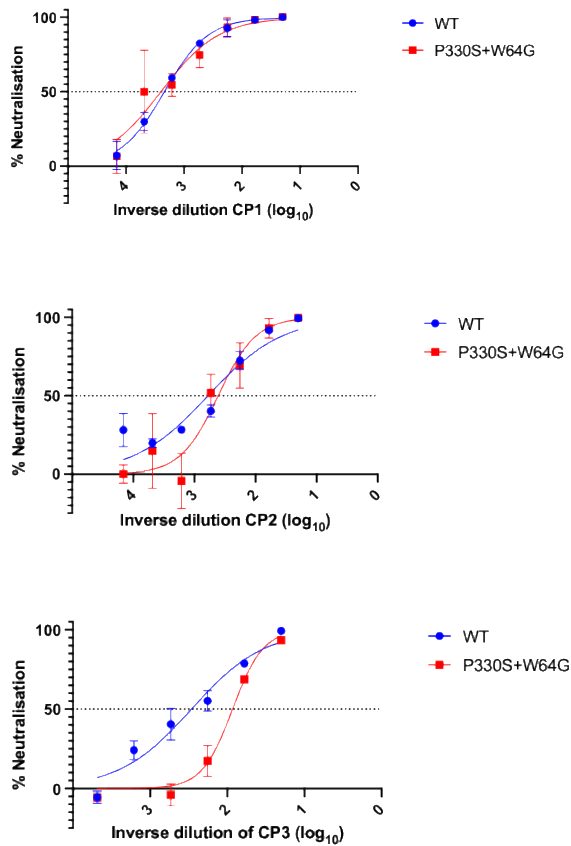
Fig 4.17. Neutralisation sensitivity of Spike WT and W64G+ P330 pseudotyped lentiviruses. Representative neutralisation curves for Spike variants against convalescent plasma units 1-3. WT- wild type Spike pseudotyped virus, mutants; W64G- tryptophan to glycine at Spike protein amino acid position 64, P330S- proline to serine at Spike protein amino acid position 330. Data points represent mean neutralisation of technical replicates and error bars represent standard error of the mean. Data are representative of two independent experiments.

Sera	WT ID50	W64G+ P330 ID50	Fold decrease
CPI	2181	2501	0.9
CP2	597.8	408.4	1.5
CP3	291.3	85.4	3.4

Table 4.6. Inhibitory dilution 50 and fold changes for W64G+ P330 spike mutations.

ID50- inhibitory dilution 50, CP- convalescent plasma, WT- wild type Spike pseudotyped virus, mutants; W64G- tryptophan to glycine at Spike protein amino acid position 64, P330S- proline to serine at Spike protein amino acid position 330. No fold decrease (green), moderate fold decrease (orange), significant fold decrease (red).

4.7 The location of D796H and △H69/V70 Spike mutations may affect antibody binding and neutralisation by allostery

The location of the Spike mutants D796H and △H69/V70 were mapped using the PyMOL Molecular Graphics System v.2.4.0 (https://github.com/schrodinger/pymol-open-source/releases), onto a previously published SARS-CoV-2 spike structure (PDB: 6ZGE). Amino acids 69 and 70 are in the NTD, in an exposed loop with a somewhat disordered structure, however in the tertiary protein structure, it is close to the receptor binding domain of the Spike protein. This may affect antibody binding at this site by a conformational change in the protein structure. D796 is located near

the base of Spike in the S2 subunit and is also in an external loop. This site may not come into contact with neutralising antibodies but may mediate escape by introduction of a charged amino acid, again, leading to a conformational change of the protein structure (Figure 4.18).
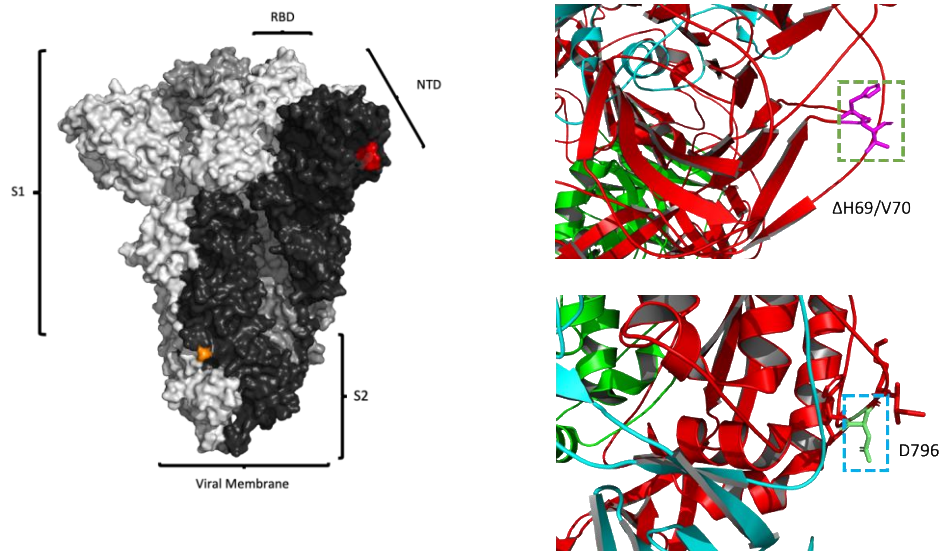
Figure 4.18. Spike mutations ΔH69/70 and D796H. Surface protein representation of the Spike homotrimer are coloured separately in shades of grey. In red are the amino acid residues H69 and V70, deleted in the N-terminal domain. In orange is D796H in the Spike S2 subunit. Close-ups are cartoon stick representation of ΔH69/Y70 (above) and D796H (below) (PDB: 6ZGE Wrobek et al., 2020). Source: Kemp and Collier et al, SARS-CoV-2 evolution during treatment of chronic infection, Nature 592, 277–282 (2021). https://doi.org/10.1038/s41586-021-03291-y.

4.8 Discussion

This in-depth analyses of *in vivo* dynamic evolution and escape, demonstrates within-host adaptation of SARS-CoV-2 to evade neutralising antibodies.

It is possible that some of the viruses sampled represent a compartmentalised population that was previously unsampled, particularly the virus population containing W64G + P330S, which emerged at high prevalence on a particularly long branch at days 93 and 95. The SGA analyses suggests that although this was a divergent cluster, it was present within the mixed population of viruses at a low level in a sample from a later timepoint. In addition, the phylogenetic tree showed that this population was rooted within the main population.

Ideally, the neutralisation assay would be replicated in a live virus system. However, Spike pseudotyped viral infection assays have been shown to have similar characteristics as assays using infectious wild type SARS-CoV-2 (Schmidt et al. 2020).

Although recent head to head analyses of long read Oxford Nanopore Technology (ONT) and short read Illumina found that they are equally good at consensus-level sequence determination (Bull et al. 2020), short read Illumina MiSeq data was used for variant calling in the mutational prevalence analyses in this study. This is because Illumina is known to have a lower sequencing error rate (Bull et al. 2020; Kim et al. 2020). However, errors introduced in sample preparation and PCR may result in missing some variants and patterns of intrahost diversity. Therefore, future

work will include using high-throughput single genome amplification from RNA, which lends itself well to exploring variations in a limited region of the genome (Ko et al. 2021), as well as reconciling the dynamic shifts in the virus population with antibody responses and mapping the immunodominant epitopes targeted by autologous antibodies to the emerging viruses detected.

Of note, no regions outside of the Spike protein were considered in this work. This is because the Spike protein is really well characterised and it is the target of neutralising antibodies following natural infection. Although other regions outside of the Spike protein might contribute to immune escape, the Spike protein is most likely directly involved in this process.

This is a study of a single case and therefore limits generalisability. The circumstances of this participant with absent B cell, suboptimal T cell response maybe applicable to a broad group of immunocompromised hosts for example the elderly. In the next chapter I will explore some factors associated with immune response to Spike directed mRNA vaccines including age and cytomegalovirus seropositivity.

CHAPTER 5: RESULTS 3- AGE-RELATED HETEROGENEITY OF SARS-COV-2 MRNA VACCINE-ELICITED RESPONSES

Key Findings:

- SARS-CoV-2 specific binding antibody levels negatively correlate with age.

- The odds of an inadequate neutralising antibody response to a spike pseudotyped virus after only one dose of the mRNA BNT162b2 vaccine was 3.7 in those ≥80 years old.

- Poor serum neutralisation of WT Spike pseudotyped virus after the first dose of mRNA BNT162b2 is overcome by the booster/second dose.

- The frequency of IFNγ secreting Spike specific T cells following only one dose of mRNA BNT162b2 vaccine is significantly lower in the ≥80 years group compared with the <80 years group. This is overcome by the booster/second dose.

- The frequency of IL-2 secreting Spike specific T cells following only one dose of mRNA BNT162b2 vaccine is significantly lower in the ≥80 years group compared with the <80 years group. This however was not increased by the booster/second dose.

- The frequency of IFNγ secreting Spike specific T cells following only one dose *of* mRNA BNT162b2 vaccine is higher in the ≥80 years group with HCMV seropositivity compared to ≥80 year who are HCMV negative.

From the previous body of work on virus escape from neutralising antibodies, it was clear that the ability of the SARS-CoV-2 Spike expressing pseudotyped viruses to escape inhibition by neutralising antibodies had implications for response to vaccine-elicited antibodies. Vaccines are a key tool to controlling the COVID-19 pandemic. In particular, mRNA vaccine had shown good efficacy in clinical trials after two doses, separated by a three or four week gap (Baden et al. 2020; Polack et al. 2020). However, the duration of protection is not known and very few participants above the age of 80 years were included in the clinical trials despite being at greatest risk of severe disease and death (Docherty et al. 2020). The Joint Committee on Vaccination and Immunisation (JCVI) recommended extending the dosing interval to 12 weeks in order to vaccinate a greater proportion of the population during a period of substantial transmission (Care 2021). *This* strategy is being considered in other countries in order to maximise first dose administration (DW. 2021; Kadire, Wachter, and Lurie 2021).

Ageing can impact responses in many ways including but not limited to reducing the magnitude of antibody response to vaccines (Wilkinson et al. 2017), shortening the persistence of antibody response (Powers and Belshe 1993), limiting the quality of antibodies via reduced affinity and breadth, limiting induction of CD4+/8+ T cell response (Westmeier et al. 2020) and limiting induction of effector memory (Gustafson et al. 2020). Data on vaccine responses, particularly in groups under-represented in clinical trial are necessary to understand the efficacy of vaccination using this regime.

Under the ethical framework of an ongoing large NIHR CRF trial called the Bioresource, a cohort study was conducted to assess real world immune responses following vaccination with mRNA-based vaccine BNT162b21 in elderly participants from the community and younger health care workers. The study aimed to investigate age as a risk factor for escape from mRNA BNT162b2 vaccine-elicited antibody neutralisation. It also aimed to characterise binding antibody responses, T cell immune responses to the vaccine by measuring SARS-CoV-2 Spike specific IFNγ and IL-2 responses in PBMCs, the B Cell repertoire, autoantibodies and inflammatory markers.

5.1 Cohort Description and study procedure

One hundred and one participants received the first dose of the BNT162b2 vaccine whilst twenty-one received their second dose, including one who was recruited after their second dose of vaccine (Table 5.1). The second dose was initially given under the manufacturer's recommended schedule, three weeks after the first but was changed during the study to 12 weeks following the JCVI's recommendation. Therefore, the remaining eighty participants are yet to receive their second dose.

Participants recruited reflected the Joint Committee on Vaccination and Immunisation (JCVI) priority groups for immunisation and included older adults and health care workers receiving the first dose of the BNT162b2 vaccine between the 14th of December 2020 to the 29th of January 2021. Each participant was seen at least 3 weeks after their first vaccine dose and up to 12 weeks after. Participants were then

seen again at 3 weeks following the second dose of the BNT162b2 vaccine and

underwent the same blood sampling procedure (Figure 5.1).
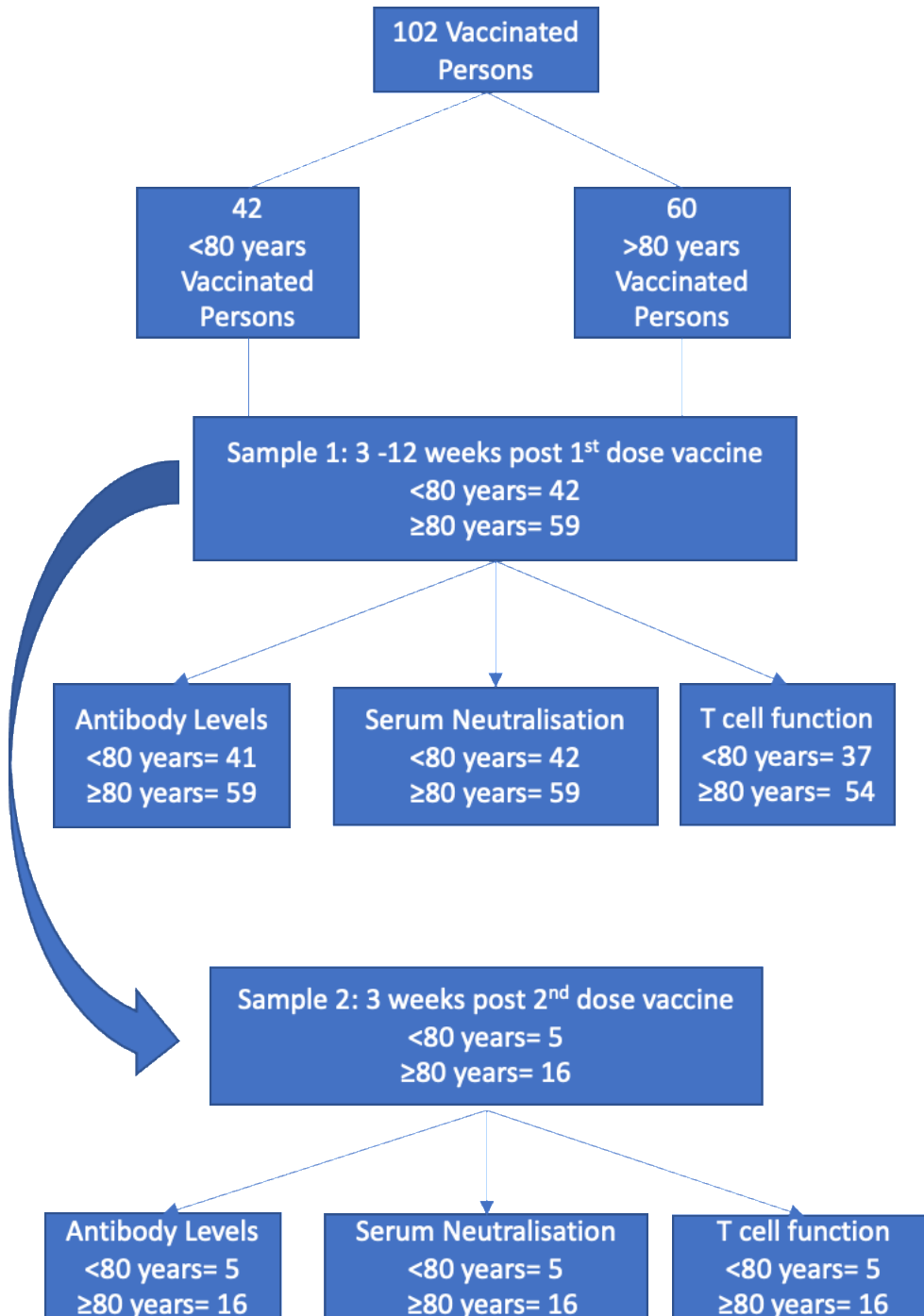
Figure 5.1: Study flow describing the procedures for participants recruited. The numbers of participants with results from each type of assay is indicated. The study was approved by the East of England – Cambridge Central Research Ethics Committee (17/EE/0025).

The median age of the cohort was 81 years (IQR 70-84) and 39% of participants were female. Participants were stratified by age into 2 exposure groups- < 80 years and ≥ 80 years. 42 (41.6%) of participants were in the <80 years group, whilst 59 (58.4%) were in the ≥ 80 years (Table 5.1). These exposure levels were chosen because the JCVI priority list included participants over the 80 years old as the highest priority to receive the available SARS-CoV-2 vaccines. In addition, there was considerable concern about the efficacy of only one dose of the vaccine to provide protection for the elderly who were at greatest risk from COVID1-9 disease during a period of high background transmission in the community. It was hypothesised that due to aging, the ≥80 years group would mount an inadequate immune response with only one dose of the vaccine and would require a booster/second for an adequate protective immune response.

Previous SARS-CoV-2 infection was determined by measuring anti-Nucleocapsid protein and anti-Spike protein IgG antibody levels in all participants. When both were elevated, the participants were deemed to have likely had exposure to pandemic SARS-CoV-2. This is because the mRNA vaccine contains the sequence for the Spike protein only. Although cross-reactivity with one of the other seasonal coronaviruses could not be excluded, it was assumed that N protein antibody

positivity could reasonably be ascribed to recent infection with SAR-CoV-2 given high background community transmission rates. 10% of <80 years group and 9% of the ≥ 80 years group were positive for both anti-N IgG & anti-S IgG and thought to have prior exposure to SARS-CoV-2. As expected, Human cytomegalovirus (HCMV) seropositivity was higher in the ≥ 80 years than in the <80 years group (54.6% vs 37.0%) (Table 5.1).

	<80 years (N=42 or n/N)	≥80 YEARS (N=60 or n/N)	p value
Female %	40.5 (17)	38.3 (23)	0.83[a]
Median age (IQR) *years*	62.5 (47.0-71.0)	83.0 (81.0-85.5)	-
Sera GMT WT (95% CI)			
dose 1	83.4 (52.0-133.7)	46.6 (33.5-64.8)[c]	0.01[b]
dose 2	651.0 (155.6-2722.9)[d]	555.1 (351.5-876.5)[e]	ns[b]
Serum ID50<20 for WT %			
dose 1	21.4 (9)	50.9 (30/59)[c]	0.003[a]
dose 2	0 (0/5)	0 (0/16)	-
Prior SARS-CoV-2 combined positive anti-N IgG & anti-S IgG %	9.8 (4/41)	8.6 (5/58)	0.85 [a]
HCMV positive %	37.0 (10/27)	54.6 (24/44)	0.15

Table 5.1: Characteristics of study participants.

[a] Chi-square test, [b] Mann-Whitney test, [c]neutralisation data unavailable for one individual, [d]neutralisation data available for 5 of 42, [e]neutralisation data available for

16 of 60, GMT- geometric mean titre, WT- wild type, ID50- (Inhibitory dilution) –
the serum dilution achieving 50% neutralisation, ns- non-significant, CI-confidence
interval, N- Nucleocapsid protein, S- Spike protein, IgG- immunoglobulin.

**5.2 Neutralisation activity of vaccine-elicited sera against SARS-CoV-2
following first and second dose mRNA BNT162b2 vaccine**

The primary outcome was inadequate vaccine-elicited serum antibody neutralisation
activity at least 3 weeks after the first dose of vaccine, defined as an inhibitory
dilution 50 (ID50) of ≤1 in 20. This dilution was chosen because although the limit
of detection of the assay in 1 in 4, there is a degree of variability at lower dilutions
which is overcome by using a more conservative cut-off. This was measured in an *in
vitro* neutralisation assay and is the dilution of serum required to inhibit 50% of
infection of a Spike pseudotyped virus (Methods 2.3.4). The pseudotyped virus was
generated by transfection of HEK293T cells with p8.91 lentivirus packaging
plasmid, pCSFLW luciferase expression plasmid and pCDNA_SARS-CoV-2_S
Spike codon optimized expression plasmid D614G. Pseudotyped viruses were
harvested 48 hours later (Methods 2.3.3). The wild type (WT) was the Spike
pseudotyped virus bearing D614G mutation in Spike protein.

Serum neutralisation curves were generated from infection of 293T cells transiently
transfected with ACE2 and TMPRSS2 plasmids and infected with WT Spike
pseudotyped virus in the presence of a limiting dilution of vaccinee sera. This was
done in duplicate. ID50s were determined from these curves using Prism 9. (Figure

5.2) Examples of serum neutralisation curves from six individuals with reduced responses after first dose of the mRNA BNT162b2 vaccine but adequate responses 3 weeks after the second dose.
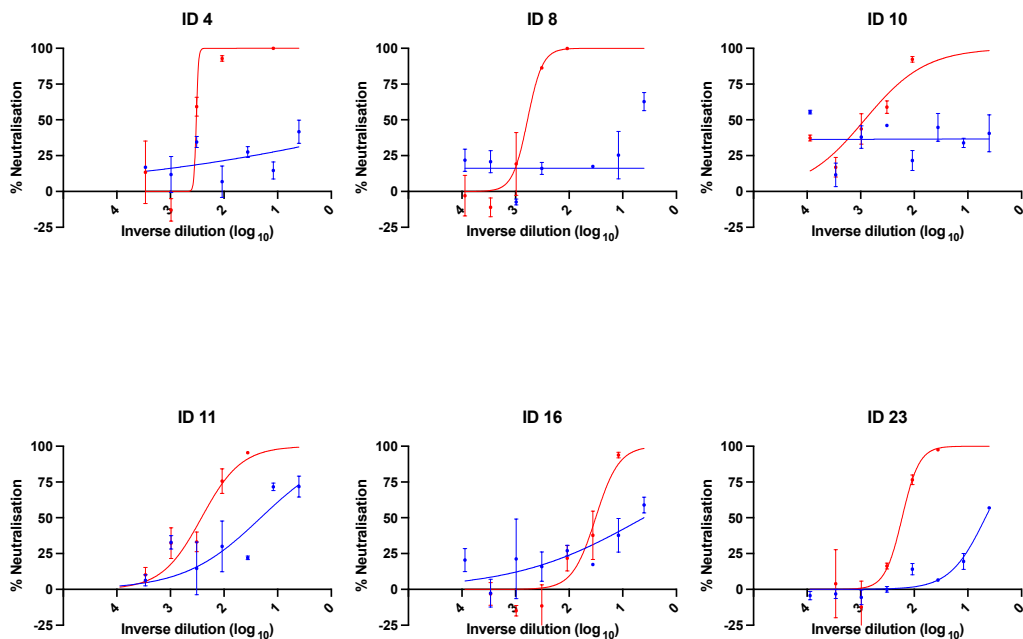
Figure 5.2 Serum neutralisation curves. From vaccine-elicited serum from six individuals after first dose (blue) and after dose 2 (red) of mRNA BNT162b2 vaccine against pseudotyped virus expressing wild type Spike.

5.2.1 Poor serum neutralisation of WT Spike pseudotyped virus after the first dose of mRNA BNT162b2 is overcome by the booster/second dose

The geometric mean neutralisation titre (GMT) at which ID50 was achieved after the first dose was lower in the ≥ 80 years than the < 80 years group [46.6 (95% CI 33.5-64.8) vs 83.4 (95% CI 52.0-133.7) $p<0.01$]. (Table 5.1, Figure 5.3A). 21% of the

<80 years group had inadequate neutralisation against the WT Spike pseudotyped virus with an ID50 of <1:20 compared with 51% of the ≥ 80 years group. After the second dose, sera from vaccinated individuals exhibited an increase in neutralising titres. The GMT increased for the <80 years group to 651.0 (95% CI 155.6-2722.9) and the ≥ 80 years group to 555.1 (95% CI 351.5-876.5) with no statistically significant difference between these two groups (Table 5.1, Figure 5.3B).

A

B
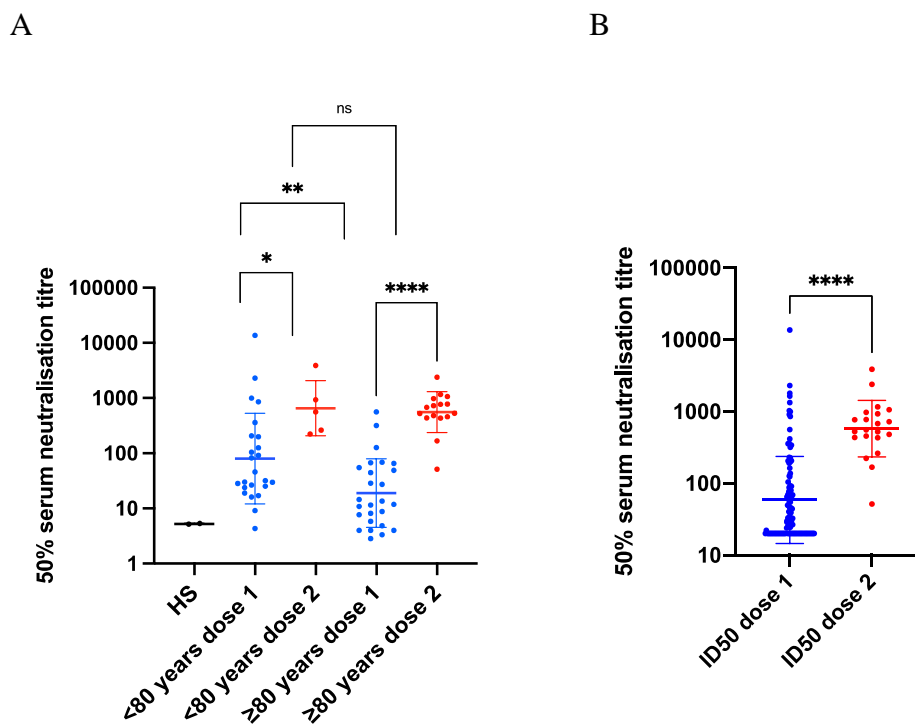
Figure 5.3. Neutralisation activity by mRNA BNT162b2 vaccine sera against SARS-CoV-2 in a Spike lentiviral pseudotyping assay expressing wild type Spike. A. Inhibitory Dilution 50 (ID50) values for individuals after dose 1 (n=101, blue) and after dose 2 (n=21, red) a comparing the two age groups; < 80 and ≥ 80 years. Geometric mean with s.d is shown. Each point is a mean of technical replicates from two experiment repeats. **B.** ID50 for all individuals after dose 1 (n=101, blue) and after dose 2 (n=21, red)**.** Mann-Whitney test was used for unpaired

comparisons and Wilcoxon matched-pairs signed rank test for paired comparisons. p-values * <0.05, ** <0.01, **** <0.0001, ns not significant, HS – human AB serum control.

Given that ageing is a continuum and stratification into < 80 and ≥ 80 years groups is arbitrary but justified given public health policy at the time, age as a continuous variable was correlated with serum neutralisation using a linear regression model. Age showed a statistically significant negative correlation with $\log_{10}$ ID50 against WT pseudotyped virus after the first but not the second dose of the BNT162b2 vaccine (Figure 5.4A, B).

A   B
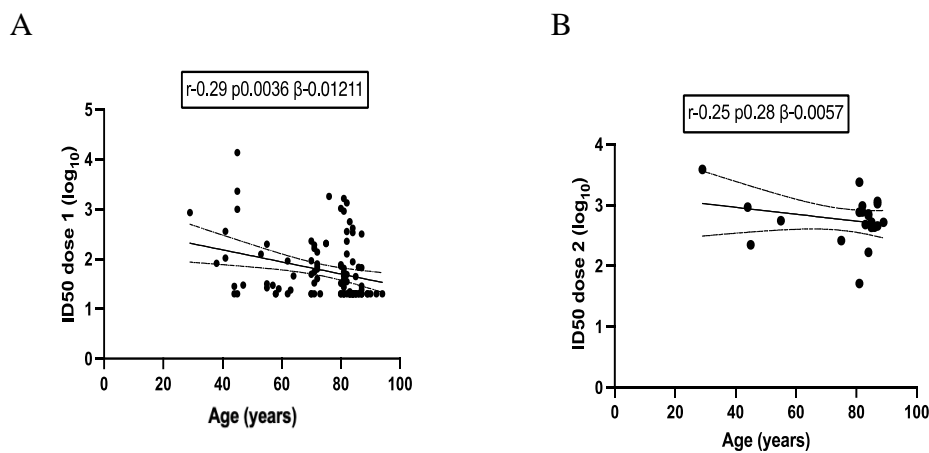
Figure 5.4. Neutralisation activity by mRNA BNT162b2 vaccine sera against SARS-CoV-2 in a Spike lentiviral pseudotyping assay expressing wild type Spike. Serum neutralisation of Spike (D614G) pseudotyped lentiviral particles (inhibitory dilution at which 50% inhibition of infection is achieved, ID50), $\log_{10}$

transformed ID50 after dose 1 (A, n=101) and dose 2 (B, n=21). r– Pearson's correlation coefficient, β slope/regression coefficient, p-value.

