Neutralizing Antibody Responses in HIV-1 Dual Infection: Lessons for Vaccine Design

Daniel James Sheward

Supervisor: Carolyn Williamson

Co-supervisor: Zenda Woodman

University of Cape Town

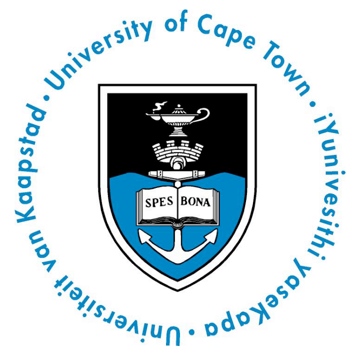
Dissertation submitted to the

Faculty of Health Sciences, University of Cape Town, South Africa

in fulfillment of the requirements for the degree of

Master of Science in Medicine.

October 2013



The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

University of Cape Town

Acknowledgements

I gratefully acknowledge bursary funding provided by the German Academic Exchange Service (DAAD) administered from the National Research Foundation of South Africa (NRF), the Poliomyelitis Research Foundation (PRF), and the University of Cape Town (UCT).

I am indebted to my supervisor, Professor Carolyn Williamson, for her unwavering patience and for giving me the freedom to pursue my interests, wherever they lead. I am truly fortunate to have found myself in your lab.

I would like to thank Zenda Woodman for her supervision and comments on this thesis, and for being a good friend. I would also like to thank Debbie Stewart for proofreading this thesis, for always looking out for me, and for generally making sure the lab doesn't implode.

I will forever be grateful to Alexis Bick, who has been my biggest supporter. Thank you for putting up with me, and for pushing me to "typey typey". This pesky thesis may never have been submitted without you.

Finally, I would like to thank my parents. Your sacrifices for my education did not go unnoticed. Thank you for always telling me I could be whatever I wanted to be. I'm sorry I picked this.

Plagiarism declaration

I know the meaning of plagiarism and declare that all of the work in this dissertation, save for that which is properly acknowledged, is my own.

I have used the Journal of Virology as the convention for citations and referencing. Each significant contribution to, and quotation in this thesis from the works of other people has been attributed and has been cited and referenced where applicable.

Signature: ____________________________

Date: ________________________________

Table of Contents

List of Figures i
List of Tables iii
List of Abbreviations iv

Abstract 1

CHAPTER 1: Background 3

Aims: 19

Objectives: 19

CHAPTER 2: Identification and genetic characterization of HIV-1 dual infection 20

2.1 Introduction: 20
2.2 Methods: 23
2.3 Results: 29
2.4 Discussion 48

CHAPTER 3: The clinical causes and consequences of multiple HIV infection 52

3.1 Introduction 52
3.2 Methods: 56
3.3 Results 59
3.4 Discussion 69

CHAPTER 4: Dual infection was not associated with greater breadth of the neutralizing antibody response 72

4.1 Introduction 72
4.2 Methods 74
4.3 Results 78
4.4 Discussion 93

References 102

Supplementary data 119

List of Figures

Figure 1.1. Targets for broadly neutralizing HIV antibodies 19

Figure 1.2. Bread-potency curves for nine broadly neutralizing HIV antibodies.. 19

Figure 1.3. Schematic of HIV recombination 24

Figure 2.1. Schematic illustrating the classification of multiplicity of infection 37

Figure 2.2. DNA distance in *envelope* at the first available HIV positive visit plotted according to the estimated time post infection. 32

Figure 2.3. Maximum likelihood phylogenetic tree illustrating representative *gp160* sequences from 15 participants initially infected with more than one virus 41

Figure 2.4. No significant evidence of recombination with unlinked viruses in seven participants infected with more than one variant 41

Figure 2.5. Evidence of divergent subgenomic regions in seven participants 42

Figure 2.6. Multiple unique recombination breakpoints detected in *gp160* over time in CAP84.... 43

Figure 2.7. Viral loads (red) and CD4 counts (blue) over time in participants CAP256 (A), CAP237 (B), and CAP281 (C), along with the estimated timing of superinfection 46

Figure 2.8. Maximum likelihood phylogeny depicting the presence of two strains separated on the phylogeny by epidemiologically unlinked viruses 47

Figure 2.9. Nucleotide highlighter plots of gp160 SGA sequences sampled recently post superinfection 47

Figure 2.10. No evidence of recombinant regions inherited from the primary virus within gp160 of the superinfecting viruses 48

Figure 2.11. Superinfecting T/F viruses have Envelopes with shorter, less glycosylated variable regions compared to viruses sampled at three years post infection 50

Figure 2.12. Highlighter plot depicting recombination in *envelope* over time in 3 superinfected participants 51

Figure 2.13. Highlighter plot depicting recombination over time in a co-infected participant, CAP137 52

Figure 2.14. Modelling a hypothetical correlate of risk of superinfection 54

Figure 3.1. Comparison of the pre-infection plasma CD4+ T cell counts between participants where infection was founded by a single virus (Single infection) and those infected with more than one virus (Multiple infection) 69

Figure 3.2. Average rate of CD4 decline in singly and dual infected participants 72

Figure 3.3. Multiplicity of infection and cumulative viral load 73

Figure 3.4. Viral load trajectories prior to- and post-superinfection 76

Figure 4.1. No significant association between HIV dual infection and the later development of neutralizing antibody breadth 88

Figure 4.2. The association between A. CD4 counts at 6 months post infection (mpi) and B. viral load setpoint and neutralizing antibody breadth 90
Figure 4.3. The association between early diversity in *envelope(env*) and the later development of neutralizing antibody breadth 92
Figure 4.4. An alignment of V2 sequences related to the primary virus (A) and the superinfecting virus (B) in CAP256, with sites under positive selection highlighted in red 95
Figure 4.5. Amino acid conservation at sites 165, 166, and 169 in highly sensitive and resistant viruses 96
Figure 4.6. The autologous neutralizing antibody response in CAP256 97
Figure 4.7. Sequence homology in V2 between the primary- (PI) and superinfecting (SI) viruses in CAP256 89
Figure 4.8. The dependence of CAP256 plasma antibodies on specific residues for neutralization of ConC over time (from 42) is overlaid onto the CAP256 viral diversification 100

List of Tables

Table 1.1. Representatives of broadly neutralizing HIV monoclonal antibodies (mAbs) 12

Table 2.1. Early *envelope* diversity at the first HIV positive visits providing evidence of multivariant or co-infection of 15 participants. 31

Table 2.2. Defining the multiplicity of infection...45

Table 3.1. Multivariate analysis testing the association between sexually transmitted infections or bacterial vaginosis, and multiple HIV infection..62

Table 3.2. Results of a multivariate model testing the association between multiplicity of infection and viral load, taking HLA alleles associated with slow and rapid progression into account. .66

Table 3.3. Results of a multivariate model testing the association between multiplicity of infection and CD4+ T cell counts, taking HLA alleles associated with slow and rapid progression into account...66

Table 3.4. Results of a Chow test comparing linear regressions of viral load and CD4 counts prior to and following superinfection in three participants. 68

Table 4.1. Results of a multivariate model testing the association between dual infection and the development of neutralizing antibody breadth taking into account set point viral load and CD4 count..82

List of Abbreviations

ACD	Acid citrate dextrose
ARV	Anti-retroviral
AIDS	Acquired Immunodeficiency Syndrome
bnAbs	Broadly neutralizing antibodies
BV	Bacterial vaginosis
CAPRISA	Center for the AIDS Programme of Research in South Africa
cDNA	Complementary deoxyribonucleic acid
CTL	Cytotoxic T lymphocyte
DEAE-Dextran	Diethylaminoethyl-Dextran
dNTP	Deoxynucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
EIA	Enzyme immunoassay
FBS	Fetal bovine serum
GALT	Gut-associated lymphoid tissue
HIV	Human Immunodeficiency Virus
HLA	Human leucocyte antigen
HMA	Heteroduplex mobility assay
ID_{50}	50% inhibitory dose
KIR	Killer-cell immunoglobulin-like receptor
LPS	Lipopolysaccharide
mAbs	Monoclonal antibodies
MHC	Major histocompatibility complex
mOPV	Monovalent oral polio vaccine
mpi	Months post infection
MRCA	Most recent common ancestor
MSM	Men who have sex with men
nAbs	Neutralizing antibodies
NGS	Next generation sequencing
NK cells	Natural killer cells
OPV	Oral polio vaccine
PCR	Polymerase chain reaction
/py	Per person year
PI	Primary infecting virus
RLU	Relative light unit
RNA	Ribonucleic Acid
RPR	Rapid Plasma Reagin test
SGA	Single genome amplification
SHIV	Chimeric SIV/HIV virus
SI	Superinfecting virus
SIV	Simian Immunodeficiency Virus
SNP	Single nucleotide polymorphism
ssSGA	Strain-specific single genome amplification
STI	Sexually transmitted infections
TCID	Tissue culture infectious dose
T/F	Transmitted/founder
TLR	Toll-like receptor
tMRCA	Time since the most recent common ancestor
tOPV	Trivalent oral polio vaccine
wpi	Weeks post infection
ypi	Years post infection

Abstract

A better understanding of how neutralizing antibody responses develop during natural HIV-1 infection, and the factors that augment them, would be invaluable in informing vaccine design and testing. The use of polyvalent immunogens is a common strategy used in an attempt to enhance the breadth of vaccine-elicited immune responses and HIV dual infection (infection by >1 distinct HIV strain) provides a unique *in vivo* model to characterize how the immune system responds to multiple HIV strains.

In the CAPRISA 002 cohort from KwaZulu Natal, South Africa eleven cases of dual infection were detected, of whom eight were co-infected at the first seropositive visit and three were superinfected (infected with a second strain following seroconversion to an initial infection). The incidence of superinfection was ~5.47% per person year (in a cohort with a primary incidence of infection of 7.2% per person year), although systematic screening to accurately estimate the incidence of superinfection was beyond the scope of this project. To establish the clinical consequences of dual infection, we compared the viral loads, and CD4+ T cell counts of dual infected participants to 18 singly infected participants from the same cohort. Neither dual infection, nor *envelope* diversity in early infection was significantly associated with accelerated disease progression. Further, a multivariate analysis suggested that infection with *Neisseria Gonorrhoeae* was associated with multiple infection, suggesting *N. Gonorrhoeae* increased the susceptibility to multiple infection, although this did not reach statistical significance (p=0.0566).

Single genome amplification (SGA) derived sequences of the early superinfecting viral quasispecies were generated. In all three superinfected participants identified, a single founder variant likely established superinfection. Similar to transmitted/founder viruses from primary HIV infection, these superinfecting viruses had significantly shorter, less glycosylated variable

regions compared to viruses sampled at three years post infection from participants in the same cohort.

Neutralization breadth data at three years post infection was available for all 11 of the dual infected participants, as well as for 16 singly infected controls (1). Comparing the breadth revealed that dual infection was not significantly associated with greater neutralizing antibody breadth, even after controlling for viral loads and CD4+ T cell counts. However, one superinfected participant, CAP256, developed a broad, and extremely potent neutralizing antibody response (2). Using computational methods, the epitope targeted by CAP256 was predicted to be highly dependent on residues 166 and 169, overlapping epitopes recognized by monoclonal antibodies PG9/PG16. This was later validated, and the epitope comprehensively mapped (2). Longitudinal sequencing of *envelope* revealed that escape from this potent and broad antibody response occurred via multiple pathways, and was accelerated by recombination. Further, we show that the targeted epitope was not conserved in the primary infecting virus, suggesting that superinfection did not focus the neutralizing antibody response onto an epitope conserved in both infecting viruses.

Taken together, these results suggest that particular sexually transmitted infections might enhance susceptibility to HIV infection. Further, they show that exposure to two divergent HIV variants is not sufficient to broaden the neutralizing antibody response, and raise doubts in the potential efficacy of polyvalent HIV immunogens. They also suggest the clinical consequences of dual infection, if any, are not consistently severe. However, larger studies will be required to definitively address the clinical and immunological consequences of HIV co- and superinfection, and to address the correlates of risk of HIV superinfection.

CHAPTER 1: Background

Disease burden and the need for an HIV vaccine

In 2012 alone, it is estimated that approximately 2.3 million people were newly infected with the Human Immunodeficiency Virus (HIV), bringing the total number of HIV infected individuals worldwide to an estimated 35.3 million (3). Although anti-retroviral drugs (ARVs) have dramatically improved the prognosis of HIV infection, they are costly, and unavailable to many who need them. Estimates are that antiretroviral coverage in low- and middle-income countries is less than 35% of those who need them (according to the 2013 WHO guidelines). Safe and effective prophylactic measures are therefore desperately needed to curb the pandemic. While microbicides, and ARVs for pre- and post-exposure prophylaxis can be effective in reducing the risk of HIV transmission (4–6), the efficacy of these approaches are highly dependent on treatment adherence. The persistently high HIV transmission rates over the duration of the pandemic despite the availability of highly protective, relatively inexpensive prophylaxis in the form of condoms stresses the need for prophylactic measures that are not dependent on continuous and active adherence as an intervention.

Vaccines remain the most effective tool available to control infectious diseases. Vaccines are largely responsible for the eradication of smallpox, the near-eradication of poliomyelitis, and for significant reductions in the incidence of other infectious diseases such as tetanus and measles. However, an effective HIV vaccine remains elusive. Except for the recent RV144 vaccine trial in Thailand which had a modest protective effect (~31%) (7), empirical efforts to date have not been able elicit protective responses against HIV. Even in the RV144 trial, although protection appears to correlate with V2-dependent antibodies (8), the mechanism of this protection is not known. Therefore, basic research into features of the immune response to HIV and their relevance to vaccine design is needed.

HIV disease course

Sequencing of viral populations shortly after transmission has revealed that in ~80% of mucosal transmissions, a single virus establishes a productive infection (9–11). Following sexual transmission and crossing of the mucosal barrier, HIV replicates locally in the mucosa for a few days, but viral RNA remains at undetectable levels in the plasma for ~5 days (95% CI: 3-8 days) (12). During this so-called “eclipse phase”, the virus is disseminated from the loci of infection to draining lymph nodes, wherein numerous activated CD4+ CCR5+ T cells, susceptible to HIV infection, are located. Thereafter, the virus spreads throughout the lymphoid tissues. The gut-associated lymphoid tissue (GALT) is a major site of HIV replication, with an abundance of activated memory T cells (13, 14). The high level of replication in the lymphoid tissues is associated with an exponential rise in viral load, which peaks ~3-4 weeks following transmission. The depletion of CD4+ T cells in the GALT is substantial, and while the number of CD4+ T cells in the plasma can rebound following the spike in viremia, CD4+T cell numbers in the GALT do not recover (13–15). Thereafter, viral load decreases and reaches a relatively stable level (set point). Importantly, the viral load at set point is a strong predictor of the rate of progression to Acquired Immunodeficiency Syndrome (AIDS) (16). In the absence of ARV treatment, with continuous viral replication the numbers of CD4+ T cells are gradually depleted, until eventually the integrity of the immune response is compromised such that it cannot adequately maintain effective immune responses and individuals are consequently susceptible to opportunistic infections, characteristic of AIDS. While some HIV-infected individuals develop AIDS soon after infection, others are able to spontaneously control virus replication without ARV treatment.

The role of cellular immune responses in the control of HIV replication

Early studies into the role of immune responses in HIV control investigated cases of HIV infected individuals who were found to naturally control viral replication. Studies into these so-called “elite controllers” (individuals able to control viral loads to below 50copies/ml in the absence of ARV therapy) and “long term non-progressors” (individuals able to control viral

loads to between 50 and 2000 copies/ml in the absence of ARV therapy) are of great interest, as correlates of these phenotypes may hold the key to understanding what kinds of immune responses provide protection against HIV replication and disease progression.

Much of the evidence to date implicates a robust response by CD8+ cytotoxic T cells (CTL responses) in the control of HIV replication. Evidence for the role of CTL responses includes the temporal correlation that exists between the onset of CTL responses in peripheral blood, and the initial decline in peak viremia during primary infection (17, 18). Mathematical modelling of the early CTL responses to HIV estimates that a single CTL response could contribute as much as 15-35% to the initial decline in viral load (17). Further evidence is provided by experiments in non-human primates, where the reduction in peak viremia was highly dependent on CD8+ cells [both CTLs and a population of natural killer (NK) cells are CD8+], as depletion of these cells with anti-CD8 monoclonal antibodies (mAbs) resulted in a loss of control of Simian Immunodeficiency Virus (SIV) replication and a lack of decline in viral load (19–21). However, depletion of CD8+ T cells was also associated with a rebound in cell proliferation, including CD4+ T cells, that may have contributed to the temporary increase in viral load.

The importance of CTL responses in the control of viral replication is supported by a strong association between particular Human Leucocyte Antigen (HLA) class I alleles and disease progression (reviewed in 22). Genome-wide association studies (GWAS) have convincingly shown, with striking concordance in several populations, that the HLA class I locus is by far the most significant host factor associated with the control of HIV viral loads (23–27). Specifically, a single-nucleotide polymorphism (SNP) in the HLA complex P5 (*HCP5)* which is in high linkage disequilibrium with HLA B*57:01, an intronic variant of the *HLA-B* gene linked to HLA B*57:03 alleles, and variants ~35kb upstream of *HLA-C*, were identified as the polymorphisms most significantly associated with viral load. This is consistent with the previously identified association of HLA B (28), and HLA B*57 alleles in particular (29, 30), with the control of HIV replication. In a more recent GWAS of a large Caucasian cohort of 2554 HIV infected individuals,

86 other SNPs were significantly associated with lower set point viral loads after correcting for multiple testing across the whole genome, and all were located within the MHC (23). There is also evidence that NK cells, through interactions of the killer cell immunoglobulin-like receptor (KIR) in association with certain HLA alleles may, in part, be responsible for the association of certain HLA alleles with improved control (31).

Epitopes targeted by CTL responses are under strong selective pressure during infection (18, 32). Mutations in, or flanking these epitopes can disrupt antigen processing or recognition via a number of mechanisms, and consequently can prevent recognition by CTLs (CTL escape) (17, 32–37). In line with the role of CTL responses in controlling viral load, escape from CTL responses has been temporally correlated with a temporary loss of immune control of HIV replication (38). Furthermore, escape from CTL responses often comes at a fitness cost that is particularly evident in the escape from responses restricted by HLA alleles associated with better control (such as the B*57 and B*81 supertypes) (39–42). However, HIV is often able to restore fitness with compensatory mutations (43, 44).

Taken together, these data implicate CTL responses in the control of HIV, and further show that certain CTL responses are more effective than others. The effectiveness of CTL responses is attributable, in part, to the functional constraints of the epitopes targeted, and consequently the fitness costs required to escape the response.

The role of neutralizing antibodies in the control of HIV replication

In contrast to CTL responses, which are first detected as early as several days before peak viral load (17), neutralizing antibodies (i.e. antibodies that are able to prevent entry into cells) are only detectable after viral load has stabilized following primary infection (usually ~12-20 weeks post infection), suggesting they are not relevant for the initial control of viremia. While the earliest B cell responses are evident within eight days following detectable plasma viremia (initially in the form of immune complexes, and approximately five days later as free plasma antibodies), they are non-neutralizing, and mathematical modeling found little or no effect of

these antibodies on the control of acute-phase viremia (45). Further, results from experimental depletion of B cells using anti-CD20 mAbs in non-human primate models have been less clear than those from T cell depletion. While B cell depletion in rhesus monkeys did not have an effect on the initial decline in peak viremia, there was an inverse correlation between neutralizing antibody titers and subsequent viral load, suggesting that neutralizing antibodies do contribute to viral control (46). B cell depleted monkeys also had an accelerated progression to AIDS (Acquired Immunodeficiency Syndrome) (47). In contrast, however, others found that B cell depletion in rhesus macaques (48), and African green monkeys (49) did not significantly affect acute or chronic SIV viral loads. A recent case study of an HIV infected individual that underwent B cell depletion (as part of treatment for lymphoplasmacytoid lymphoma) showed a temporal correlation between neutralizing titers and plasma viral load, suggesting that neutralizing antibodies may indeed play a significant role in the control of HIV replication in humans (50). However, as B cells have a multifaceted role in humoral and cellular responses, B cell depletion may have affected immune functions other than neutralizing antibody titers.

Passive transfer of neutralizing antibodies provides a more direct means of evaluating the role of neutralizing antibodies in controlling replication. Treatment of macaques with SIV-specific IgG seven days after infection resulted in lower viral load set point (51), and in humans, passive transfer of neutralizing antibodies delayed the rebound of viremia following ART interruption in some individuals (52). However, from this latter study, it has since been revealed that titers 10- to 100-fold higher than the autologous titers present were required to produce a clinical benefit (53). Passive immunotherapy of HIV-1 infected humanized mice with multiple monoclonal antibodies that each target relatively conserved epitopes was also able to suppress viral loads to below detection (54).

Neutralizing antibodies exert significant selective pressure on HIV during natural infection (55–57). The HIV Envelope glycoproteins gp120 and gp41, which form trimeric “spikes” of gp120-gp41 heterodimers, are responsible for mediating viral entry into cells, and are the only

known targets of HIV neutralizing antibodies. The neutralizing antibody response to HIV develops several months after infection (55–59), and much of the extensive genetic variation in Envelope that follows in early infection is thought to be due to escape from antibody pressure (55, 56). While the early neutralizing antibody response is often potent (55–58), Envelope is able to tolerate a large amount of diversity. Together with an error-prone reverse transcriptase, HIV can readily mutate, allowing rapid escape from these neutralizing antibody responses. This is evident in observations that most HIV variants in early infection are resistant to neutralization by contemporaneous plasma (56, 57, 59, 60). However, variants remain sensitive to plasma from later timepoints, indicating that *de novo* responses are subsequently generated against escape mutants, highlighting the dynamic nature of the neutralizing antibody response to HIV (59, 61).

Escape from antibody neutralization can occur via typical point mutations (59, 62), or insertions and deletions (63, 64). Furthermore, the HIV Envelope is extensively glycosylated (with approximately half of its mass attributable to glycans, (65)), such that much of the surface of Envelope is decorated with relatively non-immunogenic glycans (see Figure 1.1; glycans are shown in purple). These glycans are large and therefore provide a means to shield underlying epitopes. It is not surprising, therefore, that HIV also appears to utilise an "evolving glycan shield" to escape neutralizing antibody responses, whereby mutations result in shifting patterns of glycosylation (56, 63).

Escape from early neutralizing antibody responses appears to be primarily mediated by mutations in V1V2, C3V4, and to a lesser extent, V5 (reviewed in 66). While mutations in V4 and V5 can apparently mediate escape from neutralizing antibodies, they themselves do not appear to be direct targets of early neutralizing antibodies (59, 60). On the other hand, V1V2, the membrane proximal external region (MPER) of gp41, and in the case of subtype C, C3 (particularly the alpha-2 helix), appear to be major targets of the early neutralizing antibody response (58–60, 67, 68). While it is evident that V1V2 is a direct target of the neutralizing

antibody response, the ability of the V1V2 loop to shield other neutralization determinants is also well documented (64, 69). Interestingly, while anti-V3 antibodies are generated in early infection (45), they do not appear to contribute to autologous neutralization (59), potentially due to the conformational shielding of V3 in the context of the native trimer.

By characterising the specificities of the early neutralizing antibody response to HIV, it has become apparent that this early response frequently comprises of only one or a few specificities. Further, when more than one specificity were evident; they appeared sequentially (59, 60). An in depth analysis of the dynamics of the early neutralizing antibody response, escape, and viral control in one individual showed that the onset of the neutralizing antibody response was associated with a 7-fold drop in viral load (59). However, approximately seven weeks and relatively low antibody titers (ID_{60} titer > 45) were required to drive selection of escape variants. Escape from this response was temporally correlated with a four-fold increase in viral load, suggesting a temporary loss of immune control.

Taken together, these data suggest that while neutralizing antibodies may exert some control over HIV replication, titers present during natural infection are likely too low to be effective to prevent disease progression in the face of a rapidly escaping virus. Consistent with this, reducing viral load and diversity with ARV pre-treatment enabled passively transferred neutralizing antibodies to control HIV replication in humanized mice (70).

Neutralizing antibodies can prevent infection

While even potent neutralizing antibody responses do not prevent disease progression in HIV-infected individuals (71–73), there is significant evidence that suggests that neutralizing antibodies, if present at the time of challenge, could prevent infection. The most compelling of which is that passive immunotherapy can protect non-human primates against challenges with SIV, and SIV-HIV chimeric viruses (SHIV) (74–81). While initial studies utilized high antibody titers targeted against the challenge strains, collectively, these results provided a proof of concept that neutralizing antibodies can provide sterilizing immunity against primate

immunodeficiency viruses. Interestingly, a monoclonal antibody, b12, provided macaques with greater protection against SHIV challenge compared to a b12 variant with similar neutralizing activity but deficient in Fc mediated effector function (75). However, passive immunotherapy with weakly- and non-neutralizing antibodies provided little or no protection from SHIV challenge, emphasizing that vaccine research should still be focused on eliciting potent neutralizing antibodies (82), the activities of which may be enhanced by Fc mediated functions.

Studies in non-human primates have also, importantly, provided quantitative data on the levels of neutralizing antibodies that would likely be required to provide protection. While modest levels of neutralizing antibodies typical of those commonly found in HIV infected individuals could provide partial protection and increase the number of challenges required to infect macaques (75), high levels of neutralizing antibodies have been estimated to be required to provide complete protection (74, 83, 84). In some (74), although not all (75) cases, titers present at the time of challenge also correlated with the susceptibility to infection. While it is difficult to extrapolate these data directly to humans, taken together, they do suggest that a protective vaccine will likely need to elicit long-lived, high titer neutralizing antibody responses. Therefore, understanding factors that enhance B cell responses and the development of neutralizing antibodies is an important objective.

Given the extraordinary diversity inherent in the circulating HIV variants, vaccines must not only elicit relatively potent responses, but must also elicit antibodies able to neutralize a broad range of variants (broadly neutralizing antibodies) at these titers. However, the early neutralizing antibody response is invariably highly specific for an individual's own virus (autologous virus), and is unable to neutralize variants that the immune system has not been exposed to (heterologous viruses). This is perhaps unsurprising given the highly variable nature of the major targets of the early autologous neutralizing antibody response (described above). Nevertheless, a number of studies that evaluated the breadth of the neutralizing antibody response in HIV infected individuals have revealed that a subset of chronically infected

individuals possess broadly neutralizing antibodies. While the definition of breadth was not entirely consistent between studies, it is estimated that ~10-30% of HIV infected individuals develop moderate to broadly neutralizing activity during the course of natural HIV infection (1, 85–87), providing hope that a vaccine could elicit similar responses. While in some of these cases, breadth was mediated by several different specificities, each responsible in part for the breadth (88), there is also evidence that a single antibody specificity can mediate this breadth (1, 87, 89, 90). The identification of epitopes targeted by these broadly neutralizing antibodies can identify vulnerable regions in the HIV Envelope that can be targeted in vaccine design.

Targets of broadly neutralizing antibodies

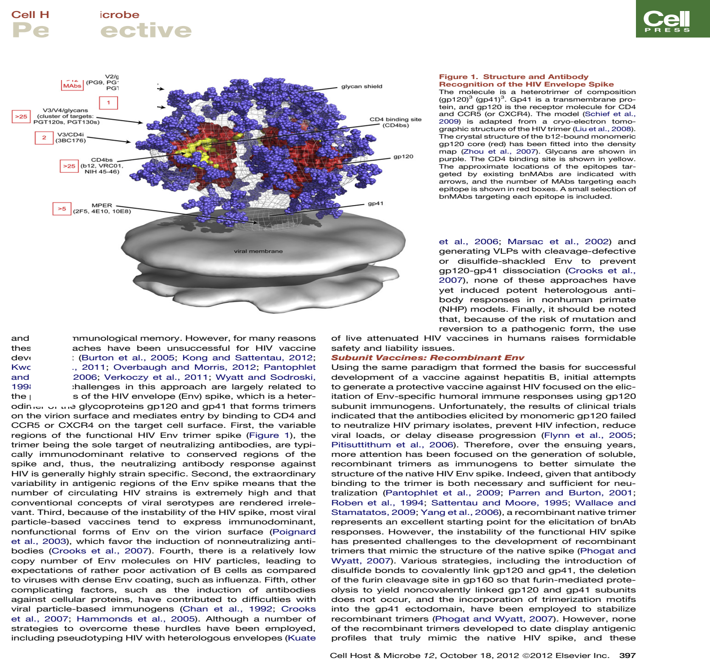
The HIV Envelope presents a formidable target for neutralizing antibodies. In addition to the extensive glycosylation of the native Envelope trimer, immunogenic non-native forms of Envelope are also present on the surface of virions that elicit non-neutralizing antibodies at the expense of neutralizing responses (91). Native trimers are few and far between, with on average only ~14 spikes on the surface of each virion (92). This scarcity of Envelope spikes on the surface of virions effectively prevents bivalent binding by antibodies, and consequently minimises avidity, which antibodies use to achieve high affinity binding, and potent neutralization (see 93). HIV Envelope is also hypervariable, such that conserved epitopes that could potentially be targets for broadly neutralizing antibodies are also relatively scarce. Regions that are highly conserved are generally conformationally masked and consequently difficult for antibodies to access.

Before 2009, only a few broadly neutralizing monoclonal antibodies had been isolated from HIV infected individuals (namely 2F5, 4E10, Z13, b12, and 2G12). Renewed interest and the advent of new high throughput techniques since have resulted in the identification of a number of new broadly neutralizing monoclonal antibodies, some of which display exceptional breadth and potency. The detailed characterisation of the epitopes targeted by these broadly neutralizing antibodies has identified conserved, accessible epitopes in the HIV-1 Envelope that

represent vulnerable regions that could be exploited by vaccines. These targets include (i) the CD4bs, (ii) a quaternary (i.e. preferentially expressed in trimeric Envelope) epitope dependent on residues in V2 and V3 (targeted by the so-called "PG9/16-like" antibodies), (iii) a cluster of glycans on the outer domain of gp120 (targeted by 2G12), (iv) an epitope in V3 including glycans (targeted by several of the PGT monoclonal antibodies; "PGT128-like"), and (v) the membrane-proximal external region (MPER) of gp41 (see Figures 1.1, 1.2 and Table 1.1). Studies that have mapped the specificities of broadly neutralizing sera have revealed that the quaternary V2, PG9/16-like epitope as well an epitope including the glycan at position 332 (overlapping both the 2G12 and PGT128-like epitope) are the most common targets of broadly neutralizing antibodies (1, 87). Importantly, in these studies, a fraction of broadly neutralizing activity could not be mapped to known specificities, indicating that at least one additional broadly neutralizing epitope awaits discovery.

Table 1.1 | Representatives of broadly neutralizing HIV monoclonal antibodies (mAbs)

mAb	Year of discovery	Target	Ref.
2F5 **4E10** **Z13e1** **10E8**	1993 1994 2001 2012	MPER	(94) (95) (96) (97)
2G12	1994	Glycan array near the base of V3	(98, 99)
b12 **HJ16**	1994 2010	CD4 binding site	(100) (101)
VRC01-VRC03 **VRC-PG04**	2010 2011	CD4 binding site	(89) (102)
PG9, PG16 **CH01-04** **PGT141-145**	2009 2011 2011	Quaternary epitope dependent on V1V2(V3), including glycans	(90) (103) (104)
PGT121-123, 125- 128, 130, 131	2011	Glycans and V3	(104)



Cell H icrobe
Pe ective

Cell PRESS

MAbs (PG9, PG PG V2/g
1
>25 V3/V4/glycans (cluster of targets: PGT120s, PGT130s)
2 V3/CD4i (3BC176)
>25 CD4bs (b12, VRC01, NIH 45-46)
>5 MPER (2F5, 4E10, 10E8)
glycan shield
CD4 binding site (CD4bs)
gp120
gp41
viral membrane

Figure 1. Structure and Antibody Recognition of the HIV Envelope Spike
The molecule is a heterotrimer of composition $(gp120)^3$ $(gp41)^3$. Gp41 is a transmembrane protein, and gp120 is the receptor molecule for CD4 and CCR5 (or CXCR4). The model (Schief et al., 2009) is adapted from a cryo-electron tomographic structure of the HIV trimer (Liu et al., 2008). The crystal structure of the b12-bound monomeric gp120 core (red) has been fitted into the density map (Zhou et al., 2007). Glycans are shown in purple. The CD4 binding site is shown in yellow. The approximate locations of the epitopes targeted by existing bnMAbs are indicated with arrows, and the number of MAbs targeting each epitope is shown in red boxes. A small selection of bnMAbs targeting each epitope is included.

and nmunological memory. However, for many reasons thes aches have been unsuccessful for HIV vaccine dev : (Burton et al., 2005; Kong and Sattentau, 2012; Kwc ., 2011; Overbaugh and Morris, 2012; Pantophlet and 2006; Verkoczy et al., 2011; Wyatt and Sodroski, 199 challenges in this approach are largely related to the s of the HIV envelope (Env) spike, which is a heterodin of the glycoproteins gp120 and gp41 that forms trimers on the virion surface and mediates entry by binding to CD4 and CCR5 or CXCR4 on the target cell surface. First, the variable regions of the functional HIV Env trimer spike (Figure 1), the trimer being the sole target of neutralizing antibodies, are typically immunodominant relative to conserved regions of the spike and, thus, the neutralizing antibody response against HIV is generally highly strain specific. Second, the extraordinary variability in antigenic regions of the Env spike means that the number of circulating HIV strains is extremely high and that conventional concepts of viral serotypes are rendered irrelevant. Third, because of the instability of the HIV spike, most viral particle-based vaccines tend to express immunodominant, nonfunctional forms of Env on the virion surface (Poignard et al., 2003), which favor the induction of nonneutralizing antibodies (Crooks et al., 2007). Fourth, there is a relatively low copy number of Env molecules on HIV particles, leading to expectations of rather poor activation of B cells as compared to viruses with dense Env coating, such as influenza. Fifth, other complicating factors, such as the induction of antibodies against cellular proteins, have contributed to difficulties with viral particle-based immunogens (Chan et al., 1992; Crooks et al., 2007; Hammonds et al., 2005). Although a number of strategies to overcome these hurdles have been employed, including pseudotyping HIV with heterologous envelopes (Kuate et al., 2006; Marsac et al., 2002) and generating VLPs with cleavage-defective or disulfide-shackled Env to prevent gp120-gp41 dissociation (Crooks et al., 2007), none of these approaches have yet induced potent heterologous antibody responses in nonhuman primate (NHP) models. Finally, it should be noted that, because of the risk of mutation and reversion to a pathogenic form, the use of live attenuated HIV vaccines in humans raises formidable safety and liability issues.

Subunit Vaccines: Recombinant Env
Using the same paradigm that formed the basis for successful development of a vaccine against hepatitis B, initial attempts to generate a protective vaccine against HIV focused on the elicitation of Env-specific humoral immune responses using gp120 subunit immunogens. Unfortunately, the results of clinical trials indicated that the antibodies elicited by monomeric gp120 failed to neutralize HIV primary isolates, prevent HIV infection, reduce viral loads, or delay disease progression (Flynn et al., 2005; Pitisuttithum et al., 2006). Therefore, over the ensuing years, more attention has been focused on the generation of soluble, recombinant trimers as immunogens to better simulate the structure of the native HIV Env spike. Indeed, given that antibody binding to the trimer is both necessary and sufficient for neutralization (Pantophlet et al., 2009; Parren and Burton, 2001; Roben et al., 1994; Sattentau and Moore, 1995; Wallace and Stamatatos, 2009; Yang et al., 2006), a recombinant native trimer represents an excellent starting point for the elicitation of bnAb responses. However, the instability of the functional HIV spike has presented challenges to the development of recombinant trimers that mimic the structure of the native spike (Phogat and Wyatt, 2007). Various strategies, including the introduction of disulfide bonds to covalently link gp120 and gp41, the deletion of the furin cleavage site in gp160 so that furin-mediated proteolysis to yield noncovalently linked gp120 and gp41 subunits does not occur, and the incorporation of trimerization motifs into the gp41 ectodomain, have been employed to stabilize recombinant trimers (Phogat and Wyatt, 2007). However, none of the recombinant trimers developed to date display antigenic profiles that truly mimic the native HIV spike, and these

Cell Host & Microbe *12*, October 18, 2012 ©2012 Elsevier Inc. 397

Figure 1.1 | Targets for broadly neutralizing HIV antibodies. A model of the HIV-1 gp160, unliganded trimer spike is shown, along with the targets of broadly neutralizing antibodies. Examples of mAbs targeting each site are shown in parentheses. The gp120 core, shown in red with the b12 epitope overlapping the CD4 binding site (CD4bs) highlighted in yellow, is fitted in an cryo-electron tomography density map (grey mesh). Estimated N-linked glycosylation is shown in purple (glycan shield). CD4bs, CD4-binding site; MPER, membrane proximal external region; CD4i, CD4 binding-induced epitope. Figure adapted from (195)**, which used a model from** (196).

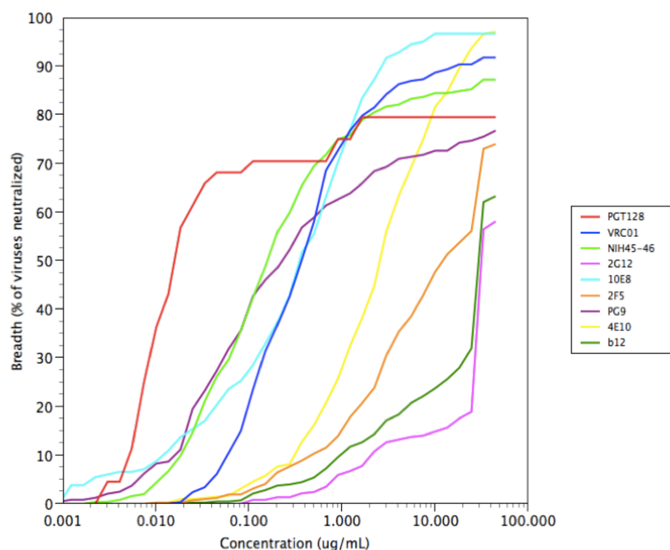


Figure 1.2 | Bread-potency curves for nine broadly neutralizing HIV antibodies. Curves were generated from published data in HIV Antibody Database (197)**. The lower the concentration required to neutralize, the more potent the antibody.**

Development of broadly neutralizing antibodies

Understanding how broadly neutralizing activity develops during natural infection, and the factors that augment this process, may provide clues as to how vaccines might be able to elicit similar responses. While the initial neutralizing antibody response is highly specific for an individual's own virus and develops several months after infection, the ability to neutralize heterologous viruses increases over time and is typically evident only years (mean ~2.5 years) after infection (1, 105). Further, although broadly neutralizing antibodies (bnAbs) can be found in individuals with low or undetectable viremia, albeit at a lower frequency (106), viral load has been consistently associated with the development of enhanced neutralization breadth in a number of cohorts (1, 72, 105, 107). The dependency on time and high viral loads suggests that chronic antigen exposure is required for the development of broadly cross-reactive antibodies. Cloning and sequencing of the immunoglobulin genes from B cells expressing broadly neutralizing antibodies has also revealed that they are typically highly altered from their genomic precursors by somatic hypermutation (89, 90, 108, 109), suggesting a role for affinity maturation in the development of cross-reactivity. Further, the broadly neutralizing antibodies characterized to date appear to each possess unusual physical features that are important determinants of their neutralization capacity, such as a long heavy chain complementarity-determining region 3 (CDR H3), sulphated tyrosines (eg. PG9, PG16, CH01-04, PGT141-145), heavy chain domain swaps [exchange of the heavy-chain variable (VH) domains on the adjacent antigen-binding arms of the antibody (eg. 2G12)], or apparent self- or poly-reactivity and membrane binding (eg. 2F5, 4E10, b12, 2G12) (reviewed in 110). This suggests that these broadly neutralizing antibodies may have been generated as a result of atypical pathways, and may account in part for the fact that most infected individuals do not elicit broadly neutralizing antibodies, as well as the appearance of these antibodies only after years of infection in those that do. Whether extensive affinity maturation and unique physical features are necessary for broad neutralization of HIV remains to be determined. However, if they are, eliciting broadly neutralizing antibodies by typical vaccination strategies will prove difficult, and may require

novel and extended vaccination regimens in addition to novel immunogens. Pre-clinical testing protocols may also need to allow for extended follow up following vaccination.

Polyvalent vaccines, and immunodominance

The use of polyvalent vaccines may be critical to an effective HIV vaccine, with the hope that exposure of the immune system to multiple HIV antigens will result in immune responses that are effective against a broader range of variants. This approach has been used successfully in vaccines against Polio, Influenza*, and Streptococcus pneumonia*.

However, one potential shortcoming in the use of polyvalent immunogens in HIV vaccines is the propensity of the human immune system to respond to only a fraction of potentially immunogenic determinants, a phenomenon known as "immunodominance". Some mechanisms responsible for influencing the immunodominance of the CTL response have been identified and include preferential antigen processing, HLA affinity and affinity for the cognate TCR, as well as competition between CD8 T cells for dendritic cells (111, 112). While, immunodominance is also a prominent feature of the neutralizing antibody response to HIV, the mechanisms are less clear than those for CTL responses. In acute infection, HIV-infected individuals develop a limited number of oligoclonal B cell responses that are initially generated against one or a few immunodominant epitopes within Envelope (59, 60). These epitopes are generally regions that can tolerate a large amount of genetic variability (and consequently can be highly variable among circulating variants), and therefore while responses to these epitopes may be effective against the variant that elicited them, they are not protective against heterologous variants (type-specific responses). It is unclear whether immunodominance will also extend to the responses to multiple Envelopes and therefore would diminish the efficacy of polyvalent HIV vaccines incorporating Envelope.

To date, vaccines have been unsuccessful in eliciting broad or potent neutralizing antibody responses. The diversity inherent in hypervariable, immunodominant epitopes in Envelope is overwhelming such that incorporating sufficient diversity in a polyvalent formulation to

provide significantly broad protection may not be feasible. This is supported by the fact that polyvalent vaccines have thus far not been able to elicit any significant neutralizing antibody titers against tier 2 viruses (representative of primary variants, as opposed to easy to neutralize laboratory-adapted " tier 1" strains) not incorporated in the vaccine, in animal models or in humans, despite eliciting responses capable of neutralizing several variants that were incorporated in the vaccine (113–119), although this may be confounded by inadequacies in pre-clinical testing. Furthermore, an eight-valent formulation was previously shown to be less immunogenic than a three-valent one, providing evidence that higher valencies may not necessarily be able to ceaselessly elicit greater specificities (120). Consequently, a key feature of an effective neutralizing antibody-based vaccine will likely rest on the ability to prevent the generation of immunodominant, type-specific responses typically seen during acute infection, and to focus the immune response onto conserved, sub-dominant epitopes such as those targeted by broadly neutralizing antibodies.

While polyvalent vaccination is being evaluated for the potential to enhance the number of specificities generated as a means to enhance breadth, the potential of polyvalent immunogens to focus the immune response to conserved epitopes remains largely, although not completely, unexplored (121, 122). Boosting of an *envelope* DNA-primed (rabbit) immune system with V3 fusion peptides resulted in >80% of the elicited neutralizing antibodies being specific for epitopes in V3 (122). Furthermore, sequential immunisation of rhesus macaques with a quasispecies partially mimicked the response to that quasispecies seen in natural infection (121). Taken together, these results suggest that it may be possible to focus the immune response on a particular epitope within Envelope, or to direct the maturation of an antibody specificity in a vaccine setting (121, 122). However, it is unclear whether the immune response to a heterologous boost will be preferentially targeted against epitopes present in the priming variants. If so, this would provide a means of eliciting responses to conserved epitopes.

Dual infection

HIV vaccine research is further complicated by the fact that the correlates of protection in humans are largely unknown. However, it seems likely that vaccines will need to elicit better responses than those generated in natural infection. This is indicated by the fact that HIV infected individuals can still be "superinfected" with a second variant after an established initial infection.

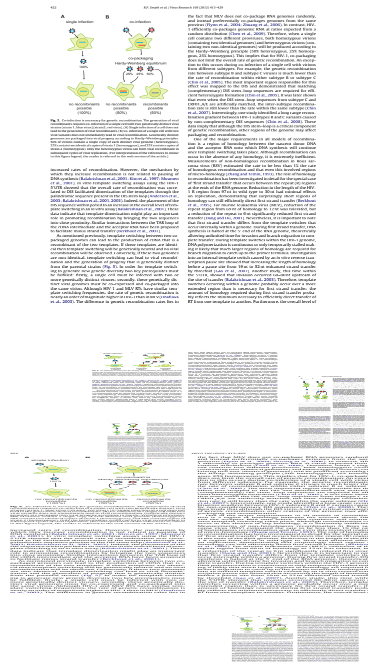
Superinfection, as opposed to co-infection where both variants are transmitted prior to seroconversion, provides an opportunity to address the correlates of protection as a virus is transmitted in the face of an on-going anti-HIV immune response. Several population-based studies found the incidence of superinfection to be similar to that of the incidence of primary infection in the same, or comparatively similar cohorts (123–125), and suggested that immune responses generated during natural infection likely provided limited or no protection against subsequent superinfection. However, other studies have reported incidences of superinfection to be less than that of primary infection (126–129).

Dual infection (infection with two or more distinct variants, regardless of timing i.e. encompassing both co- and superinfection) is also of clinical relevance, as it has previously been associated with accelerated disease progression (130, 131). As HIV is diploid (virions encapsidate two genome copies), if two variants infect the same individual, and the same cell, the genomes from each variant can subsequently be packaged into the same virion. Upon infection of another cell, during the reverse transcription of the HIV genome, reverse transcriptase can switch between templates thereby resulting in the generation of recombinant viruses, whose genomes are a mosaic of the two different variants (Figure 1.3). The potential for HIV to recombine is highlighted by the high prevalence of circulating and unique recombinant inter-subtype recombinant forms. In addition to their value in addressing the correlates of protection, and factors influencing the rate of progression to AIDS, dual infection also provides a unique opportunity to evaluate how the immune system responds to multiple,

highly diverse HIV antigens. However, no studies to date have comprehensively evaluated the immune responses in HIV dual-infected individuals.

A. **B.**

RT



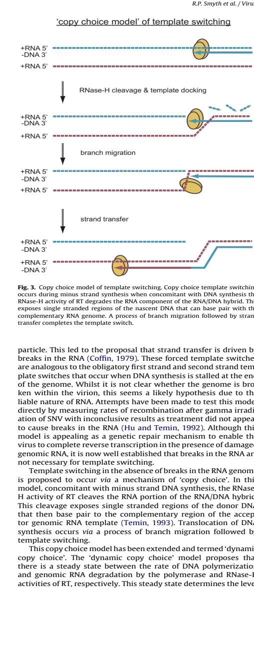


Fig. 3. Copy choice model of template switching. Copy choice template switching occurs during minus strand synthesis when concomitant with DNA synthesis the RNase-H activity of RT degrades the RNA component of the RNA/DNA hybrid. This exposes single stranded regions of the nascent DNA that can base pair with the complementary RNA genome. A process of branch migration followed by strand transfer completes the template switch.

particle. This led to the proposal that strand transfer is driven by breaks in the RNA (Coffin, 1979). These forced template switches are analogous to the obligatory first strand and second strand template switches that occur when DNA synthesis is stalled at the end of the genome. Whilst it is not clear whether the genome is broken within the virion, this seems a likely hypothesis due to the liable nature of RNA. Attempts have been made to test this model directly by measuring rates of recombination after gamma irradiation of SNV with inconclusive results as treatment did not appear to cause breaks in the RNA (Hu and Temin, 1992). Although this model is appealing as a genetic repair mechanism to enable the virus to complete reverse transcription in the presence of damaged genomic RNA, it is now well established that breaks in the RNA are not necessary for template switching.

Template switching in the absence of breaks in the RNA genome is proposed to occur *via* a mechanism of 'copy choice'. In this model, concomitant with minus strand DNA synthesis, the RNase-H activity of RT cleaves the RNA portion of the RNA/DNA hybrid. This cleavage exposes single stranded regions of the donor DNA that then base pair to the complementary region of the acceptor genomic RNA template (Temin, 1993). Translocation of DNA synthesis occurs *via* a process of branch migration followed by template switching.

This copy choice model has been extended and termed 'dynamic copy choice'. The 'dynamic copy choice' model proposes that there is a steady state between the rate of DNA polymerization and genomic RNA degradation by the polymerase and RNase-H activities of RT, respectively. This steady state determines the level of template switching (Svarovskaia et al., 2000). The model predicts that exposure of single stranded DNA increases when the RNase-H activity of RT is enhanced compared to the DNA polymerase activity. This leads to an increase in the rate of template switching by facilitating base pairing between the donor and acceptor genomic template. Conversely, a low RNase-H to DNA polymerase activity reduces availability of single stranded regions of DNA for binding to the complementary RNA genome, leading to a decrease in the rate of template switching. The 'dynamic copy choice' model can be used to explain the experimental observations that led to the development of the 'forced copy choice' and 'copy choice' models. The predictions of this model have been rigorously tested, and are generally accepted to be the most complete explanation of minus strand template switching (Delviks-Frankenberry et al., 2007; Nikolenko et al., 2004; Operario et al., 2006; Svarovskaia et al., 2000).

4.2. Genetic consequences of recombination

In contrast to the slow and steady change caused by mutation, recombination is a much more powerful evolutionary force (Fig. 4).

First, recombination facilitates the repair of viral genomes. This can be *physical* repair at genome breaks (Fig. 4A) *via* a mechanism of forced copy choice or *genetic* repair of genomes that have accumulated deleterious mutations (Fig. 4B). In the absence of recombination, organisms tend to accumulate deleterious mutations at each replication cycle *via* a process termed Muller's ratchet (Muller, 1964). The central tenants of Muller's ratchet are that (1) new mutations are more likely than reversions and (2) new mutations are more likely to be deleterious than beneficial. Consequently, with each replication cycle, the accumulation of mutations decreases the replication capacity of the organism until extinction occurs. Recombination can bypass Muller's ratchet by recreating mutation free individuals from a population of mutants (Kim et al., 2005).

Second, recombination can both *create* and *maintain* genetic diversity in a population (Fig. 4C and D). As the ability of a population to respond to natural selection is proportional to the genetic variance in that population, recombination can speed adaptation in the face of rapid evolutionary change (Fisher, 1958). During infection HIV is exposed to huge selective pressures by the host immune system. In order to survive, the virus must continually adapt to evade a rapidly changing immune response that continually generates new selective pressures through the generation of novel antibody and cytotoxic T-cell responses. This evolutionary arms race between host and parasite, is known as the 'Red Queen hypothesis' after the Red Queen in 'Through the Looking Glass' who stated that 'it takes all the running you can do to stay in the same place' (Van Valen, 1973). This need for continual evolution through natural selection for adaptation and counter-adaptation is thought to be the driving force behind the evolution of recombination. Recombination *creates* new diversity by shuffling pre-existing mutations within a population, generating complex combinations of mutations that would otherwise have to arise sequentially in an asexual organism (Fig. 4C). Recombination also *maintains* genetic diversity by breaking linkages between mutations (Fig. 4D). For example, if a viral genome containing a drug resistance mutation is targeted by a new CTL response, in the absence of recombination the CTL response will clear the beneficial mutation as it is physically linked to the genetic sequence under immune selection. Recombination ensures that the drug resistance mutation can be maintained within the population as long as it can recombine onto a genome containing an appropriate CTL escape mutation. Thus recombination ensures that different areas of the viral genome can evolve independently of one another.

Figure 1.3 | Schematic of HIV recombination. A. As HIV is diploid, co-infection of cells with two distinct virions can produce a proportion of heterozygous viruses (50%, assuming equal numbers and random pairing of each genome). Subsequent infection of cells by heterozygous virions can produce recombinant viruses. B. Copy-choice model of HIV recombination. During reverse transcription, RNAse H activity of Reverse Transcriptase (RT) degrades the RNA strand in a DNA-RNA hybrid. The resulting ssRNA can anneal to a complementary RNA leading to strand (and template) switching. Figure adapted from (198)**.**

Aims:

The use of polyvalent immunogens is a common strategy used in an attempt to enhance the breadth of vaccine-elicited immune responses and HIV dual infection provides a unique *in vivo* model to characterize how the immune system responds to multiple HIV strains. This study takes advantage of multiply HIV-infected participants to characterise the properties of the infecting viruses; the clinical consequences of dual infection; and evaluate the impact that exposure to two or more divergent HIV variants will have on the neutralizing antibody response. Taken together, these data will potentially shed light on the potential of polyvalent and prime-boost based vaccinations to influence the specificity of vaccine-elicited neutralizing antibody responses.

Objectives:

I. Identify cases of multivariant or dual HIV infection
    a. Identify and confirm cases of multivariant or dual HIV infection in the CAPRISA 002 cohort
    b. Characterise the infecting variants, and the relative timing of transmission events
    c. Estimate the incidence of superinfection based on available follow up and compare to the incidence of primary infection in the cohort.
    d. Model the statistical power to detect correlates of protection against superinfection
    e. Identify whether higher CD4+ T cell counts, or sexually transmitted infections predisposed participants to infection with multiple viruses

II. Characterise the impact of dual infection on viral diversification, and disease progression

III. Determine whether HIV superinfection promoted cross-neutralizing antibody responses

CHAPTER 2: Identification and genetic characterization of HIV-1 dual infection

2.1 Introduction:

A significant challenge facing the development of an HIV vaccine is that correlates of protection against HIV are not clear. Studies in non-human primates provide strong evidence that high titers of antibodies capable of neutralizing the challenge virus can provide protection. However, eliciting neutralizing antibodies against HIV by vaccination has proven to be a daunting challenge. Further, it is not clear whether results from these animal models will translate into humans. Results from the analysis of the RV144 vaccine trial have also suggested that non-neutralizing antibodies against V2 may have provided partial protection (8). However, large vaccine efficacy trials are time consuming and expensive, and models that increase our understanding of correlates of protection can potentially accelerate vaccine development. By definition, in HIV superinfection the secondary transmission occurs in the face of an ongoing anti-HIV immune response, and therefore provides a unique opportunity to address whether the immune responses typically elicited during natural HIV infection provide any protection against subsequent HIV infection.

Several studies have compared immunological markers (such as autologous and heterologous neutralizing antibodies, and non-neutralizing V1V2 antibodies) in superinfected participants to those who were presumably exposed, but not superinfected. In one study, three men who had sex with men (MSM), who subsequently became superinfected, were unable to neutralize two heterologous viruses (tier 1, easy to neutralize), compared to detectable neutralization of the same viruses in 9/11 (81.82%) matched, non-superinfected controls (132). Superinfected participants also displayed weaker autologous nAb responses compared to non-superinfected controls (although this was not statistically significant). Similarly, in a heterosexual cohort from Zambia (133), three superinfected participants had significantly

lower, more delayed neutralizing antibody titers to their primary virus than ten non-superinfected controls did. Together, these data suggest that stronger humoral responses to primary infection may be correlated with a reduced susceptibility to superinfection. In contrast, in a cohort of female sex workers from Mombasa, Kenya, no difference was evident in the breadth or potency of neutralizing antibody responses at the time of superinfection in six superinfected women compared to 18 matched, monoinfected controls (134).

In an attempt to establish whether the natural immune response to HIV affords any protection from superinfection, without the need to explicitly test a specific correlate, several studies have compared the experimentally determined incidence of superinfection to the incidence of primary infection in the same or similar cohorts. In a high-risk cohort of Kenyan women, superinfection was detected at a rate approximately half that of initial infection: ~4% per person year (/py), (123, 135) compared to a primary infection incidence of 8%/py (136). In a cohort of Californian MSM (124), the incidence of superinfection was estimated at 5%/py; comparable to the primary incidence in a comparable cohort (5%/py;(137, 138)). In a heterosexual cohort of 149 individuals from the Rakai district in Uganda screened with deep sequencing, the incidence of superinfection (1.44%/py) was comparable to the primary HIV incidence (1.15%/py) in the general population of Rakai (125). In contrast, two studies were not able to find any evidence of superinfection despite extensive follow up (127, 128). More recently, screening of high-risk women from Mombasa, Kenya with next generation sequencing (NGS) estimated the incidence of superinfection at less than half the primary incidence in the same cohort (126). This apparent contrast may be due, in part, to differences in study cohorts, methodology used, and the difficulties in identifying cases of superinfection. Superinfecting viral populations can potentially comprise the minor population and consequently can evade detection if the depth of sequencing is inadequate. Furthermore, recombination between variants can obscure the identification of superinfection if only small regions of the genome are sequenced, although how soon after superinfection recombination occurs is unclear.

Alternatively, immune responses elicited by primary HIV infection may afford protection against superinfection, albeit incomplete.

The utility of superinfection as a challenge model also depends on superinfection resembling primary transmission, though this has not been established. Primary HIV transmission results in a severe genetic bottleneck associated with the transmission of HIV. It is estimated that in the sexual transmission of HIV, ~80% of infections can be traced to a single founder variant (10, 11). These transmitted/founder (T/F) viruses are consistently CCR5-tropic (9, 10), and have previously been shown to have shorter, less glycosylated Envelopes compared to viruses from chronic infection (139, 140). However, the genetic bottleneck associated with superinfection, and whether the superinfecting viruses have phenotypes typical of primary viruses has not been addressed.

In this chapter, we characterize the multiplicity of infection in women from the CAPRISA002 cohort. We confirm three cases of intra-subtype superinfection, and isolate the *envelope* genes of the superinfecting T/F. Through the characterization of the properties of viruses responsible for superinfection, this chapter aims to contribute to our understanding of dual infection, and the utility of this model for understanding correlates of protection. Furthermore, it touches on the kinetics and extent of recombination following dual infection, and its potential to mislead analyses when not taken into account.

2.2 Methods:

2.2.1 Cohort

Samples were provided from participants from the CAPRISA 002 Acute Infection Study established in 2004 (141). High risk, HIV negative women from Durban, KwaZulu-Natal, South Africa were monitored monthly and enrolled upon detection of HIV infection. The demographics, self-reported sexual history and risk behavior are described elsewhere (141). HIV serostatus was determined using two rapid HIV-1 antibody tests (Determine, Abbott 89 Laboratories; Tokyo and Capillus, Trinity Biotech, Jamestown, NY). All antibody negative samples were tested for HIV-1 RNA by pooled PCR (Ampliscreen v1.5, Roche Diagnostics, Rotkreuz, Switzerland) and if positive were confirmed by an HIV enzyme immunoassay (EIA, BEP2000, Dade Behring, Marburg, Germany) and quantitating viral RNA (Cobas Amplicor HIV-1 monitor test version 1.5; Roche Diagnostics, Rotkreuz, Switzerland). Women were also enrolled from other seroincidence cohorts from Vulindlela, Kwa-Zulu-Natal, if they had a reactive antibody test within five months of a negative result, or if HIV-1 RNA was detectable by PCR in the absence of detectable HIV antibodies. The timing of infection was estimated as either the midpoint between the last antibody negative and first antibody positive visits, or 14 days prior to an RNA-positive, antibody negative sample.

Enrolled HIV positive participants were followed longitudinally, and plasma samples were taken weekly for three weeks, fortnightly until approximately three months post infection, monthly until approximately 1 year post infection, and quarterly thereafter. Plasma was stored in either EDTA, or ACD (acid citrate dextrose) to prevent coagulation, and stored at -80°C until use. Participants were initiated on ARV therapy once CD4 T cell counts fell below 200 cells/µl, and later 350 cells/µl, consistent with the prevailing South African ARV treatment guidelines. Ethical approval for this study was received from the ethics committees of the University of Cape Town, the University of KwaZulu-Natal, and the University of the Witwatersrand, and all participants provided written, informed consent.

2.2.2 PCR and Sequencing of envelope

Plasma viral RNA was extracted from 200 µl or 400 µl of plasma using either the Roche MagNApure (Roche Applied Science, Mannheim, Germany) or QIAgen viral RNA extraction kits (Qiagen, Valencia, CA, Valencia, CA). RNA was reversed-transcribed to cDNA on the same day as the RNA extraction, using Superscript III (Invitrogen, Life Technologies, Carlsbad, CA). Briefly, 50ul of extracted RNA was reverse transcribed in 1X Superscript RT buffer, with 0.2 mM of each dNTP, 1 µM of oligo-dT(20) (Integrated DNA Technologies, Germany), 5 mM Dithithreitol (Invitrogen, Life Technologies, Carlsbad, CA), 80 units of an RNAse inhibitor (RNAseOUT, Invitrogen, Life Technologies, Carlsbad, CA), and 400 units of reverse transcriptase (SuperScript III, Invitrogen, Life Technologies, Carlsbad, CA). The mixture of RNA, deoxynucleotide triphosphates (dNTPs) and oligo-dT was first heated to 65°C for 5 minutes to denature RNA secondary structures, then cooled to 4°C. The full reaction was then incubated at 45°C for 120 minutes for reverse transcription, and then heat inactivated at 75°C for 15 minutes. Reactions were then treated with 2 units of RNAse H (Invitrogen, Life Technologies, Carlsbad, CA), and digested at 37°C for 20 minutes to degrade RNA annealed to newly synthesized cDNA. The synthesized cDNA was either used immediately or stored at -60°C until use.