5.2.2 Older age is a risk factor for poor serum neutralisation of WT Spike pseudotyped virus after the first dose of mRNA BNT162b2

A logistic regression model was used to analyse the association between age and inadequate vaccine-elicited serum antibody neutralisation against WT pseudotyped virus after the first dose of the BNT162b2. The unadjusted odds ratio (OR) for achieving inadequate neutralisation against WT was 4.4 (1.7-11.1), p0.002) for participants ≥ 80 years old versus <80 years (Table 5.2). In the univariable analyses, a longer time interval since the first vaccine dose was received was also associated with inadequate neutralisation with an OR of 2.9 (95% CI 1.3-6.8) in those vaccinated 9-12 weeks prior compared with those vaccinated between 3-8 weeks prior. Sex, prior exposure to SAR-CoV-2 and HCMV serostatus were not associated with inadequate vaccine-elicited serum antibody neutralisation. In the fully adjusted model including age, sex, and time interval since the first dose of vaccine was received, the association of older age with inadequate vaccine-elicited serum antibody neutralisation was still present with an OR of 3.7 (95% CI 1.2-11.2) for participants ≥ 80 years old versus <80 years but the association with time interval since the first dose of vaccine was no longer present as it was a confounder for the association between age and inadequate vaccine-elicited serum antibody neutralisation. Other confounders such as the presence of comorbidities and a history

of immunosuppressive treatment could not be adjusted for as these data was not available for the majority of participants.

	Number	Risk ID50<20	Unadjusted OR (95% CI)	P value	Adjusted OR* (95% CI)	P value
WT						
Age group *years*						
<80	42	19.1 (8/42)	1		1	
≥ 80	59	50.9 (30/59)	4.4 (1.7-11.1)	0.002	3.7 (1.2-11.2)	0.02
Sex						
Male	61	36.1 (22/61)	1		1	
Female	40	40.0 (16/40)	1.2 (0.5- 2.7)	0.59	1.2 (0.5-2.9)	0.72
Time since dose 1 *weeks*						
3-8	68	29.4 (20/68)	1		1	
9-12	33	54.6 (18/33)	2.9 (1.2-6.8)	0.02	1.4 (0.5-3.9)	0.57
Previous SARS-CoV-2						
Negative anti-N IgG	90	36.7 (33/90)	1			
Positive anti-N IgG	9	44.4 (4/9)	1.4 (0.3-5.5)	0.65	-	
HCMV serostatus						
Negative	37	43.2 (16/37)	1			
Positive	34	38.2 (13/34)	0.8 (0.3-2.1)	0.67	-	

Table 5.2: Neutralisation in participants after the first dose of the mRNA BNT162b2 vaccine against wild type pseudotyped viruses.

*Adjusted for sex and any other variables in the table with a p value<0.1 in the univariable analyses. WT- wild type, ID50- (Inhibitory dilution) –

the serum dilution achieving 50% neutralisation, ns- non-significant, CI-confidence interval, HCMV- human cytomegalovirus.

5.3 Binding antibody responses following first and second dose mRNA BNT162b2 vaccination

5.3.1 Binding antibody levels increases in response to mRNA BNT162b2 vaccine

Binding antibodies to the WT Spike (S), Receptor Binding Domains (RBD) and Nucleocapsid (N) proteins were comprehensively measured using a particle-based assay. (Methods 2.3.5). The serology was performed by the Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge, in collaboration with Professor Rainer Doffinger. IgA, total IgG and IgG 1-4 responses were detected in unexposed healthy controls (UHC), COVID-19 convalescent sera and after both doses of vaccine (Figure 5.5A, B). The median anti-S IgG response increased between the two time points from 20 297 (IQR 8 481-27 629) MFI to 30 894 (IQR 30 341-31 180) MFI after the second dose. This was comparable with that of COVID-19 convalescent sera which was 29 711(IQR 28 621-30 238) MFI (Figure 5.5A). The same trend was observed with anti-S IgA responses (Figure 5.5B). This confirms the expectation that the booster dose of the vaccine leads to production of higher   levels of vaccine specific antibodies.

Although the median anti-N IgG was elevated after the first dose of vaccine compared with healthy controls [1 793(IQR 1 283-2 997) vs. 198 (IQR 159-375) MFI], it was significantly lower than the response flowing natural infection with SARS-CoV-2 (29 877(IQR 28 652-31 008) MFI and may indicate cross reactivity with other seasonal coronaviruses or in 9 individuals it was sufficiently elevated

along with anti-S IgG to suggest prior SARS-CoV-2 exposure (Figure 5.5A). The

same trend was observed with anti-S IgA responses (Figure 5.5B)
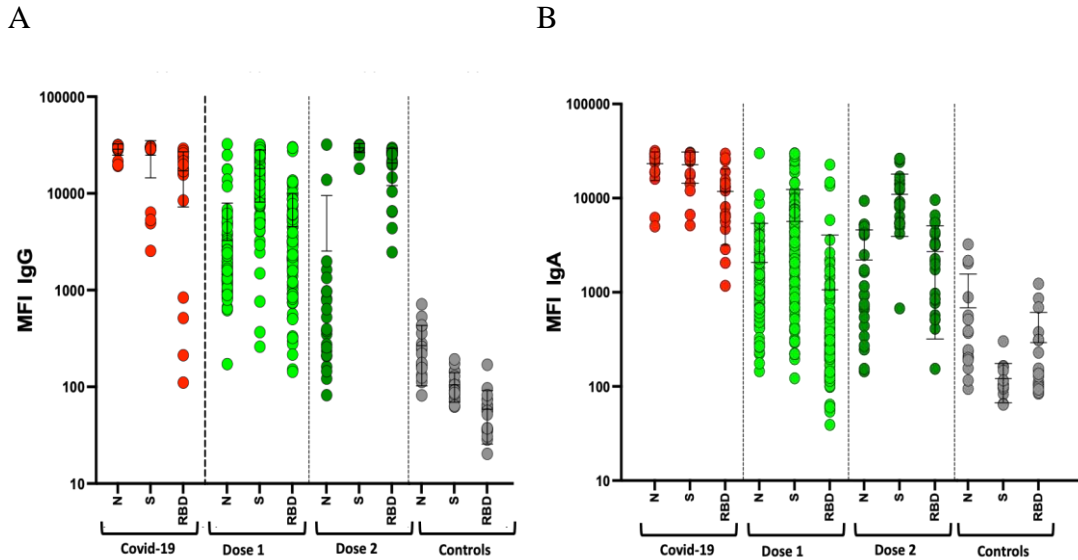
Figure 5.5. Binding antibody responses following vaccination with mRNA BNT162b2 vaccine. A. Anti- S, N, RBD IgG responses post first dose (light green, n=99) and second dose (dark green, n=21) compared to COVID-19 convalescent patients (red, n=18) and healthy controls (grey, n=18) at serum dilutions 1 in 100. **B.** Anti- S, N, RBD total IgA responses post first dose (light green, n=99) and second dose (dark green, n=21) compared to COVID-19 convalescent patients (red, n=18) and healthy controls (grey, n=18) at serum dilutions 1 in 100. Data presented are MFI – mean fluorescence intensity ± s.d. S – Spike, N – nucleocapsid, RBD – Spike receptor binding domain, Ig- immunoglobulin. Serology was done by the Department of Clinical Biochemistry and Immunology, Addenbrooke's Hospital, Cambridge, UK.

Anti-S and anti-RBD IgG 1-4 increased between vaccine doses (Figure 5.6), as is seen in natural infection with SARS-CoV-2. However, the concentration of total and subclass anti-S and anti-RBD IgGs were significantly lower in the ≥80 years group compared with the <80 years group after dose 1 but was no longer significantly different after dose 2 (Figure 5.7).
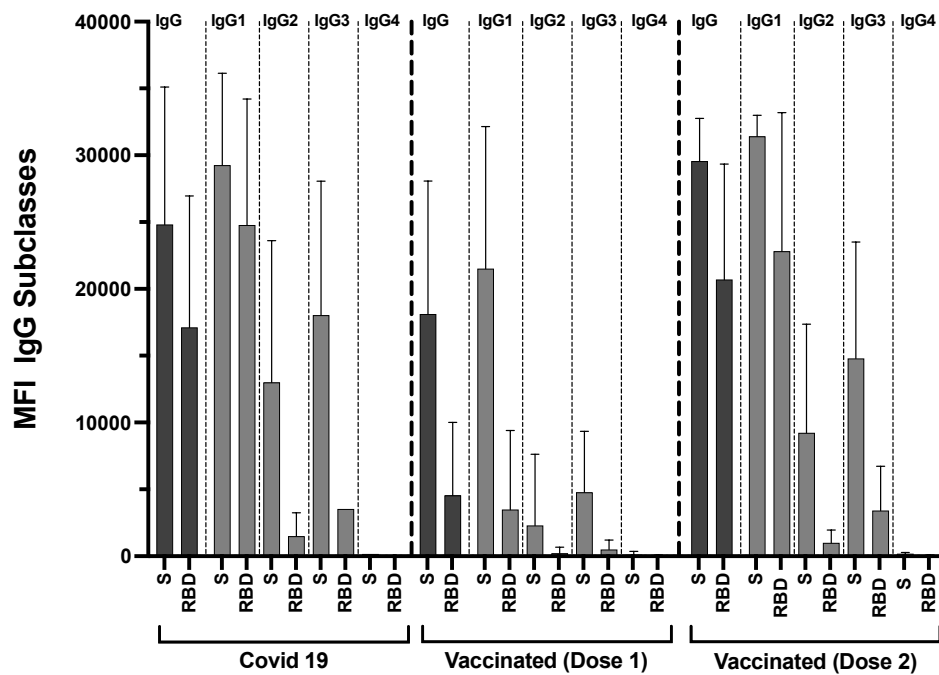
Figure 5.6. Anti-S, N, RBD IgG subclass responses post first and second dose of mRNA BNT162b2 vaccine. Covid-19 convalescent (n=19), first dose (n=99), second dose (n=21). Data presented are median with upper IQR. MFI – mean fluorescence intensity. S – Spike, N – nucleocapsid, RBD – Spike receptor binding domain, Ig- immunoglobulin.
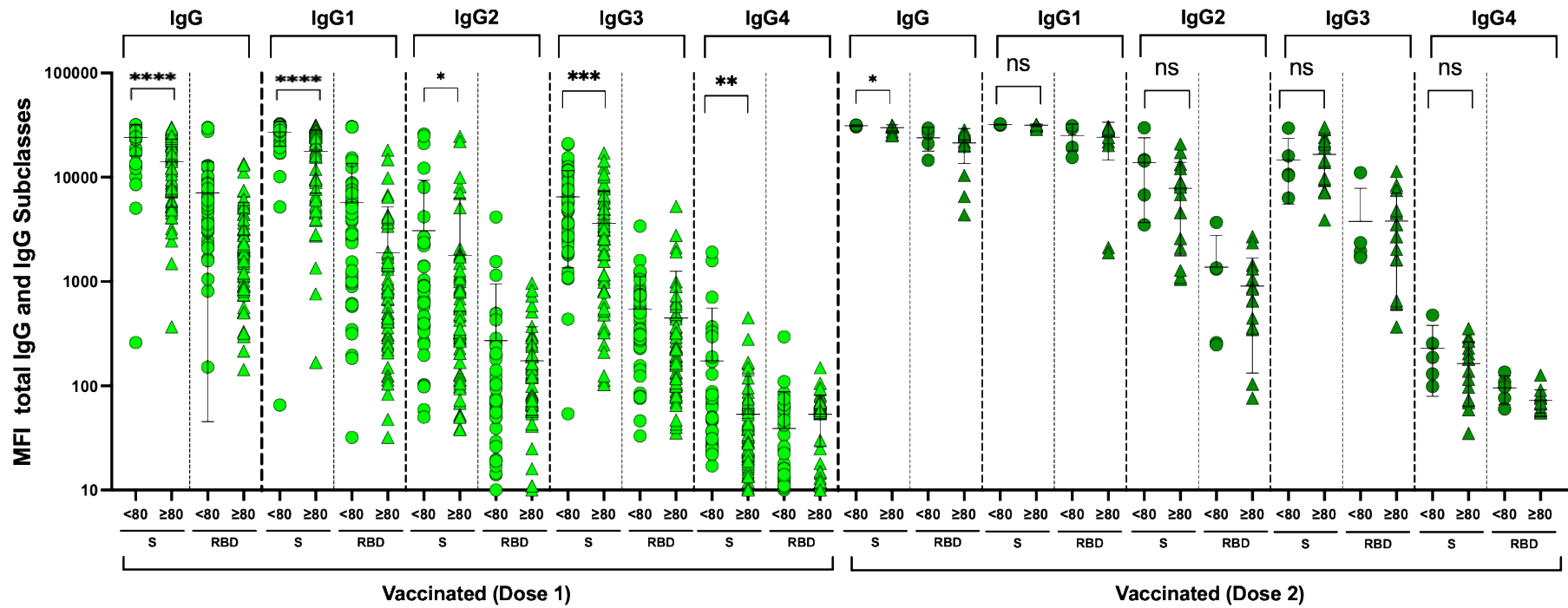
Figure 5.7. Anti-Spike and anti-RBD IgG- total and subclasses after first dose of mRNA BNT16b2 vaccine stratified by age < and ≥80 years old. MFI – mean fluorescence intensity. S – Spike, RBD – Spike receptor binding domain, Ig- immunoglobulin.

5.3.2 Binding antibody levels correlate with age

The binding antibody MFIs after the first vaccine dose were $\log_{10}$ transformed and their correlation with age was analysed with a simple linear regression model. Pearson correlation was used to determine the correlation coefficient (r) between vaccine-elicited Spike specific binding antibody response and age in years. mRNA vaccine-elicited anti-S IgA and IgG responses were significantly, negatively correlated with age (Figure 5.8A, B). All anti-S IgG subclasses were also significantly, negatively correlated with age (Figure 5.8C).
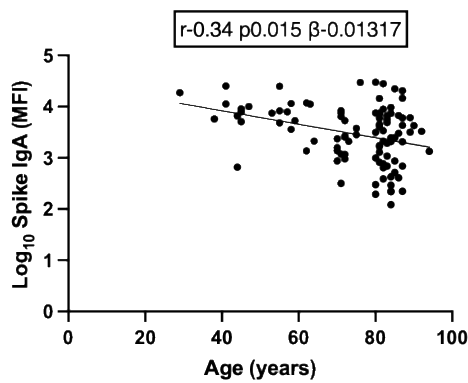
A   B
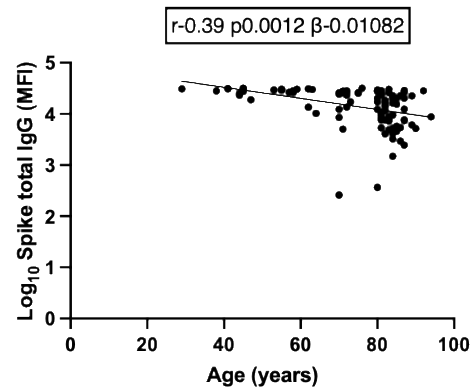
C
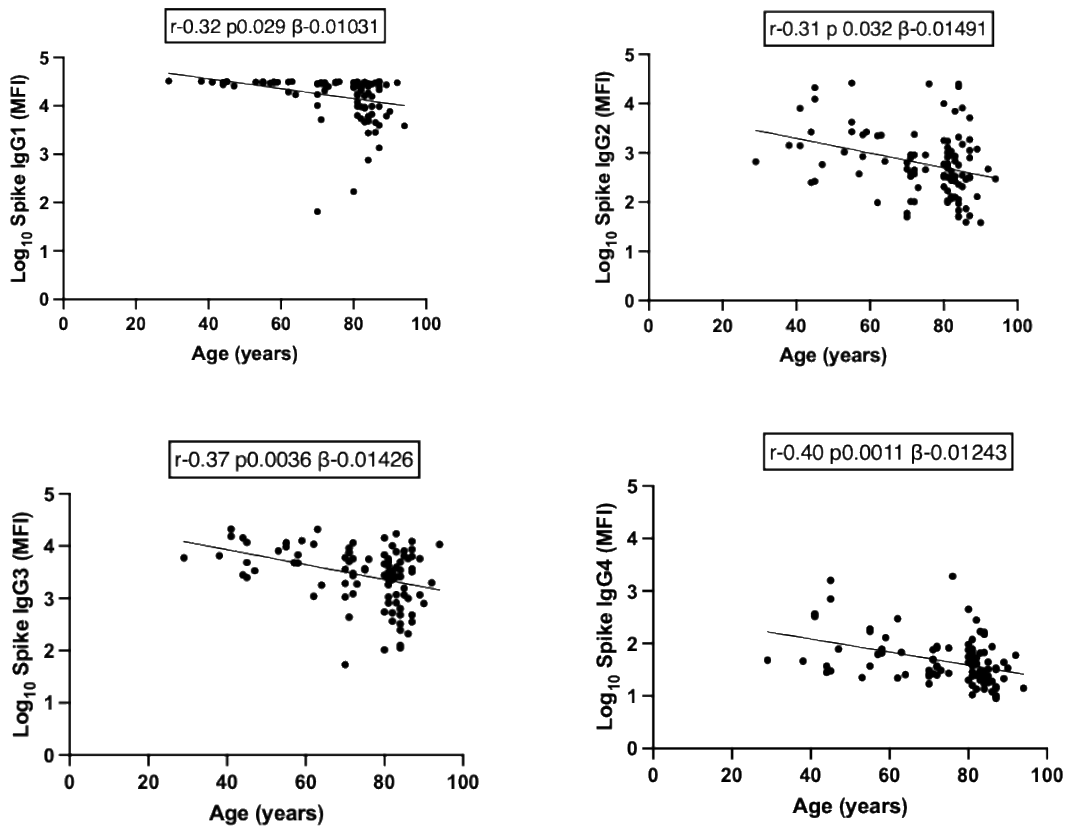
Figure 5.8. Correlation between anti-Spike Ig A and IgG binding antibody responses and age. A. Anti-S IgA. **B.** Anti-S total IgG. **C.** Anti-Spike IgG subclass 1-4. MFI – mean fluorescence intensity. S – Spike, Ig- immunoglobulin, r– Pearson's correlation coefficient, p- P value, β slope/regression coefficient.

5.3.3 Binding antibody levels correlate with serum neutralisation of SARS-CoV-2

Vaccine specific antibody responses after dose 1 of the mRNA vaccine were also correlated with ID50; the serum dilution at which 50% inhibition of WT Spike pseudotyped virus bearing D614G. The association between $\log_{10}$ transformed anti-S

IgA, total IgG and IgG1-4 subclasses and $\log_{10}$ transformed the ID50 was analysed with simple linear regression. Pearson correlation was used to determine the correlation coefficient (r) between vaccine-elicited Spike specific binding antibody response and ID50. There was a significant positive correlation between anti-S IgA, total IgG and IgG1-4 subclasses binding antibody levels and ID50. (Figure 5.9A, B, C). These results suggest that poorer neutralisation of WT Spike pseudotyped virus by vaccine-elicited antibodies in the ≥80 years group after the first dose of the BNT162b2 vaccine, is mediated by lower magnitude of neutralising antibody levels, which is in turn determined by age.

A
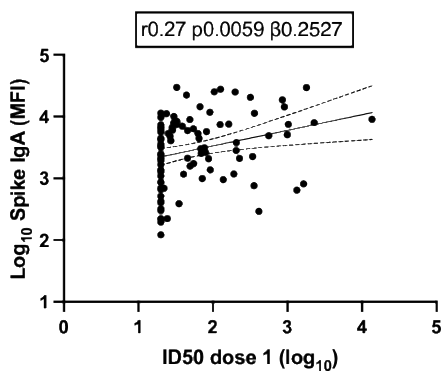
B
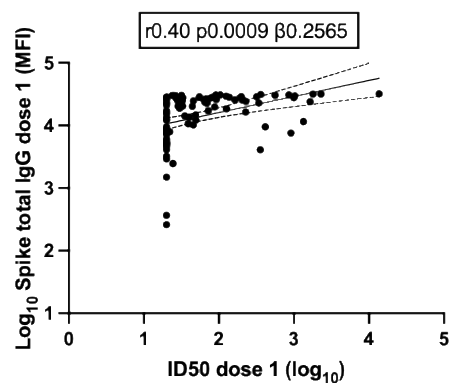
C
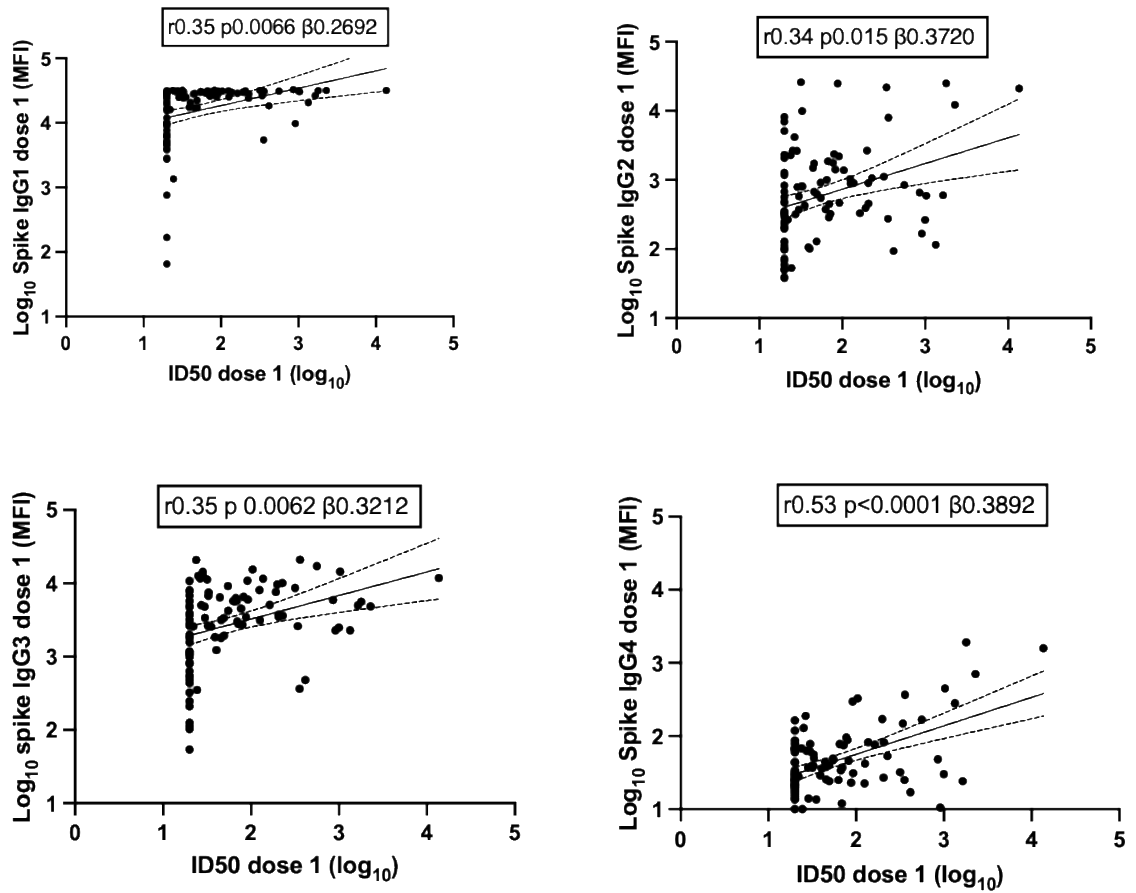
Figure 5.9. Correlation between anti-Spike Ig A and IgG binding antibody responses and serum neutralisation. A. Anti-S IgA. **B.** Anti-S total IgG. **C.** Anti-Spike IgG subclass 1-4. MFI – mean fluorescence intensity. ID50- inhibitory dilution at which 50% inhibition of infection is achieved against WT Spike pseudotyped virus, S- Spike, Ig- immunoglobulin, r– Pearson's correlation coefficient, p- P value, β slope/regression coefficient.

5.4 T cell responses following first and second dose mRNA BNT162b2 vaccine

5.4.1 Frequency of Spike specific T cells following vaccination

Induction of memory T cells is key to T cell vaccine efficacy, therefore T cell responses to SARS-CoV-2 spike protein following the first and second doses of the mRNA BNT162b2 vaccine were measured. PBMCs from vaccinees were stimulated with overlapping peptide pools to the WT SARS-CoV-2 spike or as a virus specific control, a peptide pool including Cytomegalovirus, EBV and Flu (CEF+) specific peptides. IFNγ and IL-2 FluoroSpot assay was used to count spike specific T cells. An unexposed healthy control (UHC) population was selected from PBMC biobanked between 2014-2016 which precedes the current SARS-CoV-2 pandemic, in order to provide a background control for SARS-CoV-2 spike protein responses. A positive control population was selected from biobanked PBMCs from participants with PCR confirmed SARS-CoV-2 infection.

Following the first dose of the mRNA vaccine in the < 80 years group, the frequency of IFNγ spike specific T cells was significantly larger than the responses seen in the UHC population. However, in the ≥80 years group, the IFNγ spike specific T cell responses were not different from the UHC following first dose (Figure 5.10A). Spike specific IL-2 T cell frequencies were significantly greater than the UHCs in both age groups but there was a difference between the <80 and ≥80 years groups, with the ≥80s having significantly lower IL-2 responses (Figure 5.10B).
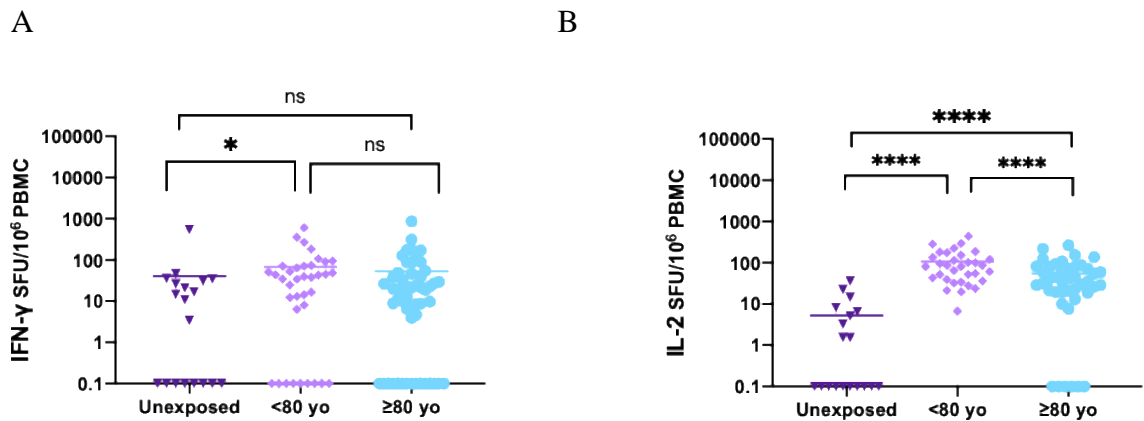
Figure 5.10: T cell responses after the first dose of the mRNA BNT162b2 vaccine. A. IFNγ FluoroSpot and **B**. IL-2 FluoroSpot. T cell responses specific to SARS-CoV-2 Spike protein peptide pool following PBMC stimulation of unexposed healthy controls (n=20) and vaccinees <80 years IFNγ (n=37), IL-2 (n= 33) and ≥80 years IFNγ (n=54), IL-2 (n=43) three weeks or more after the first doses of mRNA BNT162b2 vaccine. SFU- Spot forming unit, IFNγ- interferon gamma, IL-2- interleukin-2, PBMC- peripheral blood mononuclear cells. Mann-whitney test was used for unpaired comparisons and Wilcoxon matched-pairs signed rank test for paired comparisons. p-values * <0.05, **** <0.0001, ns not significant. Done in collaboration with Dr Mark Wills.