Envelope cassettes were amplified from the full-length cDNA in a nested PCR by Single Genome Amplification (SGA) using 0.025 units of Platinum Taq High Fidelity (Invitrogen, Life Technologies, Carlsbad, CA) per 20 µl reaction, as previously described (10, 11). Reactions were performed in 1X High Fidelity Platinum PCR buffer, with 2 mM MgSO4, 0.2 mM of each dNTP, and 0.2 µM of each primer. cDNA was serially diluted until <30% of the reactions were PCR positive, such that >80% of the amplicons likely would have been amplified from a single template (according to a Poisson distribution) (9). This approach prevents template resampling, and PCR-induced recombination that can occur if multiple templates are amplified in a reaction (9). Primers used in the outer reaction were Vif1 (5'-GGGTTTATTACAGGGACAGCAGAG -3'; HXB2 nt 4900 - 4923) and OFM19 (5'-

GCACTCAAGGCAAGCTTTATTGAGGCTTA -3'; HXB2 nt 9604 - 9632). Inner primers used were *ENV1A-Rx (5'- CACCGGCTTAGGCATCTCCTATAGCAGGAAGAA-3'; HXB2 nt 5954 - 5982) and EnvN (5'- TTGCCAATCAAGGAAGTAGCCTTGTGT -3'; HXB2 nt 9145-9171). *Env1A-Rx possesses a 5' CACC overhang that facilitates directional cloning into the pcDNA3.1 Directional cloning vector (Invitrogen, Life Technologies, Carlsbad, CA). Further, it introduces an M to I mutation that removes the Rev start codon, in order to enhance the expression of *envelope* downstream, and restore functionality in Env-pseudotyped assays of viruses that encode a Rev-Vpu fusion protein (142). Outer reaction thermal cycling conditions were as follows: initial denaturation at 94°C for 2 minutes, followed by 35 cycles of [at 94°C for 15 seconds, 55°C for 30 seconds, 68°C for 4 minutes], followed by a final extension for 1 cycle at 68°C for 20 minutes. Inner reaction thermal cycling conditions were the same as above, for 45 cycles. The products were analyzed by 1% agarose gel electrophoresis, visualized with a UVIpro Silver transilluminator (UVItech, Cambridge, UK) using GelRed (Biotium, Hayward, CA) incorporated into the loading dye.

Amplicons identified as ~ 3kb in size by agarose gel electrophoresis, were directly sequenced using an ABI3000 genetic analyser and BigDye terminator reagents (Applied Biosystems, Foster City, CA) by the Central Analytic Facility at the University of Stellenbosch, South Africa. The 12 primers used for *envelope* sequencing are tabulated in supplementary Table S2.1. Contigs were assembled using Sequencher 4.10.1 (Gene Codes, Ann Arbor, MI). All sequences were screened for contamination against a database of all sequences generated in the laboratory including all constructs used, with a pairwise DNA distance of less than 5% flagged as potential contamination.

2.2.3 Strain specific single genome amplification (ssSGA)

For strain-specific single genome amplification (ssSGA) of the CAP256 superinfecting virus, reverse transcription was performed using a primer specific for the superinfecting variant (256spR 5'-CTCCCTCTGCTGTTGGCTGCGCTCGCGC-3'; HXB2 nt 8856 - 8884). Single genome

amplification was then performed as described above (2.2.2), using the strain-specific primer as the anti-sense primer in both rounds of amplification.

2.2.4 *Sequence alignment and analysis*

Sequences were aligned relative to a curated HIV alignment via a Hidden Markov Model (HMM-align) using HIValign (http://www.hiv.lanl.gov/content/sequence/HIValign/HIValign.html), and phylogenetic trees were constructed using FastTree 2 (143), with a GTR+CAT model. Where necessary, trees were pruned for ease of visualization, using iTOL (144). Diversification was visualized using the Highlighter tool (http://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/HIGHLIGHT_POSTSCRIPT/highlighter.html), with the "matches" setting used to visualize intra-participant recombination. Recombination with an unknown, divergent variant was evaluated with RIP (http://www.hiv.lanl.gov/content/sequence/RIP/RIP.html), using a background alignment of 26 pseudo-randomly selected, subtype-matched sequences, weighted heavily for sequences from the CAPRISA cohort. Pairwise DNA distances were computed using Mega 5 (30).

2.2.5 *Estimating the time to the most recent common ancestor (tMRCA)*

In order to improve the estimates of the timing of superinfection, and to differentiate the transmission of one or multiple viruses, we estimated the time to the most recent common ancestor (tMRCA) of multiple closely related sequences shortly following the estimated time of infection/superinfection. The tMRCA was estimated with Poisson Fitter (http://www.hiv.lanl.gov/content/sequence/poisson_fitter/poisson_fitter/html) with default settings (mutation rate of $2.16x10^{-05}$ and a generation time of two days). Sequences were determined to have originated from a single transmitted/founder virus if the number of mutations conformed to a Poisson distribution, displayed a star phylogeny, and with an

estimated MRCA dating within the clinically defined window of infection (described in detail in 10).

2.2.6 Comparison of variable region length and glycosylation

For analysis, consensus sequences were derived for each participant at each timepoint according to "majority rules". Variable region lengths and the number of N-linked glycosylation sites were calculated from the alignment using an in-house script in Wolfram Mathematica 8 (Wolfram Research, Champaign, IL, USA). V1, V2, V3, and V4 loops were measured from Cysteine to Cysteine. However, for V2, the first cysteine was designated as the end of V1. V5 was measured from HXB2 position 460 to 471. Potential N-linked glycosylation sites were identified by the presence of sequons (N X S/T, where X is any amino acid except for Proline). Overlapping sequons (i.e. NNST, NNSS, NNTT, NNTS) were counted as a single N-linked glycosylation site as it is highly unlikely that both can be glycosylated. The length and glycosylation of the variable regions in acute and superinfecting viruses were compared to those of the three year sequences using a two-tailed T-test, implemented in Prism 5 (GraphPad Software, San Diego, CA).

2.2.7 Power calculations

Statistical power was calculated by simulating data in Wolfram Mathematica 8 (Wolfram Research, Champaign, IL, USA).

We modeled the risk of superinfection as:

$$\left(1 - \frac{1}{1 + e^{-kx}}\right)(b - p) + p$$

where

b = the baseline risk of infection over the sample period

p = the risk of superinfection in the maximally protected participants over the sample period

k = slope parameter

x = correlate sampled from a normal distribution with mean = μ, and standard deviation = 1.

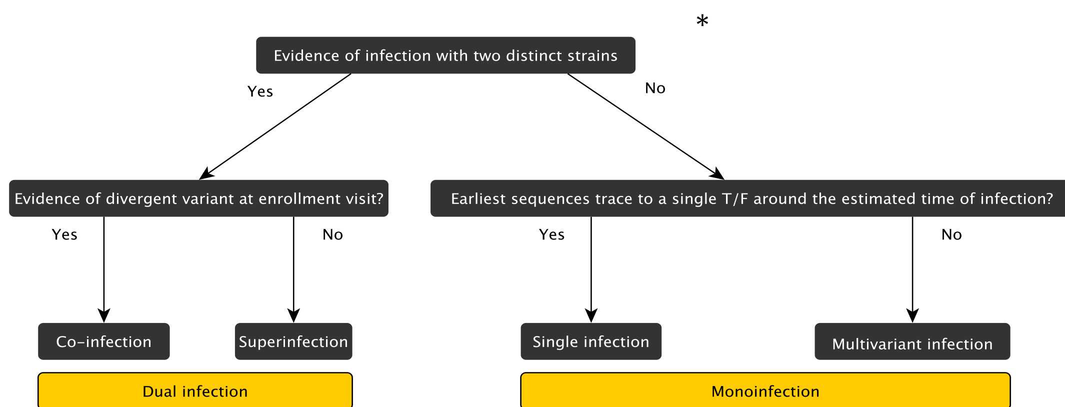
The model was developed in collaboration with Ben Murrell (UCSD, San Diego, USA). Statistical power for a Mann-Whitney U test to detect correlates of risk of superinfection was estimated from data simulated under these models, with 500 iterations.

2.3 Results:

Dual infection with two HIV-1 strains can be a consequence of either co-infection or superinfection (Table 2.1). Superinfection was defined as infection following seroconversion with a second variant that is separated on a phylogeny by epidemiologically unlinked viruses (phylogenetically unlinked). Co-infection was defined by the identification of two (or more) phylogenetically unlinked variants, or by evidence of a recombinant region of at least 150 bp inherited from a phylogenetically unlinked variant, at or before the first seropositive visit. Sequences from individuals with no evidence of dual infection (i.e. monoinfected) were further classified as infected with a single variant (single infection) or multiple closely related variants or quasispecies (multivariant infection) based on statistical modeling of early virus evolution as described by Keele et al. (10). Multivariant infection was defined by the sequences (obtained shortly following transmission) not following a star phylogeny, with the number of observed mutations not conforming to a Poisson distribution, and the estimated tMRCA of these sequences pre-dating the clinically defined date of infection.

This chapter expands on previous data which identified four cases of dual infection out of 39 participants from the CAPRISA002 cohort screened using an *envelope* C2C3 heteroduplex mobility assay (HMA) (145), and two putative cases of superinfection (CAP256 and CAP281) (114, F. Treurnicht, UCT, unpublished data). These were combined with sequences from published and ongoing studies from UCT and the National Institute for Communicable Diseases (R. Thebus unpublished data, B. Lambson unpublished data, Wibmer et al., in press, 2, 68, 118, 119). As part of these studies, *gag*, *nef* and/or *envelope* sequences were generated for 58 participants over time, of which 18 participants were followed up for approximately three years. These sequences were used to identify cases of potential dual infection (i.e. individuals who harboured viruses with high diversity inconsistent with the known duration of clinical infection, or evidence of recombination with divergent variants). For all individuals with putative dual infection, additional full-length e*nvelope* sequences were generated at the

enrollment visit and at approximately 3-, 6-, and 12-months post infection (mpi). Where dual infection was evident in one sample but not in the preceding sample, intervening timepoints were sequenced in order to identify the likely timing of the secondary/super- infection (Figure 2.7). In one case (CAP256), strain-specific primers were also designed to estimate the timing of the superinfection event, and to generate amplicons from the superinfecting virus. For this thesis, a total of 361 full-length *envelope* sequences were generated from 11 participants over time, by single genome amplification (SGA)



*Participants for whom insufficient sequence data was available are classified as "indeterminate multiplicity".

Figure 2.1 | Schematic illustrating the classification of multiplicity of infection.

2.3.1 At least eleven participants from the CAPRISA 002 cohort were dual infected

Multivariant and co-infections

To determine the multiplicity of infection, we analyzed *envelope* SGA sequences at (or before) the first HIV seropositive visit for forty-four participants. An analysis of the phylogenies of these early sequences, and the distribution of the number of mutations, identified 29 participants as harbouring homogeneous viruses consistent with single infection, and 15 participants likely infected with multiple viruses by seroconversion (Table 2.1). The maximum

DNA distances observed in sequences from participants likely infected with more than one virus (red dots), was consistently higher at comparable time post-infection than those from participants where a single virus likely established infection (black dots) (Figure 2.2).

Table 2.1 | Early *envelope* diversity at the first HIV positive visits providing evidence of multivariant or co-infection of 15 participants.

PID	Clinically estimated weeks post infection	Star phylogeny?	Poisson distribution of mutations?	Maximum *env* genetic distance	Estimated multiplicity of infection (number)	Evidence of recombination with an unlinked virus
CAP8	2	No	No	0.47%	Multivariant (2)	No
CAP177	4	No	No	3.54%	Multivariant (2)	No
CAP228	7	Yes	No	0.35%	Multivariant (2)	No
CAP69	2	No	No	3.52%	Multivariant (5)*	No
CAP136	2	No	No	1.31%	Multivariant (2)*	No
CAP222	2	No	No	3.0%	Multivariant (3)*	No
CAP224	3	No	No	2.05%	Multivariant (2)*	No
CAP260	2	No	No	2.53%	Multivariant (2)*	No
CAP37	2	No	No	9.04%	Coinfection (3)*	Yes
CAP84	4	Yes	No	2.29%	Coinfection (2)	Yes
CAP137	7	No	No	14.49%	Coinfection (2)	Yes
CAP225	3	Yes	No	4.43%	Coinfection (2)	Yes
CAP258	7	No	No	5.57%	Coinfection (3)	Yes
CAP266	7	No	No	6.24%	Coinfection (2)	Yes
CAP267	6	No	No	6.49%	Coinfection (3)	Yes

* Reference: (11)

Figure 2.2 | DNA distance in *envelope* at the first available HIV positive visit (RNA positive/antibody negative or seropositive), plotted according to the estimated time post infection. Participants where the viral evolution is consistent with a single virus founding infection are shown in black, whereas those who were likely infected by more than one virus are highlighted in red. A linear regression line is fitted to distances observed in participants where infection was likely established by a single founder virus.

Reconstructing the phylogeny from *envelope* sequences for the 15 participants initially infected with multiple viruses together with a panel of 100 randomly selected subtype C reference sequences from South Africa revealed that only sequences from CAP37 and CAP137 clustered separately from each other on the tree, separated by epidemiologically unlinked sequences. Sequences from the remaining 13 participants were monophyletic, although long branch lengths were suggestive of highly diverse infecting viruses in some cases. We therefore looked for evidence of recombination with a divergent virus in these 13 participants. Using sliding window distance measures to a panel of 25 viruses, we found no evidence of recombination with a divergent virus in six participants. However, for seven of these 13 participants, there was evidence of recombination with an unknown divergent virus over a sub-genomic window (illustrated in Figure 2.5). Partitioning each alignment according to predicted breakpoints confirmed that these subgenomic regions were separated by epidemiologically unlinked sequences on the phylogeny in five out of six cases, providing evidence of

recombination with an unlinked virus. In five of these cases, several discontiguous recombinant regions were detected. In the sixth case, CAP84, a single recombinant region was detected in only one of seven sequences at the first seropositive visit. However, multiple unique recombinant regions arose over the course of infection (Figure 2.6), suggesting the presence, below detection, of a second strain. This, together with previous data demonstrating heteroduplex banding with a C2C3 HMA (145) is highly suggestive of co-infection in this participant. In the last case, CAP177, several subgenomic regions of a virus detected in the CVL (B. Lambson, NICD, JHB, RSA, unpublished data) were genotypically divergent from the initial virus detected in the plasma. However, these viruses were still monophyletic based on these partitions (Figure 2.5). CAP177 was therefore classified as a “multivariant” infection.

In total, we classified eight participants as infected with more than one closely related virus (multivariant), of whom five were previously identified (11). Seven participants were classified as infected with more than one divergent virus by the first available seropositive visit (coinfection). In two cases of coinfection (CAP37, CAP137), viruses were phylogenetically unlinked based on full *envelope* sequences. In a further five cases, although viruses were monophyletic based on full *envelope* there was significant evidence of sub-genomic regions inherited from a phylogenetically unlinked virus, highly suggestive of recombination with a divergent virus. However, it is not clear whether recombination occurred post infection, or prior to transmission in a dual infected donor. All infections and dual infections occurred with subtype C viruses.

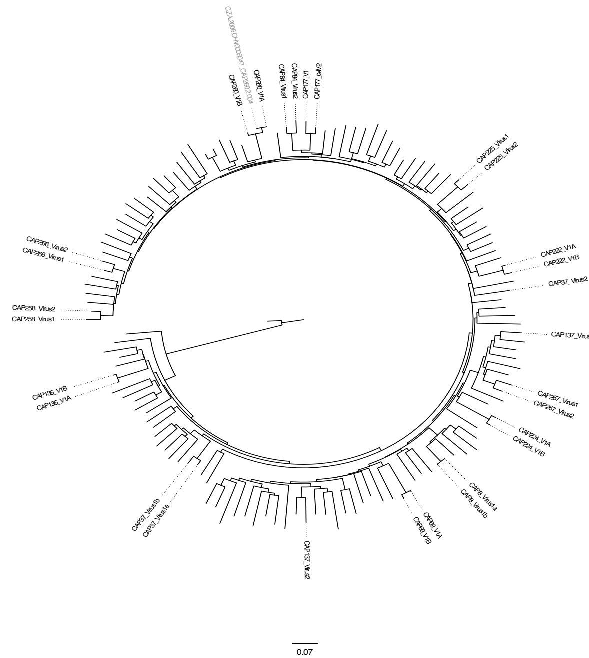


Figure 2.3 | Maximum likelihood phylogenetic tree illustrating representative *envelope* sequences from 15 participants initially infected with more than one virus in a background of 100 reference subtype C sequences from South Africa and rooted on the subtype B HXB2 reference sequence. For the 15 participants, only the sequence pairs with the greatest DNA distance are shown.

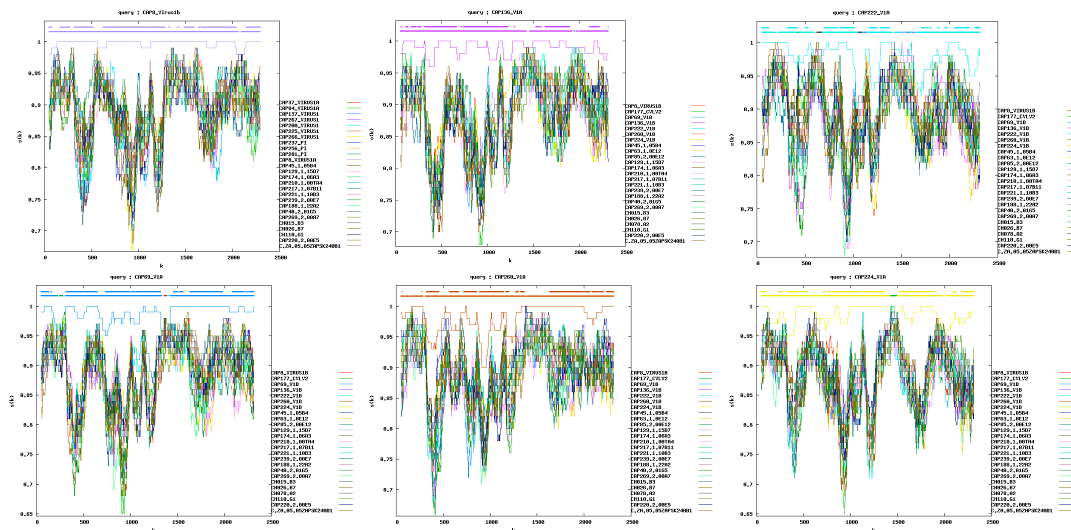


Figure 2.4 | No significant evidence of recombination with unlinked viruses in seven participants infected with more than one variant. DNA similarity to a panel of 25 subtype C viruses as well as the most distant within patient sequence was calculated using a sliding window (100bp in size) across *envelope* using RIP.

Figure 2.5 | Evidence of divergent subgenomic regions in seven participants. Distances were calculated with RIP using a sliding window of 100bp. Aligments were partitioned according to the inferred breakpoints and the phylogeny was inferred from the partition. The two most divergent within participant sequences at the first seropositive visit are highlighted in red and blue in each tree, which was inferred from the largest partition in each case.

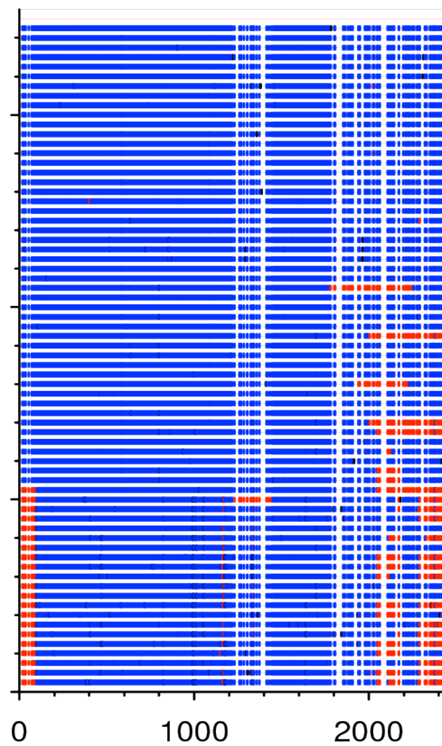


Figure 2.6 | Multiple unique recombination breakpoints detected in *envelope* over time in CAP84.

Superinfections

For three participants (namely CAP237, CAP256, and CAP281), a second phylogenetically unlinked variant was detectable (Figure 2.8), only following seroconversion (Figure 2.7), indicating likely cases of superinfection.

CAP256. CAP256 was previously identified as potentially superinfected prior to 30 weeks post infection (wpi), based on longitudinal, near full-length genome sequences (F. Treurnicht, NICD, JHB, RSA, unpublished results). To identify the superinfecting virus in CAP256, a strain-specific primer was designed based on the sequences of later recombinant viruses from 30 wpi. Strain-specific single genome amplification (ssSGA) was performed using plasma from 6, 11, 13, and 15 weeks post infection, corresponding to the earliest available sample and those prior to and at the peak of the viral load spike at 15 weeks post infection (Figure 2.7). While no amplification was evident in the samples from 6, 11, and 13 wpi, a phylogenetically unlinked *envelope* was successfully amplified at 15 wpi. The six amplicons sequenced from 15 wpi were

highly homogenous (three were identical, with the other three differing by no more than two nucleotide changes (Figure 2.9B) and no mutation common to more than one virus), suggesting that they had been recently transmitted (estimated time since MRCA of 8 days) (95% CI: 0 – 26 days). Furthermore, there was no significant evidence of recombination with the initial virus at this time point (Figure 2.10B). We therefore estimate that CAP256 was superinfected at ~14 wpi (95% CI: 11-15wpi).

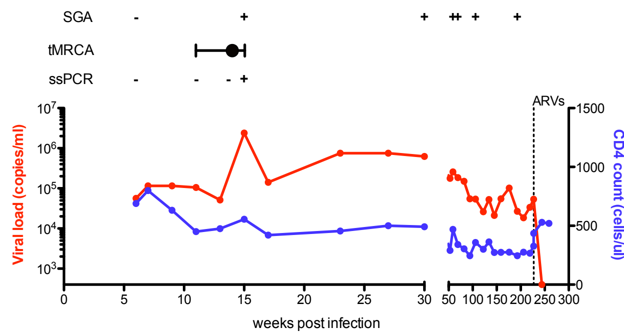
CAP237. A comparison of *envelope* SGA sequences from three years post infection in CAP237 to those at the first HIV positive visit (R. Thebus, unpublished sequences) was highly suggestive of superinfection. In order to confirm superinfection, and to identify the superinfecting *envelope* sequence in CAP237, single genome PCR amplification was performed on longitudinal plasma samples, taken at 4, 11, 24 and 54 wpi. A phylogenetically unlinked virus was detected at 11 wpi, but not at 4 wpi, consistent with superinfection. Viruses from CAP237 were separated on the phylogeny by a virus (C.ZA.2004.04ZASK13681) sampled in 2004, three years before CAP237 was initially infected (inconsistent with a transmission chain). The superinfecting *envelope* sequences were highly homogenous (Figure 2.9A), with no significant evidence of recombinant regions inherited from the initial virus (Figure 2.10A). The time since a MRCA was estimated at 8 days (95% CI: 0-24 days), suggesting superinfection likely occurred at ~ 10 wpi (95% CI: 7-11 wpi). Similarly, a strain-specific PCR of *gag,* estimated that superinfection likely occurred between 7 and 11 wpi (146).

CAP281. In the case of CAP281, gag sequences of the dominant population at 46 wpi were highly divergent from those at 26 wpi, suggestive of superinfection (146). In order to identify the superinfecting envelope, longitudinal SGAs were generated from the earliest available sample (10 wpi), and from 13, 30, 42, 46 wpi plasma samples. The superinfecting virus was detectable at 46 wpi, but not 42 wpi or earlier. Sequences at 46 wpi were highly homogenous (Figure 2.9C), with an estimated time since MRCA of 28 days (95% CI: 0-43 days), suggesting that superinfection likely occurred at ~42 wpi (95% CI: 40 – 46 wpi). There was also no

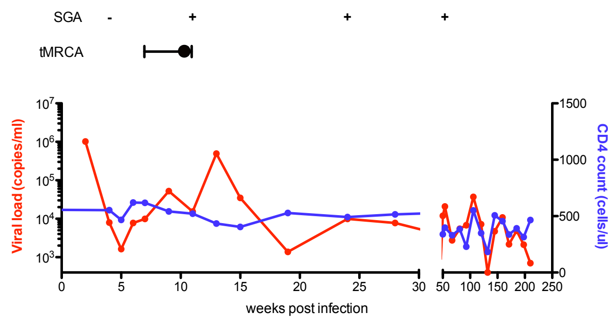
significant evidence of recombinant regions in the 46 wpi sequences inherited from the initial variant (Figure 2.10C).

In summary, in three participants two phylogenetically unlinked viruses were detectable, only following seroconversion, suggestive of superinfection (Figure 2.9). In all three cases, the estimated tMRCA indicated the secondary infection dated post-seroconversion, and was consistent with the sequencing data, and/or strain-specific SGAs (Figure 2.7). In each case, a spike in viral load (increase >10-fold) was also evident at or shortly after the estimated time of superinfection. While it is not possible to exclude the possibility of co-infection, these data strongly suggest that these three participants were superinfected, all within the first year of infection. The three superinfecting viruses from CAP256, CAP237 and CAP281 were all classified as subtype C, and differed from the primary infecting variants across *envelope* by 12.53%, 11.92%, and 13.99%, respectively.

A. CAP256



B. CAP237



C. CAP281

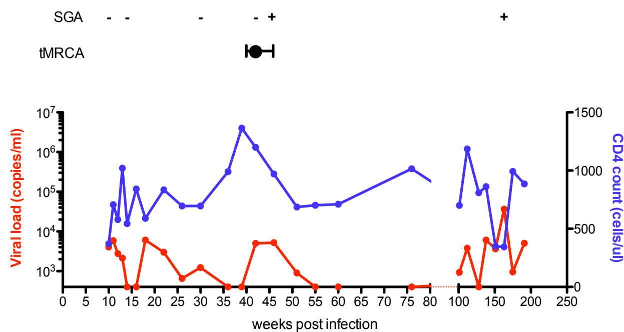


Figure 2.7 | Viral loads (red) and CD4 counts (blue) over time in participants CAP256 (A), CAP237 (B), and CAP281 (C), along with the estimated timing of superinfection. The timepoints for which single genome PCR amplification (SGA) was performed, along with the presence (+) or absence (-) of the superinfecting strain is shown. The estimated timing of superinfection based on dating of the most recent common ancestor of the superinfecting sequences (tMRCA), is also depicted, along with the 95% confidence interval for this estimate. For CAP256, PCR specific for the superinfecting variant (ssPCR) were additionally performed and the detection of the superinfecting variant is shown. ARVs, start of anti-retroviral treatment.

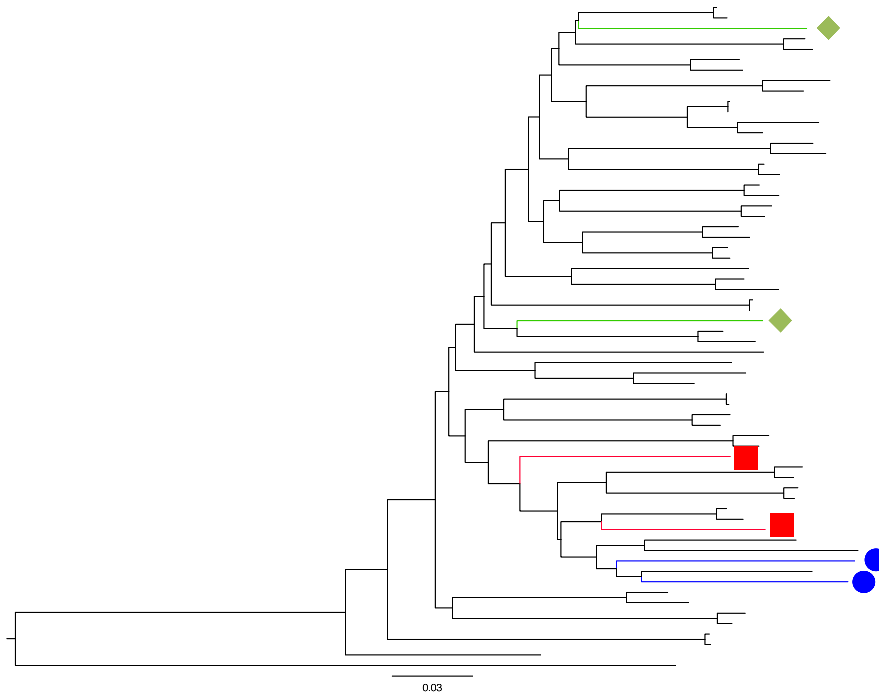


Figure 2.8 | Maximum likelihood phylogeny depicting the presence of two strains separated on the phylogeny by epidemiologically unlinked viruses in participants CAP237 (blue circles), CAP256 (red squares), and CAP281 (green diamonds). The phylogeny was reconstructed using *envelope* sequences from CAP237, CAP256, CAP281 along with 58 subtype C sequences from 30 individuals including 24 participants from the CAPRISA 002 cohort, and was rooted on the subtype B, HXB2 reference sequence.

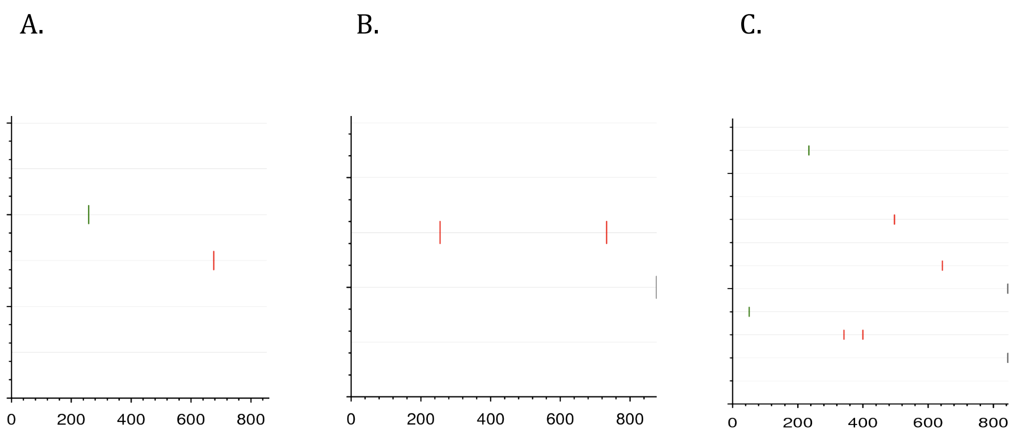


Figure 2.9 | Nucleotide highlighter plots of *envelope* SGA sequences sampled recently post superinfection depicting the high homogeneity in CAP237 (A), CAP256 (B), and CAP281 (C). Identity to the consensus sequence is depicted by a horizontal gray line and the absence of a coloured tick. Synonymous mutations are shown in red and non-synonymous mutations are shown in green. In all three cases, the superinfecting population was highly homogenous, with no two sequences harbouring the same mutations away from the consensus.

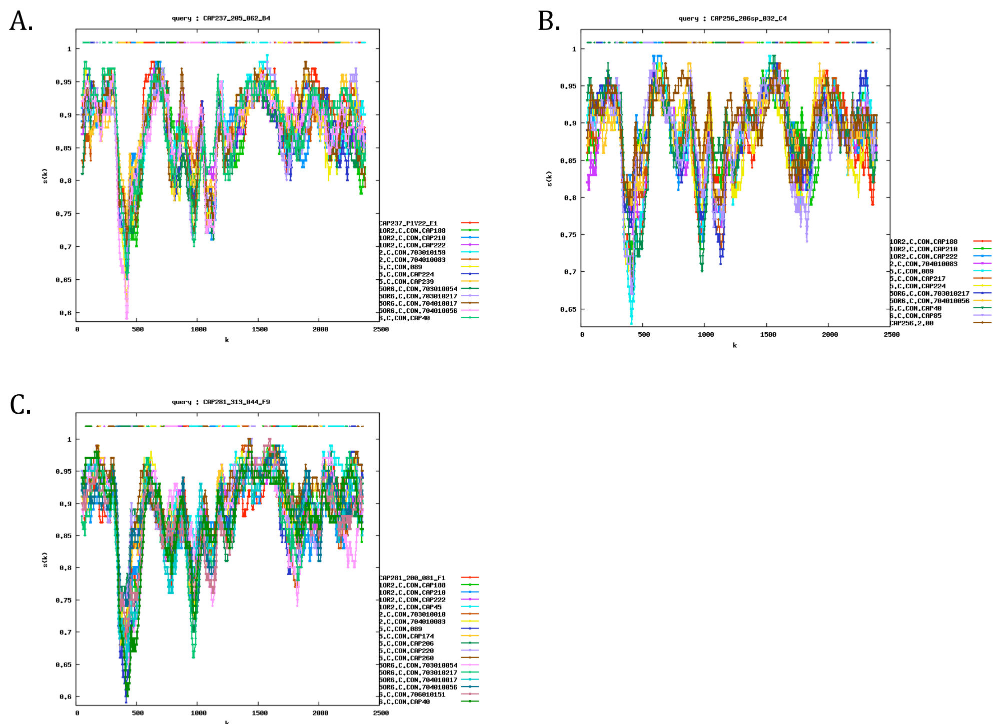


Figure 2.10 | No evidence of recombinant regions inherited from the primary virus within *envelope* of the superinfecting viruses. The similarity of the (A) CAP237, (B) CAP256 and (C) CAP281 superinfecting *envelope* sequences to a panel of 24 subtype-matched viruses, including the primary infecting viruses was evaluated. Only the plots from the most similar subset of viruses are displayed for ease of visualization. Similarity was calculated with a 100 bp sliding window using RIP. No windows within *envelope* of the superinfecting viruses were significantly more similar to the primary virus than they were to the panel viruses, suggesting the absence of recombination.

2.3.3 Superinfecting viruses have phenotypes typical of recently transmitted viruses

While a single founder variant establishes infection in ~80% of primary HIV transmissions (9–11), the multiplicity of superinfection is not known. Here we were able to accurately time superinfection to within a short window in three participants. Sequencing of the early superinfecting viral population revealed that, in all three cases, sequences exhibited a star-like phylogeny with the number of mutations arising consistent with a Poisson distribution. The inferred common ancestors also dated near to the time of superinfection. This is consistent with superinfection established by a single transmitted/founder (T/F) variant. Importantly, in all three cases of superinfection, the earliest *envelope* sequences representative of the

superinfecting viral population were highly homogenous (Figure 2.9), and showed no evidence of recombinant regions inherited from the initial virus (Figure 2.10). Therefore, sequences identical to the consensus were considered as an accurate representation of the superinfecting transmitted/founder (T/F) virus.

Recently transmitted viruses appear to possess shorter, less glycosylated variable regions than viruses from chronic infection (139, 140). In order to establish whether superinfecting founder viruses had similar characteristics, we compared the length and potential N-linked glycosylation of the variable regions to a panel of matched acute and chronic subtype C sequences from the same cohort, sampled over a similar period. Acute *envelope* SGAs were available from 21 participants in the CAPRISA002 cohort, for whom matched SGAs were available at either one year (n=5), or three years (n=16) post infection. Consistent with previous observations, viruses sampled at three years post infection had longer, potentially more glycosylated variable regions compared to those sampled shortly following transmission. Superinfecting viruses had characteristics typical of recently transmitted viruses, with length and glycosylation significantly lower than in viruses from three years post infection (p = 0.0074, p = 0.0131 respectively; Figure 2.11).

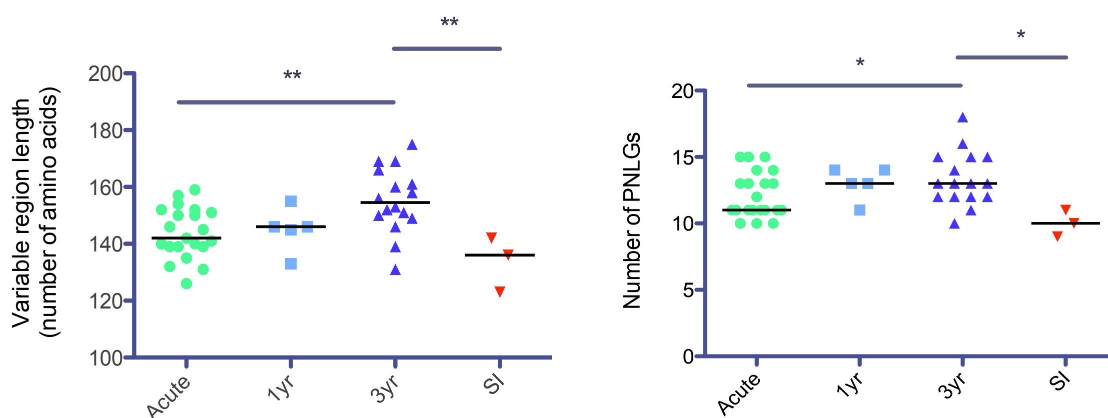


Figure 2.11 | Superinfecting T/F viruses have Envelopes with shorter, less glycosylated variable regions compared to viruses sampled at three years post infection. Each point represents the consensus sequence for each participant at that timepoint. *P<0.05; **P<0.01

2.3.2 Recombination occurred early, and frequently played a major role in generating envelope diversity.

Recombination is intrinsic to HIV replication, and is an important evolutionary mechanism employed by HIV to generate sequence diversity. Recombination is not an unconstrained process, and the extent and timing of recombination following dual infection is not clear. To assess the extent of recombination following dual infection, we used single genome amplification to sequence *envelope* over the first year of infection in seven dual infected participants (CAP37, CAP84, CAP137, CAP267, CAP237, CAP256, CAP281). Importantly, single genome amplification prevents artificial recombination due to PCR amplification.

Recombination following superinfection was evident (within *envelope*) in all three superinfected participants (Figure 2.12). As sampling in the CAPRISA002 cohort was more frequent during acute infection (weekly vs monthly/quarterly), co-infected participants offer the best cases to evaluate the early kinetics of recombination. Of the co-infected participants, CAP137 was the only case in which there was no evidence of recombinant regions shared between the two infecting variants. The absence of homology between the infecting viruses in CAP137 sampled pre-seroconversion (ie HIV RNA positive, antibody negative; an estimated two weeks post infection) is consistent with infection by two separate donors, and provides an opportunity to evaluate how soon recombination can become evident after dual infection. Analyzing *envelope* sequences (N=5) sampled five weeks later (~7 wpi), we find evidence of at least two unique recombinant viruses (Figure 2.13). Viral loads at this time point reached 360,000 copies/ml. While we cannot exclude the possibility that these recombinant variants were co-transmitted, the small number of mutations away from the putative parental virus sequences supports the hypothesis that recombination occurred (rapidly) following transmission.

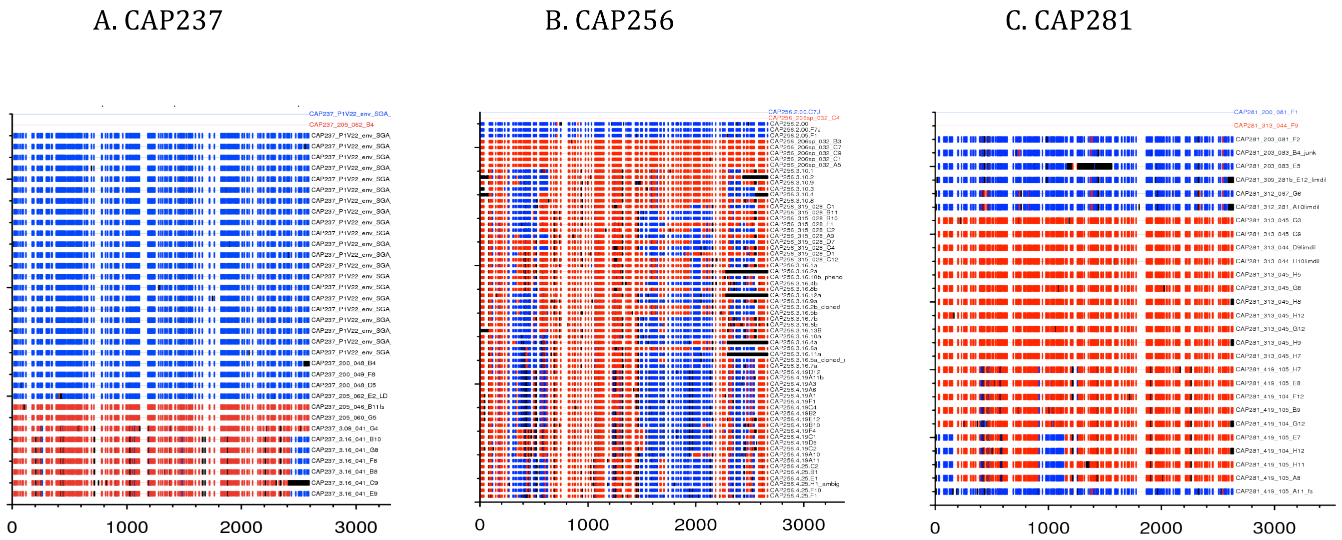


Figure 2.12 | Highlighter plot depicting recombination in *envelope* over time in 3 superinfected participants: CAP237 (A), CAP256 (B), and CAP281(C). Residues identical to the superinfecting or primary infecting viruses are show with red or blue ticks respectively. The presence of mutations not present in neither of the infecting strains are shown with black ticks, while sites conserved in both strains are shown in white.

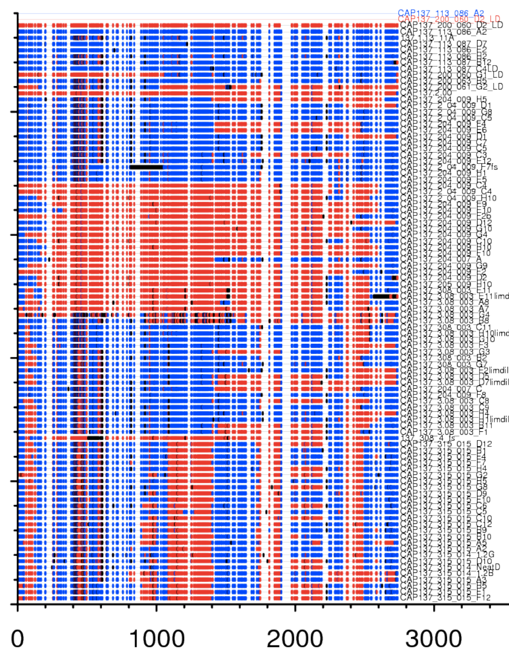


Figure 2.13 | Highlighter plot depicting recombination over time in a co-infected participant, CAP137. Residues identical to each of the infecting strains are shown with red or blue ticks. The presence of mutations not present in neither of the infecting strains are shown with black ticks, while sites conserved in both strains are shown in white.

2.3.4 Small sample sizes have little power to detect correlates of protection from superinfection

In total, 18 participants were followed up for three years, with no evidence of superinfection and were therefore classified as "singly infected" (Table 2.2). For 16 of these participants, sequence data was available from three genes (*gag, nef* and *env*) (146). For two participants, sequences were only available for two genes (*gag/nef* and *env*).

Table 2.2 | Defining the multiplicity of infection

	Mono infection		**Dual infection****	
	Single variant	Multivariant	Co-infection	Superinfection
Multiplicity	1	>1 variant	>1	>1
Timing	At transmission	At transmission	At or prior to seroconversion	Second infection following seroconversion
Virus relationship	Single virus resulting in clinical infection	>1 closely related viruses	>1 divergent viruses	>1 unlinked viruses
N	18	7 (3)#	7	3

*Indeterminate multiplicity: N=33
** CAP200 was identified as dual infected, but it is unclear whether this was as a result of co- or superinfection
# While seven participants were infected with more than one closely related virus, only three of these participants had at least two years of follow up to exclude superinfection.

In summary, a total 32 individuals had *envelope S*GA sequencing upon enrolment with follow up over at least 18 months, of whom 21 were "monoinfected" and 11 were dual infected. Of the 21 with monoinfection, three were classified as multivariant infection and 18 as single variant infection. Of the 11 cases of dual-infection identified, five were first identified as part of this thesis. Of these 11, seven were classified as co-infected, three were superinfected, and in one case the timing could not be resolved (Table 2.2). To estimate the incidence of superinfection, we considered only SGA follow up as population sequencing potentially offers little power to detect superinfection and its inclusion could therefore have resulted in

underestimating the incidence of superinfection. A total of 54.89 person years of follow up were available with SGA sequencing. We therefore conservatively estimate the incidence of superinfection as 5.47/100 person years. This is not markedly different from the primary incidence in this cohort (7.2/100 person years) (141).

To establish whether we had any power to detect correlates of protection against superinfection with three cases of superinfection and 18 singly infected controls, we modeled a hypothetical protective effect *in silico*. While we cannot know, a priori, the relationship between this hypothetical correlate and risk of infection, we can estimate the power under the most, and less favourable combination(s) of variables. The most favourable scenario (from the standpoint of power of detection) would result from complete protection above a particular threshold, but no protection for individuals with values below this threshold (i.e. Titers above "X" provide complete protection against superinfection, but any titers below "X" offer no protection) (see Figure 14A). Under a less favourable, but perhaps more biologically relevant, scenario, protection would scale with the correlate, providing 100% protection only to those with the highest values (Figure 2.14B).

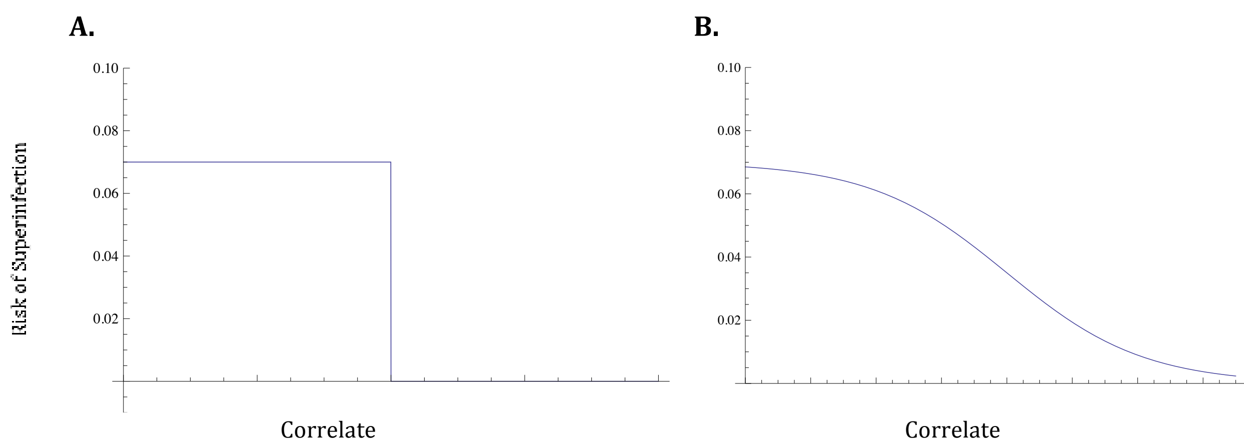


Figure 2.14 | Modelling a hypothetical correlate of risk of superinfection. A. The most statistically favourable scenario can be modeled with complete protection above a critical threshold. This relationship was modeled with $k \rightarrow \infty$, p = 0, and μ =-0.7. B. A statistically less favourable, but perhaps more biologically relevant, scenario can be modeled with k = 1, p = 0 and μ = -1.5. Under this model, protection against superinfection scales proportionally to the correlate, providing complete protection to only those individuals with the highest response.

Even under a favourable scenario (as in Figure 2.14A), a protective effect which produces a mean incidence of superinfection of 5.47%/py in a cohort with a baseline risk of 7.2%/py, screened over three years, would have ~6% power to detect this association. With a shallower slope culminating in complete protection for those participants with the highest values (as in Figure 2.14B), there is essentially no power to detect this association (i.e. equivalent to the false positive rate of 5%). Even assuming a true incidence of superinfection half that of primary infection (i.e. 3.6%/py in this cohort), under our model, a sample ~3-15 times larger (i.e. 9-45 superinfections and 54-270 monoinfected controls) would be needed to have ~80% power to detect this hypothetical correlate of protection. This also assumes that a single variable would be tested. Testing a number of candidate immunological variables will require correcting for multiple testing, further reducing power.

2.4 Discussion

Since the first case was reported in 2002 (147), cases of HIV superinfection are increasingly being identified. A case of HIV triple infection has also been described (148). The implication is that the natural immune response to HIV infection is frequently not sufficient to provide protection from infection with another HIV variant, raising concerns for vaccine design. Even if the immune responses generated during natural infection provide no protection from HIV infection, quantifying these responses at the time of superinfection will still address an important question, namely: What is not sufficient to provide protection? However, the utility of superinfection as a model of vaccine protection depends on the superinfection event accurately reflecting primary infection, though this has not previously been demonstrated. Here, we were able to show that all three cases of superinfection identified were likely established by a single virus. This is consistent with primary HIV infection, where infection is established by a single virus in ~80% of individuals (10, 11). We also show here that the superinfecting viruses had shorter, less glycosylated variable regions than variants from chronic infection; a characteristic of recently transmitted viruses. These data indicate that superinfecting viruses are genotypically similar to primary infecting viruses, providing further support for superinfection as a model to address correlates of protection against HIV.

While the natural immune response is frequently inadequate to prevent superinfection, identifying correlates of even partial protection would be invaluable for vaccine design. To address whether the immune response elicited in response to HIV infection provides *any* protection from superinfection, several studies have systematically screened for superinfection, and compared the incidence to that of primary infection in the same or similar cohort(s). Though systematic screening of CAPRISA 002 to accurately quantify the incidence of superinfection was beyond the scope of this project, within 54.89 person years of follow up with full *envelope* SGAs from 25 participants, we identified three cases of superinfection. This corresponds to an incidence of superinfection of 5.47%/py of follow up, which is not

significantly different from a primary incidence of infection of 7.2%/py (141). This is consistent with the idea that, at best, only a minority of HIV infected individuals are likely protected from superinfection. Similar observations were made from cohorts from Rakai, Uganda (heterosexual), and California, USA (MSM), where the incidence of superinfection was comparable to that of primary infection (124, 125). However, in a high-risk heterosexual cohort from Kenya, superinfection was detected at a rate approximately half that of primary infection (126, 135), and two other studies were not able to find any evidence of superinfection despite extensive follow up (127, 128). Moreover, behavioural changes following initial diagnosis can result in lower risk behavior following infection (127, 135). The reported incidences of superinfection in these (and other) studies are also likely underestimates as fluctuating viral populations can result in minority variants evading detection and recombination can result in divergent viruses appearing as monophyletic.

Through in-depth sequencing of dual infected participants, we have shown here that recombination following dual infection can be extensive. While six participants could be separated phylogenetically based on partial *envelope* sequences, in some cases DNA distances were not significantly higher than those observed for multivariant infection. These data highlight the fact that DNA distance measures and phylogenetic approaches that ignore recombination lack sensitivity to detect cases of dual infection. Further, we show that recombination can obscure the ability to detect superinfection as rapidly as seven weeks following dual infection. With this in mind, we re-analyzed sequences from a previous study that found no evidence of superinfection in 1072 person years of follow up (128). *Gag* sequences were publicly available from ten of the 37 participants with high diversity in Protease and Reverse Transcriptase (greater than 4.5% and 3% respectively). In two out of these ten pairs of *gag* sequences, we find significant evidence of recombination with an unknown divergent variant (Supplementary Figure S2.1), suggestive of superinfection. Though we cannot exclude a transmission chain (i.e. the sequences that phylogenetically separate the two viruses are not epidemiologically unlinked), these results highlight that if recombination is

not taken into account, it can obscure the identification of superinfection. Furthermore, while we identified extensive recombination in *envelope* here, breakpoints within *envelope* (apart from near the borders) occurs less frequently than elsewhere in the HIV-1 genome (149) suggesting that recombination in other genes is likely to be even more pervasive.

As cases of superinfection are comparatively rare, simply comparing the incidence of superinfection to the incidence of primary infection offers little statistical power to identify protective effects. To illustrate this, we assumed the lower incidence of superinfection compared to primary infection (5.4%/py vs 7.2%/py) in the CAPRISA002 cohort was solely due to a hypothetical correlate of protection, and calculated the power an analysis based on 21 participants screened here (three superinfections; 18 single infections) would have. While we cannot know, *a priori*, the relationship between any potential correlate and the susceptibility to superinfection, we calculated the statistical power under the most, and less favourable models. A scenario in which a fraction of individuals are completely protected, and the rest have no protection against superinfection would be the most favourable for statistical power of detection (though it may represent a biologically unlikely scenario). Even under this model, if we were able to accurately predict, and precisely measure the hypothetical correlate of protection, we would have only ~ 6% power to detect this effect. In a perhaps more biologically plausible (though still optimistic) model where protection correlates proportionally, the power was essentially zero. Even assuming an incidence of superinfection half that of primary infection (3.6%/py vs 7.2%/py) due to a protective correlate, the entire CAPRISA 002 cohort (n=62) meticulously screened over 5 years, would have only ~50% power to detect this effect. These calculations bring to light the magnitude of the cohort that will likely be required to detect a correlate of protection under the observation that (at best) only a fraction of infected individuals are protected from superinfection. The only situation we can envisage in which an association would be readily evident in these sample sizes is if the immune responses to HIV infection did in fact provide substantial protection against superinfection, but that HIV+

individuals are more likely to become superinfected than HIV- individuals are to become infected.

Consequently, any study that aims to identify correlates of protection by comparing the immune responses in superinfected participants to singly infected controls needs to analyse a large number of participants. Nevertheless, several studies have attempted to identify potential correlates of protection in a handful of superinfected individuals by comparing specific immunological markers (such as autologous and heterologous neutralizing antibodies, and non-neutralizing V1V2 antibodies) in superinfected participants to those with similar risk factors, but who did not become superinfected. Two studies observed that the superinfected participants had weaker autologous and heterologous humoral responses than non-superinfected controls, suggesting a potential role for antibodies in lowering susceptibility to superinfection (132, 133). In contrast, however, Blish *et al.* (134) found that in 6 women, superinfection occurred despite the presence of robust neutralizing antibody responses at the time of superinfection. However, our calculations suggest that these studies had limited power to detect correlates of protection.

In conclusion, we show a high frequency of dual infection and superinfection in women from Kwa-Zulu Natal, not markedly different from the incidence of primary infection in the cohort. We found that superinfection was likely established by a single transmitted/founder (T/F) virus in each case, and that this superinfecting virus had properties similar to those of a primary T/F virus. Further, we show that recombination can occur rapidly following transmission, and can confound the detection of dual infection if not taken into account appropriately. While superinfection provides a useful model to address the correlates of protection, we also show that only an analysis of a very large number of individuals will be sufficiently powered to detect a moderate protective effect.

CHAPTER 3: The clinical causes and consequences of multiple HIV infection

3.1 Introduction

Identifying factors associated with mitigation of the genetic bottleneck and susceptibility to HIV infection may provide targets for preventative measures that could slow the epidemic. Observational studies have overwhelmingly suggested that sexually transmitted infections (STIs) can enhance susceptibility to HIV infection through mucosal disruption, target cell recruitment and activation, or an imbalance of the protective lactobacilli-dominated flora (reviewed in 150, 151). Control of STIs could therefore represent an effective strategy to reduce HIV incidence. However, to date, only one of nine intervention trials has demonstrated an effect of STI treatment on the incidence of HIV (152, reviewed in 153). It is therefore important to clarify the role of STIs in modulating the susceptibility to HIV infection. Several studies that have adjusted for risk behavior, have suggested that non-ulcerative STIs or bacterial vaginosis (BV) can increase the susceptibility to HIV infection by 2-4 fold, while genital ulcer diseases (GUD) can increase HIV acquisition by 2.2-11.3 fold (150, 154, 155). However, it is difficult to fully delineate the role of the STI, as opposed to common risk factors (individuals who undertake risky sexual activities are more likely to acquire both HIV and other STIs).

Individuals with a higher susceptibility to HIV infection may be expected to be more likely to be infected with more than one virus, reflecting the "breakdown" in the transmission barrier. The multiplicity of infection in 69 participants (including 26 from the CAPRISA 002 cohort) has also previously been shown not to follow a Poisson distribution (11), indicating that multivariant infections likely do not represent multiple independent transmission events. This would be consistent with the presence of factors in the recipient that influence the risk of acquiring multiple viruses (though this could also result from other scenarios, including infections by cell-associated viruses). Indeed, the use of hormonal contraception has been

associated with both a higher risk of infection (156), and with a more diverse viral population after infection in a cohort of Kenyan female sex-workers (157) . Sagar *et al.* further found that the presence of STIs was significantly associated with a heterogenous viral population in primary infection (158). However, the heteroduplex mobility assay used to characterize the diversity of the viral population in primary infection would be insensitive to detecting infection with multiple closely related viruses. In a subsequent study that used sequencing methods to characterize the multiplicity of infection, genital tract inflammation or ulceration was also associated with the mitigation of the genetic bottleneck at transmission (159). However, these analyses were based on only five cases of multivariant infection and therefore only very strong effects could have been identified. Furthermore, only symptomatic criteria were identified in these patients, and specific STIs were not diagnosed confounding the results (155).

It is also potentially important to identify the consequences of infection with multiple divergent HIV variants or strains, which could potentially promote viral diversification and consequently accelerate the evolution of immune escape, fitness enhancements, and clinical progression. An early study of five dual infected participants revealed that all progressed to AIDS or died within 2.4 years following initial infection, compared to monoinfected individuals who typically take more than eight years to progress to AIDS (130). Similarly, in a cohort from South Africa, dual infection was associated with higher viral loads (131). Elite controllers were also reported to have lost control of viral load following superinfection (160, 161). On the other hand, in a cohort of Kenyan women, superinfected participants had significantly lower viral loads than matched monoinfected controls (162). Superinfections have also been identified in long-term non-progressors (160, 163), and superinfection of two individuals who were controlling their viral load did not accelerate disease progression (164). However, these reports were largely anecdotal. More recently, Pacold *et al.* systematically screened an MSM cohort for dual infection using deep sequencing, and identified 11 cases (165). While in this cohort, neither coinfected nor superinfected individuals had a significantly faster CD4 decline than monoinfected controls, the superinfected group had a significantly faster viral load

increase over time (165). However, while host HLA genotypes are strongly associated with disease progression (reviewed in 22), this study did not control for host HLA alleles. A recent study of 37, well characterized, subtype B infected men-who-have-sex-with-men (MSM) from Amsterdam (10 dual infected, and 27 monoinfected) controlled for HLA B alleles known to be associated with rapid progression (based on 22). In this study, dual infection was the factor most significantly associated with disease progression (as defined by rate of CD4+ T cell decline after the mean time to superinfection) (166). However, the time from initial infection to CD4 counts <350 cells/μl was not significantly different in dual infected participants compared to monoinfected controls. Furthermore, the analysis was confounded as although all monoinfected participants were infected with subtype B, four of the six superinfections in this cohort occurred with a non-B HIV subtype.

Recently, the potential of HIV superinfection to accelerate disease progression is of renewed interest as organ transplants from HIV-positive donors to HIV-positive recipients are being considered. HIV associated nephropathy is common and consequently many HIV positive patients require renal transplantation. In resource-limited settings, access to dialysis is severely restricted and HIV positive patients (like others with co-morbidity) are frequently denied treatment (167). There is frequently a shortage of donor organs and consequently many patients die while waiting for transplantation. In developing countries with a high incidence of HIV, such as in South Africa, transplants from HIV positive donors to HIV positive recipients have the potential to alleviate the situation for both HIV-positive and –negative patients. Several such renal transplants have been performed successfully in South Africa (167) , and recently there has been a petition by American doctors and scientists to repeal a federal ban on the transplantation of organs from HIV+ patients in the USA (168). However, organ donations from HIV-positive patients have the inherent risk of superinfecting the recipient and accordingly, the clinical consequences of HIV dual infection need to be established.

To better assess the impact of dual infection on disease progression, we utilized a well-characterized cohort of 62 individuals from South Africa (CAPRISA 002). The cohort is comprised only of women, and is located in an area where subtype C is almost exclusively circulating, allowing us to control for both gender and infecting subtype. Participants were screened prospectively for HIV infection, and consequently the timing of initial infection in each case is known. Furthermore, HLA genotyping has been performed on all participants. This study expands on previous work that screened 39 participants from the cohort using an *envelope* C2C3 heteroduplex mobility assay (HMA) to determine multiplicity of infection (145). In this study Woodman *et al*., detected four cases of dual infection, and found no significant difference in their CD4+ T cell counts or viral loads at 12 months post infection compared to 30 singly infected participants. Woodman *et al*. did find an association between the presence of any STI and the infection with heterogenous viruses, though this was not statistically significant. Here, with more intensive sequencing efforts, we have identified an additional seven cases of dual infection (see Chapter 2), and revisit the impact of dual infection on disease progression, controlling for potentially confounding variables. With this more sensitive sequencing approach, together with more sophisticated statistical methods, we also address whether the presence of particular STIs at the time of infection predisposed to infection with multiple viruses.

3.2 Methods:

3.2.1 Viral load and CD4+ T cell counts

Viral loads and CD4 T+ cell counts for the 62 participants were provided by the CAPRISA 002 study team. Viral load at each visit was determined using the COBAS Amplicor HIV-1 Monitor Test v1.5 (Roche Diagnostic Corporation, Indianapolis, IN Diagnostics, Rotkreuz, Switzerland) and CD4+ T cell counts were determined using flow cytometry (FACScalibur, BD Biosciences, Mountain View, CA).

As viral load can fluctuate significantly over time, even at set point, the area under the log transformed viral load curve, per unit time, from three months to one year post infection was used as an estimate of viral load set point. All participants except for one (CAP63) were antiretroviral naïve over this period. For CAP63, area under the log viral load curve per unit time was calculated from three months post infection to the initiation of treatment (approximately nine months post infection).

Infection stage was categorized according to the criteria of Fiebig *et al.* (12). Individuals who were viral RNA positive but EIA antibody negative were classified as being in stage I/II. A positive EIA antibody test with a negative Western blot (GS HIV-1 Western blot analysis kit, Bio-Rad, WA) was classified as stage III. Those with an indeterminate Western blot result were classified as in stage IV. A positive Western blot positive but without reactivity to the p31 Integrase band was classified as stage V, while a Western blot positive including a p31 band was classified as stage VI.