Spike specific T cell responses after the first and second dose of vaccines were compared with both infected cases of SARS-CoV-2 (confirmed PCR positive) and the UHCs. As was expected, IFNγ and IL2 responses in infected individuals were significantly greater than the UHCs. The frequency of IFNγ and IL2 spike specific T cells after the first dose was significantly greater than the UHCs. This increased

further following the second dose however, only the IFNγ response was statistically

significantly increased (Fig 5.11A, B).
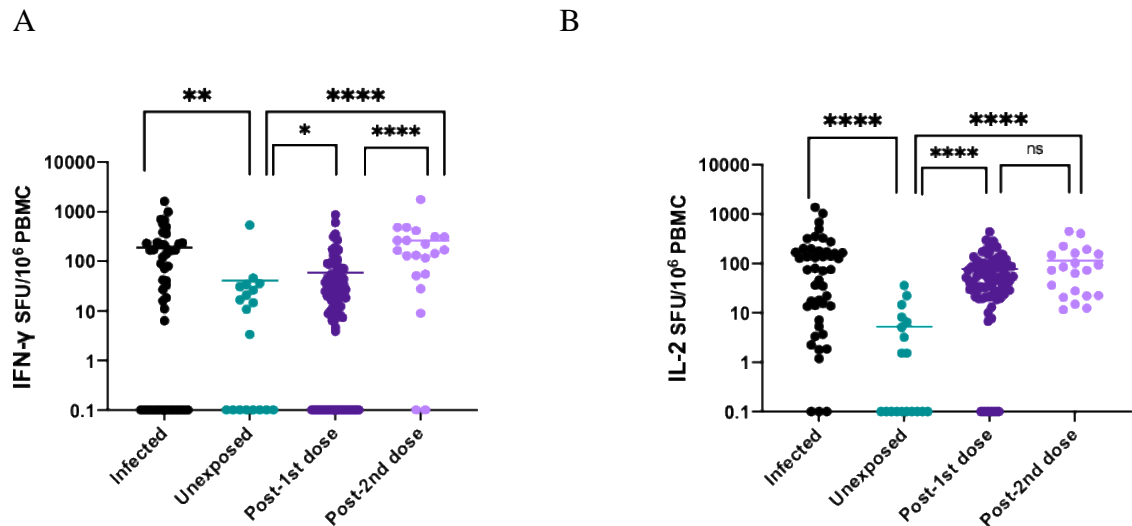
Figure 5.11: T cell responses after the first and second dose of the mRNA

BNT162b2 vaccine. A. IFNγ FluoroSpot and **B**. IL-2 FluoroSpot. T cell responses

specific to SARS-CoV-2 Spike protein peptide pool following PBMC stimulation of

infected SARS-CoV-2 cases (n=46), unexposed healthy controls (n=20) and all

vaccinees three weeks or more after the first doses IFNγ (n=92), IL-2 (n=76) and

three weeks after second IFNγ and IL-2 (n=21) of the mRNA BNT162b2 vaccine.

SFU- Spot forming unit, IFNγ- interferon gamma, IL-2- interleukin-2, PBMC-

peripheral blood mononuclear cells. Mann-whitney test was used for unpaired

comparisons and Wilcoxon matched-pairs signed rank test for paired comparisons.

p-values * <0.05, ** <0.01, **** <0.0001, ns not significant. Done in collaboration

with Dr Mark Wills.

The frequency of IFNγ and IL-2 vaccine induced spike T cells were compared between the <80 years and ≥80 years groups for paired first and second vaccine dose samples. The spike specific IFNγ responses significantly increased for both age groups following the second dose. The spike specific IL-2 responses did not significantly increase following the robust response induced by the first dose. However, the IL-2 response was significantly lower in the ≥80 years group than the < 80 years group following the first dose of the vaccine (Figure 5.12A, B).

A   B
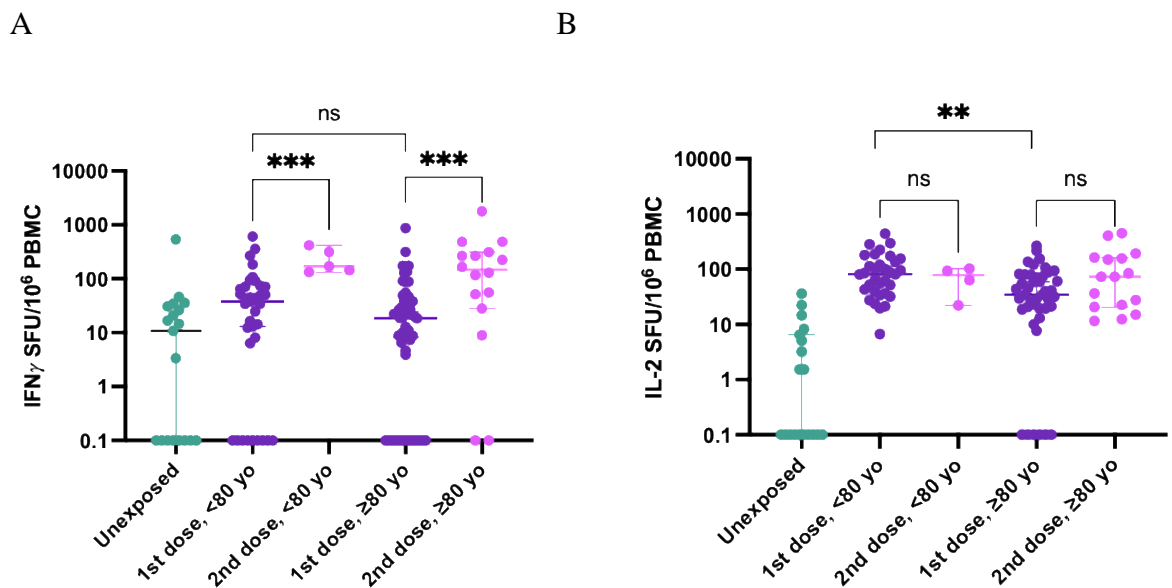
Figure 5.12: T cell responses after the first and second dose of the mRNA BNT162b2 vaccine according to age group. A. IFNγ FluoroSpot and **B**. IL-2 FluoroSpot. T cell responses specific to SARS-CoV-2 Spike protein peptide pool following PBMC stimulation of a unexposed healthy controls (n=20) and vaccinees <80 years IFNγ (n=37), IL-2 (n=33) and ≥80 years IFNγ (n=54), IL-2 (n=43) at least three weeks after the first doses and <80 years IFNγ (n=5), IL-2 (n=4) and ≥80 years IFNγ (n=16), IL-2 (n=16) three weeks after second dose of the mRNA BNT162b2

vaccine. SFU- Spot forming unit, IFNγ- interferon gamma, IL-2- interleukin-2, PBMC- peripheral blood mononuclear cells. Mann-whitney test was used for unpaired comparisons and Wilcoxon matched-pairs signed rank test for paired comparisons. p-values ** <0.01, *** <0.001, ns not significant. Done in collaboration with Dr Mark Wills.

5.4.2 Correlation of Spike specific T cell response with serum neutralisation of SARS-CoV-2

Given that T helper cells have a role in effecting extrafollicular and follicular B cell responses to antigens (Papavasiliou and Schatz 2002), it was hypothesised that Spike specific T cell response following vaccination with the mRNA vaccine may be correlated with serum neutralisation. The association between $\log_{10}$ transformed frequency of IFNγ and IL-2 vaccine induced Spike T cell responses with $\log_{10}$ transformed ID50; the serum dilution at which 50% inhibition of infection was achieved were analysed with simple linear regression. Pearson correlation was used to determine the correlation coefficient (r) between vaccine-elicited Spike specific binding antibody response and ID50. IFNγ and IL-2 vaccine induced spike T cell responses did not correlate with ID50 against WT Spike pseudotyped virus after first or second dose vaccine (Figures 5.13A-D).
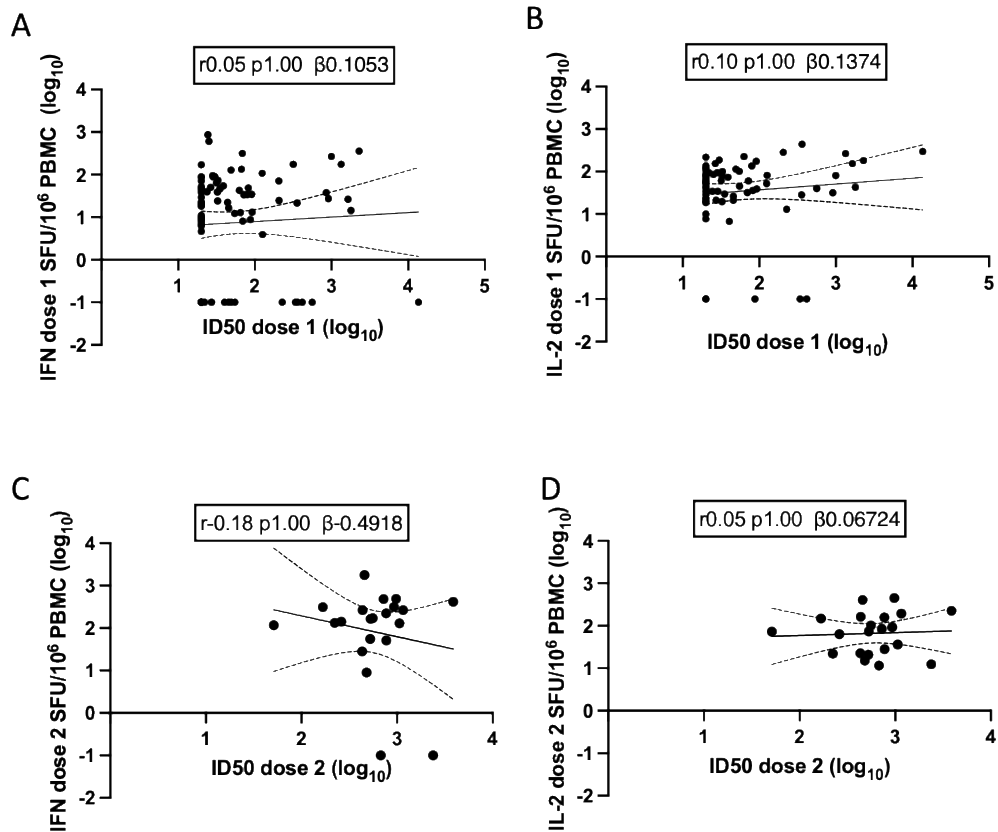
Figure 5.13. Correlation between T cell responses against SARS-CoV-2 Spike and serum neutralisation of WT Spike pseudotyped virus following dose 1 and 2 of the mRNA BNT162b2 vaccine. A, B. IFNγ (n=89) and IL2 (n=74) FluoroSpot after first dose. **C, D.** IFNγ (n=21) and IL2 (n=21) FluoroSpot after second dose. SFU- Spot forming unit, IFNγ- interferon gamma, IL-2- interleukin-2, PBMC- peripheral blood mononuclear cells, r- Pearson correlation coefficient, p- p value, β- the slope or coefficient. Bonferroni adjustment was made for multiple comparisons.

5.4 3 T cell responses to SARS-CoV-2 Spike peptide pools are specific

The PBMC from UHCs and vaccinees after the first dose of the mRNA vaccine were stimulated with a peptide pool including Cytomegalovirus, EBV and Flu (CEF+) specific peptides. The response to CEF+ peptides were comparable across age groups, indicating that differences in observed responses were vaccine specific and unlikely due to generalised suboptimal T cell responses/immune paresis (Figure 5.14).
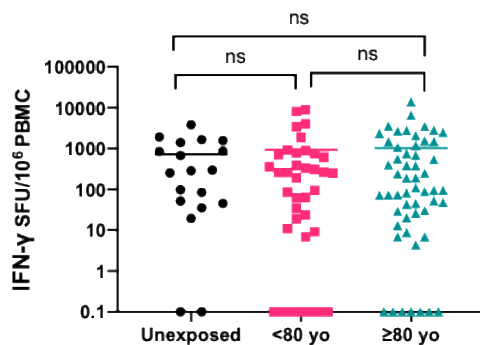
Figure 5.14. FluoroSpot IFNγ PBMC responses to peptide pool of Cytomegalovirus, Epstein Barr virus and Influenza virus (CEF). Response from unexposed healthy controls (n=20), <80yo (n=37) and >80yo (n=54) three weeks after the first doses of Pfizer BNT162b2 vaccine. SFU- Spot forming unit, IFNγ- interferon gamma, PBMC- peripheral blood mononuclear cells. Mann-whitney test was used for unpaired comparisons and Wilcoxon matched-pairs signed rank test for paired comparisons. Ns- not significant. Done in collaboration with Dr Mark Wills.

5.4.4 HCMV seropositivity maybe associated with increased frequency of IFNγ producing Spike specific Tcells

HCMV seropositivity has been associated with poor age-related responses to vaccination (van den Berg et al. 2019). HCMV serostatus was determined for UHCs and vaccinees. Given the demographic from which the study population was drawn, HCMV positivity was more common in the ≥80 years compared with the <80 years group (55% vs 37%) (Figure 5.15A). Individuals in the ≥80 years group with positive HCMV serology had significantly higher IFNγ, but not IL2, responses to SARS-CoV-2 Spike peptides compared to the <80 years group (Figure 5.15B, C). Although not significant, there was a there was a trend towards a lower ID50 in HCMV positive individuals (Figure 5.15D). HCMV positive individuals as expected has a greater response to CEF peptides than HCMV seronegative. The HCMV positive group responded to the CMV peptides in the CEF pool, while the HCMV negative group could only have responded to the EBV and Flu components. This effect the same in both age groups (Figure 5.15E).
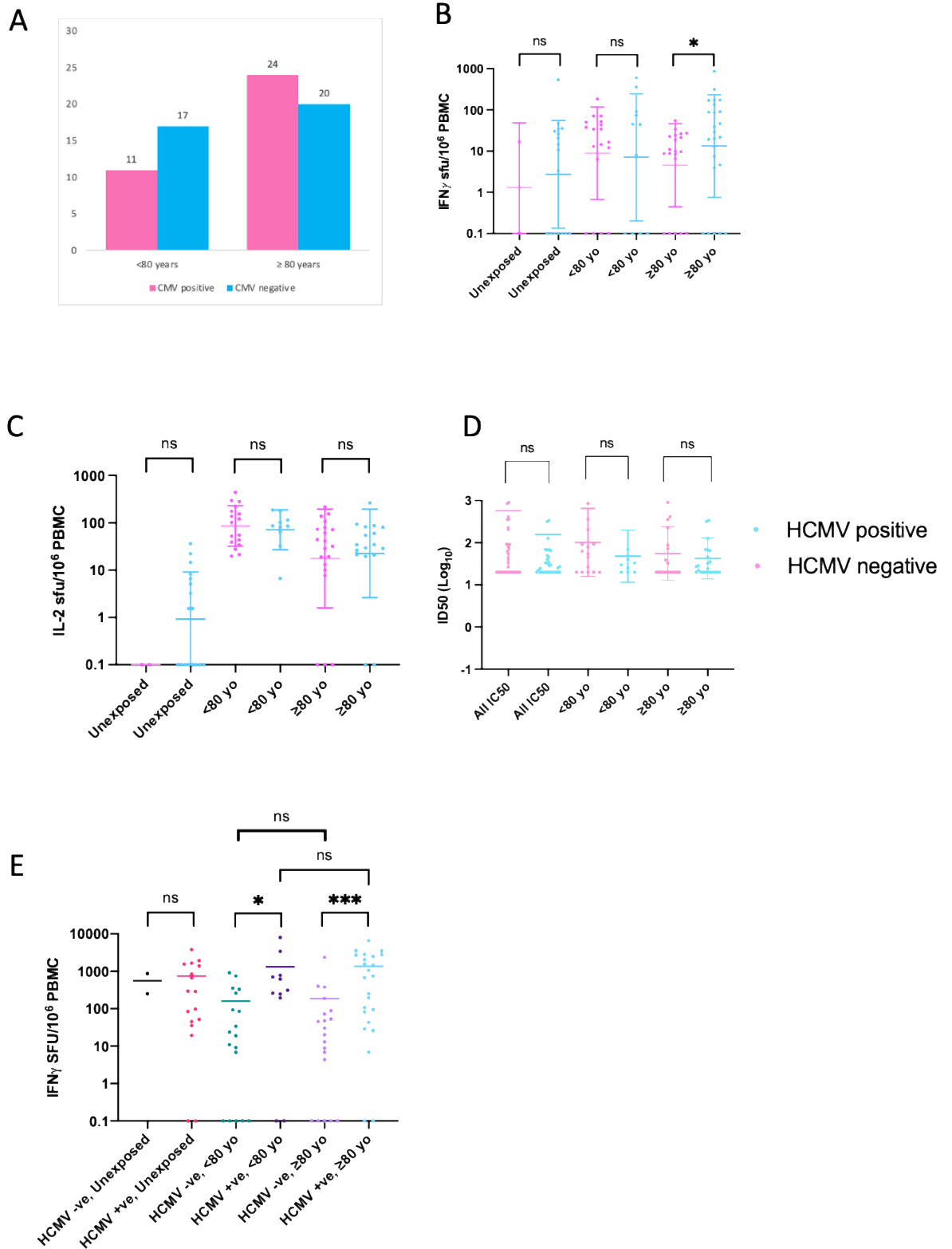
Figure 5.15. Human cytomegalovirus serostatus, T cell responses and serum neutralisation of WT Spike pseudotyped virus after dose 1 of the mRNA BNT162b2 vaccine. A. HCMV serostatus by <80 and ≥80 year groups (n=72),

HCMV seropositive (blue), HCMV seronegative (pink). **B** IFNγ FluoroSpot (n=72) and **C.** IL-2 FluoroSpot (n=64) T cell responses to SARS-CoV-2 Spike peptides after the first dose of BNT162b2 vaccine by HCMV serostatus and age group. **D.** ID50 after the first dose of BNT162b2 vaccine by HCMV serostatus and age group. ID50- serum dilution at which 50% inhibition of infection of WT Spike pseudotyped virus bearing D614G is achieved. **E.** IFNγ FluoroSpot response to CEF peptides after the first dose of BNT162b2. after the first dose. SFU- Spot forming unit, IFNγ-interferon gamma, PBMC- peripheral blood mononuclear cells. Mann-whitney test was used for unpaired comparisons and Wilcoxon matched-pairs signed rank test for paired comparisons. p-values * <0.05, *** <0.001, ns- not significant. Done in collaboration with Dr Mark Wills.

5.5 Discussion

It was hypothesised that an appropriate vaccine response would include titres of neutralising antibodies sufficient to inhibit infection in a Spike pseudotyped assay. It is still unknown what level of neutralisation offers protection. However, an arbitrary cut off of ID50 at a serum dilution of 1 in 20 was deemed an adequate response. Although the limit of detection of the assay was at a dilution of 1 in 4, I found some variability in ID50 estimates at dilutions lower than 1 in 20. In addition, vaccines should elicit a response in the T cell arm of the adaptive immune system and it was hypothesised that an adequate response to the mRNA BNT162b2 vaccine would increase the frequency of Spike-specific IFNγ and IL-12 secreting T cells compared with unvaccinated individuals. The data presented shows that the odds of an inadequate neutralising antibody response to a spike pseudotyped virus after

only one dose of the mRNA BNT162b2 vaccine was 3.7 in those ≥80 years old. This correlated with anti-Spike IgA and IgG binding antibody levels. In addition, frequency of IFNγ and IL-12secreting Spike specific T cells following only one dose of mRNA BNT162b2 vaccine was significantly lower in the ≥80 years group compared with the <80 years group. These age-related differences were overcome with a second dose of the vaccine.

In the analyses of the association between age group and inadequate neutralisation, other confounders including immune suppression and comorbidities were not adjusted for as these data were not collected. They may be other unmeasured confounders that were not adjusted for. In other to prevent selection bias, consecutive participants were recruited without exclusion and the sample size was adequate.

Important correlates of vaccine response have been shown which are age dependent. Future work would include following up these participants to look for break through SARS-CoV-2 infection, to track the kinetics of neutralisation activity and more definitively identify correlates of protection. It would be important to widen the study population to include more broadly immunocompromised individuals because as seen from the within-host evolution work presented in the previous chapter, a suboptimal immune response may drive within-in host adaptation of the virus to escape vaccine control. In addition, Spike specific B cell receptor sequencing, antibody cloning and expression would be required, as well as a broader analyses of T cell phenotype and cytokine responses would be important to further understand the processes mediating the effects shown.

CHAPTER 6: DISCUSSION

PART 1

6.1 HIV-1 CSF escape in the context of cryptococcal meningitis

HIV-1 CSF escape was uncommon in this study population and occurred in only 4% (2/44). The true prevalence of CSF escape in virologically suppressed individuals is uncertain. Estimates have ranged from 1.5 to 10 % (Eden et al. 2010; Underwood et al. 2018; van Zoest et al. 2017). The prevalence is highly influenced by the limit of detection of the HIV-1 VL assay used and the volume of CSF analysed. Studies that have done longitudinal sampling have reported that some of these events are CSF blips and on resampling persistently detectable HIV-1 VL in the CSF is uncommon (reduction from 36% to 3%) (Eden et al. 2016).

However, in neurosymptomatic individuals or people having a lumbar puncture for clinical reasons, the prevalence is higher, ranging between 6 and 18% (Kugathasan et al. 2017; Mukerji et al. 2017; Mukerji et al. 2018; Rawson et al. 2012). As yet, there are no published prevalence studies of CSF escape/discordance from sub–Saharan Africa, so the burden in this region remains unknown (Collier et al. 2018). However, one study looked at the prevalence in a population with HIV-1 subtype C in Brazil and found a combined prevalence of CSF escape/discordance of 15.3% (de Almeida et al. 2020). Although reported asymptomatic, neuropsychological assessment of this cohort found evidence of global neuropsychological deficits in participants with CSF escape/discordance compared with HIV-1 negative controls.

Overall, CSF escape is higher in symptomatic individuals and those having a clinically indicated LP, it was therefore surprising that the prevalence of CSF escape in this study was not higher.

6.2 CNS compartmentalisation and patient characteristics associated with CNS compartmentalisation in the context of CM

Several studies have shown that the presence of HIV-1 replicating independently in the CNS is associated with clinically significant neurological disease. However, it was unclear to what extent the HIV-1 variants in the CNS are an independently replicating/compartmentalised population in patients with co-infection with HIV-1 subtype C and cryptococcal meningitis.

Visual inspection of the maximum likelihood phylogenetic trees revealed a spatial population structure that clearly segregated the taxa of the CSF derived genomes from those of the plasma in 2 out of 4 participants. This was supported by the results from the compartmentalisation analyses. The prevalence of CNS compartmentalisation in this preliminary analysis was determined. The prevalence of CNS compartmentalisation was 50% (95% CI 6.8-93.2). This is small sample size and the true prevalence maybe a lot smaller or greater as indicated by the wide confidence intervals. Consistent with the data presented in this thesis CNS compartmentalisation has been found in up to 40% of HIV-1/cryptococcal meningitis co-infected participants in other studies in sub-Saharan Africa (Adewumi et al. 2020; Lange et al. 2018). This sample size was too small to do any meaningful

analyses of clinical or demographic factors associated with CNS compartmentalisation.

6.3   Phylogenetic characteristics of paired CNS and plasma genomes of HIV-1 subtype C in the context of co-infection with cryptococcal meningitis

CNS compartmentalisation was identified by visual inspection of maximum likelihood phylogenetic tree as detailed above. Statistical analyses using distance and tree-based methods were used in addition and these supported the conclusions from visual inspection. In the 2 participants with CNS compartmentalisation, the most recent common ancestor was inferred to be a plasma variant that was in circulation prior to emergence of the cluster of viruses detected on the CNS. This fits with the hypothesis that the viruses in the CNS crosses the blood brain barrier earlier in the individual's history of HIV-1 infection and establishes an independently replicating reservoir. In the other participants with intermixed viruses, this evolutionary process is not evident. Although, it is possible that all the variants in each compartment have not been sampled by SGA and the founder variant may have been missed.

The viruses in the CNS cluster are not clonal, suggesting that the virus has continued to evolve in the CNS compartment (Stam et al. 2013). One of the participants with CNS compartmentalisation was on ART for 10.6 months, it possible that poor penetration of ART into the CSF may have driven this evolution under drug pressure.

In addition, analyses of the diversity measured by average pairwise distance (APD) between genomes in plasma compared with CSF found that participants with CNS compartmentalisation had increased diversity of HIV-1 genomes in either their plasma or CSF compartment but those participants with intermixed viruses had similarly diverse viruses in both compartments. Increased diversity of CSF derived HIV-1 genomes has previously been reported (Tong et al. 2015; Liu et al. 2013). This maybe due to differential drug levels, cellular and humoral immune responses in the CNS and plasma compartments which may exert differential pressure on viruses in either compartment, leading to viral evolution and the observed diversity.

Participant 002 had received tenofovir/lamivudine/efavirenz combination therapy for 10 months but had no DRMs in the context of very high VLs, suggesting this participant was non-adherent. However, if this participant had taken some drug for a proportion of time, it is likely that the selection pressure exerted by poor ART levels may have driven a soft selective sweep and limited the diversity of the HIV-1 plasma population as observed by average pairwise distance (APD) (Feder et al. 2016a). In contrast participant 003 was not on ART and had a high diversity of viruses in the plasma according to APD.

6.4 Phenotypic properties of HIV-1 isolated from the CSF

6.4.1 Co-receptor usage

Effective HIV-1 entry and infection is dependent on primary binding to host CD4 receptor followed by secondary binding to either CCR5 or CXCR4 co-receptors. Subtype dependent, co-receptor switch from CCR5 to CXCR4 has also been shown to precede the onset of advanced HIV/AIDS. Since cryptococcal meningitis is an AIDS defining illness, it maybe expected that patients with HIV-1/CM co-infection may utilize the CXCR4 co-receptor. Furthermore, a study of patients with HIV-1/CM co-infection in South Africa using computer predictive algorithms (CPAs), identified participants with CXCR4 utilising HIV-1 *env* clones in plasma and CSF (Sojane et al. 2018). All four participants V3 loop sequences were analysed using CPAs- GenotoPheno and Web PSSM. GenotoPheno predicted that all sequences from both compartments in all participants tested, utilised the CCR5 co-receptor. In contrast, PSSM predicted compartmentalised CNS co-receptor usage, with more than half of the CSF sequences from one participant and 3% of CSF sequences from another participant utilising the CXCR4 co-receptor. Both these participants had CNS compartmentalised viruses. However, this was not consistent with the results in the phenotypic co-receptor analyses. All clones derived from CSF and plasma tested phenotypically were found to utilise the CCR5 co-receptor. Furthermore, inhibition with a CCR5 inhibitor, blocked infection and confirmed dependence of these clones on the CCR5 co-receptor.

The CPAs have their drawbacks with regards their accuracy in predicting coreceptor usage (Low et al. 2007; Sing et al. 2007). They only consider the V3 loop, whereas regions outside the V3 loop such as the V2 and gp41 have been found to also contribute to determination of coreceptor usage (Huang et al. 2008; Thielen et al. 2010). Although GenotoPheno accurately predicted coreceptor usage in this dataset, PSSM did not. PSSM may have performed more poorly here because it uses the amino acid at positions 11 and 25 and the net charge to predict co-receptor usage, which is vulnerable to indels in the alignment.

6.4.2 Drug resistant mutations (DRMs)

Persistence of HIV-1 in tissue reservoirs such as the CNS, where there is poor drug penetration may allow the virus to sequester and replicate (Beguelin et al. 2016; Collier et al. 2018; Canestri et al. 2010; Peluso et al. 2012). In addition, mutations G73S in protease, M184V and M230I in reverse transcriptase have been reported to be significantly over-represented in patients with CSF escape and CNS compartmentalisation (Fourati et al. 2014; Liu et al. 2013; Mukerji et al. 2018). The RT sequence of 31 paired CSF and plasma samples were analysed and revealed no clear evidence of compartmentalised DRMs. The frequency and type of DRMs were largely similar in both CSF and plasma compartments. This is not surprising in 55% of participants who were not receiving ART. Therefore, their viruses would not be under drug pressure. In the two participants identified with CNS compartmentalisation, one was treatment naïve whilst the other was treatment

experienced. The treatment experienced participant had no DRMs in either compartment.

6.4.3 Neutralisation sensitivity to broadly neutralising antibodies (bnAbs)

The HIV-1 Envelope is the target of neutralising antibodies. It was hypothesised that viral escape through mutation of the HIV-1 envelope glycoprotein (Env) maybe a means through which an independently replicating viral population can exist in the CNS. Given the chronic history of HIV-1 infection in this cohort it was expected that clones from both plasma and CSF maybe resistant to neutralisation bnAbs. There was heterogeneity in the neutralising effect of bnAbs on the clones of the two participants tested. An interesting finding was that the CSF clones from participant 003 were 3 to 30 -fold more sensitive to neutralisation activity of VRCO1- CD4 binding site antibody but more resistant to CCR5 inhibition by TAK-779 compared to their plasma clones. It is difficult to draw conclusions as to what these patterns of neutralisation sensitivity means because each virus is constrained by the niche in which it finds itself. However, this may indicate that these CSF clones have become adapted to binding CCR5 with higher affinity. Clones from participant 002 were more resistant to bnaAbs compared with participant 003. However, one CSF clone from participant 002 was sensitive to neutralisation activity of VRCO1. In contrast to participant 003, this clone was sensitive to inhibition by the CCR5 inhibitor.