3.2.2 STI diagnosis

Participants were screened upon enrollment by the CAPRISA 002 team for the presence of sexually transmitted infections (STIs) in vulvovaginal swab samples as previously described (155). Briefly, *Chlamydia trachomatis, Neisseria gonorrhoeae, Trichomonas vaginalis, Mycoplasma genitalium, Haemophilus ducreyi, and herpes simplex virus type 2* (HSV-2) were

identified by PCR. Genital herpes diagnosis was confirmed by enzyme immunoassay (HerpeSelect-1 and HerpeSelect-2, Focus Technologies, Cypress Hill, CA). Syphilis screening was done using the Becton Dickinson Macro-Vue RPR (rapid plasma reagin [RPR]) card test, and positive reactions were confirmed by the *T. pallidum* hemagglutination test (Omega ImmuTrep TPHA test). Agents of bacterial vaginosis (BV) were assessed by Gram staining, using Nugents criteria (169).

3.2.3 HLA genotyping

HLA genotypes of participants were kindly provided by C. Gray (UCT, South Africa). Four digit HLA class I genotyping was performed as described previously (170). Briefly, HLA-A, -B and -C typing was performed by sequencing of exons 2, 3 and 4 using AlleleSeqr kits (Atria Genetics/Abbott Laboratories, Abbott Park, IL) in DNA extracted from either PBMCs or granulocytes using the Pel-Freez DNA Isolation kit (Pel-Freez Biologicals, Inc., Rogers, AR). HLA types were inferred using Assign-SBT 3.5 (Conexio Genomics). Any ambiguities resulting from either polymorphisms outside the sequenced exons or identical heterozygote combinations were resolved using sequence-specific primers.

3.2.4 Statistical analysis

Times for CD4 counts to fall below 350 cells/µl for two consecutive visits were compared in a survival analysis using Kaplan-Meir plots, implemented in Prism 5 (Graphpad, San Diego, CA, USA). Participants were right censored when lost to follow up, or initiated onto antiretroviral therapy.

CD4+ T-cell decline was compared using a linear mixed effects model, blocking by participant, adjusting for HLA genotypes associated with slow (HLA-B*5703, HLA-B*5801) or rapid progression (HLA-B*18, HLA-B*5802, and HLA Cw0602) (as defined in 171).

For the Chow tests, data was partitioned into (i) prior to the lower bound of superinfection and (ii) post the upper bound of superinfection for each of the three superinfected participants.

Chow tests were performed on 17 singly infected participants using the same three partitions to establish the specificity of the test.

Multivariate analyses of viral load and CD4 decline, as well as Chow tests were implemented in the R statistical computing environment (172), utilizing the “nlme” (linear mixed effects models) and “gap” (Chow test) packages. Pre-infection CD4+ T cell counts were compared using an unpaired T test in Prism 5 (Graphpad, San Diego, CA, USA). The association between STIs and multiplicity of infection was tested using a logistic regression in R, with the “aod” package.

3.3 Results

The CAPRISA 002 cohort enrolled a total of 62 individuals, of which three (CAP259, CAP260, and CAP285) were lost to follow up prior to one year post infection and were excluded from our analysis. One individual (CAP63) was referred to treatment at approximately nine months post infection, according to the national guidelines at the time (CD4 count < 200 cells/μl or AIDS-related illness). The median follow up time was 1912 days (range: 259-2934 days).

In this chapter we investigated the factors that may enhance susceptibility to multiple infection, and the clinical consequences of dual infection. Defining the multiplicity of infection is described in Chapter 2. Briefly, HIV-1 "monoinfection" is defined here as infection by a single donor with either one variant (single) or multiple closely related variants (quasispecies) (multivariant), together with at least two years of follow up with no evidence of superinfection. HIV-1 dual infection here refers to individuals who were infected with highly diverse viruses that could have originated from two independent donors. Dual infection was classified in individuals who were either co-infected (infected with >1 virus prior to seroconversion) or superinfected (second infection occurred after seroconversion and the onset of the immune response) (Table 2.1). Seven participants were identified with multiple variant infection, of whom three had sufficient longitudinal follow up to classify as monoinfected. Six participants were classified as co-infected, three as superinfected, and one was dual infected with unknown timing (Table 2.1). All infections (and dual infections) occurred with HIV-1 subtype C. Viral loads and CD4+ T cell counts for all dual infected participants described here are shown in Supplementary Figure S3.1.

3.3.1 Susceptibility to infection with multiple viruses

Understanding the role of target cell availability, and the integrity of the mucosal barrier on the risk of HIV infection has important epidemiological implications. Higher numbers of target cells may represent one factor correlated with a higher risk of infection with multiple viruses. Using peripheral CD4+ T cell count as a proxy of target cells at the mucosa we compared CD4+ T cell counts at the last visit prior to HIV infection between singly infected and multiply (co-infected and multivariant transmissions) infected participants. Of the 59 participants analyzed here, pre-infection CD4+ T-cell counts were measured for 27 (12 singly infected, three infected with more than one closely related variants (multivariant), four coinfected, one superinfected, and seven of undetermined multiplicity of infection). We found no significant difference in the pre-infection CD4+ T cell counts between monoinfected (mean=1042; range: 424-2231) and multiply infected (mean=1017; range: 756-1395) participants (p=0.9068; figure 3.1) suggesting that target cell availability may not correlate with higher risk of infection with multiple viruses.

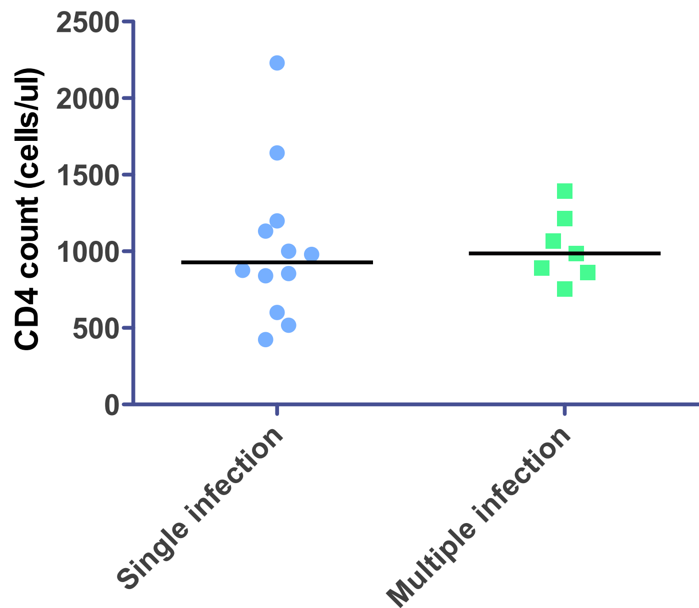


Figure 3.1 | Comparison of the pre-infection plasma CD4+ T cell counts between participants where infection was founded by a single virus (Single infection) and those infected with more than one virus (Multiple infection).

This analysis investigated target cells in the periphery and does not account for local target cell availability in the genital mucosa, which is likely influenced by local factors, including the presence of STIs and inflammation. Participants in the cohort were screened for the presence of STIs (*Trichomonis vaginalis, Neisseria gonorrhoea, Chlamydia trachomatis, Mycoplasma genitalium, Herpes simplex virus type 2* (HSV2), *Treponema pallidum* (Syphilis)), or bacterial vaginosis (BV) upon enrollment into the cohort. To evaluate whether BV or particular STIs, or a combination thereof, was associated with infection by multiple HIV viruses, we performed a logistic regression. Of all the participants, only four harboured no STIs upon enrollment. BV and *Trichomonas vaginalis* were highly prevalent, and were diagnosed in 30/40 (75%) and 32/37 (86%) of participants, respectively. No association with HSV-2 or Syphilis and multiple

HIV infection could be evaluated due to their low frequency. Of all the other STIs and BV, only *N. gonorrhoea* was strongly associated with multiple HIV infection (Table 3.2). Participants with Gonorrhoea were ~11.2 times more likely to be multiply infected with HIV (p=0.0566).

Table 3.1 | Multivariate analysis testing the association between sexually transmitted infections or bacterial vaginosis, and multiple HIV infection

	N	Frequency in single infections	Frequency in multiple infections	Odds Ratio	p
Bacterial vaginosis (BV)	30/40 (75%)	13/18 (72.2%)	8/13 (61.5%)	0.64	0.6732
T. vaginalis	32/37 (86.5%)	3/18 (16.7%)	1/13 (7.7%)	1.30	0.8496
N. gonorrhoea	7/38 (18.4%)	1/18 (5.6%)	4/13 (30.8%)	11.18	0.0566
C. trachomatis	6/38 (15.8%)	1/18 (5.6%)	2/13 (15.4%)	5.26	0.2236
M. genitalium	5/39 (12.8%)	2/18 (11.1%)	0/13 (0%)	0	0.9951
HSV-2	5/41 (12.2%)	0/18 (0%)	0/13 (0%)		NA
T. pallidum	2/42 (4.8%)	0/18 (0%)	0/13 (0%)		NA
Any STI or BV	39/43 (93%)	17/18 (94.4%)	13/13 (100%)		

3.3.2 Dual infection was not associated with accelerated disease progression

HIV-1 dual infection has previously been associated with accelerated progression to AIDS (130, 131). To determine whether infection with two divergent viruses (i.e. dual infection encompassing both co- and superinfection) was associated with an accelerated loss of CD4+ T cells over time in the CAPRISA 002 cohort, we performed a Kaplan-Meier survival analysis. We found that no significant difference existed in the number of days for CD4+ T cell counts to fall below 350 cells/µl between dual infected (N=11; median survival > 2226 days) and singly infected participants (N=18 median survival = 1582 days) (p=0.6227; Figure 3.2).

As there may be a systematic bias in selecting the monoinfected controls based on the requirement for extensive sequencing data, we also compared the disease progression between the identified dually infected participants, and the rest of the cohort. However, there was still no significant difference in the rate of CD4+ T cell decline (p=0.4030; Supplementary Figure S3.2). Of the 20 rapid progressors in the cohort (ie CD4 counts <350 cells/µl within the first year of infection), only two (10%) were dual infected (both were coinfected); lower than the incidence of dual infection in the rest of the cohort (~23%).

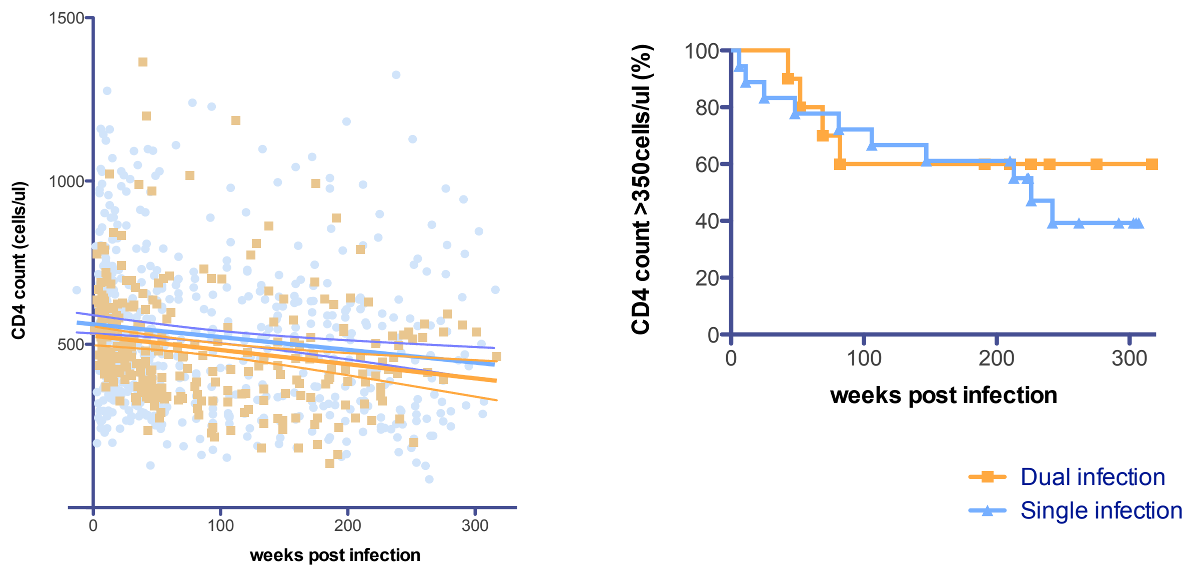


Figure 3.3 | A. Average rate of CD4 decline in singly and dual infected participants. B. Kaplan-Meier survival plot showing time for CD4+ T cell counts to fall below 350 cells/µl in singly infected (N=18) and dual infected (N=11) participants.

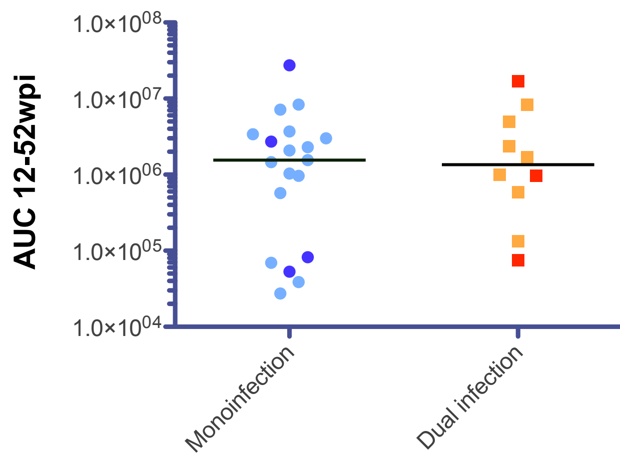


Figure 3.2 | Multiplicity of infection and cumulative viral load. In the monoinfected group. single infections are shown in light blue, while multivariant transmissions are show in navy. Within the dual infected group, superinfected participants are highlighted in red.

Viral load at set point has previously been demonstrated to correlate strongly with the rate of disease progression (16). We therefore compared set point viral loads between dual infected participants and the rest of the cohort. Using the area under the log viral load curve per unit time from three to 12 months post infection as a representation of set point viral load, we found no significant association between dual infection and a higher set point viral load (Figure 3.3; p=0.8008).

As host HLA alleles have previously been demonstrated to be strongly associated with viral load and CD4+ T cell counts, we performed a multivariate analysis. While the number of possible genotypes precludes the ability to control for these individually, we adjusted for the presence of HLA alleles previously associated with slow or rapid progression (171), in a linear mixed effects model. Even after adjusting for HLA alleles, neither the rate of CD4+ T cell decline (p=0.8642), nor the area under the viral load curve (p= 0.2028) was significantly associated with dual infection (Tables 3.2 and 3.3).

As superinfected participants are essentially infected with a single variant for an initial follow up period, we also compared the set point viral loads in the cohort to the seven co-infected participants only. However, the difference in the viral loads was still not significant (p=0.447). Viral loads were also not significantly different between co-infected participants and the 18 singly infected participants alone (p=0.5704).

Table 3.2 | Results of a multivariate model testing the association between multiplicity of infection and viral load, taking HLA alleles associated with slow and rapid progression into account.

	N	*P*	*P(multivariate)*
HLA(Slow) [a]	7	**0.496**	**0.4662**
HLA(rapid) [b]	21	**0.579**	**0.6278**
Dual infection	11	**0.344**	**0.2028**
Coinfection	7	**0.447**	**0.4308**
Superinfection	3	**0.828**	**0.9419**

Table 3.3 | Results of a multivariate model testing the association between multiplicity of infection and CD4+ T cell counts, taking HLA alleles associated with slow and rapid progression into account.

	N	*P*	**P(multivariate)**
HLA(Slow) [a]	7	0.2117	**0.6529**
HLA(rapid) [b]	21	0.0714	**0.0761**
Dual infection	11	0.8648	**0.8441**
Coinfection	7	0.7580	**0.8397**
Superinfection	3	0.7080	**0.7213**

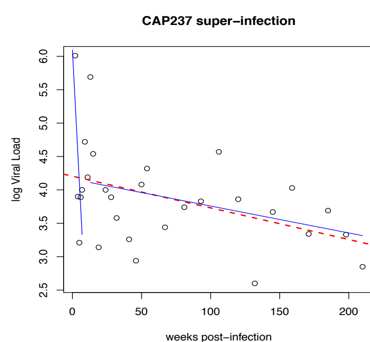
[a] HLA-B*5703, HLA-B*5801
[b] HLA-B*18, HLA-B*5802, and HLA Cw0602

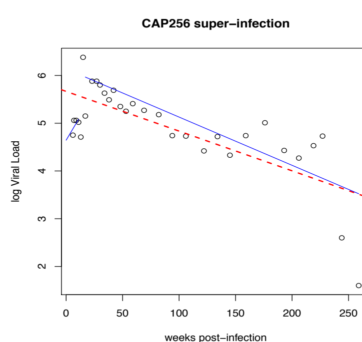
To determine whether superinfection resulted in accelerated disease progression within participants, we used a Chow Test (173) to establish whether superinfection was associated with a change in the viral load and CD4+ T cell count trajectories. The Chow test partitions the data prior to, and post superinfection, and tests whether the coefficients of the two regression lines are the same. This approach found no significant change in the CD4+ T cell count trajectories following superinfection in CAP237 and CAP256 (p=0.311, 0.49, respectively) but did find evidence of an accelerated CD4 decline in CAP281 ($p=9.07 \times 10^{-6}$) (Table 3.4). However, there was evidence of a structural change in viral load trajectory following superinfection in all three participants (p = 0.043, 0.01, 2.99×10^{-7}, for CAP237, CAP256 and CAP28,1 respectively) (Table 3.4). Comparing the slopes of the curves prior to and post superinfection revealed that in two participants (CAP237 and CAP281), there was a significantly greater viral load slope after superinfection than before (Figure 3.4A and C). However, for CAP256, the slope after superinfection was significantly less (Figure 3.4B). Furthermore, this analysis was likely

confounded by the fact that superinfection in two participants occurred early (approximately nine and 13 weeks post infection respectively), prior to viral load reaching an equilibrium set point. Therefore, as a control, we used the Chow Test on all of the singly infected participants using the same three partitions on viral load. There was significant evidence of a change in viral load trajectory, according to the Chow test, for 3/11 (27.27%) and 9/18 (50%) of singly infected participants at nine and 13 wpi, respectively (for seven participants, HIV infection was likely only detected after nine wpi and therefore the denominator for the first partition is 11). There was also significant evidence for a structural change in viral load dynamics in 18/18 (100%) participants at 40 wpi, in the absence of any evidence of superinfection. This illustrates that the Chow test has a low specificity to detect alterations in disease progression, despite being previously used for this purpose (165).

Taken together, these results suggest that dual infection did not significantly accelerate disease progression in the 11 participants described here. However, the small sample size here would not have had sufficient power to detect a modest effect (supplementary table S3.1)

A. CAP237 B. CAP256 C. CAP281





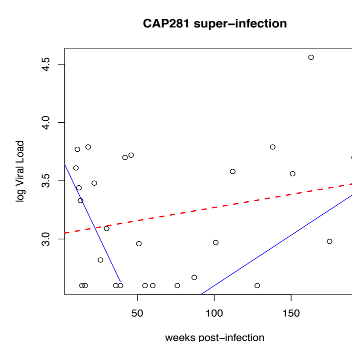


Figure 3.4 | Viral load trajectories prior to- and post-superinfection in CAP237 (A), CAP256 (B), and CAP281 (C) are shown in blue. The overall trend in viral load for each participant is fitted with a dashed red line.

Table 3.4 | Results of a Chow test comparing linear regressions of viral load and CD4 counts prior to and following superinfection in three participants.

	SI window	**P(CD4 count)**	***P(viral load)***
CAP237	09 - 11 wpi	0.311	**0.0430**
CAP256	11 - 15 wpi	0.49	**0.0138**
CAP281	40 - 46 wpi	**$9.07x10^{-6}$**	**$2.98x10^{-7}$**

3.4 Discussion

The susceptibility to infection with multiple variants may be indicative of a higher vulnerability to infection. Here we assessed the relationship of multiple HIV infection with the presence of STIs, and showed that participants with Gonorrhoea were ~11 times more likely to be infected with multiple variants than singly infected controls, though this did not reach statistical significance (p=0.0566). The association between Gonnorhoea and multiple HIV transmission is consistent with previous observations identifying Gonnorhoea as a risk factor for HIV acquisition (174, 175). Indeed, in this cohort, Gonorrhoea was associated with an approximately 4.6 times higher risk of HIV-1 infection (155). Interestingly, of all the STIs, the strongest associations with multiple transmission were detected with Gonorrhoea and Chlamydia, both of which are inflammatory and target the columnar epithelium. HSV-2 and Syphilis, which are ulcerative infections, have also been identified as strong risk factors for HIV acquisition (reviewed in 150) even after controlling for other confounding factors (176). However we could not address their role in multiple HIV transmission due to insufficient data.

Here we also sought to identify the clinical consequences of infection with more than one divergent virus. Comparing 11 cases of HIV intra-subtype C dual-infection to 48 participants from the same cohort for which we possessed no evidence of dual infection, we found no significant difference in the rate of CD4+ T cell decline, or in their viral loads, even after controlling for host HLA alleles. This is seemingly in contrast to studies in several other cohorts, where dual infection was associated with substantially accelerated disease progression (130), higher viral loads (131), and shorter time to WHO disease events (177). However, cases and controls were frequently not matched for gender, infecting subtype, or host HLA alleles, and the dates of initial infection were often unknown or estimated. The most robust study to date on the influence of dual infection on clinical progression analyzed a well-characterized cohort of 37 Caucasian men who have sex with men (MSMs), matched for age (166). In this cohort, dual infection was associated with a modest but significantly higher rate of CD4 decline, after

controlling for HLA B alleles previously associated with rapid progression and for CCR5Δ32 homozygosity. However, while all singly infected controls were infected with subtype B, four of the six superinfections occurred with a non-B subtype, including subtype A and recombinant forms AG and AE, confounding the analysis.

One approach to control for differences between hosts is to test whether superinfection accelerates disease progression within individuals. In a cohort of MSM from San Diego, USA, there was no significant difference in CD4+ T cell decline between dual infected and monoinfected participants (165) . To evaluate whether superinfection resulted in an acceleration in disease progression *within* individuals following superinfection, Pacold et al. used an econometric test (Chow test) to test for a "structural break" in viral load or CD4+ T cell dynamics at the time of superinfection (165). For four out of seven superinfected participants, a significant shift in viral load dynamics at the time of superinfection was detected by the Chow test, suggesting that superinfection can result in higher viral loads (165) . However, superinfection was not associated with a shift in CD4+ T cell count trajectory. Using the same approach, we show here that superinfection was not associated with a shift in CD4+ T cell count trajectory in two out of three superinfected participants. While we found significant evidence of structural breaks in viral load at the time of superinfection in all three superinfected participants, in two of the cases, this was due to a significant decline in the slope of viral load. This was perhaps to be expected, as in these two cases, superinfection occurred early, prior to viral load reaching set point. Importantly, we also show here that the Chow test detects significant structural breaks in the viral load trajectories at the same timepoints in many monoinfected control participants. For the three partitions on viral load, the Chow test detects significant evidence ($p<0.05$) of a structural break in viral load in 3/11 (27.27%), 9/18 (50%), and 18/18 (100%) of monoinfected controls respectively. This is perhaps unsurprising given the volatility of viral load, and the fact that the trajectory is not expected to be linear. However, it does illustrate that the Chow test does not have a high specificity as a means to evaluate the effect of superinfection on viral load, despite previously being used for this purpose (165).

While dual infection was not significantly associated with accelerated disease progression in this cohort, we acknowledge caveats in the present study. Cases of dual infection, particularly intra-subtype dual infection, can be difficult to identify, and consequently, we may have misclassified cases of dual infection as singly infected. Further, we possessed limited (only population level) sequencing data for several (N=6) "singly infected" participants in the cohort. However, as superinfection is a relatively rare event, the number of misclassified participants is expected to be low. Nevertheless, a sample size of only 11 cases of dual infection likely would not be sufficient to detect a mild effect of dual infection on disease progression. Particularly as disease progression has high variance, is complex, and can be affected by many factors both host and viral.

Due to difficulties identifying cases of dual infection, studies to date have analyzed relatively few cases, and individually, offer little power to establish the clinical consequences of dual infection. Given significant heterogeneity in factors including ethnic makeup, infecting subtype, and mode of transmission, it also may not be appropriate to pool the data from several cohorts in a meta-analysis. Nevertheless, HIV superinfection has the inherent risk that the superinfecting variant is more virulent (178), or harbours drug resistance mutations (137) and therefore still poses a public health concern. Consequently, serosorting practices should be discouraged in any event and HIV concordant partners should still protect themselves from superinfection. However, larger studies in well-characterized cohorts will be required to definitively address the clinical significance of HIV dual infection.

CHAPTER 4: Dual infection was not associated with greater breadth of the neutralizing antibody response

4.1 Introduction

Approximately 20% of HIV infected individuals develop broadly cross-neutralizing antibodies during the course of infection (1, 85, 86, 179). Identifying the factors that promote their development during natural infection may provide valuable information for the design of future vaccines. Given the tremendous diversity in circulating HIV strains, polyvalent immunogens may be necessary to provide adequate vaccine coverage. HIV dual infection, the infection with more than one HIV strain, provides a unique opportunity to characterize how the immune system responds to exposure with multiple HIV antigens, and can inform the use of polyvalent or heterologous prime-boost immunizations.

Piantadosi *et al.* (72) previously examined the correlates of neutralizing antibody (nAb) breadth, and found that in 26 women, greater *env* diversity early in infection was significantly associated with greater antibody breadth, measured at five years post infection. As viral load has been shown previously to correlate with both neutralizing antibody breadth and greater sequence diversity, this association may have been driven by viral load. Indeed, in the study by Piantadosi *et al*. (72) viral load set point was significantly correlated with breadth. However, there was no significant association between Gag diversity and breadth, suggesting that *envelope* diversity may be directly associated with breadth.

Two recent studies that specifically examined the effect of HIV dual infection on the breadth and potency of the neutralizing antibody response both suggested that dual infection is associated with the development of a broader, and potentially more potent, nAb response later on in the course infection (162, 180). In a cohort from Yaoundé, Cameroon, superinfection was associated with a significant broadening of the neutralizing antibody (180), although individuals were infected for an unknown duration, confounding the analysis. Cortez *et al.*, (162), on

comparing 12 superinfected participants to 36 monoinfected, matched controls, found a significant association between superinfection and neutralization breadth, even when adjusting for contemporaneous CD4+ T cell count and viral load. While the duration of infection was known the timings of superinfections ranged to over 5 years post infection. As breadth typically develops on average ~2.4 years post infection (105), these results are potentially confounded by broadly neutralizing antibody responses that developed prior to, and independent of superinfection. Indeed, in this cohort, participants that later went on to be superinfected already had significantly broader neutralizing antibody responses than monoinfected controls, prior to superinfection.

These results stand somewhat in contrast to the immunodominant nature of the early nAb response to HIV infection, which is largely focused on one or a few variable epitopes resulting in typically type-specific nAb responses. Should an association between dual infection and broadly neutralizing responses in natural infection be validated, this would support the use of polyvalent immunogens and prime-boost regimens as a potential means to circumvent this immunodominance. In this chapter, we explore the influence of dual infection on the neutralizing antibody response. Of the 62 participants in the CAPRISA002 cohort, plasma samples from 41 have been screened for the presence of broadly neutralizing antibodies (1). Taking advantage of these data, we evaluate whether infection and exposure to a diversity of envelopes, through either co-infection or sequential infection (superinfection) within the first year of infection, resulted in a greater likelihood of developing cross-neutralizing antibodies.

4.2 Methods

4.2.1 Sequence analysis

Sequences were aligned using Muscle (181). All amino acid positions are numbered relative to the HXB2 reference sequence. DNA distances were calculated in MEGA 5.2.2 (182), using Maximum Likelihood with gamma distributed rates among sites (*gamma* = 1). Diversity was estimated as the mean DNA distance between all sequences.

4.2.2 Detecting evidence of selection in CAP256

To take recombination between the superinfecting virus (SI) and primary infecting virus (PI) into account, sequences were partitioned into two alignments (an SI-related, and a PI-related alignment) based on the inferred recombination breakpoints using an in-house script. Breakpoints were identified by a shift in identity from one reference towards the other, and required at least two sequential polymorphisms in common with a corresponding PI/SI-related virus in order to be considered. Phylogenies for both alignments were then reconstructed using FastTree (143) with a GTR+CAT model, and rooted on the PI/SI. Signals of selective pressure were detected with MEME (episodic diverifying selection, 183) and DEPS (directional selection, 184) using the FastTree-generated trees, implemented in Hyphy (185).

4.2.3 Cloning

Select *envelope* amplicons were re-amplified from the SGA first-round products using the high fidelity, Phusion enzyme (Finnzymes, Thermo Fisher Scientific, Vantaa, Finland) with the EnvN primer (5'–TTGCCAATCAGGGAAGTAGCCTTGTGT-3'; HXB2 nt 9145 - 9171) and a forward primer containing a 5' CACC overhang allowing directional cloning (Env1Arx: 5' - CAC CGG CTT AGG CAT CTC CTA TAG CAG GAA GAA-3'; HXB2 nt 5954 - 5982). Amplification was performed under the same cycling conditions as the second round SGA, described above (see 2.2.2). Amplicons were ligated into pcDNA3.1-Topo Directional Cloning Vector (Invitrogen, Life Technologies, Carlsbad, CA), and transformed into Top10 Chemically Competent *E. Coli* cells (Invitrogen, Life Technologies, Carlsbad, CA) according to the manufacturer's instructions, and

cultured on Luria-Bertani agar supplemented with 100 µg/ml Carbenicillin (Sigma-Aldrich, St Louis, MO). Colonies were screened by PCR, using the following primers: T7 (5'-TAATACGACTCACTATAGGG-3') and Rev15 (5'-CTGCCATTTAACAGCAGTTGAGTTGA-3'; HXB2 nt 6990 - 7015). Colonies were added to 12.5 µl reactions containing 1X SuperTherm PCR buffer, 2 mM $MgCl_2$, 0.16 mM of each dNTP, 0.375 units of SuperTherm Taq (Medox Biotech, India), and 0.3 µM of each primer. Reactions were cycled under the following conditions: 94°C for 3 minutes, followed by 35 cycles of [94°C for 15 seconds, 57°C for 45 seconds, and 72°C for 2.5 minutes], with a final extension of 7 minutes at 72°C.

Large-scale plasmid purification was performed using the QIAprep Spin Miniprep, or the QIAfilter Plasmid Midi kits (Qiagen, Valencia, CA), as per the manufacturer's instructions. Plasmids were sequenced and assembled as described above (section 2.2.2) in order to ensure the clone contained no non-synonymous mutations in *envelope* relative to the SGA-derived sequence.