Data on duration of HIV-1 infection were missing however it is possible that some of these observation are driven by the duration of HIV-1 infection and viruses in

participants with a longer history of HIV-1 would have had more opportunity to acquire adaptations to allow survival in its niche such as the acquisition of more N-linked glycosylation to facilitate escape from neutralising antibodies (Wei et al. 2003).

<div align="center">

PART 2

</div>

6.5 SARS-CoV-2 escape in the context of neutralising antibodies

It is increasingly recognised that some individuals have a protracted illness with chronic infection with SARS-CoV-2 and persistent viral shedding. The case presented in this thesis exemplifies this and provides a rare insight to *in vivo* escape from neutralising antibodies. This case was intensively sampled over the course of illness and gave the opportunity to map the viral response to antiviral treatment with remdesivir and convalescent plasma (CP) in an individual with poor B and T cell immune responses. In this case the dynamics of the viral population revealed a shifting landscape of Spike mutations in response to treatment with CP. Two mutations- $\triangle$H69/V70 and D796H emerged after the receipt of CP, diminished after a period of time consistent with the wash-out of the CP and remerged after receipt of a further unit of CP. D796H was found to be 2-fold more resistant to the neutralising activity of the third unit of CP and a combination of $\triangle$H69/V70 and D796H had a similar effect. Other pairs of mutations arose and receded during the interval between CP administration, underlying a dynamic process of selective sweeps on the viral population. Interestingly, in a tissue culture system, D796H had a 3.5-fold

reduced infectivity compared to wild type (WT) and △H69/V70 had a 2-fold increased infectivity and when it occurred with D796H, the infectivity was restored to levels similar to WT. This appears to be evidence of within-host adaption following CP. It also exemplifies what is referred a "short-sighted" evolution whereby a virus acquires "transmissibility-reducing adaptations" in response to adaptive immune responses (Lythgoe et al. 2017) – so for example D796H is an adaptation to CP but it reduced infectivity in tissue culture.

Two other cases of prolonged shedding that demonstrated within-host evolution have been reported. Choi et al[10], presented an immunosuppressed individual with significant virus evolution, which including NTD deletions and RBD mutations but in the absence of SARS-CoV-2 specific antibody therapy. Avanzato et al also reported a case of another immunocompromised individual with significant shifts in genetic composition of the viral population over 105 days (Avanzato et al. 2020). In this case CP was also used.

Intriguingly, △H69/V70 is observed in the B.1.1.7 variant first reported in the UK and it appears that we are witnessing converging evolution where this deletion in the NTD maybe a common pathway to restoring detects in infectivity from other spike mutations. Other deletion and mutations in the NTD have been reported, for example △144 in B.1.1.7 and △242-4 in B.1.135 first reported in South Africa. This is of importance because there is evidence of epitopes in the NTD targeted by strongly neutralising monoclonal antibodies such as 4A8 (Chi et al. 2020; Collier, De Marco, et al. 2021).

Divergent cluster of viruses bearing W64G and P330S were identified on days 93 and 95. This probably represents a previously unsampled compartmentalised population but remained rooted within the original population and consistent with having arisen from a single underlying viral population. Other studies have reported detection of SARS-CoV-2 RNA in sites outside of the airways (Choi et al. 2020; Zhang et al. 2020).

6.6 Age-related heterogeneity to mRNA BNT162b2 vaccine against SARS-CoV-2

Protection from infectious pathogens elicited by vaccines is driven by neutralising antibodies binding to the pathogens and preventing attachment to host cells. The next line of defence is the cellular immune response that targets and clears infected cells (The Royal Society 2021). The same process applies to protection from SARS-CoV-2 infection, where studies in rhesus macaques have shown that RBD-specific antibodies levels, neutralisation titres and T cells responses to both vaccine and natural infection predict protection from SARS-CoV-2 in challenge models (Mercado et al. 2020; McMahan et al. 2021). Similarly, a broad neutralising antibody and T cell response to the mRNA and Adenovirus vectored vaccines was seen in human clinical trials (Anderson et al. 2020; Ewer et al. 2021; Sahin et al. 2020) and following natural infection in humans with SARS-CoV-2 (Peng et al. 2020; Zuo et al. 2021; Seow et al. 2020). Although no human challenge studies have taken place, it stands to reason that these are likely correlates of protection.

B and T cells must work in concert to provide effective protection from infectious pathogens (Papavasiliou and Schatz 2002). However, the quality and magnitude of responses of both these arms of the adaptive immune system are affected by ageing (Gustafson et al. 2020; Westmeier et al. 2020; Wilkinson et al. 2017). The data presented in this thesis showed that 55% of those >80 years had a suboptimal neutralising antibody response after vaccination with one dose of BNT162b2 and a lower magnitude of neutralising antibody levels compared with the < 80 years (GMT 83.4 (95% CI 52.0-133.7) vs 46.6 (95% CI 33.5-64.8) p 0.01), which was restored to similar levels after the second dose was given. Binding IgA and IgG antibody levels were correlated with neutralisation. The finding of a lower antibody response following vaccines was reported in Phase I/II trials of the mRNA vaccines (Ramasamy et al. 2020; Walsh et al. 2020) and has been seen with other vaccines (Anderson et al. 2020). The improvement in antibody levels after the second dose was also seen in mouse studies where the ChadOx1 n-CoV-19 Adenovirus vectored vaccine was used (Silva-Cayetano et al. 2021).

Similar to the B cell antibody responses, the data presented show that Spike specific T cell responses to BNT162b2 were also impaired by ageing. The frequency of Interferon gamma (IFNγ) secreting CD8+ T cells and interleukin-12 (IL-12) secreting CD4+ T cells responding to stimulation by a pool of Spike peptides were shown to be lower in the >80 years group. The IFNγ secreting T cells increased after the second dose but the IL-2 secreting T cells were not boosted further. Human vaccine studies have shown a Th1 skewed response to vaccination (Ewer et al. 2021; Zhu et al. 2020; Sahin et al. 2020; Anderson et al. 2020; Jackson et al. 2020). Similar to the data presented, Spike specific T cell response were poorer in

aged mice and improved after the second dose when the ChadOx1 n-CoV-19 Adenovirus vectored vaccine was used (Silva-Cayetano et al. 2021). In addition, the data presented showed an expected higher prevalence of cytomegalovirus (CMV) seropositivity in the >80 years but this surprisingly was associated with greater IFNγ Spike specific responses, although some studies have linked CMV seropositive to poorer immune response to vaccines (van den Berg et al. 2019).

The importance of the dual action of both arms of the adaptive immune response in protecting against SARS-COV-2 is underlined in a monkey challenge study were depletion of CD8+ T cells in the context of low neutralising antibody levels, led to inadequate protection from infection (McMahan et al. 2021). This may offer sufficient protection in the real-world setting as observational studies have demonstrated some effect against severe disease and death after 1 dose of the BNT162b2 or ChAdOx1nCoV-19 (Voysey, Costa Clemens, et al. 2021; Bernal et al. 2021; Vasileiou 2021).

6.7 Conclusion

HIV-1 and SARS-CoV-2 are highly successful viruses that have established pandemics. They successfully jumped species, overcame transmission bottlenecks and adapted to successful transmission in humans. HIV-1 has demonstrated that it highly diverse and can continue to adapt to the constraints of its niche to stablish chronic infection in its host. Similarly, we have now seen that in the absence of

adequate immune responses, SARS-CoV-2 can similarly adapt to the constraints of its niche and establish chronic infection.

Compartmentalised cellular and immune responses in the CNS and plasma compartments may exert differential pressure on HIV-1 in either compartment, leading to viral escape. Since epitopes on the HIV-1 envelope are the main target for humoral immune response (Sok and Burton 2018), variation in Env can occur due to selection pressure that render them better adapted to one compartment versus another.

CNS compartmentalisation may lead to establishment of a long-lived viral reservoir in the CNS that may have different genotypic and phenotypic properties to the viruses in the periphery, which has implications for HIV-1 treatment, vaccine and cure strategies. Persistence of drug resistant HIV-1 in latent reservoirs is not only a challenge for HIV-1 eradication but more immediately for deciding on subsequent regimens following virological failure. An understanding of the biology of CNS compartmentalisation may yield insights into how patients should be managed.

SARS-CoV-2 immune escape can be driven by suboptimal B and T cell effector function, where antibodies have little support from cytotoxic T cells responses. By understanding the within-host evolution in response to selection pressures in the immunocompromised host we can learn how to improve patient care and adapt our pandemic response tools to prevent transmission of such variants.

Immune senescence is a well-characterised (Akbar and Gilroy 2020) and impacts on vaccine response in multiple ways, including reducing the magnitude of antibody response to vaccines (Wilkinson et al. 2017), shortening the persistence of antibody response (Powers and Belshe 1993), limiting the quality of antibodies via reduced affinity and breadth, limiting induction of CD4+/8+ T cell response (Westmeier et al. 2020) and limiting induction of effector memory (Gustafson et al. 2020). Indeed vaccines are impaired/dysregulated with age (Haq and McElhaney 2014). Although the second dose was able to boost neutralising antibodies, it is unclear if the elderly and other immunosuppressed people will be adequately protected in the protracted interval between doses. Suboptimal antibody levels could also provide the optimal niche for selecting escape mutants which may in the long term jeopardies our the vaccine program's efforts. It is also possible that more vaccine resistant variants would lead to poorer outcome from a single vaccine dose (Collier, De Marco, et al. 2021).

Knowledge gained from the work presented in this thesis contributes to the field of within-host evolution of these two pandemic viruses. It explored the mechanisms underlying central nervous system escape of HIV-1 from immune control in subtype C HIV-1. It also identified functional adaptions in the Spike protein of SARS-CoV-2 that drove escape from antibody neutralisation and showed the blue-print for emergence of Variants of Concern. This work also identified the elderly as a risk group where suboptimal vaccine-elicited immune responses could compromise the vaccine program's efforts. These finding have implications for the immediate choice of therapeutic options in patients, public health and vaccination program policy, vaccine development and cure strategies.

REFERENCES

Abbasi, Jennifer. 2020. 'COVID-19 and mRNA Vaccines—First Large Test for a New Approach', *JAMA*, 324: 1125-27.

Adewumi, O. M., E. Dukhovlinova, N. Y. Shehu, S. Zhou, O. D. Council, M. O. Akanbi, B. Taiwo, A. Ogunniyi, K. Robertson, C. Kanyama, M. C. Hosseinipour, and R. Swanstrom. 2020. 'HIV-1 Central Nervous System Compartmentalization and Cytokine Interplay in Non-Subtype B HIV-1 Infections in Nigeria and Malawi', *AIDS Res Hum Retroviruses*, 36: 490-500.

Akbar, Arne N., and Derek W. Gilroy. 2020. 'Aging immunity may exacerbate COVID-19', *Science*, 369: 256-57.

Alexander, D. J., and I. H. Brown. 2009. 'History of highly pathogenic avian influenza', *Rev Sci Tech*, 28: 19-38.

Andersen, K. G., A. Rambaut, W. I. Lipkin, E. C. Holmes, and R. F. Garry. 2020. 'The proximal origin of SARS-CoV-2', *Nat Med*, 26: 450-52.

Anderson, Evan J., Nadine G. Rouphael, Alicia T. Widge, Lisa A. Jackson, Paul C. Roberts, Mamodikoe Makhene, James D. Chappell, Mark R. Denison, Laura J. Stevens, Andrea J. Pruijssers, Adrian B. McDermott, Britta Flach, Bob C. Lin, Nicole A. Doria-Rose, Sijy O'Dell, Stephen D. Schmidt, Kizzmekia S. Corbett, Phillip A. Swanson, Marcelino Padilla, Kathy M. Neuzil, Hamilton Bennett, Brett Leav, Mat Makowski, Jim Albert, Kaitlyn Cross, Venkata Viswanadh Edara, Katharine Floyd, Mehul S. Suthar, David R. Martinez, Ralph Baric, Wendy Buchanan, Catherine J. Luke, Varun K. Phadke, Christina A. Rostad, Julie E. Ledgerwood, Barney S. Graham, and John H. Beigel. 2020. 'Safety and Immunogenicity of SARS-CoV-2 mRNA-1273 Vaccine in Older Adults', *New England Journal of Medicine*, 383: 2427-38.

Araujo, L. A., and S. E. Almeida. 2013. 'HIV-1 diversity in the envelope glycoproteins: implications for viral entry inhibition', *Viruses*, 5: 595-604.

Arrildt, K. T., C. C. LaBranche, S. B. Joseph, E. N. Dukhovlinova, W. D. Graham, L. H. Ping, G. Schnell, C. B. Sturdevant, L. P. Kincer, M. Mallewa, R. S. Heyderman, A. V. Rie, M. S. Cohen, S. Spudich, R. W. Price, D. C. Montefiori, and R. Swanstrom. 2015. 'Phenotypic Correlates of HIV-1 Macrophage Tropism', *J Virol*, 89: 11294-311.

Avanzato, Victoria A, M Jeremiah Matson, Stephanie N Seifert, Rhys Pryce, Brandi N Williamson, Sarah L Anzick, Kent Barbian, Seth D Judson, Elizabeth R Fischer, and Craig Martens. 2020. 'Case Study: Prolonged infectious SARS-CoV-2 shedding from an asymptomatic immunocompromised cancer patient', *Cell*.

Baden, L. R., H. M. El Sahly, B. Essink, K. Kotloff, S. Frey, R. Novak, D. Diemert, S. A. Spector, N. Rouphael, C. B. Creech, J. McGettigan, S. Kehtan, N. Segall, J. Solis, A. Brosz, C. Fierro, H. Schwartz, K. Neuzil, L. Corey, P. Gilbert, H. Janes, D. Follmann, M. Marovich, J. Mascola, L. Polakowski, J. Ledgerwood, B. S. Graham, H. Bennett, R. Pajon, C. Knightly, B. Leav, W. Deng, H. Zhou, S. Han, M. Ivarsson, J. Miller, T. Zaks, and Cove Study Group. 2020. 'Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine', *N Engl J Med*.

Barre-Sinoussi, F, JC Chermann, F Rey, MT Nugeyre, S Chamaret, J Gruest, C Dauguet, C Axler-Blin, F Vezinet-Brun, C Rouzioux, W Rozenbaum, and L

Montagnier. 1983. 'Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)', *Science*, 220: 868-71.

Beguelin, C., M. Vazquez, M. Bertschi, S. Yerly, D. de Jong, K. Gutbrod, A. Rauch, and A. Cusini. 2016. 'Viral Escape in the Central Nervous System with Multidrug-Resistant Human Immunodeficiency Virus-1', *Open Forum Infect Dis*, 3: ofv210.

Beigel, John H., Kay M. Tomashek, Lori E. Dodd, Aneesh K. Mehta, Barry S. Zingman, Andre C. Kalil, Elizabeth Hohmann, Helen Y. Chu, Annie Luetkemeyer, Susan Kline, Diego Lopez de Castilla, Robert W. Finberg, Kerry Dierberg, Victor Tapson, Lanny Hsieh, Thomas F. Patterson, Roger Paredes, Daniel A. Sweeney, William R. Short, Giota Touloumi, David Chien Lye, Norio Ohmagari, Myoung-don Oh, Guillermo M. Ruiz-Palacios, Thomas Benfield, Gerd Fätkenheuer, Mark G. Kortepeter, Robert L. Atmar, C. Buddy Creech, Jens Lundgren, Abdel G. Babiker, Sarah Pett, James D. Neaton, Timothy H. Burgess, Tyler Bonnett, Michelle Green, Mat Makowski, Anu Osinusi, Seema Nayak, and H. Clifford Lane. 2020. 'Remdesivir for the Treatment of Covid-19 — Final Report', *New England Journal of Medicine*, 383: 1813-26.

Berkhout, B., R. H. Silverman, and K. T. Jeang. 1989. 'Tat trans-activates the human immunodeficiency virus through a nascent RNA target', *Cell*, 59: 273-82.

Bernal, Jamie Lopez, Nick Andrews, Charlotte Gower, Julia Stowe, Chris Robertson, Elise Tessier, Ruth Simmons, Simon Cottrell, Richard Roberts, Mark O'Doherty, Kevin Brown, Claire Cameron, Diane Stockton, Jim McMenamin, and Mary Ramsay. 2021. 'Early effectiveness of COVID-19 vaccination with BNT162b2 mRNA vaccine and ChAdOx1 adenovirus vector vaccine on symptomatic disease, hospitalisations and mortality in older adults in England', *medRxiv*: 2021.03.01.21252652.

Besson, G. J., C. M. Lalama, R. J. Bosch, R. T. Gandhi, M. A. Bedison, E. Aga, S. A. Riddler, D. K. McMahon, F. Hong, and J. W. Mellors. 2014. 'HIV-1 DNA decay dynamics in blood during more than a decade of suppressive antiretroviral therapy', *Clin Infect Dis*, 59: 1312-21.

Bieniasz, P. D. 2009. 'The cell biology of HIV-1 virion genesis', *Cell Host Microbe*, 5: 550-8.

Bingham, R., N. Ahmed, P. Rangi, M. Johnson, M. Tyrer, and J. Green. 2011. 'HIV encephalitis despite suppressed viraemia: a case of compartmentalized viral escape', *Int J STD AIDS*, 22: 608-9.

Borrow, P., H. Lewicki, X. Wei, M. S. Horwitz, N. Peffer, H. Meyers, J. A. Nelson, J. E. Gairin, B. H. Hahn, M. B. Oldstone, and G. M. Shaw. 1997. 'Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus', *Nat Med*, 3: 205-11.

Bruner, K. M., A. J. Murray, R. A. Pollack, M. G. Soliman, S. B. Laskey, A. A. Capoferri, J. Lai, M. C. Strain, S. M. Lada, R. Hoh, Y. C. Ho, D. D. Richman, S. G. Deeks, J. D. Siliciano, and R. F. Siliciano. 2016. 'Defective proviruses rapidly accumulate during acute HIV-1 infection', *Nat Med*, 22: 1043-9.

Bull, Rowena A., Thiruni N. Adikari, James M. Ferguson, Jillian M. Hammond, Igor Stevanovski, Alicia G. Beukers, Zin Naing, Malinna Yeang, Andrey Verich, Hasindu Gamaarachchi, Ki Wook Kim, Fabio Luciani, Sacha Stelzer-Braid,

John-Sebastian Eden, William D. Rawlinson, Sebastiaan J. van Hal, and Ira W. Deveson. 2020. 'Analytical validity of nanopore sequencing for rapid SARS-CoV-2 genome analysis', *Nature Communications*, 11: 6272.

Burton, D. R., and J. R. Mascola. 2015. 'Antibody responses to envelope glycoproteins in HIV-1 infection', *Nat Immunol*, 16: 571-6.

Canestri, A., F. X. Lescure, S. Jaureguiberry, A. Moulignier, C. Amiel, A. G. Marcelin, G. Peytavin, R. Tubiana, G. Pialoux, and C. Katlama. 2010. 'Discordance between cerebral spinal fluid and plasma HIV replication in patients with neurological symptoms who are receiving suppressive antiretroviral therapy', *Clin Infect Dis*, 50: 773-8.

Care, Department of Health and Social. 2021. "Optimising the COVID-19 vaccination programme for maximum short-term impact." In.

Centers for Disease, Control. 1981. 'Pneumocystis pneumonia--Los Angeles', *MMWR Morb Mortal Wkly Rep*, 30: 250-2.

Chang, C. C., R. Kangethe, S. Omarjee, K. Hiramen, B. Gosnell, K. Sojane, M. S. Moosa, S. R. Lewin, M. A. French, and T. Ndung'u. 2017. 'Relationship of Human Immunodeficiency Virus Viral Load in Cerebrospinal Fluid and Plasma in Patients Co-infected With Cryptococcal Meningitis', *Open Forum Infect Dis*, 4: ofx032.

Chi, Xiangyang, Renhong Yan, Jun Zhang, Guanying Zhang, Yuanyuan Zhang, Meng Hao, Zhe Zhang, Pengfei Fan, Yunzhu Dong, Yilong Yang, Zhengshan Chen, Yingying Guo, Jinlong Zhang, Yaning Li, Xiaohong Song, Yi Chen, Lu Xia, Ling Fu, Lihua Hou, Junjie Xu, Changming Yu, Jianmin Li, Qiang Zhou, and Wei Chen. 2020. 'A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2', *Science*, 369: 650-55.

Choi, B., M. C. Choudhary, J. Regan, J. A. Sparks, R. F. Padera, X. Qiu, I. H. Solomon, H. H. Kuo, J. Boucau, K. Bowman, U. D. Adhikari, M. L. Winkler, A. A. Mueller, T. Y. Hsu, M. Desjardins, L. R. Baden, B. T. Chan, B. D. Walker, M. Lichterfeld, M. Brigl, D. S. Kwon, S. Kanjilal, E. T. Richardson, A. H. Jonsson, G. Alter, A. K. Barczak, W. P. Hanage, X. G. Yu, G. D. Gaiha, M. S. Seaman, M. Cernadas, and J. Z. Li. 2020. 'Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host', *N Engl J Med*, 383: 2291-93.

Christo, P. P., D. B. Greco, A. W. Aleixo, and J. A. Livramento. 2007. 'Factors influencing cerebrospinal fluid and plasma HIV-1 RNA detection rate in patients with and without opportunistic neurological disease during the HAART era', *BMC Infect Dis*, 7: 147.

Cilliers, T., J. Nhlapo, M. Coetzer, D. Orlovic, T. Ketas, W. C. Olson, J. P. Moore, A. Trkola, and L. Morris. 2003. 'The CCR5 and CXCR4 coreceptors are both used by human immunodeficiency virus type 1 primary isolates from subtype C', *J Virol*, 77: 4449-56.

Ciuffi, Angela, Manuel Llano, Eric Poeschla, Christian Hoffmann, Jeremy Leipzig, Paul Shinn, Joseph R. Ecker, and Frederic Bushman. 2005. 'A role for LEDGF/p75 in targeting HIV DNA integration', *Nature Medicine*, 11: 1287-89.

Collier, D. A., L. Haddow, J. Brijkumar, M. S. Moosa, L. Benjamin, and R. K. Gupta. 2018. 'HIV Cerebrospinal Fluid Escape and Neurocognitive Pathology in the Era of Combined Antiretroviral Therapy: What Lies Beneath the Tip of the Iceberg in Sub-Saharan Africa?', *Brain Sci*, 8.

Collier, Dami A., Anna De Marco, Isabella A. T. M. Ferreira, Bo Meng, Rawlings P. Datir, Alexandra C. Walls, Steven A. Kemp, Jessica Bassi, Dora Pinto, Chiara Silacci-Fregni, Siro Bianchi, M. Alejandra Tortorici, John Bowen, Katja Culap, Stefano Jaconi, Elisabetta Cameroni, Gyorgy Snell, Matteo S. Pizzuto, Alessandra Franzetti Pellanda, Christian Garzoni, Agostino Riva, Stephen Baker, Gordon Dougan, Christoph Hess, Nathalie Kingston, Paul J. Lehner, Paul A. Lyons, Nicholas J. Matheson, Willem H. Owehand, Caroline Saunders, Charlotte Summers, James E. D. Thaventhiran, Mark Toshner, Michael P. Weekes, Ashlea Bucke, Jo Calder, Laura Canna, Jason Domingo, Anne Elmer, Stewart Fuller, Julie Harris, Sarah Hewitt, Jane Kennet, Sherly Jose, Jenny Kourampa, Anne Meadows, Criona O'Brien, Jane Price, Cherry Publico, Rebecca Rastall, Carla Ribeiro, Jane Rowlands, Valentina Ruffolo, Hugo Tordesillas, Ben Bullman, Benjamin J. Dunmore, Stuart Fawke, Stefan Gräf, Josh Hodgson, Christopher Huang, Kelvin Hunter, Emma Jones, Ekaterina Legchenko, Cecilia Matara, Jennifer Martin, Federica Mescia, Ciara O'Donnell, Linda Pointon, Nicole Pond, Joy Shih, Rachel Sutcliffe, Tobias Tilly, Carmen Treacy, Zhen Tong, Jennifer Wood, Marta Wylot, Laura Bergamaschi, Ariana Betancourt, Georgie Bower, Chiara Cossetti, Aloka De Sa, Madeline Epping, Richard Grenfell, Andrew Hinch, Oisin Huhn, Sarah Jackson, Isobel Jarvis, Daniel Lewis, Joe Marsden, Francesca Nice, Georgina Okecha, Ommar Omarjee, Marianne Perera, Nathan Richoz, Veronika Romashova, Natalia Savinykh Yarkoni, Rahul Sharma, Luca Stefanucci, Jonathan Stephens, Mateusz Strezlecki, Lori Turner, Eckart M. D. D. De Bie, Katherine Bunclark, Masa Josipovic, Michael Mackay, Sabrina Rossi, Mayurun Selvan, Sarah Spencer, Cissy Yong, Ali Ansaripour, Alice Michael, Lucy Mwaura, Caroline Patterson, Gary Polwarth, Petra Polgarova, Giovanni di Stefano, Codie Fahey, Rachel Michel, Sze-How Bong, Jerome D. Coudert, Elaine Holmes, John Allison, Helen Butcher, Daniela Caputo, Debbie Clapham-Riley, Eleanor Dewhurst, Anita Furlong, Barbara Graves, Jennifer Gray, Tasmin Ivers, Mary Kasanicki, Emma Le Gresley, Rachel Linger, Sarah Meloy, Francesca Muldoon, Nigel Ovington, Sofia Papadia, Isabel Phelan, Hannah Stark, Kathleen E. Stirrups, Paul Townsend, Neil Walker, Jennifer Webster, Anne Elmer, Nathalie Kingston, Barbara Graves, Laura E. McCoy, Kenneth G. C. Smith, John R. Bradley, Nigel Temperton, Lourdes Ceron-Gutierrez, Gabriela Barcenas-Morales, Samuel C. Robson, Nicholas J. Loman, Thomas R. Connor, Tanya Golubchik, Rocio T. Martinez Nunez, Catherine Ludden, Sally Corden, Ian Johnston, David Bonsall, Colin P. Smith, Ali R. Awan, Giselda Bucca, M. Estee Torok, Kordo Saeed, Jacqui A. Prieto, David K. Jackson, William L. Hamilton, Luke B. Snell, Catherine Moore, Ewan M. Harrison, Sonia Goncalves, Derek J. Fairley, Matthew W. Loose, Joanne Watkins, Rich Livett, Samuel Moses, Roberto Amato, Sam Nicholls, Matthew Bull, Darren L. Smith, Jeff Barrett, David M. Aanensen, Martin D. Curran, Surendra Parmar, Dinesh Aggarwal, James G. Shepherd, Matthew D. Parker, Sharon Glaysher, Matthew Bashton, Anthony P. Underwood, Nicole Pacchiarini, Katie F. Loveson, Alessandro M. Carabelli, Kate E. Templeton, Cordelia F. Langford, John Sillitoe, Thushan I. de Silva, Dennis Wang, Dominic Kwiatkowski, Andrew Rambaut, Justin O'Grady, Simon Cottrell,

Matthew T. G. Holden, Emma C. Thomson, Husam Osman, Monique Andersson, Anoop J. Chauhan, Mohammed O. Hassan-Ibrahim, Mara Lawniczak, Alex Alderton, Meera Chand, Chrystala Constantinidou, Meera Unnikrishnan, Alistair C. Darby, Julian A. Hiscox, Steve Paterson, Inigo Martincorena, David L. Robertson, Erik M. Volz, Andrew J. Page, Oliver G. Pybus, Andrew R. Bassett, Cristina V. Ariani, Michael H. Spencer Chapman, Kathy K. Li, Rajiv N. Shah, Natasha G. Jesudason, Yusri Taha, Martin P. McHugh, Rebecca Dewar, Aminu S. Jahun, Claire McMurray, Sarojini Pandey, James P. McKenna, Andrew Nelson, Gregory R. Young, Clare M. McCann, Scott Elliott, Hannah Lowe, Ben Temperton, Sunando Roy, Anna Price, Sara Rey, Matthew Wyles, Citiid-Nihr BioResource Covid-Collaboration The, investigators Principal, Crf, nurses volunteer research, logistics Sample, processing Sample, acquisition data, collection Clinical data, I. C. U. Royal Papworth Hospital, I. C. U. Addenbrooke's Hospital, Cambridge, Trust Peterborough Foundation, Anpc, Medicine Centre for Molecular, Therapeutics Innovative, Nihr BioResource, Covid-Genomics U. K. Consortium The, leadership Funding acquisition, metadata curation project administration samples supervision, sequencing logistics, software analysis, tools analysis, visualization, leadership Funding acquisition, metadata curation project administration samples supervision, sequencing logistics, analysis, software, tools analysis, Leadership, metadata curation project administration samples supervision, sequencing logistics, software analysis, tools analysis, visualization, metadata curation samples Funding acquisition, sequencing logistics, software analysis, tools analysis, visualization, leadership Funding acquisition, metadata curation project administration supervision, samples, logistics, leadership Funding acquisition, metadata curation samples supervision, logistics, sequencing, analysis, leadership Funding acquisition, project administration samples supervision, logistics, sequencing, analysis, leadership Funding acquisition, sequencing supervision, software analysis, tools analysis, visualization, samples Funding acquisition, sequencing logistics, software analysis, tools analysis, visualization, Leadership, metadata curation project administration samples supervision, logistics, sequencing, analysis, Leadership, metadata curation project administration samples supervision, logistics, visualization, Leadership, metadata curation project administration sequencing supervision, analysis, software, tools analysis, project administration samples Metadata curation, sequencing logistics, analysis, software, tools analysis, project administration samples Metadata curation, sequencing logistics, analysis, visualization, leadership Funding acquisition, metadata curation supervision, samples, logistics, leadership Funding acquisition, project administration supervision, samples, logistics, Leadership, metadata curation project administration supervision, samples, logistics, Leadership, metadata curation samples supervision, logistics, sequencing, analysis, Leadership, metadata curation samples supervision, logistics, software, tools analysis, Leadership, metadata curation samples supervision, logistics, visualization, Leadership, metadata curation sequencing supervision, analysis, software, tools analysis, Leadership, project administration

samples supervision, logistics, sequencing, analysis, Leadership, sequencing supervision, software analysis, tools analysis, visualization, project administration samples Metadata curation, logistics, sequencing, analysis, project administration sequencing Metadata curation, analysis, software, tools analysis, samples Metadata curation, sequencing logistics, analysis, visualization, sequencing Metadata curation, software analysis, tools analysis, visualization, sequencing Project administration, software analysis, tools analysis, visualization, leadership Funding acquisition, supervision, curation metadata, leadership Funding acquisition, supervision, administration project, leadership Funding acquisition, supervision, sequencing, analysis, Leadership, metadata curation supervision, sequencing, analysis, Leadership, project administration supervision, samples, logistics, Leadership, project administration supervision, sequencing, analysis, Leadership, samples supervision, logistics, sequencing, analysis, Leadership, sequencing supervision, analysis, software, tools analysis, Leadership, sequencing supervision, analysis, visualization, project administration Metadata curation, samples, logistics, project administration Metadata curation, sequencing, analysis, samples Metadata curation, logistics, sequencing, analysis, samples Metadata curation, logistics, visualization, sequencing Metadata curation, analysis, software, and tools analysis. 2021. 'Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies', *Nature*.

Collier, Dami A., Isabella A.T.M. Ferreira, Rawlings Datir, Prasanti Kotagiri, Eleanor Lim, Bo Meng, Laura Bergamaschi, Anne Elmer, Nathalie Kingston, Barbara Graves, Emma Le Gresley Laura McCoy, Kenneth GC Smith, John R. Bradley, Paul A. Lyons, Paulina Cortes-Acevedo, Lourdes Ceron-Gutierrez, Gabriela Barcenas-Morales, Eoin McKinney, Rainer Doffinger, Mark Wills, and Ravindra K. Gupta. 2021. 'Age-related heterogeneity in immune responses to SARS-CoV-2 vaccine BNT162b2', *medRxiv*: 2021.02.03.21251054.

Corman, V. M., D. Muth, D. Niemeyer, and C. Drosten. 2018. 'Hosts and Sources of Endemic Human Coronaviruses', *Adv Virus Res*, 100: 163-88.

D'Arc, M., A. Ayouba, A. Esteban, G. H. Learn, V. Boué, F. Liegeois, L. Etienne, N. Tagg, F. H. Leendertz, C. Boesch, N. F. Madinda, M. M. Robbins, M. Gray, A. Cournil, M. Ooms, M. Letko, V. A. Simon, P. M. Sharp, B. H. Hahn, E. Delaporte, E. Mpoudi Ngole, and M. Peeters. 2015. 'Origin of the HIV-1 group O epidemic in western lowland gorillas', *Proc Natl Acad Sci U S A*, 112: E1343-52.

Das, K., S. E. Martinez, J. D. Bauman, and E. Arnold. 2012. 'HIV-1 reverse transcriptase complex with DNA and nevirapine reveals non-nucleoside inhibition mechanism', *Nat Struct Mol Biol*, 19: 253-9.

Davies, Nicholas G., Sam Abbott, Rosanna C. Barnard, Christopher I. Jarvis, Adam J. Kucharski, James D. Munday, Carl A. B. Pearson, Timothy W. Russell, Damien C. Tully, Alex D. Washburne, Tom Wenseleers, Amy Gimma, William Waites, Kerry L. M. Wong, Kevin van Zandvoort, Justin D. Silverman, Karla Diaz-Ordaz, Ruth Keogh, Rosalind M. Eggo, Sebastian Funk, Mark Jit, Katherine E. Atkins, and W. John Edmunds. 2021. 'Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England', *Science*, 372: eabg3055.

Davies, Nicholas G., Christopher I. Jarvis, Kevin van Zandvoort, Samuel Clifford, Fiona Yueqian Sun, Sebastian Funk, Graham Medley, Yalda Jafari, Sophie R. Meakin, Rachel Lowe, Matthew Quaife, Naomi R. Waterlow, Rosalind M. Eggo, Jiayao Lei, Mihaly Koltai, Fabienne Krauer, Damien C. Tully, James D. Munday, Alicia Showering, Anna M. Foss, Kiesha Prem, Stefan Flasche, Adam J. Kucharski, Sam Abbott, Billy J. Quilty, Thibaut Jombart, Alicia Rosello, Gwenan M. Knight, Mark Jit, Yang Liu, Jack Williams, Joel Hellewell, Kathleen O'Reilly, Yung-Wai Desmond Chan, Timothy W. Russell, Simon R. Procter, Akira Endo, Emily S. Nightingale, Nikos I. Bosse, C. Julian Villabona-Arenas, Frank G. Sandmann, Amy Gimma, Kaja Abbas, William Waites, Katherine E. Atkins, Rosanna C. Barnard, Petra Klepac, Hamish P. Gibbs, Carl A. B. Pearson, Oliver Brady, W. John Edmunds, Nicholas P. Jewell, Karla Diaz-Ordaz, Ruth H. Keogh, and Cmmid Covid- Working Group. 2021. 'Increased mortality in community-tested cases of SARS-CoV-2 lineage B.1.1.7', *Nature*.

de Almeida, Sérgio Monteiro, Indianara Rotta, Ana Paula de Pereira, Bin Tang, Anya Umlauf, Cléa Elisa Lopes Ribeiro, Scott Letendre, Ronald J. Ellis, and H. I. V. Neurobehavioral Research Center Group. 2020. 'Cerebrospinal fluid pleocytosis as a predictive factor for CSF and plasma HIV RNA discordance and escape', *Journal of NeuroVirology*, 26: 241-51.

Delobel, P., K. Sandres-Saune, M. Cazabat, F. E. L'Faqihi, C. Aquilina, M. Obadia, C. Pasquier, B. Marchou, P. Massip, and J. Izopet. 2005. 'Persistence of distinct HIV-1 populations in blood monocytes and naive and memory CD4 T cells during prolonged suppressive HAART', *AIDS*, 19: 1739-50.

Deng, K., M. Pertea, A. Rongvaux, L. Wang, C. M. Durand, G. Ghiaur, J. Lai, H. L. McHugh, H. Hao, H. Zhang, J. B. Margolick, C. Gurer, A. J. Murphy, D. M. Valenzuela, G. D. Yancopoulos, S. G. Deeks, T. Strowig, P. Kumar, J. D. Siliciano, S. L. Salzberg, R. A. Flavell, L. Shan, and R. F. Siliciano. 2015. 'Broad CTL response is required to clear latent HIV-1 due to dominance of escape mutations', *Nature*, 517: 381-5.

Docherty, A. B., E. M. Harrison, C. A. Green, H. E. Hardwick, R. Pius, L. Norman, K. A. Holden, J. M. Read, F. Dondelinger, G. Carson, L. Merson, J. Lee, D. Plotkin, L. Sigfrid, S. Halpin, C. Jackson, C. Gamble, P. W. Horby, J. S. Nguyen-Van-Tam, A. Ho, C. D. Russell, J. Dunning, P. J. Openshaw, J. K. Baillie, M. G. Semple, and Isaric C. investigators. 2020. 'Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study', *BMJ*, 369: m1985.

Dong, E., H. Du, and L. Gardner. 2020. 'An interactive web-based dashboard to track COVID-19 in real time', *Lancet Infect Dis*, 20: 533-34.

Dragic, T., V. Litwin, G. P. Allaway, S. R. Martin, Y. Huang, K. A. Nagashima, C. Cayanan, P. J. Maddon, R. A. Koup, J. P. Moore, and W. A. Paxton. 1996. 'HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5', *Nature*, 381: 667-73.

DW. 2021. "COVID vaccines: Will Germany follow the UK's example?" In.

Eden, A., D. Fuchs, L. Hagberg, S. Nilsson, S. Spudich, B. Svennerholm, R. W. Price, and M. Gisslen. 2010. 'HIV-1 viral escape in cerebrospinal fluid of subjects on suppressive antiretroviral treatment', *J Infect Dis*, 202: 1819-25.

Eden, A., S. Nilsson, L. Hagberg, D. Fuchs, H. Zetterberg, B. Svennerholm, and M. Gisslen. 2016. 'Asymptomatic Cerebrospinal Fluid HIV-1 Viral Blips and Viral Escape During Antiretroviral Therapy: A Longitudinal Study', *J Infect Dis*, 214: 1822-25.

Edler, Daniel, Johannes Klein, Alexandre Antonelli, and Daniele Silvestro. 2019. 'raxmlGUI 2.0 beta: a graphical interface and toolkit for phylogenetic analyses using RAxML', *bioRxiv*: 800912.

Ewer, Katie J., Jordan R. Barrett, Sandra Belij-Rammerstorfer, Hannah Sharpe, Rebecca Makinson, Richard Morter, Amy Flaxman, Daniel Wright, Duncan Bellamy, Mustapha Bittaye, Christina Dold, Nicholas M. Provine, Jeremy Aboagye, Jamie Fowler, Sarah E. Silk, Jennifer Alderson, Parvinder K. Aley, Brian Angus, Eleanor Berrie, Sagida Bibi, Paola Cicconi, Elizabeth A. Clutterbuck, Irina Chelysheva, Pedro M. Folegatti, Michelle Fuskova, Catherine M. Green, Daniel Jenkin, Simon Kerridge, Alison Lawrie, Angela M. Minassian, Maria Moore, Yama Mujadidi, Emma Plested, Ian Poulton, Maheshi N. Ramasamy, Hannah Robinson, Rinn Song, Matthew D. Snape, Richard Tarrant, Merryn Voysey, Marion E. E. Watson, Alexander D. Douglas, Adrian V. S. Hill, Sarah C. Gilbert, Andrew J. Pollard, Teresa Lambe, Aabidah Ali, Elizabeth Allen, Megan Baker, Eleanor Barnes, Nicola Borthwick, Amy Boyd, Charlie Brown-O'Sullivan, Joshua Burgoyne, Nicholas Byard, Ingrid Cabrera Puig, Federica Cappuccini, Jee-Sun Cho, Paola Cicconi, Elizabeth Clark, Wendy E. M. Crocker, Mehreen S. Datoo, Hannah Davies, Susanna Jane Dunachie, Nick J. Edwards, Sean C. Elias, Julie Furze, Ciaran Gilbride, Stephanie A. Harris, Susanne H. C. Hodgson, Mimi M. Hou, Susan Jackson, Kathryn Jones, Reshma Kailath, Lloyd King, Colin W. Larkworthy, Yuanyuan Li, Amelia M. Lias, Aline Linder, Samuel Lipworth, Raquel Lopez Ramon, Meera Madhavan, Emma Marlow, Julia L. Marshall, Alexander J. Mentzer, Hazel Morrison, Andrés Noé, Dimitra Pipini, David Pulido-Gomez, Fernando Ramos Lopez, Adam John Ritchie, Indra Rudiansyah, Helen Sanders, Adam Shea, Sarah Silk, Alexandra J. Spencer, Rachel Tanner, Yrene Themistocleous, Merin Thomas, Nguyen Tran, Adam Truby, Cheryl Turner, Nicola Turner, Marta Ulaszewska, Andrew T. Worth, Lucy Kingham-Page, Marco Polo Peralta Alvarez, Rachel Anslow, Louise Bates, Kirsten Beadon, Rebecca Beckley, Amy Beveridge, Else Margreet Bijker, Luke Blackwell, Jamie Burbage, Susana Camara, Melanie Carr, Rachel Colin-Jones, Rachel Cooper, Christina J. Cunningham, Tesfaye Demissie, Claudio Di Maso, Naomi Douglas, Rachael Drake-Brockman, Ruth Elizabeth Drury, Katherine R. W. Emary, Sally Felle, Shuo Feng, Karen J. Ford, Emma Francis, Lara Gracie, Joseph Hamlyn, Brama Hanumunthadu, Daisy Harrison, Thomas C. Hart, Sophia Hawkins, Jennifer Hill, Elizabeth Howe, Nicola Howell, Elizabeth Jones, Jade Keen, Sarah Kelly, David Kerr, Liaquat Khan, Jasmin Kinch, Stanislava Koleva, Emily A. Lees, Alice Lelliott, Xinxue Liu, Spyridoula Marinou, Joanne McEwan, Ella Morey, Gertraud Morshead, Jilly Muller, Claire Munro, Sarah Murphy, Philomena Mweu, Elizabeth Nuthall, Katie O'Brien, Daniel O'Connor, Peter John O'Reilly, Blanché Oguti, Piper Osborne, Nelly Owino, Kaye Parker, Katja Pfafferott, Samuel Provstgaard-Morys, Helen Ratcliffe, Thomas Rawlinson, Sarah Rhead, Hannah Roberts, Katherine

Sanders, Laura Silva-Reyes, Catherine C. Smith, David J. Smith, Anna Szigeti, Tonia M. Thomas, Amber Thompson, Susan Tonks, Rachel Varughes, Iason Vichos, Laura Walker, Caroline White, Rachel White, Xin Li Yao, Christopher P. Conlon, John Frater, Liliana Cifuentes, Ioana Baleanu, Emma Bolam, Elena Boland, Tanja Brenner, Brad E. Damratoski, Chandra Datta, Omar El Muhanna, Richard Fisher, Pablo Galian-Rubio, Gina Hodges, Frederic Jackson, Shuchang Liu, Lisa Loew, Roisin Morgans, Susan Jane Morris, Vicki Olchawski, Catarina Oliveria, Helena Parracho, Emilia Reyes Pabon, Abdessamad Tahiri-Alaoui, Keja Taylor, Paul Williams, Dalila Zizi, Edward H. Arbe-Barnes, Philip Baker, Alexander Batten, Charlotte Downing, Jonathan Drake, Marcus Rex English, John Aaron Henry, Poppy Iveson, Annabel Killen, Thomas B. King, Jessica P. J. Larwood, Gary Mallett, Kushal Mansatta, Neginsadat Mirtorabi, Maia Patrick-Smith, James Perring, Kajal Radia, Sophie Roche, Ella Schofield, Rebecca te Water Naude, James Towner, Natalie Baker, Kevin R. Bewley, Emily Brunt, Karen R. Buttigieg, Sue Charlton, Naomi S. Coombes, Michael J. Elmore, Kerry Godwin, Bassam Hallis, Daniel Knott, Lorna McInroy, Imam Shaik, Kelly Thomas, Julia A. Tree, Caitlin L. Blundell, Michelangelo Cao, Dearbhla Kelly, Donal T. Skelly, Andreas Themistocleous, Tao Dong, Samantha Field, Elizabeth Hamilton, Elizabeth Kelly, Paul Klenerman, Julian C. Knight, Yolanda Lie, Christos Petropoulos, Cynthia Sedik, Terri Wrin, Gretchen Meddaugh, Yanchun Peng, Gavin Screaton, Elizabeth Stafford, and Covid Vaccine Trial Group the Oxford. 2021. 'T cell and antibody responses induced by a single dose of ChAdOx1 nCoV-19 (AZD1222) vaccine in a phase 1/2 clinical trial', *Nature Medicine*, 27: 270-78.

Feder, A. F., S. Y. Rhee, S. P. Holmes, R. W. Shafer, D. A. Petrov, and P. S. Pennings. 2016a. 'More effective drugs lead to harder selective sweeps in the evolution of drug resistance in HIV-1', *Elife*, 5.

Feder, Alison F., Soo-Yon Rhee, Susan P. Holmes, Robert W. Shafer, Dmitri A. Petrov, and Pleuni S. Pennings. 2016b. 'More effective drugs lead to harder selective sweeps in the evolution of drug resistance in HIV-1', *eLife*, 5: e10670.

Filipowicz, A. R., C. M. McGary, G. E. Holder, A. A. Lindgren, E. M. Johnson, C. Sugimoto, M. J. Kuroda, and W. K. Kim. 2016. 'Proliferation of Perivascular Macrophages Contributes to the Development of Encephalitic Lesions in HIV-Infected Humans and in SIV-Infected Macaques', *Sci Rep*, 6: 32900.

Fourati, S., S. Lambert-Niclot, C. Soulie, M. Wirden, I. Malet, M. A. Valantin, R. Tubiana, A. Simon, C. Katlama, G. Carcelain, V. Calvez, and A. G. Marcelin. 2014. 'Differential impact of APOBEC3-driven mutagenesis on HIV evolution in diverse anatomical compartments', *AIDS*, 28: 487-91.

Frampton, D., T. Rampling, A. Cross, H. Bailey, J. Heaney, M. Byott, R. Scott, R. Sconza, J. Price, M. Margaritis, M. Bergstrom, M. J. Spyer, P. B. Miralhes, P. Grant, S. Kirk, C. Valerio, Z. Mangera, T. Prabhahar, J. Moreno-Cuesta, N. Arulkumaran, M. Singer, G. Y. Shin, E. Sanchez, S. M. Paraskevopoulou, D. Pillay, R. A. McKendry, M. Mirfenderesky, C. F. Houlihan, and E. Nastouli. 2021. '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*.

Fun, Axel, Annemarie M. J. Wensing, Jens Verheyen, and Monique Nijhuis. 2012. 'Human Immunodeficiency Virus gag and protease: partners in resistance', *Retrovirology*, 9: 63.

Gall, A., B. Ferns, C. Morris, S. Watson, M. Cotten, M. Robinson, N. Berry, D. Pillay, and P. Kellam. 2012. 'Universal amplification, next-generation sequencing, and assembly of HIV-1 genomes', *J Clin Microbiol*, 50: 3838-44.

Ganor, Y., F. Real, A. Sennepin, C. A. Dutertre, L. Prevedel, L. Xu, D. Tudor, B. Charmeteau, A. Couedel-Courteille, S. Marion, A. R. Zenak, J. P. Jourdain, Z. Zhou, A. Schmitt, C. Capron, E. A. Eugenin, R. Cheynier, M. Revol, S. Cristofari, A. Hosmalin, and M. Bomsel. 2019. 'HIV-1 reservoirs in urethral macrophages of patients under suppressive antiretroviral therapy', *Nat Microbiol*, 4: 633-44.

Gao, X., A. Bashirova, A. K. Iversen, J. Phair, J. J. Goedert, S. Buchbinder, K. Hoots, D. Vlahov, M. Altfeld, S. J. O'Brien, and M. Carrington. 2005. 'AIDS restriction HLA allotypes target distinct intervals of HIV-1 pathogenesis', *Nat Med*, 11: 1290-2.

Garrett, M. E., H. L. Itell, K. H. D. Crawford, R. Basom, J. D. Bloom, and J. Overbaugh. 2020. 'Phage-DMS: A Comprehensive Method for Fine Mapping of Antibody Epitopes', *iScience*, 23: 101622.

Gelman, B. B., J. G. Lisinicchia, S. Morgello, E. Masliah, D. Commins, C. L. Achim, H. S. Fox, D. L. Kolson, I. Grant, E. Singer, C. T. Yiannoutsos, S. Sherman, G. Gensler, D. J. Moore, T. Chen, and V. M. Soukup. 2013. 'Neurovirological correlation with HIV-associated neurocognitive disorders and encephalitis in a HAART-era cohort', *J Acquir Immune Defic Syndr*, 62: 487-95.

Gordon, Anthony C., Paul R. Mouncey, Farah Al-Beidh, Kathryn M. Rowan, Alistair D. Nichol, Yaseen M. Arabi, Djillali Annane, Abi Beane, Wilma van Bentum-Puijk, Lindsay R. Berry, Zahra Bhimani, Marc J.M. Bonten, Charlotte A. Bradbury, Frank M. Brunkhorst, Adrian Buzgau, Allen C. Cheng, Michelle A. Detry, Eamon J. Duffy, Lise J. Estcourt, Mark Fitzgerald, Herman Goossens, Rashan Haniffa, Alisa M. Higgins, Thomas E. Hills, Christopher M. Horvat, Francois Lamontagne, Patrick R. Lawler, Helen L. Leavis, Kelsey M. Linstrum, Edward Litton, Elizabeth Lorenzi, John C. Marshall, Florian B. Mayr, Danny McAuley, Anna McGlothlin, Shay P McGuinness, Bryan J. McVerry, Stephanie K. Montgomery, Susan C. Morpeth, Srinivas Murthy, Katrina Orr, Rachael L. Parke, Jane C. Parker, Asad E. Patanwala, Ville Pettilä, Emma Rademaker, Marlene S. Santos, Christina T. Saunders, Christopher W. Seymour, Manu Shankar-Hari, Wendy I. Sligl, Alexis F. Turgeon, Anne M. Turner, Frank L. van de Veerdonk, Ryan Zarychanski, Cameron Green, Roger J. Lewis, Derek C. Angus, Colin J. McArthur, Scott Berry, Steve A. Webb, and Lennie P.G. Derde. 2021. 'Interleukin-6 Receptor Antagonists in Critically Ill Patients with Covid-19 – Preliminary report', *medRxiv*: 2021.01.07.21249390.

Gorry, P. R., G. Bristol, J. A. Zack, K. Ritola, R. Swanstrom, C. J. Birch, J. E. Bell, N. Bannert, K. Crawford, H. Wang, D. Schols, E. De Clercq, K. Kunstman, S. M. Wolinsky, and D. Gabuzda. 2001. 'Macrophage tropism of human immunodeficiency virus type 1 isolates from brain and lymphoid tissues

predicts neurotropism independent of coreceptor specificity', *J Virol*, 75: 10073-89.

Gorry, P. R., J. Taylor, G. H. Holm, A. Mehle, T. Morgan, M. Cayabyab, M. Farzan, H. Wang, J. E. Bell, K. Kunstman, J. P. Moore, S. M. Wolinsky, and D. Gabuzda. 2002. 'Increased CCR5 affinity and reduced CCR5/CD4 dependence of a neurovirulent primary human immunodeficiency virus type 1 isolate', *J Virol*, 76: 6277-92.

Gottlieb, Robert L., Ajay Nirula, Peter Chen, Joseph Boscia, Barry Heller, Jason Morris, Gregory Huhn, Jose Cardona, Bharat Mocherla, Valentina Stosor, Imad Shawa, Princy Kumar, Andrew C. Adams, Jacob Van Naarden, Kenneth L. Custer, Michael Durante, Gerard Oakley, Andrew E. Schade, Timothy R. Holzer, Philip J. Ebert, Richard E. Higgs, Nicole L. Kallewaard, Janelle Sabo, Dipak R. Patel, Paul Klekotka, Lei Shen, and Daniel M. Skovronsky. 2021. '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*, 325: 632-44.

Gray, F., F. X. Lescure, H. Adle-Biassette, M. Polivka, S. Gallien, G. Pialoux, and A. Moulignier. 2013. 'Encephalitis with infiltration by CD8+ lymphocytes in HIV patients receiving combination antiretroviral treatment', *Brain Pathol*, 23: 525-33.

Greaney, A. J., T. N. Starr, P. Gilchuk, S. J. Zost, E. Binshtein, A. N. Loes, S. K. Hilton, J. Huddleston, R. Eguia, K. H. D. Crawford, A. S. Dingens, R. S. Nargi, R. E. Sutton, N. Suryadevara, P. W. Rothlauf, Z. Liu, S. P. J. Whelan, R. H. Carnahan, J. E. Crowe, Jr., and J. D. Bloom. 2021. 'Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition', *Cell Host Microbe*, 29: 44-57 e9.

Gupta, Ravindra K., Sultan Abdul-Jawad, Laura E. McCoy, Hoi Ping Mok, Dimitra Peppa, Maria Salgado, Javier Martinez-Picado, Monique Nijhuis, Annemarie M. J. Wensing, Helen Lee, Paul Grant, Eleni Nastouli, Jonathan Lambert, Matthew Pace, Fanny Salasc, Christopher Monit, Andrew J. Innes, Luke Muir, Laura Waters, John Frater, Andrew M. L. Lever, Simon G. Edwards, Ian H. Gabriel, and Eduardo Olavarria. 2019. 'HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem-cell transplantation', *Nature*, 568: 244-48.

Gupta-Wright, A., K. Fielding, J. J. van Oosterhout, M. Alufandika, D. J. Grint, E. Chimbayo, J. Heaney, M. Byott, E. Nastouli, H. C. Mwandumba, E. L. Corbett, and R. K. Gupta. 2020. 'Virological failure, HIV-1 drug resistance, and early mortality in adults admitted to hospital in Malawi: an observational cohort study', *Lancet HIV*, 7: e620-e28.

Gustafson, C. E., C. Kim, C. M. Weyand, and J. J. Goronzy. 2020. 'Influence of immune aging on vaccine responses', *J Allergy Clin Immunol*, 145: 1309-21.

Haaland, R. E., P. A. Hawkins, J. Salazar-Gonzalez, A. Johnson, A. Tichacek, E. Karita, O. Manigart, J. Mulenga, B. F. Keele, G. M. Shaw, B. H. Hahn, S. A. Allen, C. A. Derdeyn, and E. Hunter. 2009. 'Inflammatory genital infections mitigate a severe genetic bottleneck in heterosexual transmission of subtype A and C HIV-1', *PLoS Pathog*, 5: e1000274.

Haddad, D. N., C. Birch, T. Middleton, D. E. Dwyer, A. L. Cunningham, and N. K. Saksena. 2000. 'Evidence for late stage compartmentalization of HIV-1 resistance mutations between lymph node and peripheral blood mononuclear cells', *AIDS*, 14: 2273-81.

Haq, K., and J. E. McElhaney. 2014. 'Immunosenescence: Influenza vaccination and the elderly', *Curr Opin Immunol*, 29: 38-42.

Heaton, R. K., D. B. Clifford, D. R. Franklin, Jr., S. P. Woods, C. Ake, F. Vaida, R. J. Ellis, S. L. Letendre, T. D. Marcotte, J. H. Atkinson, M. Rivera-Mindt, O. R. Vigil, M. J. Taylor, A. C. Collier, C. M. Marra, B. B. Gelman, J. C. McArthur, S. Morgello, D. M. Simpson, J. A. McCutchan, I. Abramson, A. Gamst, C. Fennema-Notestine, T. L. Jernigan, J. Wong, I. Grant, and Charter Group. 2010. 'HIV-associated neurocognitive disorders persist in the era of potent antiretroviral therapy: CHARTER Study', *Neurology*, 75: 2087-96.

Heaton, R. K., D. R. Franklin, R. J. Ellis, J. A. McCutchan, S. L. Letendre, S. Leblanc, S. H. Corkran, N. A. Duarte, D. B. Clifford, S. P. Woods, A. C. Collier, C. M. Marra, S. Morgello, M. R. Mindt, M. J. Taylor, T. D. Marcotte, J. H. Atkinson, T. Wolfson, B. B. Gelman, J. C. McArthur, D. M. Simpson, I. Abramson, A. Gamst, C. Fennema-Notestine, T. L. Jernigan, J. Wong, I. Grant, Charter Group, and Hnrc Group. 2011. 'HIV-associated neurocognitive disorders before and during the era of combination antiretroviral therapy: differences in rates, nature, and predictors', *J Neurovirol*, 17: 3-16.

Hogan, C. A., J. Iles, E. H. Frost, G. Giroux, O. Cassar, A. Gessain, M. J. Dion, V. Ilunga, A. Rambaut, AÉ Yengo-Ki-Ngimbi, F. Behets, O. G. Pybus, and J. Pépin. 2016. 'Epidemic History and Iatrogenic Transmission of Blood-borne Viruses in Mid-20th Century Kinshasa', *J Infect Dis*, 214: 353-60.