4.2.4 Cell lines

TZM-bl (JC53-bl) cells, engineered by J. Kappes and X. Wu, were obtained from the NIH AIDS Research and Reference Reagent Program (cat# 8129). TZM-bl cells are an adherent, HeLa-derived cell line that has been modified to constitutively express CD4, CCR5 and CXCR4. Further, they contain an integrated luciferase gene under the tight regulatory control of an HIV LTR (a Tat-responsive promoter). HEK293T cells were obtained from George Shaw (University of Alabama, Birmingham, AL). Both cell lines were maintained in Dulbecco's Modified Eagle Medium (Gibco, Life Technologies, Carlsbad, CA), containing 4.5 g/L glucose, L-glutamine, sodium pyruvate, and supplemented with 50 µg/ml Gentamicin (Sigma-Aldrich, St Louis, MO), 25 mM HEPES (Sigma-Aldrich, St Louis, MO), and 10% heat inactivated Fetal Bovine Serum (FBS) (Biochrom, Cambridge, UK). Cells were cultured at 37 ºC in a humidified incubator with 5% CO_2, and monolayers were disrupted at confluence by treatment with 0.25% Trypsin in 1mM EDTA (Sigma-Aldrich, St Louis, MO).

4.2.5 Neutralization assays

Env-pseudotyped viruses were generated by co-transfecting Envelope plasmids with pSG3 Δ Env (56) into HEK293T cells using Fugene 6 (Applied Science, Indianapolis, IA) according to the manufacturer's instructions. Pseudoviruses were harvested from the supernatant 48 hours following transfection, filtered through a 0.45 μm filter (Millipore, Merck, Billerica, MA), made up to 20% FBS, and stored at -80°C until use. Neutralization assays were performed as described previously (58). Briefly, plasma was heat inactivated at 56°C for 60 minutes in order to remove complement activity, and serially diluted 3-fold in a 96 well plate, beginning at a 1:45 dilution. Pseudotyped viruses were then thawed in an ambient temperature water bath and diluted by a factor that would result in ~50,000 relative light units (RLU) per well upon infection of TZM-bl cells for 48 hours in the absence of plasma (as previously determined by serial dilutions of that stock), added to each well, and incubated at 37°C for 1 hour. Thereafter, 10,000 TZMbl cells, supplemented with 20 μg/ml DEAE-Dextran (Sigma-Aldrich, St Louis, MO), were added to each well (final concentration of 8 μg/ml DEAE Dextran), and incubated at 37°C for 48 hours. Controls included cells unexposed to virus or plasma (cell control), and cells exposed to virus in the absence of plasma (virus control). Cells were analysed for Luciferase expression using Bright-Glo (Promega, Madison, WI) as a substrate and quantifying the luminescence produced, in RLUs, using a Luminometer (Promega, Madison, WI), as per manufacturer's instructions. Neutralization was measured as the reduction in luminescence relative to the virus control, and titers were calculated as the reciprocal plasma dilution required to reduce the RLUs by 50% (ID_{50}).

Neutralization breadth for the cohort was previously characterized, and quantified as the proportion of heterologous viruses from a multi-subtype pseudovirus panel each plasma sample was able to neutralize at ID_{50} titers >45 (1). Neutralization breadth at three years post infection (ypi) for nine dual-infected participants was compared to that of 16 monoinfected participants for whom data was available.

4.2.6 Statistical analysis

Antibody breadth was compared between groups using a Mann-Whitney test (for two groups) or a Kruskal-Wallis test with Dunn's multiple comparison post tests (for multiple groups), implemented in Prism 5.0b (Graphpad Software, Sand Diego, CA). The relationships between *envelope* diversity, set point viral load, CD4 counts and neutralizing antibody breadth were assessed with a linear regression. The multivariate analysis of breadth at three years post infection was modeled with a linear mixed effects model in R, using the nlme package.

4.3 Results

In this chapter, we firstly evaluated whether infection and exposure to a diversity of envelopes, through either co-infection or superinfection within the first year of infection, resulted in a greater likelihood of developing cross-neutralizing antibodies. The second component of our work investigated the autologous neutralizing antibody responses to different infecting strains or variants in individuals with dual infection. To do this analysis, we utilized published neutralizing breadth data, determined at approximately three years post infection, available for 41 participants from the CAPRISA 002 cohort (1) . In total, 322 sequences were analyzed from 17 participants (two multivariant, five co-infected, and ten singly infected) (see Chapter 2). The mutagenesis and neutralization assays for CAP256 were performed by P. Moore, M. Madiga, and M. Nonyane (NICD, JHB, South Africa).

4.3.1 Dual infection was not associated with a broader neutralizing antibody response

To determine whether dual infection was associated with the development of broadly neutralizing antibodies, we partitioned participants according to the multiplicity of infection. Out of the 41 participants with available heterologous neutralization data, we possessed sequencing data sufficient to estimate the multiplicity of infection for 27. We classified six of these participants as co-infected, and three as superinfected (see Chapter 2). One participant was dual infected, although the timing of the secondary infection remains unclear (CAP200). Seventeen of the 41 participants, were infected with a single virus, and had no evidence of superinfection over at least two (N=7) or three (N=10) years of infection by SGA (N=11) or population sequencing (N=6).

The extent of neutralization breadth was estimated using a pseudovirus panel consisting of 44 heterologous envelopes encompassing subtypes A, B, and C (1). The median breadth against heterologous pseudoviruses was 11% for the 17 singly infected participants, compared to 10% and 5% for the co-infected (N=6) and superinfected (N=3) participants, respectively (Figure 4.1). Three of the 17 (17.7%) singly participants developed responses capable of neutralizing

>40% of the heterologous panel at ID_{50} titers greater than 45. Of the dual infected participants identified, one superinfected participant, CAP256, developed an extremely potent, broadly cross-neutralizing antibody response, and was able to neutralize 32/42 (~76%) of viruses, frequently at ID_{50} titers exceeding 10,000 (1, 2). However, except for CAP256, none of the dual-infected participants were able to neutralize more than 25% of the panel at ID50 titers >45. Plasma samples from CAP237, and CAP281 who, like CAP256, were superinfected, were unable to neutralize more than 5% of heterologous viruses tested (1).

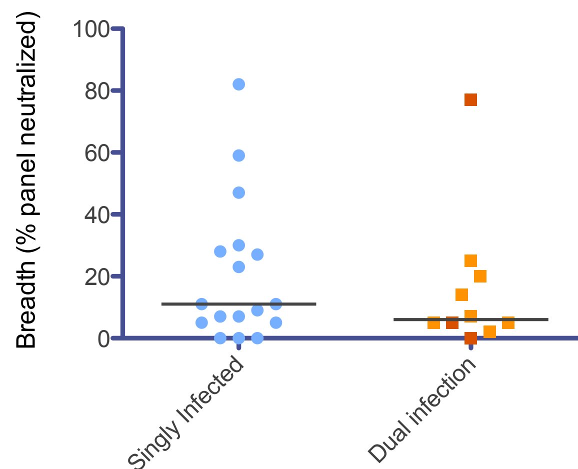


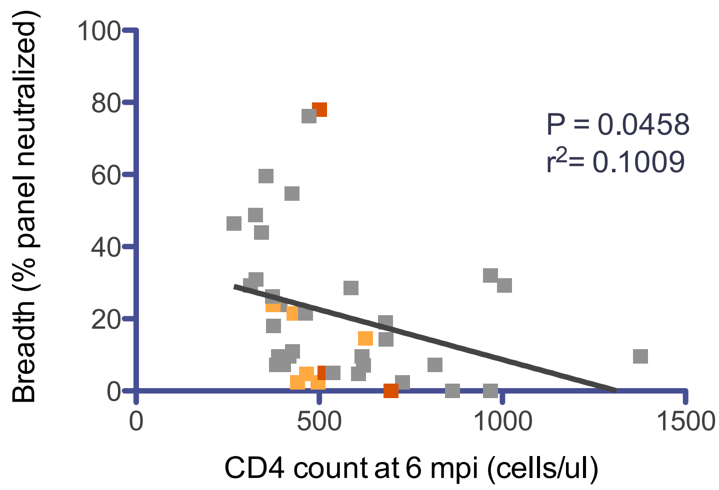
Figure 4.1 | No significant association between HIV dual infection and the later development of neutralizing antibody breadth, estimated as the percent of a 44-pseudovirus panel neutralized at ID_{50} titers >45. Within the dual infected group, superinfected participants are highlighted in red.

We found no association between dual infection (whether co-infection, or superinfection) and a broader neutralizing antibody response at three years post infection compared to single infection (p=0.4489). To determine whether the timing of the second infection influenced the breadth of the response, participants were also grouped into single infections, co-infections, and

superinfections. However, there remained no significant difference between either of the dual infected groups, and the monoinfected participants (p=0.8371).

As viral load and CD4 count have previously been demonstrated to be significantly associated with the development of breadth, we performed a multivariate analysis. In this cohort, Gray et al. found that both viral loads and CD4 counts at six months post infection were significantly associated with neutralization breadth (1). However, we found that the area under the load viral load curve from 12 to 52 weeks post infection was more predictive of breadth than viral load at six months post infection was. We therefore used this representation of viral load set point in the multivariate model. Dual infected participants did not possess more breadth, on average, than participants with similar viral loads or CD4 counts (Figure 4.2). After adjusting for viral load and CD4 counts, there was still no significant association between dual infection and the development of cross-neutralizing antibodies in this cohort (p=0.2386; Table 4.1). If anything, dual infection, was associated with less neutralization breadth, on average, taking viral load and CD4 counts into account (B<0, p=0.1196).

A.



B.

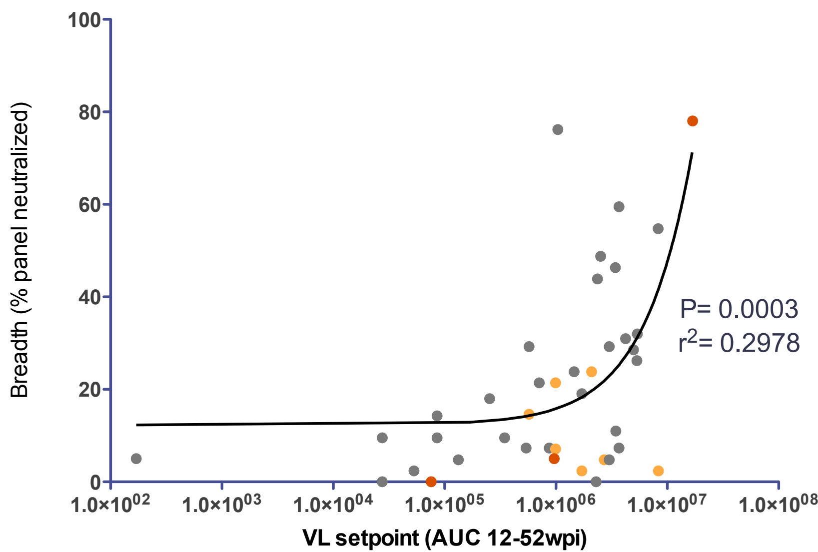


Figure 4.2 | The association between A. CD4 counts at 6 months post infection (mpi) and B. viral load setpoint and neutralizing antibody breadth. Viral load setpoint was estimated as the area under the viral load curve from 12 to 52 weeks post infection. Coinfected participants are highlighted in orange, and superinfected participants are highlighted in red.

Table 4.1 | Results of a multivariate model testing the association between dual infection and the development of neutralizing antibody breadth taking into account set point viral load and CD4 count

	N	*P*	P(multivariate)**
CD4 count*	41	0.0652	**0.4046 (0.4303)**
VL setpoint#	41	0.0018	**0.0086 (0.0061)**
Dual infection	10	0.4359	**(0.1196)**
Coinfection	6	0.2819	**0.1604**
Superinfection	3	0.5458	**0.6388**

*CD4+T cell counts measured at 6 months post infection
#Estimated using the area under the viral load curve from 12 to 52 weeks post infection.
** The P values for the multivariate model when grouping coinfections and superinfections together under "Dual infection" are shown in parentheses.

In other cohorts greater *envelope* diversity early in infection was found to be significantly associated with greater antibody breadth (72). We therefore investigated if the extent of early *envelope* genetic diversity was correlated with the later development of cross-neutralizing antibodies. In total, 322 sequences were analyzed from 17 participants (two multivariant, five co-infected, and ten singly infected). Sequences were sampled a median of two weeks post infection (range: 1-9 wpi), and cases of superinfection were excluded. Analyzing a median of 21 sequences per participant (range: 7-24), we estimated early *envelope* diversity as the mean DNA distance. We found no significant correlation between early *envelope* diversity and the later development of cross-reactive antibodies (p=0.4825; Figure 4.3).

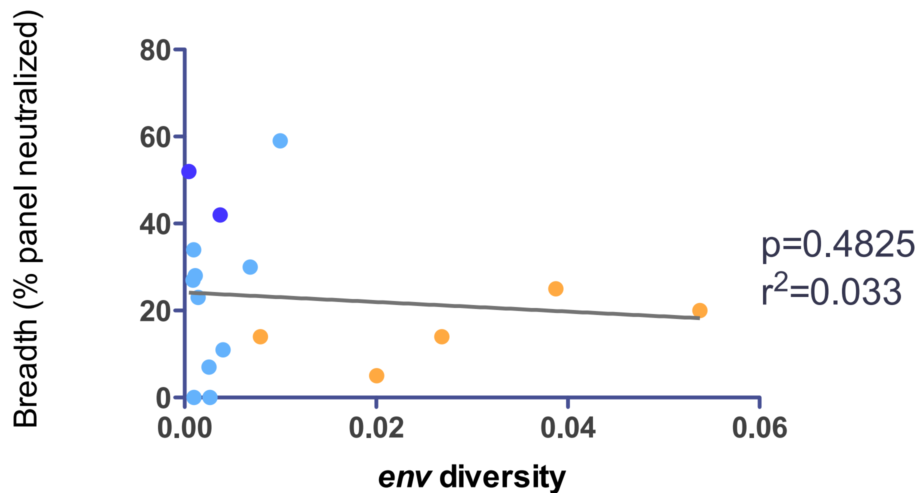


Figure 4.3 | The association between early diversity in *envelope(env)* and the later development of neutralizing antibody breadth. Diversity was plotted as the mean DNA distance between all single genome amplicons sampled near the time of seroconversion. Co-infected participants are shown in orange, while singly infected participants are shown in light blue. Participants infected with multiple, closely related viruses (multivariant transmissions) are shown in dark blue.

4.3.2 Broadly neutralizing antibodies in CAP256 target an epitope including residues in V2 of the superinfecting virus that is not fully conserved in the primary infecting virus

CAP256, a superinfected participant, developed an extremely potent, broadly cross-neutralizing antibody response, neutralizing 77% of heterologous viruses tested, often at titers exceeding 10,000 (1). Identifying the target of the cross-reactive antibody response in CAP256 may shed light on whether superinfection may have promoted this broad and potent response. Moreover, the identification of the epitope(s) targeted by these broadly cross-neutralizing antibodies could potentially identify vulnerable regions in the HIV Envelope that can be targeted with vaccines. We therefore sought to map the epitope(s) targeted by cross-neutralizing antibodies present in CAP256 plasma.

To elucidate responses to the primary and superinfecting virus, and to map the epitope(s) targeted by CAP256 plasma we generated a functional *envelope* clone of the T/F superinfecting virus described in Chapter 2. This clone was provided to P. Moore *et al.* (NICD, JHB, South

Africa) for analysis. The neutralizing antibody response to the primary infecting virus in CAP256 arose by ~23 wpi, reaching titers of approximately 300, and persisted at approximately this level over the next three years (2) (Figure 4.6). While superinfection occurred at approximately 11-13 wpi, titers to the PI were not boosted after this window. Similar to the response to the PI, neutralizing ID_{50} titers against the SI first reached detectable levels by ~23 wpi, and reached titers of ~300 at 30 wpi. However, while titers against the PI persisted at approximately this level, the nAb response against the SI became considerably more potent, rising at 42 wpi and reaching extremely high titers of >40,000 by 69 weeks of infection (Figure 4.7). The rise in titer to the superinfecting virus correlated with the onset of a novel response targeting an epitope within V1V2 that arose by 30 wpi (Figure 4.6)(2). A V1V2-dependent response was also shown to mediate neutralization of the PI after 42 wpi (186).

To aid in the effort to map the fine specificity of the cross-neutralizing antibody response in CAP256, we made use of the *envelope* sequences of the viruses used in the heterologous virus panel. We first identified residues that were highly conserved in the viruses that were sensitive to neutralization by CAP256 serum, but were variable in the resistant viruses. While we were aware that escape from neutralizing antibodies can occur by conformation shielding and mediated by mutations potentially distal to the actual epitope, we initially identified whether any of the sites that were conserved in the sensitive viruses were also under positive selection over time. As CAP256 was superinfected, we partitioned the sequences according to the inferred recombination breakpoints, and performed a selection analysis on the two alignments. Sites in V2 were under strong selective pressure in both the primary- and superinfecting viral sequences. Of the sites experiencing strong positive selection, three (namely HXB2 positions 165, 166, and 169) were conserved in sensitive viruses, but less so in resistant viruses, suggesting that they were potentially contributing to the formation of the targeted epitope (Figure 4.5A). CAP256 plasma had also been evaluated against a panel of over 160 heterologous pseudoviruses as part of the Neutralization Serotype Discovery Project (D. Montefiori *et al.*, unpublished data). Taking advantage of these data, we next compared the

amino acid frequencies between the highly sensitive (ID_{50}> 5000) and resistant (ID_{50} < 50) viruses. Supporting the role of residues 165, 166, and 169 in the neutralization by CAP256 serum, these sites were significantly more conserved in the highly sensitive viruses compared to the resistant viruses (Figure 4.5B).

```
                          160       170       180
                     |....|....|....|....|....|.
PI                   KNCSFNTITEVRDKQKKEYALFYRLD
SI                   ......AT..L...K...........
CAP256.06wpi.F7J     ..........................
CAP256.06wpi.C7J     ..........................
CAP256.13wpi.F1      ..........................
CAP256.30wpi.1       ..........................
CAP256.30wpi.2       ..........................
CAP256.30wpi.9       ..........................
CAP256.30wpi.3       ..........................
CAP256.30wpi.4       ..........................
CAP256.30wpi.8       ................R.........
CAP256.53wpi.A9      .......T..................
CAP256.53wpi.C1      ...........T......H.......
CAP256.53wpi.F1      .....D....................
CAP256.53wpi.C2      .....D....................
CAP256.59wpi.8b      ................E.........
CAP256.59wpi.12a     ..........................
CAP256.59wpi.9a      .......T..................
CAP256.59wpi.2b_cloned .....D....................
CAP256.59wpi.5b      .......T..................
CAP256.59wpi.6b      .......T..................
CAP256.59wpi.13B     .......T..................
CAP256.59wpi.10a     .......T..................
CAP256.59wpi.4a      .......T..................
CAP256.59wpi.6a      .....T....................
CAP256.59wpi.11a     ..........................
CAP256.59wpi.5a      ..........................
CAP256.59wpi.7a      ..........................
CAP256.59wpi.A11     ................N.........
CAP256.176wpi.C2     ...........S....N......K..
CAP256.176wpi.B1     ...........S....N......K..
CAP256.176wpi.E1     ...........S....N.........
CAP256.176wpi.H1     ...........S....N.........
CAP256.176wpi.F10    ...........S....N.........
```

```
                          160       170       180
                     |....|....|....|....|....|.
SI                   KNCSFNATTELRDKKKKEYALFYRLD
CAP256.15wpi.B3      ..........................
CAP256.15wpi.C7      ..........................
CAP256.15wpi.C9      ..........................
CAP256.15wpi.C1      ..........................
CAP256.15wpi.A5      ..........................
CAP256.53wpi.D7      ..............I...........
CAP256.53wpi.D1      ..............I...........
CAP256.59wpi.1a      .....S........I...........
CAP256.59wpi.2a      ........S.....E...........
CAP256.59wpi.7b      ......T.......I...........
CAP256.94wpi.D12     ......T....K..I...........
CAP256.94wpi.A11b    ......I...V...I...........
CAP256.94wpi.A3      ......T....K..I...........
CAP256.94wpi.A6      ......T....K..I...........
CAP256.94wpi.A1      ......T....K..I...........
CAP256.94wpi.F1      ......T....K..I...........
CAP256.94wpi.C4      ......T....K..I...........
CAP256.94wpi.B2      ......T....K..I...........
CAP256 94wpi.E12     ......T....K..I...........
CAP256.94wpi.B10     ......T....K..I...........
CAP256.94wpi.F       ..........VK..I...........
CAP256.94wpi.C1      ..........VK..I...........
CAP256.94wpi.D6      ..........VK..I...........
CAP256.94wpi.C2      ..........VK..I...........
CAP256.94wpi.A10     ..........VK..Q...........
CAP256.176wpi.F1     ..........V...E...........
```

Figure 4.4 | An alignment of V2 sequences related to the primary virus (A) and the superinfecting virus (B) in CAP256, with sites under positive selection highlighted in red. wpi, weeks post infection.

A. B.

	165	166	169
Highly Sensitive			
ConC	L	R	K
C. CAP206	I	.	.
C. CAP210	.	.	.
C. CAP228	I	.	.
C. CAP239	.	.	.
C. DU422	.	.	.
C. DU156	.	.	.
C. ZM233	.	.	.
C. 53M12	.	.	.
Resistant			
C. CAP85	.	.	E
C. CAP88	I	.	.
C. CAP244	.	W	.
C. DU151	I	.	.
C. DU172	I	T	.
B. CAAN	M	S	M
B. AC10	M	.	M
B. RHPA	I	.	V
B. THRO	V	.	V
B. SC42	I	.	V
B. TRO	I	.	V
B. REJO	P	.	I

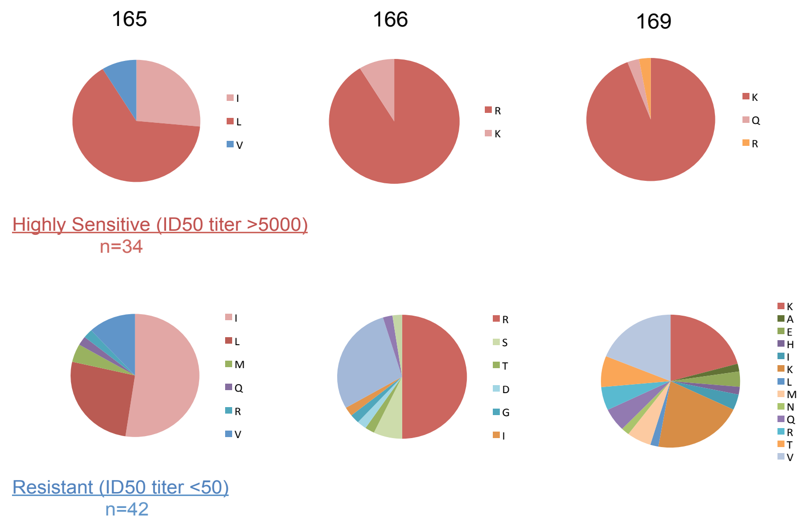


Figure 4.5 | Amino acid conservation at sites 165, 166, and 169 in highly sensitive and resistant viruses from the (A) 44 virus panel initially used to assess breadth (1)**, and validated on the large NSDP pseudovirus panel (B) (D. Montefiori *et al*., unpublished data)**

Mutations at residues 165, 166, and 169, introduced by site-directed mutagenesis, were subsequently shown by P. Moore *et al.* (2) to reduce sensitivity to neutralization, confirming their role in neutralization by CAP256 serum. The introduction of an alanine into a consensus subtype C clone (ConC) at position 165 (L165A) and at position 166 (R166A) reduced titers to CAP256 serum by 14.6 and 102.3 fold respectively. A lysine to glutamate mutation at position 169 (K169E) reduced titers by 78.5 fold. Further mapping efforts utilizing scanning mutagenesis also identified a number of adjacent residues, including sites 159, 160, 165, 167, and 171, though 166 and 169 were the major determinants of sensitivity in heterologous viruses (2). Mapping of the potent V1V2 response against the superinfecting virus found that, similar to the broadly neutralizing response, it was also highly dependent on the residues at 162, 166, 169, and 171 (2), suggesting that this may represent the precursor of the broadly neutralizing response.

To establish whether the likely precursor of the broadly neutralizing antibody response arose to the PI, the SI, or an epitope conserved in both viruses, we analyzed the sequences of the two infecting viruses in conjunction with the autologous neutralizing antibody response. The PI differed from the SI sequence at positions 161, 162, 165, and 169 (Figure 4.7). These mutations explain the relative resistance of the PI compared to the SI. However, they also reveal that the nAb response in CAP256 did not target an epitope fully conserved in both infecting viruses, but rather was likely specific for the superinfecting virus.

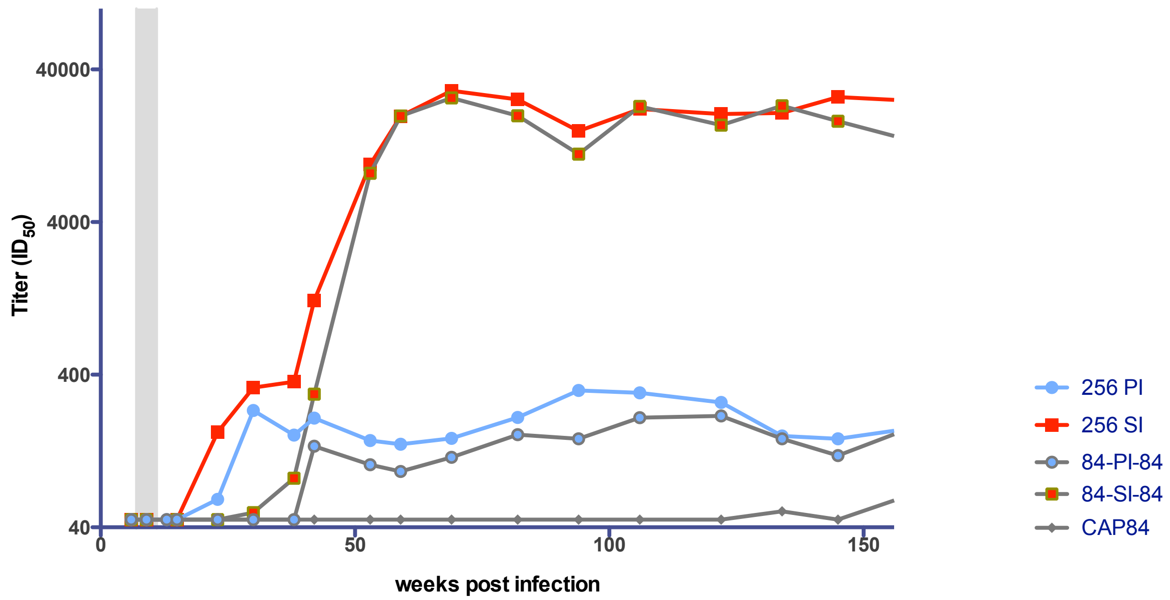


Figure 4.6 | The autologous neutralizing antibody response in CAP256. The estimated window of superinfection (7-11 wpi) is shaded in gray. PI, primary infecting virus; SI, superinfecting virus. In order to differentiate a V2 response (and therefore a potential precursor of the cross-neutralizing response), the neutralization of the PI and SI is overlaid onto the neutralization of chimeric envelopes consisting of a resistant heterologous Envelope (CAP84) with the V1V2 replaced with that of either the PI (84-PI-84) or SI (84-SI-84). Data provided by P. Moore (NICD, JHB, South Africa)

	*	*					*	**	*	*	**		*
	159	160	161	162	163	164	165	166	167	168	169	170	171
SI	F	N	A	T	T	E	L	R	D	K	K	K	K
PI	.	.	T	I	.	.	V	.	.	.	Q	.	.

Figure 4.7 | Sequence homology in V2 between the primary- (PI) and superinfecting (SI) viruses in CAP256. Sites identified in Moore et al., (2011) as import for cross-neutralization are highlighted with a * (mutations at this site resulted in a 5-15 fold reduction in titers) or ** (mutations at this site resulted in >75-fold reduction in titers)

4.3.4 Escape from autologous neutralizing antibodies and the relationship with heterologous neutralization

Next, in order to address whether viral evolution, recombination, and escape may have driven the antibody response in CAP256 towards more conserved residues (and whether superinfection contributed to this), we characterized the evolution of the viral quasispecies in CAP256 over time. An analysis of 26 SGA sequences, together with 40 sequences provided by P. Moore, revealed that recombination within *envelope* was widespread (see Figure 2.12B). By six months post initial infection (~three months post superinfection) the sampled viral population was comprised entirely of recombinants. Within gp120, sequences originating from the SI dominated with the exception of V1V2, which was predominantly inherited from the PI, and contained several mutations that confer resistance to the nAb response. By six months pi, recombinant viruses harbouring the V1V2 of the relatively neutralization resistant PI dominated. As the V2 region of the PI harboured several mutations sufficient for escape (Figure 4.7) and was significantly more resistant to neutralization, this suggests that recombination was used as an initial mechanism to provide (at least partial) escape from the extremely potent V2 response. Consistent with this, titers to the recombinant virus cloned at six months pi were shown by P. Moore to be ~20 fold lower than to the SI (186).

Partitioning the alignment according to the recombination breakpoints, we were able to identify mutations and signals of selective pressure within the recombinant sequences. An analysis of the residues known to be important for neutralization revealed that, in addition to mutations introduced by recombination, multiple point mutations were introduced (Figure 4.4). V2, and specifically residues 165, 166, and 169, and 171 were under strong selection pressure, consistent with escape from the potent V2 antibody response (see Supplementary Figure S4.1). By 39 months post infection, two distinct viral populations were evident. One cluster contained the backbone V2 of the PI with additional mutations at positions 166 and 171. The other contained the V2 from the SI, and harboured mutations at positions 165 and 169.

Taken together, these results indicate that despite an extremely potent, and later broad, neutralizing antibody response, multiple pathways to escape were possible. They also provide evidence that recombination can potentially accelerate escape, allowing the introduction of multiple escape mutations simultaneously.

As the broadly neutralizing response in CAP256 was associated with subtle changes in epitope specificity (2), we sought to identify whether the patterns of escape in the autologous virus preceded changes in the antibody specificity. We know from prior neutralization studies that heterologous neutralization by CAP256 serum becomes less dependent on residues 165, 171 and to a lesser extent 167 over time (2). Analyzing the autologous sequence data from CAP256 (Figure 4.8) revealed that D167 was conserved over time in all of the autologous sequences, indicating that escape at that position was not responsible for driving the specificity of the antibody response away from this residue. While the SI possessed a leucine at 165 and the PI possessed a valine, no other mutations were observed at position 165 either. Interestingly, a K171N mutation arose by two years post infection and became dominant by three years, coincident with the neutralizing response becoming less dependent on 171 after two years. However, the response also remained critically dependent on residues 166 and 169 (2), despite the dominance of K169I/Q/E and R166K/S mutants as early as 30 weeks and two years post infection respectively (Figure 4.8).