Horby, Peter W, Guilherme Pessoa-Amorim, Leon Peto, Christopher E Brightling, Rahuldeb Sarkar, Koshy Thomas, Vandana Jeebun, Abdul Ashish, Redmond Tully, David Chadwick, Muhammad Sharafat, Richard Stewart, Banu Rudran, J Kenneth Baillie, Maya H Buch, Lucy C Chappell, Jeremy N Day, Saul N Furst, Thomas Jaki, Katie Jeffery, Edmund Juszczak, Wei Shen Lim, Alan Montgomery, Andrew Mumford, Kathryn Rowan, Guy Thwaites, Marion Mafham, Richard Haynes, and Martin J Landray. 2021. 'Tocilizumab in patients admitted to hospital with COVID-19 (RECOVERY): preliminary results of a randomised, controlled, open-label, platform trial', *medRxiv*: 2021.02.11.21249258.

Huang, W., S. H. Eshleman, J. Toma, S. Fransen, E. Stawiski, E. E. Paxinos, J. M. Whitcomb, A. M. Young, D. Donnell, F. Mmiro, P. Musoke, L. A. Guay, J. B. Jackson, N. T. Parkin, and C. J. Petropoulos. 2007. 'Coreceptor tropism in human immunodeficiency virus type 1 subtype D: high prevalence of CXCR4 tropism and heterogeneous composition of viral populations', *J Virol*, 81: 7885-93.

Huang, W., J. Toma, S. Fransen, E. Stawiski, J. D. Reeves, J. M. Whitcomb, N. Parkin, and C. J. Petropoulos. 2008. 'Coreceptor tropism can be influenced by amino acid substitutions in the gp41 transmembrane subunit of human immunodeficiency virus type 1 envelope protein', *J Virol*, 82: 5584-93.

Hütter, G., D. Nowak, M. Mossner, S. Ganepola, A. Müssig, K. Allers, T. Schneider, J. Hofmann, C. Kücherer, O. Blau, I. W. Blau, W. K. Hofmann, and E. Thiel.

2009. 'Long-term control of HIV by CCR5 Delta32/Delta32 stem-cell transplantation', *N Engl J Med*, 360: 692-8.

Jackson, Lisa A., Evan J. Anderson, Nadine G. Rouphael, Paul C. Roberts, Mamodikoe Makhene, Rhea N. Coler, Michele P. McCullough, James D. Chappell, Mark R. Denison, Laura J. Stevens, Andrea J. Pruijssers, Adrian McDermott, Britta Flach, Nicole A. Doria-Rose, Kizzmekia S. Corbett, Kaitlyn M. Morabito, Sijy O'Dell, Stephen D. Schmidt, Phillip A. Swanson, Marcelino Padilla, John R. Mascola, Kathleen M. Neuzil, Hamilton Bennett, Wellington Sun, Etza Peters, Mat Makowski, Jim Albert, Kaitlyn Cross, Wendy Buchanan, Rhonda Pikaart-Tautges, Julie E. Ledgerwood, Barney S. Graham, and John H. Beigel. 2020. 'An mRNA Vaccine against SARS-CoV-2 — Preliminary Report', *New England Journal of Medicine*, 383: 1920-31.

Jacques, D. A., W. A. McEwan, L. Hilditch, A. J. Price, G. J. Towers, and L. C. James. 2016. 'HIV-1 uses dynamic capsid pores to import nucleotides and fuel encapsidated DNA synthesis', *Nature*, 536: 349-53.

Jaimes, Javier A., Jean K. Millet, and Gary R. Whittaker. 2020. 'Proteolytic Cleavage of the SARS-CoV-2 Spike Protein and the Role of the Novel S1/S2 Site', *iScience*, 23: 101212.

Jensen, M. A., M. Coetzer, A. B. van 't Wout, L. Morris, and J. I. Mullins. 2006. 'A reliable phenotype predictor for human immunodeficiency virus type 1 subtype C based on envelope V3 sequences', *J Virol*, 80: 4698-704.

Jones, Kate E., Nikkita G. Patel, Marc A. Levy, Adam Storeygard, Deborah Balk, John L. Gittleman, and Peter Daszak. 2008. 'Global trends in emerging infectious diseases', *Nature*, 451: 990-93.

Joseph, S. B., K. T. Arrildt, C. B. Sturdevant, and R. Swanstrom. 2015. 'HIV-1 target cells in the CNS', *J Neurovirol*, 21: 276-89.

Joseph, S. B., L. P. Kincer, N. M. Bowman, C. Evans, M. J. Vinikoor, C. K. Lippincott, M. Gisslen, S. Spudich, P. Menezes, K. Robertson, N. Archin, A. Kashuba, J. J. Eron, R. W. Price, and R. Swanstrom. 2019. 'Human Immunodeficiency Virus Type 1 RNA Detected in the Central Nervous System (CNS) After Years of Suppressive Antiretroviral Therapy Can Originate from a Replicating CNS Reservoir or Clonally Expanded Cells', *Clin Infect Dis*, 69: 1345-52.

Joseph, S. B., B. Lee, and R. Swanstrom. 2014. 'Affinofile Assay for Identifying Macrophage-Tropic HIV-1', *Bio Protoc*, 4.

Kadire, Siri R., Robert M. Wachter, and Nicole Lurie. 2021. 'Delayed Second Dose versus Standard Regimen for Covid-19 Vaccination', *New England Journal of Medicine*, 384: e28.

Kalil, Andre C., Thomas F. Patterson, Aneesh K. Mehta, Kay M. Tomashek, Cameron R. Wolfe, Varduhi Ghazaryan, Vincent C. Marconi, Guillermo M. Ruiz-Palacios, Lanny Hsieh, Susan Kline, Victor Tapson, Nicole M. Iovine, Mamta K. Jain, Daniel A. Sweeney, Hana M. El Sahly, Angela R. Branche, Justino Regalado Pineda, David C. Lye, Uriel Sandkovsky, Anne F. Luetkemeyer, Stuart H. Cohen, Robert W. Finberg, Patrick E.H. Jackson, Babafemi Taiwo, Catharine I. Paules, Henry Arguinchona, Nathaniel Erdmann, Neera Ahuja, Maria Frank, Myoung-don Oh, Eu-Suk Kim, Seow Y. Tan, Richard A. Mularski, Henrik Nielsen, Philip O. Ponce, Barbara S. Taylor, LuAnn Larson, Nadine G. Rouphael, Youssef Saklawi, Valeria D. Cantos, Emily R. Ko, John J. Engemann,

Alpesh N. Amin, Miki Watanabe, Joanne Billings, Marie-Carmelle Elie, Richard T. Davey, Timothy H. Burgess, Jennifer Ferreira, Michelle Green, Mat Makowski, Anabela Cardoso, Stephanie de Bono, Tyler Bonnett, Michael Proschan, Gregory A. Deye, Walla Dempsey, Seema U. Nayak, Lori E. Dodd, and John H. Beigel. 2020. 'Baricitinib plus Remdesivir for Hospitalized Adults with Covid-19', *New England Journal of Medicine*, 384: 795-807.

Kalyaanamoorthy, Subha, Bui Quang Minh, Thomas K. F. Wong, Arndt von Haeseler, and Lars S. Jermiin. 2017. 'ModelFinder: fast model selection for accurate phylogenetic estimates', *Nature Methods*, 14: 587-89.

Kaslow, R. A., M. Carrington, R. Apple, L. Park, A. Muñoz, A. J. Saah, J. J. Goedert, C. Winkler, S. J. O'Brien, C. Rinaldo, R. Detels, W. Blattner, J. Phair, H. Erlich, and D. L. Mann. 1996. 'Influence of combinations of human major histocompatibility complex genes on the course of HIV-1 infection', *Nat Med*, 2: 405-11.

Kearney, M., F. Maldarelli, W. Shao, J. B. Margolick, E. S. Daar, J. W. Mellors, V. Rao, J. M. Coffin, and S. Palmer. 2009. 'Human immunodeficiency virus type 1 population genetics and adaptation in newly infected individuals', *J Virol*, 83: 2715-27.

Keele, B. F., E. E. Giorgi, J. F. Salazar-Gonzalez, J. M. Decker, K. T. Pham, M. G. Salazar, C. Sun, T. Grayson, S. Wang, H. Li, X. Wei, C. Jiang, J. L. Kirchherr, F. Gao, J. A. Anderson, L. H. Ping, R. Swanstrom, G. D. Tomaras, W. A. Blattner, P. A. Goepfert, J. M. Kilby, M. S. Saag, E. L. Delwart, M. P. Busch, M. S. Cohen, D. C. Montefiori, B. F. Haynes, B. Gaschen, G. S. Athreya, H. Y. Lee, N. Wood, C. Seoighe, A. S. Perelson, T. Bhattacharya, B. T. Korber, B. H. Hahn, and G. M. Shaw. 2008. 'Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection', *Proc Natl Acad Sci U S A*, 105: 7552-7.

Kemp, S. A., D. A. Collier, R. P. Datir, Iatm Ferreira, S. Gayed, A. Jahun, M. Hosmillo, C. Rees-Spear, P. Mlcochova, I. U. Lumb, D. J. Roberts, A. Chandra, N. Temperton, K. Sharrocks, E. Blane, Y. Modis, K. E. Leigh, J. A. G. Briggs, M. J. van Gils, K. G. C. Smith, J. R. Bradley, C. Smith, R. Doffinger, L. Ceron-Gutierrez, G. Barcenas-Morales, D. D. Pollock, R. A. Goldstein, A. Smielewska, J. P. Skittrall, T. Gouliouris, I. G. Goodfellow, E. Gkrania-Klotsas, C. J. R. Illingworth, L. E. McCoy, and R. K. Gupta. 2021. 'SARS-CoV-2 evolution during treatment of chronic infection', *Nature*, 592: 277-82.

Kim, Ki Wook, Ira W Deveson, Chi Nam I Pang, Malinna Yeang, Zin Naing, Thiruni Adikari, Jillian M Hammond, Igor Stevanovski, Alicia G Beukers, and Andrey Verich. 2020. 'Respiratory viral co-infections among SARS-CoV-2 cases confirmed by virome capture sequencing'.

Ko, Sung Hee, Elham Bayat Mokhtari, Prakriti Mudvari, Sydney Stein, Christopher D. Stringham, Danielle Wagner, Sabrina Ramelli, Marcos J. Ramos-Benitez, Jeffrey R. Strich, Richard T. Davey, Jr., Tongqing Zhou, John Misasi, Peter D. Kwong, Daniel S. Chertow, Nancy J. Sullivan, and Eli A. Boritz. 2021. 'High-throughput, single-copy sequencing reveals SARS-CoV-2 spike variants coincident with mounting humoral immunity during acute COVID-19', *PLOS Pathogens*, 17: e1009431.

Koyanagi, Y., S. Miles, R. T. Mitsuyasu, J. E. Merrill, H. V. Vinters, and I. S. Chen. 1987. 'Dual infection of the central nervous system by AIDS viruses with distinct cellular tropisms', *Science*, 236: 819-22.

Kugathasan, R., D. A. Collier, L. J. Haddow, K. El Bouzidi, S. G. Edwards, J. D. Cartledge, R. F. Miller, and R. K. Gupta. 2017. 'Diffuse White Matter Signal Abnormalities on Magnetic Resonance Imaging Are Associated With Human Immunodeficiency Virus Type 1 Viral Escape in the Central Nervous System Among Patients With Neurological Symptoms', *Clin Infect Dis*, 64: 1059-65.

Kumar, S., G. Stecher, M. Li, C. Knyaz, and K. Tamura. 2018. 'MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms', *Mol Biol Evol*, 35: 1547-49.

Kwon, H., N. Pelletier, C. DeLuca, P. Genin, S. Cisternas, R. Lin, M. A. Wainberg, and J. Hiscott. 1998. 'Inducible expression of IkappaBalpha repressor mutants interferes with NF-kappaB activity and HIV-1 replication in Jurkat T cells', *J Biol Chem*, 273: 7431-40.

Kwong, P. D., R. Wyatt, J. Robinson, R. W. Sweet, J. Sodroski, and W. A. Hendrickson. 1998. 'Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody', *Nature*, 393: 648-59.

Lange, Camille, Maura Manion, Kathy Hullsiek, Brandon Keele, Helene Highbarger, David Meya, David Boulware, and Frank Maldarelli. 2018. 'P-A12 Discordant HIV populations with discordant V3 tropism in CSF and Plasma: Implications for establishing HIV reservoirs', *JAIDS Journal of Acquired Immune Deficiency Syndromes*, 77: 56.

Leitner, T., and J. Albert. 1999. 'The molecular clock of HIV-1 unveiled through analysis of a known transmission history', *Proc Natl Acad Sci U S A*, 96: 10752-7.

Letendre, S. 2011. 'Central nervous system complications in HIV disease: HIV-associated neurocognitive disorder', *Top Antivir Med*, 19: 137-42.

Li, Heng. 2018. 'Minimap2: pairwise alignment for nucleotide sequences', *Bioinformatics*, 34: 3094-100.

Li, Y., D. Y. Lai, H. N. Zhang, H. W. Jiang, X. Tian, M. L. Ma, H. Qi, Q. F. Meng, S. J. Guo, Y. Wu, W. Wang, X. Yang, D. W. Shi, J. B. Dai, T. Ying, J. Zhou, and S. C. Tao. 2020. 'Linear epitopes of SARS-CoV-2 spike protein elicit neutralizing antibodies in COVID-19 patients', *Cell Mol Immunol*, 17: 1095-97.

Libster, Romina, Gonzalo Pérez Marc, Diego Wappner, Silvina Coviello, Alejandra Bianchi, Virginia Braem, Ignacio Esteban, Mauricio T. Caballero, Cristian Wood, Mabel Berrueta, Aníbal Rondan, Gabriela Lescano, Pablo Cruz, Yvonne Ritou, Valeria Fernández Viña, Damián Álvarez Paggi, Sebastián Esperante, Adrián Ferreti, Gastón Ofman, Álvaro Ciganda, Rocío Rodriguez, Jorge Lantos, Ricardo Valentini, Nicolás Itcovici, Alejandra Hintze, M. Laura Oyarvide, Candela Etchegaray, Alejandra Neira, Ivonne Name, Julieta Alfonso, Rocío López Castelo, Gisela Caruso, Sofía Rapelius, Fernando Alvez, Federico Etchenique, Federico Dimase, Darío Alvarez, Sofía S. Aranda, Clara Sánchez Yanotti, Julián De Luca, Sofía Jares Baglivo, Sofía Laudanno, Florencia Nowogrodzki, Ramiro Larrea, María Silveyra, Gabriel Leberzstein, Alejandra Debonis, Juan Molinos, Miguel González, Eduardo Perez, Nicolás

Kreplak, Susana Pastor Argüello, Luz Gibbons, Fernando Althabe, Eduardo Bergel, and Fernando P. Polack. 2021. 'Early High-Titer Plasma Therapy to Prevent Severe Covid-19 in Older Adults', *New England Journal of Medicine*, 384: 610-18.

Liu, L., Y. Zhang, F. Wei, Q. Zhao, X. Wang, L. Yuan, N. Li, and D. Chen. 2013. 'Discordant genotypic resistance and HIV-1 genetic diversity from paired plasma and cerebrospinal fluid samples in Chinese settings', *J Neurovirol*, 19: 131-6.

Lodi, P. J., J. A. Ernst, J. Kuszewski, A. B. Hickman, A. Engelman, R. Craigie, G. M. Clore, and A. M. Gronenborn. 1995. 'Solution structure of the DNA binding domain of HIV-1 integrase', *Biochemistry*, 34: 9826-33.

Logunov, D. Y., I. V. Dolzhikova, D. V. Shcheblyakov, A. I. Tukhvatulin, O. V. Zubkova, A. S. Dzharullaeva, A. V. Kovyrshina, N. L. Lubenets, D. M. Grousova, A. S. Erokhova, A. G. Botikov, F. M. Izhaeva, O. Popova, T. A. Ozharovskaya, I. B. Esmagambetov, I. A. Favorskaya, D. I. Zrelkin, D. V. Voronina, D. N. Shcherbinin, A. S. Semikhin, Y. V. Simakova, E. A. Tokarskaya, D. A. Egorova, M. M. Shmarov, N. A. Nikitenko, V. A. Gushchin, E. A. Smolyarchuk, S. K. Zyryanov, S. V. Borisevich, B. S. Naroditsky, and A. L. Gintsburg. 2021. 'Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: an interim analysis of a randomised controlled phase 3 trial in Russia', *Lancet*, 397: 671-81.

Loman, Nick, Will Rowe, and Andrew Rambaut. 2020. "nCoV-2019 novel coronavirus bioinformatics protocol." In.: v1.

Lorenzo-Redondo, R., H. R. Fryer, T. Bedford, E. Y. Kim, J. Archer, S. L. Kosakovsky Pond, Y. S. Chung, S. Penugonda, J. G. Chipman, C. V. Fletcher, T. W. Schacker, M. H. Malim, A. Rambaut, A. T. Haase, A. R. McLean, and S. M. Wolinsky. 2016. 'Persistent HIV-1 replication maintains the tissue reservoir during therapy', *Nature*, 530: 51-56.

Lorenzo-Redondo, R., H. R. Fryer, T. Bedford, E. Y. Kim, J. Archer, S. L. K. Pond, Y. S. Chung, S. Penugonda, J. Chipman, C. V. Fletcher, T. W. Schacker, M. H. Malim, A. Rambaut, A. T. Haase, A. R. McLean, and S. M. Wolinsky. 2016. 'Persistent HIV-1 replication maintains the tissue reservoir during therapy', *Nature*, 530: 51-56.

Los Alamos HIV Sequence Database. 'http://www.hiv.lanl.gov/', Accessed 4th March 2019. https://www.hiv.lanl.gov/content/sequence/HelpDocs/subtypes-more.html.

Low, A. J., W. Dong, D. Chan, T. Sing, R. Swanstrom, M. Jensen, S. Pillai, B. Good, and P. R. Harrigan. 2007. 'Current V3 genotyping algorithms are inadequate for predicting X4 co-receptor usage in clinical isolates', *AIDS*, 21: F17-24.

Lu, Shuai, Xi-xiu Xie, Lei Zhao, Bin Wang, Jie Zhu, Ting-rui Yang, Guang-wen Yang, Mei Ji, Cui-ping Lv, Jian Xue, Er-hei Dai, Xi-ming Fu, Dong-qun Liu, Lun Zhang, Sheng-jie Hou, Xiao-lin Yu, Yu-ling Wang, Hui-xia Gao, Xue-han Shi, Chang-wen Ke, Bi-xia Ke, Chun-guo Jiang, and Rui-tian Liu. 2021. 'The immunodominant and neutralization linear epitopes for SARS-CoV-2', *Cell Reports*, 34: 108666.

Lumby, C. K., L. Zhao, J. Breuer, and C. J. Illingworth. 2020. 'A large effective population size for established within-host influenza virus infection', *Elife*, 9.

Lythgoe, K. A., A. Gardner, O. G. Pybus, and J. Grove. 2017. 'Short-Sighted Virus Evolution and a Germline Hypothesis for Chronic Viral Infections', *Trends Microbiol*, 25: 336-48.

Madhi, Shabir A., Vicky Baillie, Clare L. Cutland, Merryn Voysey, Anthonet L. Koen, Lee Fairlie, Sherman D. Padayachee, Keertan Dheda, Shaun L. Barnabas, Qasim Ebrahim Bhorat, Carmen Briner, Gaurav Kwatra, Khatija Ahmed, Parvinder Aley, Sutika Bhikha, Jinal N. Bhiman, As'ad Ebrahim Bhorat, Jeanine du Plessis, Aliasgar Esmail, Marisa Groenewald, Elizea Horne, Shi-Hsia Hwa, Aylin Jose, Teresa Lambe, Matt Laubscher, Mookho Malahleha, Masebole Masenya, Mduduzi Masilela, Shakeel McKenzie, Kgaogelo Molapo, Andrew Moultrie, Suzette Oelofse, Faeezah Patel, Sureshnee Pillay, Sarah Rhead, Hylton Rodel, Lindie Rossouw, Carol Taoushanis, Houriiyah Tegally, Asha Thombrayil, Samuel van Eck, Constantinos Kurt Wibmer, Nicholas M. Durham, Elizabeth J Kelly, Tonya L Villafana, Sarah Gilbert, Andrew J Pollard, Tulio de Oliveira, Penny L. Moore, Alex Sigal, and Alane Izu. 2021. 'Safety and efficacy of the ChAdOx1 nCoV-19 (AZD1222) Covid-19 vaccine against the B.1.351 variant in South Africa', *medRxiv*: 2021.02.10.21251247.

Mancebo, H. S., G. Lee, J. Flygare, J. Tomassini, P. Luu, Y. Zhu, J. Peng, C. Blau, D. Hazuda, D. Price, and O. Flores. 1997. 'P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro', *Genes Dev*, 11: 2633-44.

Mansky, L. M., and H. M. Temin. 1995. 'Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase', *J Virol*, 69: 5087-94.

Martin, Marcel. 2011. 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet. journal*, 17: 10-12.

McMahan, Katherine, Jingyou Yu, Noe B. Mercado, Carolin Loos, Lisa H. Tostanoski, Abishek Chandrashekar, Jinyan Liu, Lauren Peter, Caroline Atyeo, Alex Zhu, Esther A. Bondzie, Gabriel Dagotto, Makda S. Gebre, Catherine Jacob-Dolan, Zhenfeng Li, Felix Nampanya, Shivani Patel, Laurent Pessaint, Alex Van Ry, Kelvin Blade, Jake Yalley-Ogunro, Mehtap Cabus, Renita Brown, Anthony Cook, Elyse Teow, Hanne Andersen, Mark G. Lewis, Douglas A. Lauffenburger, Galit Alter, and Dan H. Barouch. 2021. 'Correlates of protection against SARS-CoV-2 in rhesus macaques', *Nature*, 590: 630-34.

Melby, T., M. Despirito, R. Demasi, G. Heilek-Snyder, M. L. Greenberg, and N. Graham. 2006. 'HIV-1 coreceptor use in triple-class treatment-experienced patients: baseline prevalence, correlates, and relationship to enfuvirtide response', *J Infect Dis*, 194: 238-46.

Mercado, N. B., R. Zahn, F. Wegmann, C. Loos, A. Chandrashekar, J. Yu, J. Liu, L. Peter, K. McMahan, L. H. Tostanoski, X. He, D. R. Martinez, L. Rutten, R. Bos, D. van Manen, J. Vellinga, J. Custers, J. P. Langedijk, T. Kwaks, M. J. G. Bakkers, D. Zuijdgeest, S. K. Rosendahl Huber, C. Atyeo, S. Fischinger, J. S. Burke, J. Feldman, B. M. Hauser, T. M. Caradonna, E. A. Bondzie, G. Dagotto, M. S. Gebre, E. Hoffman, C. Jacob-Dolan, M. Kirilova, Z. Li, Z. Lin, S. H. Mahrokhian, L. F. Maxfield, F. Nampanya, R. Nityanandam, J. P. Nkolola, S. Patel, J. D. Ventura, K. Verrington, H. Wan, L. Pessaint, A. Van Ry, K. Blade, A. Strasbaugh, M. Cabus, R. Brown, A. Cook, S. Zouantchangadou, E. Teow,

Lythgoe, K. A., A. Gardner, O. G. Pybus, and J. Grove. 2017. 'Short-Sighted Virus Evolution and a Germline Hypothesis for Chronic Viral Infections', *Trends Microbiol*, 25: 336-48.

Madhi, Shabir A., Vicky Baillie, Clare L. Cutland, Merryn Voysey, Anthonet L. Koen, Lee Fairlie, Sherman D. Padayachee, Keertan Dheda, Shaun L. Barnabas, Qasim Ebrahim Bhorat, Carmen Briner, Gaurav Kwatra, Khatija Ahmed, Parvinder Aley, Sutika Bhikha, Jinal N. Bhiman, As'ad Ebrahim Bhorat, Jeanine du Plessis, Aliasgar Esmail, Marisa Groenewald, Elizea Horne, Shi-Hsia Hwa, Aylin Jose, Teresa Lambe, Matt Laubscher, Mookho Malahleha, Masebole Masenya, Mduduzi Masilela, Shakeel McKenzie, Kgaogelo Molapo, Andrew Moultrie, Suzette Oelofse, Faeezah Patel, Sureshnee Pillay, Sarah Rhead, Hylton Rodel, Lindie Rossouw, Carol Taoushanis, Houriiyah Tegally, Asha Thombrayil, Samuel van Eck, Constantinos Kurt Wibmer, Nicholas M. Durham, Elizabeth J Kelly, Tonya L Villafana, Sarah Gilbert, Andrew J Pollard, Tulio de Oliveira, Penny L. Moore, Alex Sigal, and Alane Izu. 2021. 'Safety and efficacy of the ChAdOx1 nCoV-19 (AZD1222) Covid-19 vaccine against the B.1.351 variant in South Africa', *medRxiv*: 2021.02.10.21251247.

Mancebo, H. S., G. Lee, J. Flygare, J. Tomassini, P. Luu, Y. Zhu, J. Peng, C. Blau, D. Hazuda, D. Price, and O. Flores. 1997. 'P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro', *Genes Dev*, 11: 2633-44.

Mansky, L. M., and H. M. Temin. 1995. 'Lower in vivo mutation rate of human immunodeficiency virus type 1 than that predicted from the fidelity of purified reverse transcriptase', *J Virol*, 69: 5087-94.

Martin, Marcel. 2011. 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *EMBnet. journal*, 17: 10-12.

McMahan, Katherine, Jingyou Yu, Noe B. Mercado, Carolin Loos, Lisa H. Tostanoski, Abishek Chandrashekar, Jinyan Liu, Lauren Peter, Caroline Atyeo, Alex Zhu, Esther A. Bondzie, Gabriel Dagotto, Makda S. Gebre, Catherine Jacob-Dolan, Zhenfeng Li, Felix Nampanya, Shivani Patel, Laurent Pessaint, Alex Van Ry, Kelvin Blade, Jake Yalley-Ogunro, Mehtap Cabus, Renita Brown, Anthony Cook, Elyse Teow, Hanne Andersen, Mark G. Lewis, Douglas A. Lauffenburger, Galit Alter, and Dan H. Barouch. 2021. 'Correlates of protection against SARS-CoV-2 in rhesus macaques', *Nature*, 590: 630-34.

Melby, T., M. Despirito, R. Demasi, G. Heilek-Snyder, M. L. Greenberg, and N. Graham. 2006. 'HIV-1 coreceptor use in triple-class treatment-experienced patients: baseline prevalence, correlates, and relationship to enfuvirtide response', *J Infect Dis*, 194: 238-46.

Mercado, N. B., R. Zahn, F. Wegmann, C. Loos, A. Chandrashekar, J. Yu, J. Liu, L. Peter, K. McMahan, L. H. Tostanoski, X. He, D. R. Martinez, L. Rutten, R. Bos, D. van Manen, J. Vellinga, J. Custers, J. P. Langedijk, T. Kwaks, M. J. G. Bakkers, D. Zuijdgeest, S. K. Rosendahl Huber, C. Atyeo, S. Fischinger, J. S. Burke, J. Feldman, B. M. Hauser, T. M. Caradonna, E. A. Bondzie, G. Dagotto, M. S. Gebre, E. Hoffman, C. Jacob-Dolan, M. Kirilova, Z. Li, Z. Lin, S. H. Mahrokhian, L. F. Maxfield, F. Nampanya, R. Nityanandam, J. P. Nkolola, S. Patel, J. D. Ventura, K. Verrington, H. Wan, L. Pessaint, A. Van Ry, K. Blade, A. Strasbaugh, M. Cabus, R. Brown, A. Cook, S. Zouantchangadou, E. Teow,

H. Andersen, M. G. Lewis, Y. Cai, B. Chen, A. G. Schmidt, R. K. Reeves, R. S. Baric, D. A. Lauffenburger, G. Alter, P. Stoffels, M. Mammen, J. Van Hoof, H. Schuitemaker, and D. H. Barouch. 2020. 'Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques', *Nature*, 586: 583-88.

Meredith, Luke W., William L. Hamilton, Ben Warne, Charlotte J. Houldcroft, Myra Hosmillo, Aminu S. Jahun, Martin D. Curran, Surendra Parmar, Laura G. Caller, Sarah L. Caddy, Fahad A. Khokhar, Anna Yakovleva, Grant Hall, Theresa Feltwell, Sally Forrest, Sushmita Sridhar, Michael P. Weekes, Stephen Baker, Nicholas Brown, Elinor Moore, Ashley Popay, Iain Roddick, Mark Reacher, Theodore Gouliouris, Sharon J. Peacock, Gordon Dougan, M. Estée Török, and Ian Goodfellow. 2020. 'Rapid implementation of SARS-CoV-2 sequencing to investigate cases of health-care associated COVID-19: a prospective genomic surveillance study', *The Lancet Infectious Diseases*, 20: 1263-72.

Merindol, Natacha, Mohamed El-Far, Mohamed Sylla, Nasser Masroori, Caroline Dufour, Jia-xin Li, Pearl Cherry, Mélodie B. Plourde, Cécile Tremblay, and Lionel Berthoux. 2018. 'HIV-1 capsids from B27/B57+ elite controllers escape Mx2 but are targeted by TRIM5α, leading to the induction of an antiviral state', *PLOS Pathogens*, 14: e1007398.

Merk, A., and S. Subramaniam. 2013. 'HIV-1 envelope glycoprotein structure', *Curr Opin Struct Biol*, 23: 268-76.

Migueles, S. A., M. S. Sabbaghian, W. L. Shupert, M. P. Bettinotti, F. M. Marincola, L. Martino, C. W. Hallahan, S. M. Selig, D. Schwartz, J. Sullivan, and M. Connors. 2000. 'HLA B*5701 is highly associated with restriction of virus replication in a subgroup of HIV-infected long term nonprogressors', *Proc Natl Acad Sci U S A*, 97: 2709-14.

Minh, Bui Quang, Minh Anh Thi Nguyen, and Arndt von Haeseler. 2013. 'Ultrafast Approximation for Phylogenetic Bootstrap', *Molecular biology and evolution*, 30: 1188-95.

Minh, Bui Quang, Heiko A Schmidt, Olga Chernomor, Dominik Schrempf, Michael D Woodhams, Arndt von Haeseler, and Robert Lanfear. 2020. 'IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era', *Molecular biology and evolution*, 37: 1530-34.

Mlcochova, P., D. Collier, A. Ritchie, S. M. Assennato, M. Hosmillo, N. Goel, B. Meng, K. Chatterjee, V. Mendoza, N. Temperton, L. Kiss, L. C. James, K. A. Ciazynska, X. Xiong, J. A. Briggs, J. A. Nathan, F. Mescia, L. Bergamaschi, H. Zhang, P. Barmpounakis, N. Demeris, R. Skells, P. A. Lyons, J. Bradley, S. Baker, J. P. Allain, K. G. Smith, R. Bousfield, M. Wilson, D. Sparkes, G. Amoroso, E. Gkrania-Klotsas, S. Hardwick, A. Boyle, I. Goodfellow, R. K. Gupta, and Citiid-Nihr Covid BioResource Collaboration. 2020. 'Combined point of care nucleic acid and antibody testing for SARS-CoV-2 following emergence of D614G Spike Variant', *Cell Rep Med*: 100099.

Mlcochova, P., K. A. Sutherland, S. A. Watters, C. Bertoli, R. A. de Bruin, J. Rehwinkel, S. J. Neil, G. M. Lenzi, B. Kim, A. Khwaja, M. C. Gage, C. Georgiou, A. Chittka, S. Yona, M. Noursadeghi, G. J. Towers, and R. K. Gupta. 2017. 'A G1-like state allows HIV-1 to bypass SAMHD1 restriction in macrophages', *EMBO J*, 36: 604-16.

Molloy, S. F., C. Kanyama, R. S. Heyderman, A. Loyse, C. Kouanfack, D. Chanda, S. Mfinanga, E. Temfack, S. Lakhi, S. Lesikari, A. K. Chan, N. Stone, N. Kalata, N. Karunaharan, K. Gaskell, M. Peirse, J. Ellis, C. Chawinga, S. Lontsi, J. G. Ndong, P. Bright, D. Lupiya, T. Chen, J. Bradley, J. Adams, C. van der Horst, J. J. van Oosterhout, V. Sini, Y. N. Mapoure, P. Mwaba, T. Bicanic, D. G. Lalloo, D. Wang, M. C. Hosseinipour, O. Lortholary, S. Jaffar, T. S. Harrison, and Acta Trial Study Team. 2018. 'Antifungal Combinations for Treatment of Cryptococcal Meningitis in Africa', *N Engl J Med*, 378: 1004-17.

Montano, M. A., C. P. Nixon, and M. Essex. 1998. 'Dysregulation through the NF-kappaB enhancer and TATA box of the human immunodeficiency virus type 1 subtype E promoter', *J Virol*, 72: 8446-52.

Moore, M. D., and W. S. Hu. 2009. 'HIV-1 RNA dimerization: It takes two to tango', *AIDS Rev*, 11: 91-102.

Moyle, G. J., A. Wildfire, S. Mandalia, H. Mayer, J. Goodrich, J. Whitcomb, and B. G. Gazzard. 2005. 'Epidemiology and predictive factors for chemokine receptor use in HIV-1 infection', *J Infect Dis*, 191: 866-72.

Moyo, Sikhulile, Eduan Wilkinson, Alain Vandormael, Rui Wang, Jia Weng, Kenanao P. Kotokwe, Simani Gaseitsiwe, Rosemary Musonda, Joseph Makhema, Max Essex, Susan Engelbrecht, Tulio de Oliveira, and Vladimir Novitsky. 2017. 'Pairwise diversity and tMRCA as potential markers for HIV infection recency', *Medicine*, 96: e6041-e41.

Mukerji, S. S., V. Misra, D. Lorenz, A. M. Cervantes-Arslanian, J. Lyons, S. Chalkias, A. Wurcel, D. Burke, N. Venna, S. Morgello, I. J. Koralnik, and D. Gabuzda. 2017. 'Temporal Patterns and Drug Resistance in CSF Viral Escape Among ART-Experienced HIV-1 Infected Adults', *J Acquir Immune Defic Syndr*, 75: 246-55.

Mukerji, S. S., V. Misra, D. R. Lorenz, H. Uno, S. Morgello, D. Franklin, R. J. Ellis, S. Letendre, and D. Gabuzda. 2018. 'Impact of Antiretroviral Regimens on CSF Viral Escape in a Prospective Multicohort Study of ART-Experienced HIV-1 Infected Adults in the United States', *Clin Infect Dis*.

Nao, N., J. Yamagishi, H. Miyamoto, M. Igarashi, R. Manzoor, A. Ohnuma, Y. Tsuda, W. Furuyama, A. Shigeno, M. Kajihara, N. Kishida, R. Yoshida, and A. Takada. 2017. 'Genetic Predisposition To Acquire a Polybasic Cleavage Site for Highly Pathogenic Avian Influenza Virus Hemagglutinin', *mBio*, 8.

Nextstrain. 2021. 'Genomic epidemiology of novel coronavirus - Global subsampling, Global clock'.

Nightingale, S., A. M. Geretti, A. Beloukas, M. Fisher, A. Winston, L. Else, M. Nelson, S. Taylor, A. Ustianowski, J. Ainsworth, R. Gilson, L. Haddow, E. Ong, V. Watson, C. Leen, J. Minton, F. Post, M. Pirmohamed, T. Solomon, and S. Khoo. 2016. 'Discordant CSF/plasma HIV-1 RNA in patients with unexplained low-level viraemia', *J Neurovirol*, 22: 852-60.

Nightingale, S., B. D. Michael, M. Fisher, A. Winston, M. Nelson, S. Taylor, A. Ustianowski, J. Ainsworth, R. Gilson, L. Haddow, E. Ong, C. Leen, J. Minton, F. Post, A. Beloukas, R. Borrow, M. Pirmohamed, A. M. Geretti, S. Khoo, and T. Solomon. 2016. 'CSF/plasma HIV-1 RNA discordance even at low levels is associated with up-regulation of host inflammatory mediators in CSF', *Cytokine*, 83: 139-46.

Nobusawa, Eri, and Katsuhiko Sato. 2006. 'Comparison of the Mutation Rates of Human Influenza A and B Viruses', *Journal of Virology*, 80: 3675-78.

Novovax. 2021. "Novavax COVID-19 Vaccine Demonstrates 89.3% Efficacy in UK Phase 3 Trial." In.

Örd, Mihkel, Ilona Faustova, and Mart Loog. 2020. 'The sequence at Spike S1/S2 site enables cleavage by furin and phospho-regulation in SARS-CoV2 but not in SARS-CoV1 or MERS-CoV', *Scientific Reports*, 10: 16944.

Papathanasopoulos, M. A., T. Cilliers, L. Morris, J. L. Mokili, W. Dowling, D. L. Birx, and F. E. McCutchan. 2002. 'Full-length genome analysis of HIV-1 subtype C utilizing CXCR4 and intersubtype recombinants isolated in South Africa', *AIDS Res Hum Retroviruses*, 18: 879-86.

Papavasiliou, F. N., and D. G. Schatz. 2002. 'Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity', *Cell*, 109 Suppl: S35-44.

Parker, J., A. Rambaut, and O. G. Pybus. 2008. 'Correlating viral phenotypes with phylogeny: accounting for phylogenetic uncertainty', *Infect Genet Evol*, 8: 239-46.

Passos, D. O., M. Li, R. Yang, S. V. Rebensburg, R. Ghirlando, Y. Jeon, N. Shkriabai, M. Kvaratskhelia, R. Craigie, and D. Lyumkis. 2017. 'Cryo-EM structures and atomic model of the HIV-1 strand transfer complex intasome', *Science*, 355: 89-92.

Peluso, M. J., F. Ferretti, J. Peterson, E. Lee, D. Fuchs, A. Boschini, M. Gisslen, N. Angoff, R. W. Price, P. Cinque, and S. Spudich. 2012. 'Cerebrospinal fluid HIV escape associated with progressive neurologic dysfunction in patients on antiretroviral therapy with well controlled plasma viral load', *AIDS*, 26: 1765-74.

Peng, Yanchun, Alexander J. Mentzer, Guihai Liu, Xuan Yao, Zixi Yin, Danning Dong, Wanwisa Dejnirattisai, Timothy Rostron, Piyada Supasa, Chang Liu, César López-Camacho, Jose Slon-Campos, Yuguang Zhao, David I. Stuart, Guido C. Paesen, Jonathan M. Grimes, Alfred A. Antson, Oliver W. Bayfield, Dorothy E. D. P. Hawkins, De-Sheng Ker, Beibei Wang, Lance Turtle, Krishanthi Subramaniam, Paul Thomson, Ping Zhang, Christina Dold, Jeremy Ratcliff, Peter Simmonds, Thushan de Silva, Paul Sopp, Dannielle Wellington, Ushani Rajapaksa, Yi-Ling Chen, Mariolina Salio, Giorgio Napolitani, Wayne Paes, Persephone Borrow, Benedikt M. Kessler, Jeremy W. Fry, Nikolai F. Schwabe, Malcolm G. Semple, J. Kenneth Baillie, Shona C. Moore, Peter J. M. Openshaw, M. Azim Ansari, Susanna Dunachie, Eleanor Barnes, John Frater, Georgina Kerr, Philip Goulder, Teresa Lockett, Robert Levin, Yonghong Zhang, Ronghua Jing, Ling-Pei Ho, Eleanor Barnes, Danning Dong, Tao Dong, Susanna Dunachie, John Frater, Philip Goulder, Georgina Kerr, Paul Klenerman, Guihai Liu, Andrew McMichael, Giorgio Napolitani, Graham Ogg, Yanchun Peng, Mariolina Salio, Xuan Yao, Zixi Yin, J. Kenneth Baillie, Paul Klenerman, Alexander J. Mentzer, Shona C. Moore, Peter J. M. Openshaw, Malcolm G. Semple, David I. Stuart, Lance Turtle, Richard J. Cornall, Christopher P. Conlon, Paul Klenerman, Gavin R. Screaton, Juthathip Mongkolsapaya, Andrew McMichael, Julian C. Knight, Graham Ogg, Tao Dong, T. cell Consortium Oxford Immunology Network Covid-19 Response,

and Isaric C. Investigators. 2020. 'Broad and strong memory CD4+ and CD8+ T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19', *Nature Immunology*, 21: 1336-45.

Pepin, Jacques. 2011. *The Origins of AIDS* (Cambridge University Press: Cambridge).

Pillai, S. K., S. L. Pond, Y. Liu, B. M. Good, M. C. Strain, R. J. Ellis, S. Letendre, D. M. Smith, H. F. Gunthard, I. Grant, T. D. Marcotte, J. A. McCutchan, D. D. Richman, and J. K. Wong. 2006. 'Genetic attributes of cerebrospinal fluid-derived HIV-1 env', *Brain*, 129: 1872-83.

Polack, F. P., S. J. Thomas, N. Kitchin, J. Absalon, A. Gurtman, S. Lockhart, J. L. Perez, G. Perez Marc, E. D. Moreira, C. Zerbini, R. Bailey, K. A. Swanson, S. Roychoudhury, K. Koury, P. Li, W. V. Kalina, D. Cooper, R. W. Frenck, Jr., L. L. Hammitt, O. Tureci, H. Nell, A. Schaefer, S. Unal, D. B. Tresnan, S. Mather, P. R. Dormitzer, U. Sahin, K. U. Jansen, W. C. Gruber, and C. Clinical Trial Group. 2020. 'Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine', *N Engl J Med*, 383: 2603-15.

Pommier, Y., A. A. Johnson, and C. Marchand. 2005. 'Integrase inhibitors to treat HIV/AIDS', *Nat Rev Drug Discov*, 4: 236-48.

Powers, D. C., and R. B. Belshe. 1993. 'Effect of age on cytotoxic T lymphocyte memory as well as serum and local antibody responses elicited by inactivated influenza virus vaccine', *J Infect Dis*, 167: 584-92.

Rahimy, E., F. Y. Li, L. Hagberg, D. Fuchs, K. Robertson, D. J. Meyerhoff, H. Zetterberg, R. W. Price, M. Gisslen, and S. Spudich. 2017. 'Blood-Brain Barrier Disruption Is Initiated During Primary HIV Infection and Not Rapidly Altered by Antiretroviral Therapy', *J Infect Dis*, 215: 1132-40.

Ramasamy, Maheshi N., Angela M. Minassian, Katie J. Ewer, Amy L. Flaxman, Pedro M. Folegatti, Daniel R. Owens, Merryn Voysey, Parvinder K. Aley, Brian Angus, Gavin Babbage, Sandra Belij-Rammerstorfer, Lisa Berry, Sagida Bibi, Mustapha Bittaye, Katrina Cathie, Harry Chappell, Sue Charlton, Paola Cicconi, Elizabeth A. Clutterbuck, Rachel Colin-Jones, Christina Dold, Katherine R. W. Emary, Sofiya Fedosyuk, Michelle Fuskova, Diane Gbesemete, Catherine Green, Bassam Hallis, Mimi M. Hou, Daniel Jenkin, Carina C. D. Joe, Elizabeth J. Kelly, Simon Kerridge, Alison M. Lawrie, Alice Lelliott, May N. Lwin, Rebecca Makinson, Natalie G. Marchevsky, Yama Mujadidi, Alasdair P. S. Munro, Mihaela Pacurar, Emma Plested, Jade Rand, Thomas Rawlinson, Sarah Rhead, Hannah Robinson, Adam J. Ritchie, Amy L. Ross-Russell, Stephen Saich, Nisha Singh, Catherine C. Smith, Matthew D. Snape, Rinn Song, Richard Tarrant, Yrene Themistocleous, Kelly M. Thomas, Tonya L. Villafana, Sarah C. Warren, Marion E. E. Watson, Alexander D. Douglas, Adrian V. S. Hill, Teresa Lambe, Sarah C. Gilbert, Saul N. Faust, Andrew J. Pollard, Jeremy Aboagye, Kelly Adams, Aabidah Ali, Elizabeth R. Allen, Lauren Allen, Jennifer L. Allison, Foteini Andritsou, Rachel Anslow, Edward H. Arbe-Barnes, Megan Baker, Natalie Baker, Philip Baker, Ioana Baleanu, Debbie Barker, Eleanor Barnes, Jordan R. Barrett, Kelly Barrett, Louise Bates, Alexander Batten, Kirsten Beadon, Rebecca Beckley, Duncan Bellamy, Adam Berg, Laura Bermejo, Eleanor Berrie, Amy Beveridge, Kevin Bewley, Else M. Bijker, Geeta Birch, Luke Blackwell, Heather Bletchly, Caitlin L. Blundell, Susannah R. Blundell, Emma Bolam, Elena Boland, Daan

Bormans, Nicola Borthwick, Konstantinos Boukas, Thomas Bower, Francesca Bowring, Amy Boyd, Tanja Brenner, Phillip Brown, Charlie Brown-O'Sullivan, Scott Bruce, Emily Brunt, Jamie Burbage, Joshua Burgoyne, Karen R. Buttigieg, Nicholas Byard, Ingrid Cabera Puig, Susana Camara, Michelangelo Cao, Federica Cappuccini, Melanie Carr, Miles W. Carroll, Paul Cashen, Ana Cavey, Jim Chadwick, Ruth Challis, David Chapman, David Charles, Irina Chelysheva, Jee-Sun Cho, Liliana Cifuentes, Elizabeth Clark, Sarah Collins, Christopher P. Conlon, Naomi S. Coombes, Rachel Cooper, Cushla Cooper, Wendy E. M. Crocker, Sarah Crosbie, Dan Cullen, Christina Cunningham, Fiona Cuthbertson, Brad E. Datoo, Lynne Dando, Mehreen S. Datoo, Chandrabali Datta, Hannah Davies, Sarah Davies, Elizabeth J. Davis, Judith Davis, David Dearlove, Tesfaye Demissie, Stefania Di Marco, Claudio Di Maso, Danielle DiTirro, Claire Docksey, Tao Dong, Francesca R. Donnellan, Naomi Douglas, Charlotte Downing, Jonathan Drake, Rachael Drake-Brockman, Ruth E. Drury, Susanna J. Dunachie, Christopher J. Edwards, Nick J. Edwards, Omar El Muhanna, Sean C. Elias, Ryan S. Elliott, Michael J. Elmore, Marcus Rex English, Sally Felle, Shuo Feng, Carla Ferreira Da Silva, Samantha Field, Richard Fisher, Carine Fixmer, Karen J. Ford, Jamie Fowler, Emma Francis, John Frater, Julie Furze, Pablo Galian-Rubio, Celine Galloway, Harriet Garlant, Madita Gavrila, Felicity Gibbons, Karyna Gibbons, Ciaran Gilbride, Hardeep Gill, Kerry Godwin, Katherine Gordon-Quayle, Giacomo Gorini, Lyndsey Goulston, Caroline Grabau, Lara Gracie, Nichola Graham, Nicola Greenwood, Oliver Griffiths, Gaurav Gupta, Elizabeth Hamilton, Brama Hanumunthadu, Stephanie A. Harris, Tara Harris, Daisy Harrison, Thomas C. Hart, Birgit Hartnell, Louise Haskell, Sophia Hawkins, John Aaron Henry, Macarena Hermosin Herrera, David Hill, Jennifer Hill, Gina Hodges, Susanne H. C. Hodgson, Katie L. Horton, Elizabeth Howe, Nicola Howell, Jessica Howes, Ben Huang, Jonathan Humphreys, Holly E. Humphries, Poppy Iveson, Frederic Jackson, Susan Jackson, Sam Jauregui, Helen Jeffers, Bryony Jones, Christine E. Jones, Elizabeth Jones, Kathryn Jones, Amar Joshi, Reshma Kailath, Jade Keen, Dearbhla M. Kelly, Sarah Kelly, Debbie Kelly, David Kerr, Liaquat Khan, Baktash Khozoee, Annabel Killen, Jasmin Kinch, Lloyd D. W. King, Thomas B. King, Lucy Kingham, Paul Klenerman, Julian C. Knight, Daniel Knott, Stanislava Koleva, Gail Lang, Colin W. Larkworthy, Jessica P. J. Larwood, Rebecca Law, Arlene Lee, Kim Y. N. Lee, Emily A. Lees, Stephanie Leung, Yuanyuan Li, Amelia M. Lias, Aline Linder, Samuel Lipworth, Shuchang Liu, Xinxue Liu, Stephanie Lloyd, Lisa Loew, Raquel Lopez Ramon, Meera Madhavan, David O. Mainwaring, Garry Mallett, Kushal Mansatta, Spyridoula Marinou, Phedra Marius, Emma Marlow, Paula Marriott, Julia L. Marshall, Jane Martin, Shauna Masters, Joanne McEwan, Joanna L. McGlashan, Lorna McInroy, Nicky McRobert, Clare Megson, Alexander J. Mentzer, Neginsadat Mirtorabi, Celia Mitton, Maria Moore, Marni Moran, Ella Morey, Róisín Morgans, Susan J. Morris, Hazel Morrison Morrison, Gertraud Morshead, Richard Morter, Nathifa A. Moya, Ekta Mukhopadhyay, Jilly Muller, Claire Munro, Sarah Murphy, Philomena Mweu, Andrés Noé, Fay L. Nugent, Katie O'Brien, Daniel O'Connor, Blanché Oguti, Victoria Olchawski, Catarina Oliveira, Peter John O'Reilly, Piper Osborne,

Lydia Owen, Nelly Owino, Panagiotis Papageorgiou, Helena Parracho, Karen Parsons, Bhumika Patel, Maia Patrick-Smith, Yanchun Peng, Elizabeth J. Penn, Marco Polo Peralta-Alvarez, James Perring, Christos Petropoulos, Daniel J. Phillips, Dimitra Pipini, Samuel Pollard, Ian Poulton, Danny Pratt, Laura Presland, Pamela C. Proud, Samuel Provstgaard-Morys, Sophie Pueschel, David Pulido, Ria Rabara, Kajal Radia, Durga Rajapaska, Fernando Ramos Lopez, Helen Ratcliffe, Sara Rayhan, Byron Rees, Emilia Reyes Pabon, Hannah Roberts, Isla Robertson, Sophie Roche, Christine S. Rollier, Rossana Romani, Zoe Rose, Indra Rudiansyah, Sabeha Sabheha, Stephannie Salvador, Helen Sanders, Katherine Sanders, Iman Satti, Chloe Sayce, Annina B. Schmid, Ella Schofield, Gavin Screaton, Cynthia Sedik, Samiullah Seddiqi, Rameswara R. Segireddy, Beatrice Selby, Imam Shaik, Hannah R. Sharpe, Robert Shaw, Adam Shea, Sarah Silk, Laura Silva-Reyes, Donal T. Skelly, David J. Smith, Daniel C. Smith, Nicholas Smith, Alexandra J. Spencer, Louise Spoors, Elizabeth Stafford, Imogen Stamford, Lisa Stockdale, David Stockley, Lisa V. Stockwell, Matthew Stokes, Louise H. Strickland, Arabella Stuart, Sulaiman Sulaiman, Eloise Summerton, Zoe Swash, Anna Szigeti, Abdessamad Tahiri-Alaoui, Rachel Tanner, Iona Taylor, Keja Taylor, Ursula Taylor, Rebecca te Water Naude, Andreas Themistocleous, Merin Thomas, Tonia M. Thomas, Amber Thompson, Kevin Thompson, Viv Thornton-Jones, Lan Tinh, Adriana Tomic, Susan Tonks, James Towner, Nguyen Tran, Julian A. Tree, Adam Truby, Cheryl Turner, Richard Turner, Marta Ulaszewska, Rachel Varughese, Dennis Verbart, Marije K. Verheul, Iason Vichos, Laura Walker, Matthew E. Wand, Bridget Watkins, Jessica Welch, Alison J. West, Caroline White, Rachel White, Paul Williams, Mark Woodyer, Andrew T. Worth, Daniel Wright, Terri Wrin, Xin Li Yao, Diana-Andreea Zbarcea, and Dalila Zizi. 2020. 'Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime-boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial', *The Lancet*, 396: 1979-93.

Rambaut, Andrew, Edward C. Holmes, Áine O'Toole, Verity Hill, John T. McCrone, Christopher Ruis, Louis du Plessis, and Oliver G. Pybus. 2020. 'A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology', *Nature Microbiology*, 5: 1403-07.

Rasaiyaah, J., C. P. Tan, A. J. Fletcher, A. J. Price, C. Blondeau, L. Hilditch, D. A. Jacques, D. L. Selwood, L. C. James, M. Noursadeghi, and G. J. Towers. 2013. 'HIV-1 evades innate immune recognition through specific cofactor recruitment', *Nature*, 503: 402-05.

Ratnamohan, V. M., A. L. Cunningham, and W. D. Rawlinson. 1998. 'Removal of inhibitors of CSF-PCR to improve diagnosis of herpesviral encephalitis', *J Virol Methods*, 72: 59-65.

Rausch, J. W., and S. F. Le Grice. 2015. 'HIV Rev Assembly on the Rev Response Element (RRE): A Structural Perspective', *Viruses*, 7: 3053-75.

Rawson, T., D. Muir, N. E. Mackie, L. J. Garvey, A. Everitt, and A. Winston. 2012. 'Factors associated with cerebrospinal fluid HIV RNA in HIV infected subjects undergoing lumbar puncture examination in a clinical setting', *J Infect*, 65: 239-45.

Royal, W., 3rd, M. Cherner, J. Carr, A. G. Habib, A. Akomolafe, A. Abimiku, M. Charurat, J. Farley, A. Oluyemisi, I. Mamadu, J. Johnson, R. Ellis, J. A. McCutchan, I. Grant, and W. A. Blattner. 2012. 'Clinical features and preliminary studies of virological correlates of neurocognitive impairment among HIV-infected individuals in Nigeria', *J Neurovirol*, 18: 191-9.

Rupp, Stephanie, Philippe Ambata, Victor Narat, and Tamara Giles-Vernick. 2016. 'Beyond the Cut Hunter: A Historical Epidemiology of HIV Beginnings in Central Africa', *EcoHealth*, 13: 661-71.

Sacktor, N., R. H. Lyles, R. Skolasky, C. Kleeberger, O. A. Selnes, E. N. Miller, J. T. Becker, B. Cohen, J. C. McArthur, and Aids Cohort Study Multicenter. 2001. 'HIV-associated neurologic disease incidence changes:: Multicenter AIDS Cohort Study, 1990-1998', *Neurology*, 56: 257-60.

Sacktor, N., N. Nakasujja, R. L. Skolasky, M. Rezapour, K. Robertson, S. Musisi, E. Katabira, A. Ronald, D. B. Clifford, O. Laeyendecker, and T. C. Quinn. 2009. 'HIV subtype D is associated with dementia, compared with subtype A, in immunosuppressed individuals at risk of cognitive impairment in Kampala, Uganda', *Clin Infect Dis*, 49: 780-6.

Sadoff, Jerald, Glenda Gray, An Vandebosch, Vicky Cárdenas, Georgi Shukarev, Beatriz Grinsztejn, Paul A. Goepfert, Carla Truyers, Hein Fennema, Bart Spiessens, Kim Offergeld, Gert Scheper, Kimberly L. Taylor, Merlin L. Robb, John Treanor, Dan H. Barouch, Jeffrey Stoddard, Martin F. Ryser, Mary A. Marovich, Kathleen M. Neuzil, Lawrence Corey, Nancy Cauwenberghs, Tamzin Tanner, Karin Hardt, Javier Ruiz-Guiñazú, Mathieu Le Gars, Hanneke Schuitemaker, Johan Van Hoof, Frank Struyf, and Macaya Douoguih. 2021. 'Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19', *New England Journal of Medicine*.

Sahin, Ugur, Alexander Muik, Evelyna Derhovanessian, Isabel Vogler, Lena M. Kranz, Mathias Vormehr, Alina Baum, Kristen Pascal, Jasmin Quandt, Daniel Maurus, Sebastian Brachtendorf, Verena Lörks, Julian Sikorski, Rolf Hilker, Dirk Becker, Ann-Kathrin Eller, Jan Grützner, Carsten Boesler, Corinna Rosenbaum, Marie-Cristine Kühnle, Ulrich Luxemburger, Alexandra Kemmer-Brück, David Langer, Martin Bexon, Stefanie Bolte, Katalin Karikó, Tania Palanche, Boris Fischer, Armin Schultz, Pei-Yong Shi, Camila Fontes-Garfias, John L. Perez, Kena A. Swanson, Jakob Loschko, Ingrid L. Scully, Mark Cutler, Warren Kalina, Christos A. Kyratsous, David Cooper, Philip R. Dormitzer, Kathrin U. Jansen, and Özlem Türeci. 2020. 'COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses', *Nature*, 586: 594-99.