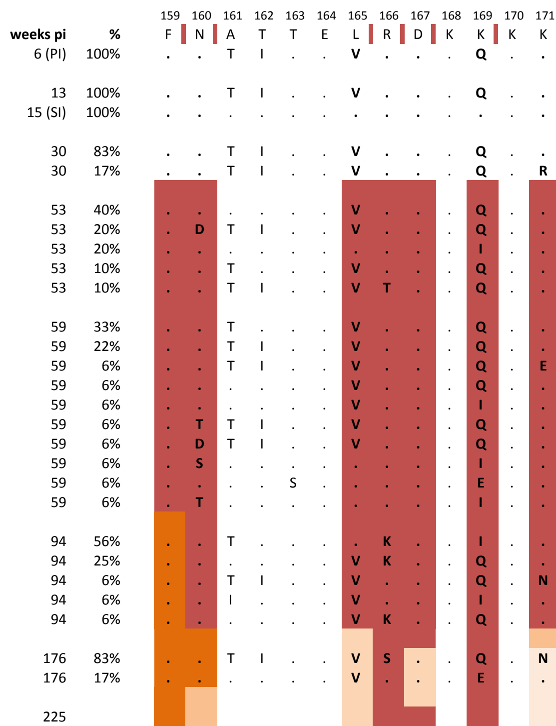


Figure 4.8 | The dependence of CAP256 plasma antibodies on specific residues for neutralization of ConC over time (from 42) is overlaid onto CAP256 viral diversification. Time proceeds from top to bottom, with a "." indicating the residue is identical to that of the highly sensitive superinfecting virus (SI). The proportion of the population with a particular sequence (%) is estimated from the number of SGAs with that sequence at each time point. SI, superinfecting virus; PI, primary infecting virus.

4.4 Discussion

We show here that in the CAPRISA 002 cohort, dual infection within the first year of infection was not consistently associated with the development of a broader neutralizing antibody response, even after adjusting for differences in CD4+ T cell counts and viral load at set point. There was also no significant association between early *envelope* diversity and the later development of breadth.

In contrast, Piantadosi *et al.* (72) observed that *envelope* diversity in early infection was associated with greater antibody breadth at five years post infection. More specifically, two other studies have previously suggested that superinfection *does* broaden the neutralizing antibody response (162, 180). Cortez et al., (162) compared the breadth of 12 superinfected participants to that of 36 monoinfected, matched controls, and found that superinfection was associated with broader neutralizing antibody responses. However, the timing of superinfection ranged from approximately two months post infection to over five years post infection. More importantly, given that breadth arises on average ~2.4 years post infection (1, 105), including superinfections that occur late in infection is potentially confounded by cross-neutralizing antibodies that developed prior to and independent of any superinfection. Cortez et al. attempted to control for this by evaluating the breadth prior to superinfection. While the participants that later went on to become superinfected already had significantly higher breadth than their singly infected counterparts prior to any superinfection (p=0.019), the difference was no longer statistically significant (p=0.067) when Cortez *et al.* then controlled for contemporaneous viral load, which they note was significantly different between groups. While higher viral loads have been associated with greater neutralization breadth (here, and previously), the superinfected group analyzed by Cortez et al had significantly *lowe*r viral loads than the singly infected group at the pre-superinfection timepoint. Nevertheless, it is incorrect to suggest that a p value of 0.067 constitutes evidence that no difference existed between the groups prior to superinfection. Cortez *et al.* did use a multivariate model (conditional Poisson

regression) to evaluate differences in breadth, "controlling for" breadth prior to superinfection (as well as contemporaneous viral loads and CD4 counts). While the association between superinfection and breadth remained statistically significant, the model accounts for a linear correlation between the breadth before and breadth after superinfection. However, as participants were not sampled at the same times post infection, and as breadth has a strong temporal component, there was no significant correlation between breadth before and breadth after superinfection. Simply comparing the broadening of the response (i.e. breadth at the post-superinfection time point that was not already accounted for at the pre-superinfection time point) with a T-test, the difference between the superinfected and singly infected groups is not statistically significant (p=0.6417 using breadth scores; p=0.8136 using percent of the panel neutralized), suggesting the result may not be a robust one.

It is also important to consider the effect of dual infection on the neutralizing antibody response mechanistically, and the *magnitude* of the effect. Antibody breadth is critically dependent on targeting an epitope that is highly conserved across variants. We must therefore differentiate between a dominant, an additive, or a synergistic effect of dual infection on the antibody response. A dominant effect would result in the targeting of one variant at the expense of responses to the other. An additive effect would result in two, typically type-specific (at least at first) responses. If dual infection instead redirected the immune response towards a more conserved epitope, this would represent a synergistic effect. Assuming no other differences other than the multiplicity of infection, in the case of dominance dual infected individuals would be expected to have a similar likelihood of developing BCN antibodies to singly infected individuals. If the effect is additive, we would expect dual infected participants to be approximately two times more likely to develop breadth (as the response to each variant would have a chance of developing cross reactivity). A synergistic effect would result in dual infected individuals having odds of developing breadth much greater than two times that of those who are monoinfected. While an additive effect is relevant, as eliciting several responses to several epitopes may be necessary for adequate vaccine coverage, a synergistic response is no doubt the

most desirable scenario. Ignoring any methodological limitations, Cortez *et al.* estimated that superinfected participants had an odds ratio of developing breadth of 1.51 times that of matched monoinfected participants; consistent with an additive effect rather than a synergistic one.

In a second study that addressed the influence of superinfection on antibody breadth in a cohort from Yaoundé, Cameroon, superinfection was associated with a significant broadening of the neutralizing antibody response (180). Broadening of the neutralizing antibody response was evident in three out of four dual-infected participants, compared to none of 23 monoinfected participants (180). Using an IC_{50} titer >80 as the definition for sensitivity, only 1/23 (4.35%) of the monoinfected participants were able to neutralize more than one of the four heterologous, primary viruses. This is in contrast to breadth in three out of the four superinfected participants. However, time post infection was unknown for all subjects and therefore duration of infection could not be controlled for. Furthermore, a panel including only four heterologous primary viruses was used to assess breadth. Nevertheless, these results are enticing, and would be consistent with a synergistic effect. However, the mechanism by which superinfection may promote breadth still needs to be addressed.

Even if dual infection did not consistently enhance breadth in the CAPRISA cohort, it remains possible that superinfection may have been a factor in the development of breadth in CAP256. We can envisage several mechanisms by which dual infection could potentially promote the development of broadly neutralizing responses (Table 4.3).

Table 4.3 | Potential mechanisms by which superinfection could promote antibody breadth

- Indirectly through higher viral loads, and lower CD4+ T cell counts
- Increasing the number of antibody specificities elicited
- Promote antibody responses to conserved epitopes
- By modulating or accelerating antibody maturation
- Modulating antigen presentation, or B cell activation in the context of sequential exposure

Given that dual infection has previously been associated with higher viral loads and lower CD4 counts, and these in turn have been associated with antibody breadth, it is possible that effect may be an indirect one. However, viral loads and CD4 counts in the Yaoundé cohort (where dual infection was significantly correlated with breadth) were not significantly different between superinfected and monoinfected participants, though participants could not be matched for duration of infection. In the CAPRISA002 cohort described here, dual infection was also not associated with significantly higher viral loads, or lower CD4 counts (see Figures 3.2 and 3.3). CAP256 (the one superinfected participant that developed breadth) did have the highest viral loads of the three cases of superinfection. However, even if superinfection does promote breadth indirectly via effects on viral load and CD4 count, this would provide limited new information for future vaccination strategies.

Breadth in CAP256 was also mediated almost exclusively by antibodies dependent on V1V2 (2), indicating that superinfection did not promote breadth by increasing the number of antibody specificities. In any case, immunodominant epitopes in HIV Envelope can be hypervariable, and consequently increasing the number of antibody specificities elicited may

never accumulate to the point of any significant breadth. As such, antibody breadth is likely critically dependent on targeting an epitope that is highly conserved across variants.

Superinfection with a second, antigenically divergent virus could preferentially boost pre-existing responses to an epitope conserved in both infecting viruses. Such a mechanism would be expected to enhance antibody responses to conserved epitopes, fostering breadth. With the aid of computational predictions, cross-neutralizing antibodies in CAP256 were mapped to residues in V2, overlapping that targeted by the mAbs PG9 and PG16 (specifically positions 159, 160, 165, 166, 167, 169, and 171). Similarly, an early autologous response to the superinfecting virus was also highly dependent on residues 162, 166, 169, and 171. This, together with the fact that breadth in CAP256 developed gradually over time, and was associated with subtle changes in specificity (2), suggests that this autologous V1V2 response may represent the precursor of the broadly neutralizing response detected later. However, the cross-neutralizing response in CAP256 did not target an epitope fully conserved in both infecting viruses. The PI virus contained several mutations within the targeted epitope (Figure 4.7), and was ~100-fold more resistant to autologous neutralization than the SI virus (Figure 4.6). Interestingly, after 42 wpi, the primary virus was also neutralized by a predominantly V1V2-dependent antibody response. Evaluating the cross-reactivity of the response at this time revealed that it was able to cross-neutralize 6/42 (14%) of heterologous viruses tested (2). This suggests that the neutralization of the PI may have been the result of the cross-reactivity of the response produced against the SI, rather than the immune system preferentially targeting an epitope present in both infecting variants. While the epitope sequence was not conserved, it remains possible that structural homology existed in the epitopes within the PI and SI, and in fact the neutralizing was directed towards a single, shared epitope. Alternatively, the PI may have been neutralized by a distinct clonal population of B cells, also targeting an epitope in V1V2. However, the isolation of a mAb from CAP256 that neutralizes both the SI and the PI (Doria-Rose *et al.,* in submission) suggests that this neutralization was mediated by the same clonal population of B cells. Importantly, as this V1V2-dependent response arose by 30 wpi, this provides evidence of an HIV neutralizing

antibody response that began to develop cross-reactivity only 12 weeks after the response was first detectable.

A significant consequence of dual infection is the potential for recombination and accelerated diversification. We show here that dual infection and recombination can produce substantial diversity in *envelope*. Recombination was also used as a mechanism to introduce multiple escape mutations simultaneously into the targeted epitope. While it is conceivable that this potential for accelerated diversification may promote or accelerate the maturation of antibody responses towards breadth, no association between viral diversification and changes in the antibody specificity were readily apparent in CAP256 (Figure 4.8).

It is also possible that the altered immunological environment as a result of primary HIV infection can modulate the immune response to a superinfecting virus. Responses to antigens as part of immune complexes are qualitatively and quantitatively different compared to responses to antigen alone, indicating that antibodies can have influential immunomodulatory effects. Immunisation with immune complexes has frequently been demonstrated to enhance immunogenicity compared to antigen alone (reviewed in 187). This is relevant here, as initial exposure to the superinfecting variants would have occurred in the presence of anti-HIV antibodies generated by the initial infection. Humoral responses are also dependent on TLR activation (188), and TLR ligands including LPS and ssRNA are present in abundance during natural HIV infection, and could profoundly affect subsequent B cell responses (reviewed in 189). Whether primary HIV infection modulates the immune response to the superinfecting virus warrants further study.

Taken together, however, we find no evidence here that superinfection promoted the development of antibody breadth in CAP256, though we cannot exclude this possibility. In contrast to the data presented by Powell et al. (180), we find no evidence that HIV co- or superinfection generally promotes the development of broadly neutralizing antibodies. It is potentially important that the participants described here were dual infected with two subtype

C viruses, as opposed to individuals described in Powell et al. who were infected with two different subtypes. It is possible that dual infection with more divergent viruses can promote cross-reactivity better than dual infection with more similar viruses can. Consistent with this, Cortez et al., (162) found that inter-subtype dual infections had greater neutralization breadth than intra-subtype dual infections (162). The effects of immunodominance are likely greater the more similar the antigens are, and the more CD4 helper epitopes they have in common. Consequently, inter-subtype dual infection can conceivably promote a strong(er) additive effect. However, it is unclear whether intersubtype dual infection can promote neutralizing antibody responses towards conserved epitopes.

Even if dual infection is convincingly demonstrated not to broaden the neutralizing antibody response, understanding how the immune system responds to multiple HIV immunogens during natural infection is relevant to the use of polyvalent vaccines, and could potentially identify avenues that could be exploited with vaccination.

In conclusion, taken together, these observations suggest that polyvalent immunogens may not necessarily be able to broaden immune responses to heterologous variants, although they may elicit responses to each of the included immunogens. However, we acknowledge limitations in the current study. Cases of superinfection are difficult to detect. Consequently, there remains a possibility that dual infected participants could have been misclassified as singly infected. However, all participants who developed broadly neutralizing antibodies in the CAPRISA cohort were subject to extensive sequencing efforts as part of ongoing studies. If anything, the bias in the sequencing would have enriched for detection of superinfection in the broad neutralizers. Nevertheless, the relatively small number of participants analyzed here potentially offers little statistical power to detect an effect. Larger studies will be required to comprehensively address how co- and superinfection modulate the immune response. It is also necessary to note that natural infection may not accurately reflect the vaccine setting. During natural infection, there is ongoing HIV replication and a significant immune dysfunction that

would not be present in the vaccine setting. However, while dual infection may not perfectly recapitulate the vaccine setting, apart from animal models and vaccine trials, dual infection potentially provides the best model available. Moreover, these results are consistent with responses seen in vaccinated macaques where polyvalent vaccination elicited, at best, antibodies capable of neutralizing several isolates from the vaccine, but few heterologous variants; indicative of multiple, type-specific responses (114). Polyvalent vaccination of rhesus macaques was able to elicit responses capable of neutralizing heterologous variants not part of the vaccine, albeit at very low titers, and was not evaluated against tier two viruses (115). The same polyvalent and mosaic vaccines that enhanced the breadth of CTL responses were not effective at enhancing the breadth of the neutralizing antibody response, and did not elicit responses capable of neutralizing any tier two viruses tested (113). In humans, a penta-valent HIV vaccine resulted in neutralizing responses to all five homologous variants, as well as several heterologous, but easy to neutralize, tier 1 viruses. However, most vaccinated individuals were unable to neutralize tier two viruses (117). Similarly, the quadrivalent HPV vaccine offers little cross-protection against serotypes not included in the vaccine (190). However, in the case of influenza, a DNA prime, followed by a heterologous boost with an inactivated virus was able to elicit cross-reactive neutralizing antibodies in mice and ferrets (191). Perhaps the most rigorous comparison between monovalent and polyvalent vaccines can be found with Polio vaccines. The earliest Polio vaccines were monovalent, but by the early 1960s, these were phased out in favour of a trivalent formulation incorporating each of the three Polio serotypes. Comparing the seroconversion rates in individuals receiving the monovalent vaccines to those who received the same doses in a trivalent formulation revealed that the monovalent vaccines are significantly more immunogenic (192, 193). A single dose of the trivalent OPV (tOPV) produces immunity to all three serotypes in ~50% of recipients. In comparison, the monovalent vaccines (mOPV) elicit immunity in >90% of recipients after a single dose with type 1 or 2, and in ~85% of recipients after a single dose with type 3. Seroconversion rates after a single dose of the mOPV are also better than after even two doses of the tOPV (reviewed in 194).

Taken together, these data suggest that multivalent immunogens may "interfere" with antibody responses to each other, perhaps due to immunodominance. Indeed, these studies also revealed a consistent predominance of responses to Polio type 2. Even the "balanced" formulation of the trivalent OPV (10:1:3 ratio of types 1:2:3) produces preferential seroconversion to serotype 2 (though this could potentially result from different growth characteristics or heat stability of the three attenuated strains (194)). Understanding the factors influencing immunodominance in the neutralizing antibody response will likely be critical to an effective HIV vaccine and warrants future work.

References

1. **Gray ES, Madiga MC, Hermanus T, Moore PL, Wibmer CK, Tumba NL, Werner L, Mlisana K, Sibeko S, Williamson C, Abdool Karim SS, Morris L**. 2011. The neutralization breadth of HIV-1 develops incrementally over four years and is associated with CD4+ T cell decline and high viral load during acute infection. J. Virol. **85**:4828–40.

2. **Moore PL, Gray ES, Sheward D, Madiga M, Ranchobe N, Lai Z, Honnen WJ, Nonyane M, Tumba N, Hermanus T, Sibeko S, Mlisana K, Abdool Karim SS, Williamson C, Pinter A, Morris L, others**. 2011. Potent and broad neutralization of HIV-1 subtype C by plasma antibodies targeting a quaternary epitope including residues in the V2 loop. J. Virol. **85**:3128–41.

3. **UNAIDS**. 2013. UNAIDS report on the global AIDS epidemic.

4. **Abdool Karim Q, Abdool Karim SS, Frohlich J a, Grobler AC, Baxter C, Mansoor LE, Kharsany ABM, Sibeko S, Mlisana KP, Omar Z, Gengiah TN, Maarschalk S, Arulappan N, Mlotshwa M, Morris L, Taylor D**. 2010. Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science **329**:1168–74.

5. **Choopanya K, Martin M, Suntharasamai P, Sangkum U, Mock P a, Leethochawalit M, Chiamwongpaet S, Kitisin P, Natrujirote P, Kittimunkong S, Chuachoowong R, Gvetadze RJ, McNicholl JM, Paxton L a, Curlin ME, Hendrix CW, Vanichseni S**. 2013. Antiretroviral prophylaxis for HIV infection in injecting drug users in Bangkok, Thailand (the Bangkok Tenofovir Study): a randomised, double-blind, placebo-controlled phase 3 trial. Lancet **381**:2083–90.

6. **Baeten JM, Donnell D, Ndase P, Mugo NR, Campbell JD, Wangisi J, Tappero JW, Bukusi E a, Cohen CR, Katabira E, Ronald A, Tumwesigye E, Were E, Fife KH, Kiarie J, Farquhar C, John-Stewart G, Kakia A, Odoyo J, Mucunguzi A, Nakku-Joloba E, Twesigye R, Ngure K, Apaka C, Tamooh H, Gabona F, Mujugira A, Panteleeff D, Thomas KK, Kidoguchi L, Krows M, Revall J, Morrison S, Haugen H, Emmanuel-Ogier M, Ondrejcek L, Coombs RW, Frenkel L, Hendrix C, Bumpus NN, Bangsberg D, Haberer JE, Stevens WS, Lingappa JR, Celum C**. 2012. Antiretroviral prophylaxis for HIV prevention in heterosexual men and women. N. Engl. J. Med. **367**:399–410.

7. **Rerks-Ngarm S, Pitisuttithum P, Nitayaphan S, Kaewkungwal J, Chiu J, Paris R, Premsri N, Namwat C, de Souza M, Adams E, Benenson M, Gurunathan S, Tartaglia J, McNeil JG, Francis DP, Stablein D, Birx DL, Chunsuttiwat S, Khamboonruang C, Thongcharoen P, Robb ML, Michael NL, Kunasol P, Kim JH**. 2009. Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N. Engl. J. Med. **361**:2209–20.

8. **Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, Evans DT, Montefiori DC, Karnasuta C, Sutthent R, Liao H-X, DeVico AL, Lewis GK, Williams C, Pinter A, Fong Y, Janes H, DeCamp A, Huang Y, Rao M, Billings E, Karasavvas N, Robb ML, Ngauy V, de Souza MS, Paris R, Ferrari G, Bailer RT, Soderberg KA, Andrews C, Berman PW, Frahm N, De Rosa SC, Alpert MD, Yates NL, Shen X, Koup RA, Pitisuttithum P, Kaewkungwal J, Nitayaphan S, Rerks-Ngarm S, Michael NL, Kim JH**. 2012. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N. Engl. J. Med. **366**:1275–86.

9. **Salazar-Gonzalez JF, Salazar MG, Keele BF, Learn GH, Giorgi EE, Li H, Decker JM, Wang S, Baalwa J, Kraus MH, Parrish NF, Shaw KS, Guffey MB, Bar KJ, Davis KL, Ochsenbauer-Jambor C, Kappes JC, Saag MS, Cohen MS, Mulenga J, Derdeyn C a, Allen S, Hunter E, Markowitz M, Hraber P, Perelson AS, Bhattacharya T, Haynes BF, Korber BT, Hahn BH, Shaw GM**. 2009. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. J. Exp. Med. **206**:1273–89.

10. **Keele BF, Giorgi EE, Salazar-Gonzalez JF, Decker JM, Pham KT, Salazar MG, Sun C, Grayson T, Wang S, Li H, Wei X, Jiang C, Kirchherr JL, Gao F, Anderson J a, Ping L-H, Swanstrom R, Tomaras GD, Blattner W a, Goepfert P a, Kilby JM, Saag MS, Delwart EL, Busch MP, Cohen MS, Montefiori DC, Haynes BF, Gaschen B, Athreya GS, Lee HY, Wood N, Seoighe C, Perelson AS, Bhattacharya T, Korber BT, Hahn BH, Shaw GM**. 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.

11. **Abrahams M-R, Anderson JA, Giorgi EE, Seoighe C, Mlisana K, Ping L-H, Athreya GS, Treurnicht FK, Keele BF, Wood N, Salazar-Gonzalez JF, Bhattacharya T, Chu H, Hoffman I, Galvin S, Mapanje C, Kazembe P, Thebus R, Fiscus S, Hide W, Cohen MS, Karim SA, Haynes BF, Shaw GM, Hahn BH, Korber BT, Swanstrom R, Williamson C**. 2009. Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. J. Virol. **83**:3556–67.

12. **Fiebig EW, Wright DJ, Rawal BD, Garrett PE, Schumacher RT, Peddada L, Heldebrant C, Smith R, Conrad A, Kleinman SH, Busch MP**. 2003. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. AIDS **17**:1871–9.

13. **Brenchley JM, Schacker TW, Ruff LE, Price D a, Taylor JH, Beilman GJ, Nguyen PL, Khoruts A, Larson M, Haase AT, Douek DC**. 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. J. Exp. Med. **200**:749–59.

14. **Guadalupe M, Reay E, Sankaran S, Prindiville T, Flamm J, Mcneil A, Dandekar S**. 2003. Severe CD4+ T-Cell Depletion in Gut Lymphoid Tissue during Primary Human Immunodeficiency Virus Type 1 Infection and Substantial Delay in Restoration following Highly Active Antiretroviral Therapy. J. Virol. **77**:11708–11717.

15. **Mehandru S, Poles M a, Tenner-Racz K, Manuelli V, Jean-Pierre P, Lopez P, Shet A, Low A, Mohri H, Boden D, Racz P, Markowitz M**. 2007. Mechanisms of gastrointestinal CD4+ T-cell depletion during acute and early human immunodeficiency virus type 1 infection. J. Virol. **81**:599–612.

16. **Lyles RH, Mun A, Yamashita TE, Bazmi H, Detels R, Rinaldo CR, Margolick JB, Phair JP, Mellors JW, Aids M**. 2000. Natural History of Human Immunodeficiency Virus Type 1 Viremia after Seroconversion and Proximal to AIDS in a Large Cohort of Homosexual Men. J. Infect. Dis. **181**:872–880.

17. **Goonetilleke N, Liu MKP, Salazar-Gonzalez JF, Ferrari G, Giorgi E, Ganusov V V, Keele BF, Learn GH, Turnbull EL, Salazar MG, Weinhold KJ, Moore S, Letvin N, Haynes BF, Cohen MS, Hraber P, Bhattacharya T, Borrow P, Perelson AS, Hahn BH, Shaw GM, Korber BT, McMichael AJ**. 2009. The first T cell response to transmitted/founder virus contributes to the control of acute viremia in HIV-1 infection. J. Exp. Med. **206**:1253–72.

18. **Borrow P, Lewicki H, Hahn BH, Shaw GM, Oldstone MB**. 1994. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. J. Virol. **68** :6103–6110.

19. **Jin X, Bauer DE, Tuttleton SE, Lewin S, Gettie A, Blanchard J, Irwin CE, Safrit JT, Mittler J, Weinberger L, Kostrikis LG, Zhang L, Perelson AS, Ho DD**. 1999. Dramatic Rise in Plasma Viremia after CD8+ T Cell Depletion in Simian Immunodeficiency Virus–infected Macaques J. Exp. Med. **189** :991–998.

20. **Matano T, Shibata R, Siemon C, Connors M, Lane HC, Martin MA**. 1998. Administration of an Anti-CD8 Monoclonal Antibody Interferes with the Clearance of Chimeric Simian/Human Immunodeficiency Virus during Primary Infections of Rhesus Macaques. J. Virol. **72** :164–169.

21. **Schmitz JE, Kuroda MJ, Santra S, Sasseville VG, Simon MA, Lifton MA, Racz P, Tenner-Racz K, Dalesandro M, Scallon BJ, Ghrayeb J, Forman MA, Montefiori DC, Rieber EP, Letvin NL, Reimann KA**. 1999. Control of Viremia in Simian Immunodeficiency Virus Infection by CD8+ Lymphocytes. Sci. **283** :857–860.

22. **Goulder PJR, Watkins DI**. 2008. Impact of MHC class I diversity on immune control of immunodeficiency virus replication. Nat. Rev. Immunol. **8**:619–30.

23. **Fellay J, Ge D, Shianna K V, Colombo S, Ledergerber B, Cirulli ET, Urban TJ, Zhang K, Gumbs CE, Smith JP, Castagna A, Cozzi-Lepri A, De Luca A, Easterbrook P, Günthard HF, Mallal S, Mussini C, Dalmau J, Martinez-Picado J, Miro JM, Obel N, Wolinsky SM, Martinson JJ, Detels R, Margolick JB, Jacobson LP, Descombes P, Antonarakis SE, Beckmann JS, O'Brien SJ, Letvin NL, McMichael AJ, Haynes BF, Carrington M, Feng S, Telenti A, Goldstein DB**. 2009. Common genetic variation and the control of HIV-1 in humans. PLoS Genet. **5**:e1000791.

24. **Fellay J, Shianna K V, Ge D, Colombo S, Ledergerber B, Weale M, Zhang K, Gumbs C, Castagna A, Cossarizza A, Cozzi-Lepri A, De Luca A, Easterbrook P, Francioli P, Mallal S, Martinez-Picado J, Miro JM, Obel N, Smith JP, Wyniger J, Descombes P, Antonarakis SE, Letvin NL, McMichael AJ, Haynes BF, Telenti A, Goldstein DB**. 2007. A whole-genome association study of major determinants for host control of HIV-1. Science (80-.). **317**:944–7.

25. **Limou S, Le Clerc S, Coulonges C, Carpentier W, Dina C, Delaneau O, Labib T, Taing L, Sladek R, Deveau C, Ratsimandresy R, Montes M, Spadoni J-L, Lelièvre J-D, Lévy Y, Therwath A, Schächter F, Matsuda F, Gut I, Froguel P, Delfraissy J-F, Hercberg S, Zagury J-F**. 2009. Genomewide association study of an AIDS-nonprogression cohort emphasizes the role played by HLA genes (ANRS Genomewide Association Study 02). J. Infect. Dis. **199**:419–26.

26. **Pelak K, Goldstein DB, Walley NM, Fellay J, Ge D, Shianna K V, Gumbs C, Gao X, Maia JM, Cronin KD, Hussain SK, Carrington M, Michael NL, Weintrob AC**. 2010. Host determinants of HIV-1 control in African Americans. J. Infect. Dis. **201**:1141–9.

27. **Le Clerc S, Limou S, Coulonges C, Carpentier W, Dina C, Taing L, Delaneau O, Labib T, Sladek R, Group AG, Deveau C, Guillemain H, Ratsimandresy R, Montes M, Spadoni J-L, Therwath A, Schächter F, Matsuda F, Gut I, Lelièvre J-D, Lévy Y, Froguel P, Delfraissy J-F, Hercberg S, Zagury J-F**. 2009. Genomewide Association Study of a Rapid Progression Cohort Identifies New Susceptibility Alleles for AIDS (ANRS Genomewide Association Study 03). J. Infect. Dis. **200** :1194–1201.

28. **Kiepiela P, Leslie AJ, Honeyborne I, Ramduth D, Thobakgale C, Chetty S, Rathnavalu P, Moore C, Pfafferott KJ, Hilton L, Zimbwa P, Moore S, Allen T, Brander C, Addo MM, Altfeld M, James I, Mallal S, Bunce M, Barber LD, Szinger J, Day C, Klenerman P, Mullins J, Korber B, Coovadia HM, Walker BD, Goulder PJR**. 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. Nature **432**:769–75.

29. **Altfeld M, Addo MM, Rosenberg ES, Hecht FM, Lee PK, Vogel M, Yu XG, Draenert R, Johnston MN, Strick D, Allen TM, Feeney ME, Kahn JO, Sekaly RP, Levy JA, Rockstroh JK, Goulder PJR, Walker BD**. 2003. Influence of HLA-B57 on clinical presentation and viral control during acute HIV-1 infection. AIDS **17**.

30. **Migueles SA, Sabbaghian MS, Shupert WL, Bettinotti MP, Marincola FM, Martino L, Hallahan CW, Selig SM, Schwartz D, Sullivan J, Connors M**. 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. **97** :2709–2714.

31. **Alter G, Altfeld M**. 2009. NK cells in HIV-1 infection: evidence for their role in the control of HIV-1 infection. J. Intern. Med. **265**:29–42.

32. **Borrow P, Lewicki H, Wei X, Horwitz MS, Peffer N, Meyers H, Nelson JA, Gairin JE, Hahn BH, Oldstone MBA, Shaw GM**. 1997. Antiviral pressure exerted by HIV-l-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. Nat. Med. **3**:205–211.

33. **Price DA, Goulder PJR, Klenerman P, Sewell AK, Easterbrook PJ, Troop M, Bangham CRM, Phillips RE**. 1997. Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. Proc. Natl. Acad. Sci. **94** :1890–1895.

34. **Kelleher AD, Long C, Holmes EC, Allen RL, Wilson J, Conlon C, Workman C, Shaunak S, Olson K, Goulder P, Brander C, Ogg G, Sullivan JS, Dyer W, Jones I, McMichael AJ, Rowland-Jones S, Phillips RE**. 2001. Clustered mutations in HIV-1 gag are consistently required for escape from HLA-B27-restricted cytotoxic T lymphocyte responses. J. Exp. Med. **193**:375–86.

35. **Jones NA, Wei X, Flower DR, Wong M, Michor F, Saag MS, Hahn BH, Nowak MA, Shaw GM, Borrow P**. 2004. Determinants of human immunodeficiency virus type 1 escape from the primary CD8+ cytotoxic T lymphocyte response. J. Exp. Med. **200**:1243–56.

36. **Allen TM, Altfeld M, Geer SC, Kalife ET, Moore C, O'sullivan KM, Desouza I, Feeney ME, Eldridge RL, Maier EL, Kaufmann DE, Lahaie MP, Reyor L, Tanzi G, Johnston MN, Brander C, Draenert R, Rockstroh JK, Jessen H, Rosenberg ES, Mallal SA, Walker BD**. 2005. Selective escape from CD8+ T-cell responses represents a major driving force of human immunodeficiency virus type 1 (HIV-1) sequence diversity and reveals constraints on HIV-1 evolution. J. Virol. **79**:13239–49.

37. **Frater a J, Brown H, Oxenius a, Günthard HF, Hirschel B, Robinson N, Leslie a J, Payne R, Crawford H, Prendergast a, Brander C, Kiepiela P, Walker BD, Goulder PJR, McLean a, Phillips RE**. 2007. Effective T-cell responses select human immunodeficiency virus mutants and slow disease progression. J. Virol. **81**:6742–51.

38. **Goulder PJR, Phillips RE, Colbert RA, McAdam S, Ogg G, Nowak MA, Giangrande P, Luzzi G, Morgana B, Edwards A, McMichael AJ, Rowland-Jones S**. 1997. Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. Nat. Med. **3**:212–217.