Salama, Carlos, Jian Han, Linda Yau, William G. Reiss, Benjamin Kramer, Jeffrey D. Neidhart, Gerard J. Criner, Emma Kaplan-Lewis, Rachel Baden, Lavannya Pandit, Miriam L. Cameron, Julia Garcia-Diaz, Victoria Chávez, Martha Mekebeb-Reuter, Ferdinando Lima de Menezes, Reena Shah, Maria F. González-Lara, Beverly Assman, Jamie Freedman, and Shalini V. Mohan. 2020. 'Tocilizumab in Patients Hospitalized with Covid-19 Pneumonia', *New England Journal of Medicine*, 384: 20-30.

Salazar-Gonzalez, J. F., E. Bailes, K. T. Pham, M. G. Salazar, M. B. Guffey, B. F. Keele, C. A. Derdeyn, P. Farmer, E. Hunter, S. Allen, O. Manigart, J. Mulenga, J. A.

Anderson, R. Swanstrom, B. F. Haynes, G. S. Athreya, B. T. Korber, P. M. Sharp, G. M. Shaw, and B. H. Hahn. 2008. 'Deciphering human immunodeficiency virus type 1 transmission and early envelope diversification by single-genome amplification and sequencing', *J Virol*, 82: 3952-70.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. 'DNA sequencing with chain-terminating inhibitors', *Proc Natl Acad Sci U S A*, 74: 5463-7.

Schmidt, Fabian, Yiska Weisblum, Frauke Muecksch, Hans-Heinrich Hoffmann, Eleftherios Michailidis, Julio C. C. Lorenzi, Pilar Mendoza, Magdalena Rutkowska, Eva Bednarski, Christian Gaebler, Marianna Agudelo, Alice Cho, Zijun Wang, Anna Gazumyan, Melissa Cipolla, Marina Caskey, Davide F. Robbiani, Michel C. Nussenzweig, Charles M. Rice, Theodora Hatziioannou, and Paul D Bieniasz. 2020. 'Measuring SARS-CoV-2 neutralizing antibody activity using pseudotyped and chimeric viruses': 2020.06.08.140871.

Schnell, G., S. Joseph, S. Spudich, R. W. Price, and R. Swanstrom. 2011. 'HIV-1 replication in the central nervous system occurs in two distinct cell types', *PLoS Pathog*, 7: e1002286.

Seow, Jeffrey, 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. 2020. 'Longitudinal observation and decline of neutralizing antibody responses in the three months following SARS-CoV-2 infection in humans', *Nature Microbiology*, 5: 1598-607.

Shafer, R. W. 2006. 'Rationale and uses of a public HIV drug-resistance database', *J Infect Dis*, 194 Suppl 1: S51-8.

Sharp, P. M., and B. H. Hahn. 2011. 'Origins of HIV and the AIDS pandemic', *Cold Spring Harb Perspect Med*, 1: a006841.

Shehu-Xhilaga, M., S. M. Crowe, and J. Mak. 2001. 'Maintenance of the Gag/Gag-Pol ratio is important for human immunodeficiency virus type 1 RNA dimerization and viral infectivity', *J Virol*, 75: 1834-41.

Shereen, M. A., S. Khan, A. Kazmi, N. Bashir, and R. Siddique. 2020. 'COVID-19 infection: Origin, transmission, and characteristics of human coronaviruses', *J Adv Res*, 24: 91-98.

Shu, Yuelong, and John McCauley. 2017. 'GISAID: Global initiative on sharing all influenza data - from vision to reality', *Euro Surveill*, 22: 30494.

Sigal, A., and D. Baltimore. 2012. 'As good as it gets? The problem of HIV persistence despite antiretroviral drugs', *Cell Host Microbe*, 12: 132-8.

Silva-Cayetano, A., W. S. Foster, S. Innocentin, S. Belij-Rammerstorfer, A. J. Spencer, O. T. Burton, S. Fra-Bidó, J. Le Lee, N. Thakur, C. Conceicao, D. Wright, J. Barrett, N. Evans-Bailey, C. Noble, D. Bailey, A. Liston, S. C. Gilbert, T. Lambe,

and M. A. Linterman. 2021. 'A booster dose enhances immunogenicity of the COVID-19 vaccine candidate ChAdOx1 nCoV-19 in aged mice', *Med (N Y)*, 2: 243-62.e8.

Simioni, S., M. Cavassini, J. M. Annoni, A. Rimbault Abraham, I. Bourquin, V. Schiffer, A. Calmy, J. P. Chave, E. Giacobini, B. Hirschel, and R. A. Du Pasquier. 2010. 'Cognitive dysfunction in HIV patients despite long-standing suppression of viremia', *AIDS*, 24: 1243-50.

Simonovich, Ventura A., Leandro D. Burgos Pratx, Paula Scibona, María V. Beruto, Marcelo G. Vallone, Carolina Vázquez, Nadia Savoy, Diego H. Giunta, Lucía G. Pérez, Marisa del L. Sánchez, Andrea Vanesa Gamarnik, Diego S. Ojeda, Diego M. Santoro, Pablo J. Camino, Sebastian Antelo, Karina Rainero, Gabriela P. Vidiella, Erica A. Miyazaki, Wanda Cornistein, Omar A. Trabadelo, Fernando M. Ross, Mariano Spotti, Gabriel Funtowicz, Walter E. Scordo, Marcelo H. Losso, Inés Ferniot, Pablo E. Pardo, Eulalia Rodriguez, Pablo Rucci, Julieta Pasquali, Nora A. Fuentes, Mariano Esperatti, Gerardo A. Speroni, Esteban C. Nannini, Alejandra Matteaccio, Hernán G. Michelangelo, Dean Follmann, H. Clifford Lane, and Waldo H. Belloso. 2020. 'A Randomized Trial of Convalescent Plasma in Covid-19 Severe Pneumonia', *New England Journal of Medicine*, 384: 619-29.

Sing, T., A. J. Low, N. Beerenwinkel, O. Sander, P. K. Cheung, F. S. Domingues, J. Büch, M. Däumer, R. Kaiser, T. Lengauer, and P. R. Harrigan. 2007. 'Predicting HIV coreceptor usage on the basis of genetic and clinical covariates', *Antivir Ther*, 12: 1097-106.

Sodeik, B., M. W. Ebersold, and A. Helenius. 1997. 'Microtubule-mediated transport of incoming herpes simplex virus 1 capsids to the nucleus', *J Cell Biol*, 136: 1007-21.

Sojane, K., R. T. Kangethe, C. C. Chang, M. S. Moosa, S. R. Lewin, M. A. French, and T. Ndung'u. 2018. 'Individuals with HIV-1 Subtype C Infection and Cryptococcal Meningitis Exhibit Viral Genetic Intermixing of HIV-1 Between Plasma and Cerebrospinal Fluid and a High Prevalence of CXCR4-Using Variants', *AIDS Res Hum Retroviruses*, 34: 607-20.

Sok, D., and D. R. Burton. 2018. 'Recent progress in broadly neutralizing antibodies to HIV', *Nat Immunol*, 19: 1179-88.

Song, H., E. E. Giorgi, V. V. Ganusov, F. Cai, G. Athreya, H. Yoon, O. Carja, B. Hora, P. Hraber, E. Romero-Severson, C. Jiang, X. Li, S. Wang, H. Li, J. F. Salazar-Gonzalez, M. G. Salazar, N. Goonetilleke, B. F. Keele, D. C. Montefiori, M. S. Cohen, G. M. Shaw, B. H. Hahn, A. J. McMichael, B. F. Haynes, B. Korber, T. Bhattacharya, and F. Gao. 2018. 'Tracking HIV-1 recombination to resolve its contribution to HIV-1 evolution in natural infection', *Nat Commun*, 9: 1928.

Stam, A. J., M. Nijhuis, W. M. van den Bergh, and A. M. Wensing. 2013. 'Differential genotypic evolution of HIV-1 quasispecies in cerebrospinal fluid and plasma: a systematic review', *AIDS Rev*, 15: 152-61.

Starr, T. N., A. J. Greaney, S. K. Hilton, D. Ellis, K. H. D. Crawford, A. S. Dingens, M. J. Navarro, J. E. Bowen, M. A. Tortorici, A. C. Walls, N. P. King, D. Veesler, and J. D. Bloom. 2020. 'Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding', *Cell*, 182: 1295-310 e20.

Sturdevant, C. B., S. B. Joseph, G. Schnell, R. W. Price, R. Swanstrom, and S. Spudich. 2015. 'Compartmentalized replication of R5 T cell-tropic HIV-1 in the central nervous system early in the course of infection', *PLoS Pathog*, 11: e1004720.

Sundquist, W. I., and H. G. Kräusslich. 2012. 'HIV-1 assembly, budding, and maturation', *Cold Spring Harb Perspect Med*, 2: a006924.

Tang, N., H. Bai, X. Chen, J. Gong, D. Li, and Z. Sun. 2020. 'Anticoagulant treatment is associated with decreased mortality in severe coronavirus disease 2019 patients with coagulopathy', *J Thromb Haemost*, 18: 1094-99.

Tao, Y., M. Shi, C. Chommanard, K. Queen, J. Zhang, W. Markotter, I. V. Kuzmin, E. C. Holmes, and S. Tong. 2017. 'Surveillance of Bat Coronaviruses in Kenya Identifies Relatives of Human Coronaviruses NL63 and 229E and Their Recombination History', *J Virol*, 91.

Teunissen, C. E., H. Tumani, J. L. Bennett, F. S. Berven, L. Brundin, M. Comabella, D. Franciotta, J. L. Federiksen, J. O. Fleming, R. Furlan, R. Q. Hintzen, S. G. Hughes, C. R. Jimenez, M. H. Johnson, J. Killestein, E. Krasulova, J. Kuhle, M. C. Magnone, A. Petzold, C. Rajda, K. Rejdak, H. K. Schmidt, V. van Pesch, E. Waubant, C. Wolf, F. Deisenhammer, G. Giovannoni, and B. Hemmer. 2011. 'Consensus Guidelines for CSF and Blood Biobanking for CNS Biomarker Studies', *Mult Scler Int*, 2011: 246412.

The RECOVERY Collaborative Group. 2020. 'Dexamethasone in Hospitalized Patients with Covid-19', *New England Journal of Medicine*, 384: 693-704.

The Royal Society. 2021. "Why we know vaccines work." In.

Thielen, A., N. Sichtig, R. Kaiser, J. Lam, P. R. Harrigan, and T. Lengauer. 2010. 'Improved prediction of HIV-1 coreceptor usage with sequence information from the second hypervariable loop of gp120', *J Infect Dis*, 202: 1435-43.

Tong, C. Y., S. Costelloe, J. Hubb, J. Mullen, S. O'Shea, M. Marta, R. Kulasegaram, and S. Rackstraw. 2015. 'Deep Sequencing of HIV-1 in Cerebrospinal Fluid', *Clin Infect Dis*, 61: 1022-5.

UNAIDS. 2020. "Global HIV & AIDS statistics — 2020 fact sheet." In.

Underwood, J., J. H. Cole, R. Leech, D. J. Sharp, A. Winston, and Charter group. 2018. 'Multivariate Pattern Analysis of Volumetric Neuroimaging Data and Its Relationship With Cognitive Function in Treated HIV Disease', *J Acquir Immune Defic Syndr*, 78: 429-36.

V'kovski, Philip, Annika Kratzel, Silvio Steiner, Hanspeter Stalder, and Volker Thiel. 2021. 'Coronavirus biology and replication: implications for SARS-CoV-2', *Nature Reviews Microbiology*, 19: 155-70.

van den Berg, S. P. H., K. Warmink, J. A. M. Borghans, M. J. Knol, and D. van Baarle. 2019. 'Effect of latent cytomegalovirus infection on the antibody response to influenza vaccination: a systematic review and meta-analysis', *Med Microbiol Immunol*, 208: 305-21.

van Zoest, R. A., J. Underwood, D. De Francesco, C. A. Sabin, J. H. Cole, F. W. Wit, M. W. A. Caan, N. A. Kootstra, D. Fuchs, H. Zetterberg, Cblm Majoie, P. Portegies, A. Winston, D. J. Sharp, M. Gisslen, P. Reiss, and Aids Collaboration Comorbidity in Relation to. 2017. 'Structural Brain Abnormalities in Successfully Treated HIV Infection: Associations With Disease and Cerebrospinal Fluid Biomarkers', *J Infect Dis*, 217: 69-81.

Vasileiou, Eleftheria & Simpson, Colin & Robertson, Chris & Shi, Ting & Kerr, Steven & Agrawal, Utkarsh & Akbari, Ashley & Bedston, Stuart & Beggs, Jillian & Bradley, Declan & Chuter, Antony & Lusignan, Simon & Docherty, Annemarie & Ford, David & Hobbs, FD & Joy, Mark & Katikireddi, Srinivasa & Marple, James & McCowan, Colin & Sheikh, Aziz. 2021. 'Effectiveness of First Dose of COVID-19 Vaccines Against Hospital Admissions in Scotland: National Prospective Cohort Study of 5.4 Million People.'.

Volz, Erik, Swapnil Mishra, Meera Chand, Jeffrey C. Barrett, Robert Johnson, Lily Geidelberg, Wes R Hinsley, Daniel J Laydon, Gavin Dabrera, Áine O'Toole, Roberto Amato, Manon Ragonnet-Cronin, Ian Harrison, Ben Jackson, Cristina V. Ariani, Olivia Boyd, Nicholas J Loman, John T McCrone, Sónia Gonçalves, David Jorgensen, Richard Myers, Verity Hill, David K. Jackson, Katy Gaythorpe, Natalie Groves, John Sillitoe, Dominic P. Kwiatkowski, Seth Flaxman, Oliver Ratmann, Samir Bhatt, Susan Hopkins, Axel Gandy, Andrew Rambaut, and Neil M Ferguson. 2021. 'Transmission of SARS-CoV-2 Lineage B.1.1.7 in England: Insights from linking epidemiological and genetic data', *medRxiv*: 2020.12.30.20249034.

Voysey, M., S. A. C. Clemens, S. A. Madhi, L. Y. Weckx, P. M. Folegatti, P. K. Aley, B. Angus, V. L. Baillie, S. L. Barnabas, Q. E. Bhorat, S. Bibi, C. Briner, P. Cicconi, A. M. Collins, R. Colin-Jones, C. L. Cutland, T. C. Darton, K. Dheda, C. J. A. Duncan, K. R. W. Emary, K. J. Ewer, L. Fairlie, S. N. Faust, S. Feng, D. M. Ferreira, A. Finn, A. L. Goodman, C. M. Green, C. A. Green, P. T. Heath, C. Hill, H. Hill, I. Hirsch, S. H. C. Hodgson, A. Izu, S. Jackson, D. Jenkin, C. C. D. Joe, S. Kerridge, A. Koen, G. Kwatra, R. Lazarus, A. M. Lawrie, A. Lelliott, V. Libri, P. J. Lillie, R. Mallory, A. V. A. Mendes, E. P. Milan, A. M. Minassian, A. McGregor, H. Morrison, Y. F. Mujadidi, A. Nana, P. J. O'Reilly, S. D. Padayachee, A. Pittella, E. Plested, K. M. Pollock, M. N. Ramasamy, S. Rhead, A. V. Schwarzbold, N. Singh, A. Smith, R. Song, M. D. Snape, E. Sprinz, R. K. Sutherland, R. Tarrant, E. C. Thomson, M. E. Torok, M. Toshner, D. P. J. Turner, J. Vekemans, T. L. Villafana, M. E. E. Watson, C. J. Williams, A. D. Douglas, A. V. S. Hill, T. Lambe, S. C. Gilbert, A. J. Pollard, and Covid Vaccine Trial Group Oxford. 2021. '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.

Voysey, M., S. A. Costa Clemens, S. A. Madhi, L. Y. Weckx, P. M. Folegatti, P. K. Aley, B. Angus, V. L. Baillie, S. L. Barnabas, Q. E. Bhorat, S. Bibi, C. Briner, P. Cicconi, E. A. Clutterbuck, A. M. Collins, C. L. Cutland, T. C. Darton, K. Dheda, C. Dold, C. J. A. Duncan, K. R. W. Emary, K. J. Ewer, A. Flaxman, L. Fairlie, S. N. Faust, S. Feng, D. M. Ferreira, A. Finn, E. Galiza, A. L. Goodman, C. M. Green, C. A. Green, M. Greenland, C. Hill, H. C. Hill, I. Hirsch, A. Izu, D. Jenkin, C. C. D. Joe, S. Kerridge, A. Koen, G. Kwatra, R. Lazarus, V. Libri, P. J. Lillie, N. G. Marchevsky, R. P. Marshall, A. V. A. Mendes, E. P. Milan, A. M. Minassian, A. McGregor, Y. F. Mujadidi, A. Nana, S. D. Padayachee, D. J. Phillips, A. Pittella, E. Plested, K. M. Pollock, M. N. Ramasamy, A. J. Ritchie, H. Robinson, A. V. Schwarzbold, A. Smith, R. Song, M. D. Snape, E. Sprinz, R. K. Sutherland, E. C. Thomson, M. E. Torok, M. Toshner, D. P. J. Turner, J.

Vekemans, T. L. Villafana, T. White, C. J. Williams, A. D. Douglas, A. V. S. Hill, T. Lambe, S. C. Gilbert, A. J. Pollard, and Covid Vaccine Trial Group Oxford. 2021. 'Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: a pooled analysis of four randomised trials', *Lancet*, 397: 881-91.

Walker, B. D., S. Chakrabarti, B. Moss, T. J. Paradis, T. Flynn, A. G. Durno, R. S. Blumberg, J. C. Kaplan, M. S. Hirsch, and R. T. Schooley. 1987. 'HIV-specific cytotoxic T lymphocytes in seropositive individuals', *Nature*, 328: 345-8.

Walsh, Edward E., Robert W. Frenck, Ann R. Falsey, Nicholas Kitchin, Judith Absalon, Alejandra Gurtman, Stephen Lockhart, Kathleen Neuzil, Mark J. Mulligan, Ruth Bailey, Kena A. Swanson, Ping Li, Kenneth Koury, Warren Kalina, David Cooper, Camila Fontes-Garfias, Pei-Yong Shi, Özlem Türeci, Kristin R. Tompkins, Kirsten E. Lyke, Vanessa Raabe, Philip R. Dormitzer, Kathrin U. Jansen, Uğur Şahin, and William C. Gruber. 2020. 'Safety and Immunogenicity of Two RNA-Based Covid-19 Vaccine Candidates', *New England Journal of Medicine*, 383: 2439-50.

Wang, Y., D. Zhang, G. Du, R. Du, J. Zhao, Y. Jin, S. Fu, L. Gao, Z. Cheng, Q. Lu, Y. Hu, G. Luo, K. Wang, Y. Lu, H. Li, S. Wang, S. Ruan, C. Yang, C. Mei, Y. Wang, D. Ding, F. Wu, X. Tang, X. Ye, Y. Ye, B. Liu, J. Yang, W. Yin, A. Wang, G. Fan, F. Zhou, Z. Liu, X. Gu, J. Xu, L. Shang, Y. Zhang, L. Cao, T. Guo, Y. Wan, H. Qin, Y. Jiang, T. Jaki, F. G. Hayden, P. W. Horby, B. Cao, and C. Wang. 2020. 'Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial', *Lancet*, 395: 1569-78.

Wei, X., J. M. Decker, S. Wang, H. Hui, J. C. Kappes, X. Wu, J. F. Salazar-Gonzalez, M. G. Salazar, J. M. Kilby, M. S. Saag, N. L. Komarova, M. A. Nowak, B. H. Hahn, P. D. Kwong, and G. M. Shaw. 2003. 'Antibody neutralization and escape by HIV-1', *Nature*, 422: 307-12.

Weinreich, David M., Sumathi Sivapalasingam, Thomas Norton, Shazia Ali, Haitao Gao, Rafia Bhore, Bret J. Musser, Yuhwen Soo, Diana Rofail, Joseph Im, Christina Perry, Cynthia Pan, Romana Hosain, Adnan Mahmood, John D. Davis, Kenneth C. Turner, Andrea T. Hooper, Jennifer D. Hamilton, Alina Baum, Christos A. Kyratsous, Yunji Kim, Amanda Cook, Wendy Kampman, Anita Kohli, Yessica Sachdeva, Ximena Graber, Bari Kowal, Thomas DiCioccio, Neil Stahl, Leah Lipsich, Ned Braunstein, Gary Herman, and George D. Yancopoulos. 2020. 'REGN-COV2, a Neutralizing Antibody Cocktail, in Outpatients with Covid-19', *New England Journal of Medicine*, 384: 238-51.

Westmeier, J., K. Paniskaki, Z. Karaköse, T. Werner, K. Sutter, S. Dolff, M. Overbeck, A. Limmer, J. Liu, X. Zheng, T. Brenner, M. M. Berger, O. Witzke, M. Trilling, M. Lu, D. Yang, N. Babel, T. Westhoff, U. Dittmer, and G. Zelinskyy. 2020. 'Impaired Cytotoxic CD8(+) T Cell Response in Elderly COVID-19 Patients', *mBio*, 11.

WHO Solidarity Trial Consortium. 2020. 'Repurposed Antiviral Drugs for Covid-19 — Interim WHO Solidarity Trial Results', *New England Journal of Medicine*, 384: 497-511.

Wilkin, T. J., Z. Su, D. R. Kuritzkes, M. Hughes, C. Flexner, R. Gross, E. Coakley, W. Greaves, C. Godfrey, P. R. Skolnik, J. Timpone, B. Rodriguez, and R. M.

Gulick. 2007. 'HIV type 1 chemokine coreceptor use among antiretroviral-experienced patients screened for a clinical trial of a CCR5 inhibitor: AIDS Clinical Trial Group A5211', *Clin Infect Dis*, 44: 591-5.

Wilkinson, K., Y. Wei, A. Szwajcer, R. Rabbani, R. Zarychanski, A. M. Abou-Setta, and S. M. Mahmud. 2017. 'Efficacy and safety of high-dose influenza vaccine in elderly adults: A systematic review and meta-analysis', *Vaccine*, 35: 2775-80.

World Health Organisation. 2018. "Updated recommendations on first-line and second-line antiretroviral

regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV:

interim guidelines. Supplement to the 2016 consolidated guidelines on the use of antiretroviral

drugs for treating and preventing HIV infection." In. Geneva: World Health Organisation

Worobey, M., M. Gemmel, D. E. Teuwen, T. Haselkorn, K. Kunstman, M. Bunce, J. J. Muyembe, J. M. Kabongo, R. M. Kalengayi, E. Van Marck, M. T. Gilbert, and S. M. Wolinsky. 2008. 'Direct evidence of extensive diversity of HIV-1 in Kinshasa by 1960', *Nature*, 455: 661-4.

Worobey, M., T. D. Watts, R. A. McKay, M. A. Suchard, T. Granade, D. E. Teuwen, B. A. Koblin, W. Heneine, P. Lemey, and H. W. Jaffe. 2016. '1970s and 'Patient 0' HIV-1 genomes illuminate early HIV/AIDS history in North America', *Nature*, 539: 98-101.

Wrobel, Antoni G., Donald J. Benton, Pengqi Xu, Chloë Roustan, Stephen R. Martin, Peter B. Rosenthal, John J. Skehel, and Steven J. Gamblin. 2020. 'SARS-CoV-2 and bat RaTG13 spike glycoprotein structures inform on virus evolution and furin-cleavage effects', *Nature Structural & Molecular Biology*, 27: 763-67.

Wu, F., S. Zhao, B. Yu, Y. M. Chen, W. Wang, Z. G. Song, Y. Hu, Z. W. Tao, J. H. Tian, Y. Y. Pei, M. L. Yuan, Y. L. Zhang, F. H. Dai, Y. Liu, Q. M. Wang, J. J. Zheng, L. Xu, E. C. Holmes, and Y. Z. Zhang. 2020. 'A new coronavirus associated with human respiratory disease in China', *Nature*, 579: 265-69.

Wu, K., G. Peng, M. Wilken, R. J. Geraghty, and F. Li. 2012. 'Mechanisms of host receptor adaptation by severe acute respiratory syndrome coronavirus', *J Biol Chem*, 287: 8904-11.

Wu, Xueling, Charlene Wang, Sijy O'Dell, Yuxing Li, Brandon F. Keele, Zhongjia Yang, Hiromi Imamichi, Nicole Doria-Rose, James A. Hoxie, Mark Connors, George M. Shaw, Richard T. Wyatt, and John R. Mascola. 2012. 'Selection Pressure on HIV-1 Envelope by Broadly Neutralizing Antibodies to the Conserved CD4-Binding Site', *Journal of Virology*, 86: 5844-56.

Wu, Z., Y. Hu, M. Xu, Z. Chen, W. Yang, Z. Jiang, M. Li, H. Jin, G. Cui, P. Chen, L. Wang, G. Zhao, Y. Ding, Y. Zhao, and W. Yin. 2021. 'Safety, tolerability, and immunogenicity of an inactivated SARS-CoV-2 vaccine (CoronaVac) in healthy adults aged 60 years and older: a randomised, double-blind, placebo-controlled, phase 1/2 clinical trial', *Lancet Infect Dis*.

Xia, Shuai, Qiaoshuai Lan, Shan Su, Xinling Wang, Wei Xu, Zezhong Liu, Yun Zhu, Qian Wang, Lu Lu, and Shibo Jiang. 2020. 'The role of furin cleavage site in SARS-CoV-2 spike protein-mediated membrane fusion in the presence or absence of trypsin', *Signal Transduction and Targeted Therapy*, 5: 92.

Xiong, X., K. Qu, K. A. Ciazynska, M. Hosmillo, A. P. Carter, S. Ebrahimi, Z. Ke, S. H. W. Scheres, L. Bergamaschi, G. L. Grice, Y. Zhang, Citiid-Nihr Covid-BioResource Collaboration, J. A. Nathan, S. Baker, L. C. James, H. E. Baxendale, I. Goodfellow, R. Doffinger, and J. A. G. Briggs. 2020. 'A thermostable, closed SARS-CoV-2 spike protein trimer', *Nat Struct Mol Biol*, 27: 934-41.

Zarate, S., S. L. Pond, P. Shapshak, and S. D. Frost. 2007. 'Comparative study of methods for detecting sequence compartmentalization in human immunodeficiency virus type 1', *J Virol*, 81: 6643-51.

Zarychanski, Ryan. 2021. 'Therapeutic Anticoagulation in Critically Ill Patients with Covid-19 – Preliminary Report', *medRxiv*: 2021.03.10.21252749.

Zhang, Yuhao, Xiuchao Geng, Yanli Tan, Qiang Li, Can Xu, Jianglong Xu, Liangchao Hao, Zhaomu Zeng, Xianpu Luo, Fulin Liu, and Hong Wang. 2020. 'New understanding of the damage of SARS-CoV-2 infection outside the respiratory system', *Biomedicine & Pharmacotherapy*, 127: 110195.

Zhou, Peng, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, Yan Zhu, Bei Li, Chao-Lin Huang, Hui-Dong Chen, Jing Chen, Yun Luo, Hua Guo, Ren-Di Jiang, Mei-Qin Liu, Ying Chen, Xu-Rui Shen, Xi Wang, Xiao-Shuang Zheng, Kai Zhao, Quan-Jiao Chen, Fei Deng, Lin-Lin Liu, Bing Yan, Fa-Xian Zhan, Yan-Yi Wang, Geng-Fu Xiao, and Zheng-Li Shi. 2020. 'A pneumonia outbreak associated with a new coronavirus of probable bat origin', *Nature*, 579: 270-73.

Zhu, F. C., X. H. Guan, Y. H. Li, J. Y. Huang, T. Jiang, L. H. Hou, J. X. Li, B. F. Yang, L. Wang, W. J. Wang, S. P. Wu, Z. Wang, X. H. Wu, J. J. Xu, Z. Zhang, S. Y. Jia, B. S. Wang, Y. Hu, J. J. Liu, J. Zhang, X. A. Qian, Q. Li, H. X. Pan, H. D. Jiang, P. Deng, J. B. Gou, X. W. Wang, X. H. Wang, and W. Chen. 2020. 'Immunogenicity and safety of a recombinant adenovirus type-5-vectored COVID-19 vaccine in healthy adults aged 18 years or older: a randomised, double-blind, placebo-controlled, phase 2 trial', *Lancet*, 396: 479-88.

Zuo, Jianmin, Alexander C. Dowell, Hayden Pearce, Kriti Verma, Heather M. Long, Jusnara Begum, Felicity Aiano, Zahin Amin-Chowdhury, Bassam Hallis, Lorrain Stapley, Ray Borrow, Ezra Linley, Shazaad Ahmad, Ben Parker, Alex Horsley, Gayatri Amirthalingam, Kevin Brown, Mary E. Ramsay, Shamez Ladhani, and Paul Moss. 2021. 'Robust SARS-CoV-2-specific T cell immunity is maintained at 6 months following primary infection', *Nature Immunology*.