39. **Troyer RM, McNevin J, Liu Y, Zhang SC, Krizan RW, Abraha A, Tebit DM, Zhao H, Avila S, Lobritz M a, McElrath MJ, Le Gall S, Mullins JI, Arts EJ**. 2009. Variable fitness impact of HIV-1 escape mutations to cytotoxic T lymphocyte (CTL) response. PLoS Pathog. **5**:e1000365.

40. **Crawford H, Lumm W, Leslie A, Schaefer M, Boeras D, Prado JG, Tang J, Farmer P, Ndung'u T, Lakhi S, Gilmour J, Goepfert P, Walker BD, Kaslow R, Mulenga J, Allen S, Goulder PJR, Hunter E**. 2009. Evolution of HLA-B*5703 HIV-1 escape mutations in HLA-B*5703-positive individuals and their transmission recipients. J. Exp. Med. **206**:909–21.

41. **Martinez-Picado J, Prado JG, Fry EE, Pfafferott K, Leslie A, Chetty S, Thobakgale C, Honeyborne I, Crawford H, Matthews P, Pillay T, Rousseau C, Mullins JI, Brander C, Walker BD, Stuart DI, Kiepiela P, Goulder P**. 2006. Fitness cost of escape mutations in p24 Gag in association with control of human immunodeficiency virus type 1. J. Virol. **80**:3617–23.

42. **Leslie AJ, Pfafferott KJ, Chetty P, Draenert R, Addo MM, Feeney M, Tang Y, Holmes EC, Allen T, Prado JG, Altfeld M, Brander C, Dixon C, Ramduth D, Jeena P, Thomas SA, St John A, Roach TA, Kupfer B, Luzzi G, Edwards A, Taylor G, Lyall H, Tudor-Williams G, Novelli V, Martinez-Picado J, Kiepiela P, Walker BD, Goulder PJR**. 2004. HIV evolution: CTL escape mutation and reversion after transmission. Nat. Med. **10**:282–9.

43. **Crawford H, Prado JG, Leslie A, Hué S, Honeyborne I, Reddy S, van der Stok M, Mncube Z, Brander C, Rousseau C, Mullins JI, Kaslow R, Goepfert P, Allen S, Hunter E, Mulenga J,**

Kiepiela P, Walker BD, Goulder PJR. 2007. Compensatory mutation partially restores fitness and delays reversion of escape mutation within the immunodominant HLA-B*5703-restricted Gag epitope in chronic human immunodeficiency virus type 1 infection. J. Virol. **81**:8346–51.

44. **Brockman MA, Schneidewind A, Lahaie M, Schmidt A, Miura T, Desouza I, Ryvkin F, Derdeyn CA, Allen S, Hunter E, Mulenga J, Goepfert PA, Walker BD, Allen TM**. 2007. Escape and compensation from early HLA-B57-mediated cytotoxic T-lymphocyte pressure on human immunodeficiency virus type 1 Gag alter capsid interactions with cyclophilin A. J. Virol. **81**:12608–18.

45. **Tomaras GD, Yates NL, Liu P, Qin L, Fouda GG, Chavez LL, Decamp AC, Parks RJ, Ashley VC, Lucas JT, Cohen M, Eron J, Hicks CB, Liao H-X, Self SG, Landucci G, Forthal DN, Weinhold KJ, Keele BF, Hahn BH, Greenberg ML, Morris L, Karim SSA, Blattner W a, Montefiori DC, Shaw GM, Perelson AS, Haynes BF**. 2008. Initial B-cell responses to transmitted human immunodeficiency virus type 1: virion-binding immunoglobulin M (IgM) and IgG antibodies followed by plasma anti-gp41 antibodies with ineffective control of initial viremia. J. Virol. **82**:12449–63.

46. **Schmitz E, Kuroda MJ, Santra S, Simon MA, Lifton MA, Lin W, Khunkhun R, Piatak M, Lifson JD, Grosschupff G, Gelman RS, Racz P, Tenner-racz K, Mansfield KA, Letvin NL, Montefiori DC, Reimann KA**. 2003. Effect of Humoral Immune Responses on Controlling Viremia during Primary Infection of Rhesus Monkeys with Simian Immunodeficiency Virus. J. Virol. **77**:2165–2173.

47. **Miller CJ, Genescà M, Abel K, Montefiori D, Forthal D, Bost K, Li J, Favre D, McCune JM**. 2007. Antiviral antibodies are necessary for control of simian immunodeficiency virus replication. J. Virol. **81**:5024–35.

48. **Gaufin T, Gautam R, Kasheta M, Ribeiro R, Ribka E, Barnes M, Pattison M, Tatum C, MacFarland J, Montefiori D, Kaur A, Pandrea I, Apetrei C**. 2009. Limited ability of humoral immune responses in control of viremia during infection with SIVsmmD215 strain. Blood **113**:4250–61.

49. **Gaufin T, Pattison M, Gautam R, Stoulig C, Dufour J, MacFarland J, Mandell D, Tatum C, Marx MH, Ribeiro RM, Montefiori D, Apetrei C, Pandrea I**. 2009. Effect of B-cell depletion on viral replication and clinical outcome of simian immunodeficiency virus infection in a natural host. J. Virol. **83**:10347–57.

50. **Huang K-HG, Bonsall D, Katzourakis A, Thomson EC, Fidler SJ, Main J, Muir D, Weber JN, Frater AJ, Phillips RE, Pybus OG, Goulder PJR, McClure MO, Cooke GS, Klenerman P**. 2010. B-cell depletion reveals a role for antibodies in the control of chronic HIV-1 infection. Nat. Commun. **1**:102.

51. **Yamamoto H, Kawada M, Takeda A, Igarashi H, Matano T**. 2007. Post-Infection Immunodeficiency Virus Control by Neutralizing Antibodies. PLoS One **2**:e540.

52. **Trkola A, Kuster H, Rusert P, Joos B, Fischer M, Leemann C, Manrique A, Huber M, Rehr M, Oxenius A, Weber R, Stiegler G, Vcelar B, Katinger H, Aceto L, Günthard HF**. 2005. Delay of HIV-1 rebound after cessation of antiretroviral therapy through passive transfer of human neutralizing antibodies. Nat. Med. **11**:615–22.

53. **Trkola A, Kuster H, Rusert P, von Wyl V, Leemann C, Weber R, Stiegler G, Katinger H, Joos B, Günthard HF**. 2008. In vivo efficacy of human immunodeficiency virus neutralizing antibodies: estimates for protective titers. J. Virol. **82**:1591–9.

54. **Klein F, Halper-Stromberg A, Horwitz J a., Gruell H, Scheid JF, Bournazos S, Mouquet H, Spatz L a., Diskin R, Abadir A, Zang T, Dorner M, Billerbeck E, Labitt RN, Gaebler C,**

Marcovecchio PM, Incesu R-B, Eisenreich TR, Bieniasz PD, Seaman MS, Bjorkman PJ, Ravetch J V., Ploss A, Nussenzweig MC. 2012. HIV therapy by a combination of broadly neutralizing antibodies in humanized mice. Nature **492**:118–22.

55. **Richman DD, Wrin T, Little SJ, Petropoulos CJ**. 2003. Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proc. Natl. Acad. Sci. U. S. A. **100**:4144–9.

56. **Wei X, Decker JM, Wang S, Hui H, Kappes JC, Wu X, Salazar-Gonzalez JF, Salazar MG, Kilby JM, Saag MS, Komarova NL, Nowak M a, Hahn BH, Kwong PD, Shaw GM**. 2003. Antibody neutralization and escape by HIV-1. Nature **422**:307–12.

57. **Bunnik EM, Pisas L, van Nuenen AC, Schuitemaker H**. 2008. Autologous neutralizing humoral immunity and evolution of the viral envelope in the course of subtype B human immunodeficiency virus type 1 infection. J. Virol. **82**:7932–41.

58. **Gray ES, Moore PL, Choge I a, Decker JM, Bibollet-Ruche F, Li H, Leseka N, Treurnicht F, Mlisana K, Shaw GM, Karim SSA, Williamson C, Morris L**. 2007. Neutralizing antibody responses in acute human immunodeficiency virus type 1 subtype C infection. J. Virol. **81**:6187–96.

59. **Moore PL, Ranchobe N, Lambson BE, Gray ES, Cave E, Abrahams M-R, Bandawe G, Mlisana K, Abdool Karim SS, Williamson C, Morris L**. 2009. Limited neutralizing antibody specificities drive neutralization escape in early HIV-1 subtype C infection. PLoS Pathog. **5**:e1000598.

60. **Rong R, Li B, Lynch RM, Haaland RE, Murphy MK, Mulenga J, Allen S a, Pinter A, Shaw GM, Hunter E, Robinson JE, Gnanakaran S, Derdeyn C a**. 2009. Escape from autologous neutralizing antibodies in acute/early subtype C HIV-1 infection requires multiple pathways. PLoS Pathog. **5**:e1000594.

61. **Moore PL, Gray ES, Wibmer CK, Bhiman JN, Nonyane M, Sheward DJ, Hermanus T, Bajimaya S, Tumba NL, Abrahams M-R, Lambson BE, Ranchobe N, Ping L, Ngandu N, Karim QA, Karim SSA, Swanstrom RI, Seaman MS, Williamson C, Morris L, Abdool Karim Q, Abdool Karim SS, others**. 2012. Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. Nat. Med. **18**:1–6.

62. **Frost SDW, Wrin T, Smith DM, Kosakovsky Pond SL, Liu Y, Paxinos E, Chappey C, Galovich J, Beauchaine J, Petropoulos CJ, Little SJ, Richman DD**. 2005. Neutralizing antibody responses drive the evolution of human immunodeficiency virus type 1 envelope during recent HIV infection. Proc. Natl. Acad. Sci. U. S. A. **102**:18514–9.

63. **Sagar M, Wu X, Lee S, Overbaugh J**. 2006. Human immunodeficiency virus type 1 V1-V2 envelope loop sequences expand and add glycosylation sites over the course of infection, and these modifications affect antibody neutralization sensitivity. J. Virol. **80**:9586–98.

64. **Rong R, Bibollet-Ruche F, Mulenga J, Allen S, Blackwell JL, Derdeyn C a**. 2007. Role of V1V2 and other human immunodeficiency virus type 1 envelope domains in resistance to autologous neutralization during clade C infection. J. Virol. **81**:1350–9.

65. **Doores KJ, Bonomelli C, Harvey DJ, Vasiljevic S, Dwek RA, Burton DR, Crispin M, Scanlan CN**. 2010. Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. Proc. Natl. Acad. Sci. U. S. A. **107**:13800–5.

66. **Moore PL, Gray ES, Morris L**. 2009. Specificity of the autologous neutralizing antibody response. Curent Oppinion HIV AIDS **4**:358–363.

67. **Gray ES, Madiga MC, Moore PL, Mlisana K, Abdool Karim SS, Binley JM, Shaw GM, Mascola JR, Morris L**. 2009. Broad neutralization of human immunodeficiency virus type 1 mediated by plasma antibodies against the gp41 membrane proximal external region. J. Virol. **83**:11265–74.

68. **Lynch RM, Rong R, Boliar S, Sethi A, Li B, Mulenga J, Allen S, Robinson JE, Gnanakaran S, Derdeyn C a**. 2011. The B cell response is redundant and highly focused on V1V2 during early subtype C infection in a Zambian seroconverter. J. Virol. **85**:905–15.

69. **Pinter A, Honnen WJ, He Y, Gorny MK, Zolla-Pazner S, Kayman SC**. 2004. The V1/V2 Domain of gp120 Is a Global Regulator of the Sensitivity of Primary Human Immunodeficiency Virus Type 1 Isolates to Neutralization by Antibodies Commonly Induced upon Infection. J. Virol. **78** :5205–5215.

70. **Horwitz J a, Halper-Stromberg A, Mouquet H, Gitlin AD, Tretiakova A, Eisenreich TR, Malbec M, Gravemann S, Billerbeck E, Dorner M, Büning H, Schwartz O, Knops E, Kaiser R, Seaman MS, Wilson JM, Rice CM, Ploss A, Bjorkman PJ, Klein F, Nussenzweig MC**. 2013. HIV-1 suppression and durable control by combining single broadly neutralizing antibodies and antiretroviral drugs in humanized mice. Proc. Natl. Acad. Sci. U. S. A. **110**:16538–16543.

71. **Cecilia D, Kleeberger C, Muñoz A, Giorgi J V, Zolla-Pazner S**. 1999. A longitudinal study of neutralizing antibodies and disease progression in HIV-1-infected subjects. J. Infect. Dis. **179**:1365–74.

72. **Piantadosi A, Panteleeff D, Blish C a, Baeten JM, Jaoko W, McClelland RS, Overbaugh J**. 2009. Breadth of neutralizing antibody response to human immunodeficiency virus type 1 is affected by factors early in infection but does not influence disease progression. J. Virol. **83**:10269–74.

73. **Euler Z, van Gils MJ, Bunnik EM, Phung P, Schweighardt B, Wrin T, Schuitemaker H**. 2010. Cross-reactive neutralizing humoral immunity does not protect from HIV type 1 disease progression. J. Infect. Dis. **201**:1045–53.

74. **Parren PWHI, Marx PA, Hessell ANNJ, Luckay A, Harouse J, Cheng-mayer C, Moore JP, Burton DR**. 2001. Antibody Protects Macaques against Vaginal Challenge with a Pathogenic R5 Simian / Human Immunodeficiency Virus at Serum Levels Giving Complete Neutralization In Vitro. J. Virol. **75**:8340–8347.

75. **Hessell AJ, Poignard P, Hunter M, Hangartner L, Tehrani DM, Bleeker WK, Parren PWHI, Marx PA, Burton DR**. 2009. Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. Nat. Med. **15**:951–4.

76. **Hessell AJ, Rakasz EG, Poignard P, Hangartner L, Landucci G, Forthal DN, Koff WC, Watkins DI, Burton DR**. 2009. Broadly neutralizing human anti-HIV antibody 2G12 is effective in protection against mucosal SHIV challenge even at low serum neutralizing titers. PLoS Pathog. **5**:e1000433.

77. **Mascola JR, Stiegler G, Vancott TC, Katinger H, Carpenter CB, Hanson CE, Beary H, Hayes D, Frankel SS, Birx DL, Lewis MG**. 2000. Protection of macaques against vaginal transmission of a pathogenic HIV-1/SIV chimeric virus by passive infusion of neutralizing antibodies. Nat. Med. **6**:207–210.

78. **Shibata R, Igarashi T, Haigwood N, Buckler-White A, Ogert R, Ross W, Willey R, Cho MW, Martin MA**. 1999. Neutralizing antibody directed against the HIV-1 envelope glycoprotein can completely block HIV-1/SIV chimeric virus infections of macaque monkeys. Nat. Med. **5**:204–10.

79. **Baba TW, Liska V, Hofmann-Lehmann R, Vlasak J, Xu W, Ayehunie S, Cavacini LA, Posner MR, Katinger H, Stiegler G, Bernacky BJ, Rizvi TA, Schmidt R, Hill LR, Keeling ME, Lu Y, Wright JE, Chou TC, Ruprecht RM**. 2000. Human neutralizing monoclonal antibodies of the IgG1

subtype protect against mucosal simian-human immunodeficiency virus infection. Nat. Med. **6**:200–6.

80. **Emini EA, Schleif WA, Nunberg JH, Conley AJ, Eda Y, Tokiyoshi S, Putney SD, Matsushita S, Cobb KE, Jett CM**. 1992. Prevention of HIV-1 infection in chimpanzees by gp120 V3 domain-specific monoclonal antibody. Nature **355**:728–30.

81. **Conley A, Kessler JA I, Boots L, McKenna P, Schleif W, Emini E, Mark GE I, Katinger H, Cobb E, Lunceford S, Rouse S, Murthy K**. 1996. The consequence of passive administration of an anti-human immunodeficiency virus type 1 neutralizing monoclonal antibody before challenge of chimpanzees with a primary virus isolate. J. Virol. **70**:6751–6758.

82. **Burton DR, Hessell AJ, Keele BF, Klasse PJ, Ketas T a, Moldt B, Dunlop DC, Poignard P, Doyle L a, Cavacini L, Veazey RS, Moore JP**. 2011. Limited or no protection by weakly or nonneutralizing antibodies against vaginal SHIV challenge of macaques compared with a strongly neutralizing antibody. Proc. Natl. Acad. Sci. U. S. A. **108**:1–6.

83. **Nishimura Y, Igarashi T, Haigwood N, Sadjadpour R, Plishka RJ, Buckler-white A, Shibata R, Martin MA**. 2002. Determination of a Statistically Valid Neutralization Titer in Plasma That Confers Protection against Simian-Human Immunodeficiency Virus Challenge following Passive Transfer of High-Titered Neutralizing Antibodies. J. Virol. **76**:2123–2130.

84. **Willey R, Nason MC, Nishimura Y, Follmann D a, Martin M a**. 2010. Neutralizing antibody titers conferring protection to macaques from a simian/human immunodeficiency virus challenge using the TZM-bl assay. AIDS Res. Hum. Retroviruses **26**:89–98.

85. **Sather DN, Armann J, Ching LK, Mavrantoni A, Sellhorn G, Caldwell Z, Yu X, Wood B, Self S, Kalams S, Stamatatos L**. 2009. Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection. J. Virol. **83**:757–69.

86. **Simek MD, Rida W, Priddy FH, Pung P, Carrow E, Laufer DS, Lehrman JK, Boaz M, Tarragona-Fiol T, Miiro G, Birungi J, Pozniak A, McPhee D a, Manigart O, Karita E, Inwoley A, Jaoko W, Dehovitz J, Bekker L-G, Pitisuttithum P, Paris R, Walker LM, Poignard P, Wrin T, Fast PE, Burton DR, Koff WC**. 2009. Human immunodeficiency virus type 1 elite neutralizers: individuals with broad and potent neutralizing activity identified by using a high-throughput neutralization assay together with an analytical selection algorithm. J. Virol. **83**:7337–48.

87. **Walker LM, Simek MD, Priddy F, Gach JS, Wagner D, Zwick MB, Phogat SK, Poignard P, Burton DR**. 2010. A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. PLoS Pathog. **6**.

88. **Scheid JF, Mouquet H, Feldhahn N, Seaman MS, Velinzon K, Pietzsch J, Ott RG, Anthony RM, Zebroski H, Hurley A, Phogat A, Chakrabarti B, Li Y, Connors M, Pereyra F, Walker BD, Wardemann H, Ho D, Wyatt RT, Mascola JR, Ravetch J V, Nussenzweig MC**. 2009. Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature **458**:636–40.

89. **Wu X, Yang Z-Y, Li Y, Hogerkorp C-M, Schief WR, Seaman MS, Zhou T, Schmidt SD, Wu L, Xu L, Longo NS, McKee K, O'Dell S, Louder MK, Wycuff DL, Feng Y, Nason M, Doria-Rose N, Connors M, Kwong PD, Roederer M, Wyatt RT, Nabel GJ, Mascola JR**. 2010. Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science (80-.). **329**:856–61.

90. **Walker LM, Phogat SK, Chan-Hui P-Y, Wagner D, Phung P, Goss JL, Wrin T, Simek MD, Fling S, Mitcham JL, Lehrman JK, Priddy FH, Olsen O a, Frey SM, Hammond PW, Kaminsky S, Zamb T,**

Moyle M, Koff WC, Poignard P, Burton DR. 2009. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science (80-.). **326**:285–9.

91. **Poignard P, Moulard M, Golez E, Vivona V, Franti M, Venturini S, Wang M, Parren PWHI, Burton DR**. 2003. Heterogeneity of Envelope Molecules Expressed on Primary Human Immunodeficiency Virus Type 1 Particles as Probed by the Binding of Neutralizing and Nonneutralizing Antibodies. J. Virol. **77**:353–365.

92. **Zhu P, Liu J, Bess J, Chertova E, Lifson JD, Grisé H, Ofek G a, Taylor K a, Roux KH**. 2006. Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature **441**:847–52.

93. **Klein JS, Bjorkman PJ**. 2010. Few and far between: how HIV may be evading antibody avidity. PLoS Pathog. **6**:e1000908.

94. **Muster T, Steindl F, Purtscher M, Trkola a, Klima a, Himmler G, Rüker F, Katinger H**. 1993. A conserved neutralizing epitope on gp41 of human immunodeficiency virus type 1. J. Virol. **67**:6642–7.

95. **Buchacher A, Predl R, Strutzenberger K, Steinfellner W, Trkola A, Purtscher M, Gruber G, Tauer C, Steindl F, Jungbauer A**. 1994. Generation of human monoclonal antibodies against HIV-1 proteins; electrofusion and Epstein-Barr virus transformation for peripheral blood lymphocyte immortalization. AIDS Res. Hum. Retroviruses **10**:359–69.

96. **Zwick MB, Labrijn AF, Wang M, Spenlehauer C, Saphire EO, Binley JM, Moore JP, Stiegler G, Katinger H, Burton DR, Parren PWHI**. 2001. Broadly Neutralizing Antibodies Targeted to the Membrane-Proximal External Region of Human Immunodeficiency Virus Type 1 Glycoprotein gp41 Broadly Neutralizing Antibodies Targeted to the Membrane-Proximal External Region of Human Immunodeficiency Virus. J. Virol. **75**:10892–10905.

97. **Huang J, Ofek G, Laub L, Louder MK, Doria-Rose N a., Longo NS, Imamichi H, Bailer RT, Chakrabarti B, Sharma SK, Alam SM, Wang T, Yang Y, Zhang B, Migueles S a., Wyatt R, Haynes BF, Kwong PD, Mascola JR, Connors M**. 2012. Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature **491**:406–12.

98. **Calarese D a, Scanlan CN, Zwick MB, Deechongkit S, Mimura Y, Kunert R, Zhu P, Wormald MR, Stanfield RL, Roux KH, Kelly JW, Rudd PM, Dwek R a, Katinger H, Burton DR, Wilson I a**. 2003. Antibody domain exchange is an immunological solution to carbohydrate cluster recognition. Science **300**:2065–71.

99. **Trkola a, Purtscher M, Muster T, Ballaun C, Buchacher A, Sullivan N, Srinivasan K, Sodroski J, Moore JP, Katinger H**. 1996. Human monoclonal antibody 2G12 defines a distinctive neutralization epitope on the gp120 glycoprotein of human immunodeficiency virus type 1. J. Virol. **70**:1100–8.

100. **Burton DR, Pyati J, Koduri R, Sharp SJ, Thornton GB, Parren PW, Sawyer LS, Hendry RM, Dunlop N, Nara PL, et al.** 1994. Efficient neutralization of primary isolates of HIV-1 by a recombinant human monoclonal antibody. Sci. **266** :1024–1027.

101. **Corti D, Langedijk JPM, Hinz A, Seaman MS, Vanzetta F, Fernandez-Rodriguez BM, Silacci C, Pinna D, Jarrossay D, Balla-Jhagjhoorsingh S, Willems B, Zekveld MJ, Dreja H, O'Sullivan E, Pade C, Orkin C, Jeffs S a, Montefiori DC, Davis D, Weissenhorn W, McKnight A, Heeney JL, Sallusto F, Sattentau QJ, Weiss R a, Lanzavecchia A**. 2010. Analysis of memory B cell responses and isolation of novel monoclonal antibodies with neutralizing breadth from HIV-1-infected individuals. PLoS One **5**:e8805.

102. **Falkowska E, Ramos A, Feng Y, Zhou T, Moquin S, Walker LM, Wu X, Seaman MS, Wrin T, Kwong PD, Wyatt RT, Mascola JR, Poignard P, Burton DR**. 2012. PGV04, an HIV-1 gp120 CD4

binding site antibody, is broad and potent in neutralization but does not induce conformational changes characteristic of CD4. J. Virol. **86**:4394–403.

103. **Bonsignori M, Hwang K-K, Chen X, Tsao C-Y, Morris L, Gray E, Marshall DJ, Crump J a, Kapiga SH, Sam NE, Sinangil F, Pancera M, Yongping Y, Zhang B, Zhu J, Kwong PD, O'Dell S, Mascola JR, Wu L, Nabel GJ, Phogat S, Seaman MS, Whitesides JF, Moody MA, Kelsoe G, Yang X, Sodroski J, Shaw GM, Montefiori DC, Kepler TB, Tomaras GD, Alam SM, Liao H-X, Haynes BF**. 2011. Analysis of a clonal lineage of HIV-1 envelope V2/V3 conformational epitope-specific broadly neutralizing antibodies and their inferred unmutated common ancestors. J. Virol. **85**:9998–10009.

104. **Walker LM, Huber M, Doores KJ, Falkowska E, Pejchal R, Julien J-P, Wang S-K, Ramos A, Chan-Hui P-Y, Moyle M, Mitcham JL, Hammond PW, Olsen O a., Phung P, Fling S, Wong C-H, Phogat S, Wrin T, Simek MD, Koff WC, Wilson I a., Burton DR, Poignard P**. 2011. Broad neutralization coverage of HIV by multiple highly potent antibodies. 0. Nature **477**:466–70.

105. **Mikell I, Sather DN, Kalams S a, Altfeld M, Alter G, Stamatatos L**. 2011. Characteristics of the earliest cross-neutralizing antibody response to HIV-1. PLoS Pathog. **7**:e1001251.

106. **Medina-Ramírez M, Sánchez-Merino V, Sánchez-Palomino S, Merino-Mansilla a, Ferreira CB, Pérez I, González N, Alvarez a, Alcocer-González JM, García F, Gatell JM, Alcamí J, Yuste E**. 2011. Broadly cross-neutralizing antibodies in HIV-1 patients with undetectable viremia. J. Virol. **85**:5804–13.

107. **Doria-Rose N a**. 2010. HIV neutralizing antibodies: clinical correlates and implications for vaccines. J. Infect. Dis. **201**:981–3.

108. **Pancera M, McLellan JS, Wu X, Zhu J, Changela A, Schmidt SD, Yang Y, Zhou T, Phogat S, Mascola JR, Kwong PD**. 2010. Crystal structure of PG16 and chimeric dissection with somatically related PG9: structure-function analysis of two quaternary-specific antibodies that effectively neutralize HIV-1. J. Virol. **84**:8098–110.

109. **Zhou T, Georgiev I, Wu X, Yang Z-Y, Dai K, Finzi A, Kwon Y Do, Scheid JF, Shi W, Xu L, Yang Y, Zhu J, Nussenzweig MC, Sodroski J, Shapiro L, Nabel GJ, Mascola JR, Kwong PD**. 2010. Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science **329**:811–7.

110. **Kwong PD, Mascola JR**. 2012. Human Antibodies that Neutralize HIV-1: Identification, Structures, and B Cell Ontogenies. Immunity **37**:412–25.

111. **Tenzer S, Wee E, Burgevin A, Stewart-Jones G, Friis L, Lamberth K, Chang C, Harndahl M, Weimershaus M, Gerstoft J, Akkad N, Klenerman P, Fugger L, Jones EY, McMichael AJ, Buus S, Schild H, van Endert P, Iversen AKN**. 2009. Antigen processing influences HIV-specific cytotoxic T lymphocyte immunodominance. Nat. Immunol. **10**:636–46.

112. **Willis R a, Kappler JW, Marrack PC**. 2006. CD8 T cell competition for dendritic cells in vivo is an early event in activation. Proc. Natl. Acad. Sci. U. S. A. **103**:12063–8.

113. **Barouch DH, O'Brien KL, Simmons NL, King SL, Abbink P, Maxfield LF, Sun Y-H, La Porte A, Riggs AM, Lynch DM, Clark SL, Backus K, Perry JR, Seaman MS, Carville A, Mansfield KG, Szinger JJ, Fischer W, Muldoon M, Korber B**. 2010. Mosaic HIV-1 vaccines expand the breadth and depth of cellular immune responses in rhesus monkeys. Nat. Med. **16**:319–23.

114. **Cho MW, Kim YB, Lee MK, Gupta KC, Ross W, Plishka R, Buckler-White A, Igarashi T, Theodore T, Byrum R, Kemp C, Montefiori DC, Martin MA**. 2001. Polyvalent envelope glycoprotein vaccine elicits a broader neutralizing antibody response but is unable to provide

sterilizing protection against heterologous Simian/human immunodeficiency virus infection in pigtailed macaques. J. Virol. **75**:2224–34.

115. **Pal R, Wang S, Kalyanaraman VS, Nair BC, Whitney S, Keen T, Hocker L, Hudacik L, Rose N, Cristillo A, Mboudjeka I, Shen S, Wu-Chou T-H, Montefiori D, Mascola J, Lu S, Markham P**. 2005. Polyvalent DNA prime and envelope protein boost HIV-1 vaccine elicits humoral and cellular responses and controls plasma viremia in rhesus macaques following rectal challenge with an R5 SHIV isolate. J. Med. Primatol. **34**:226–236.

116. **Pal R, Kalyanaraman VS, Nair BC, Whitney S, Keen T, Hocker L, Hudacik L, Rose N, Mboudjeka I, Shen S, Wu-Chou T-H, Montefiori D, Mascola J, Markham P, Lu S**. 2006. Immunization of rhesus macaques with a polyvalent DNA prime/protein boost human immunodeficiency virus type 1 vaccine elicits protective antibody response against simian human immunodeficiency virus of R5 phenotype. Virology **348**:341–53.

117. **Wang S, Kennedy JS, West K, Montefiori DC, Coley S, Lawrence J, Shen S, Green S, Rothman AL, Ennis F a., Arthos J, Pal R, Markham P, Lu S**. 2008. Cross-subtype antibody and cellular immune responses induced by a polyvalent DNA prime-protein boost HIV-1 vaccine in healthy human volunteers. Vaccine **26**:3947–3957.

118. **Seaman MS, Leblanc DF, Grandpre LE, Bartman MT, Montefiori DC, Letvin NL, Mascola JR**. 2007. Standardized assessment of NAb responses elicited in rhesus monkeys immunized with single- or multi-clade HIV-1 envelope immunogens. Virology **367**:175–86.

119. **Spearman P, Lally M a, Elizaga M, Montefiori D, Tomaras GD, McElrath MJ, Hural J, De Rosa SC, Sato A, Huang Y, Frey SE, Sato P, Donnelly J, Barnett S, Corey LJ**. 2011. A trimeric, V2-deleted HIV-1 envelope glycoprotein vaccine elicits potent neutralizing antibodies but limited breadth of neutralization in human volunteers. J. Infect. Dis. **203**:1165–73.

120. **Wang S, Pal R, Mascola JR, Chou T-HW, Mboudjeka I, Shen S, Liu Q, Whitney S, Keen T, Nair BC, Kalyanaraman VS, Markham P, Lu S**. 2006. Polyvalent HIV-1 Env vaccine formulations delivered by the DNA priming plus protein boosting approach are effective in generating neutralizing antibodies against primary human immunodeficiency virus type 1 isolates from subtypes A, B, C, D and E. Virology **350**:34–47.

121. **Malherbe DC, Doria-Rose N a, Misher L, Beckett T, Puryear WB, Schuman JT, Kraft Z, O'Malley J, Mori M, Srivastava I, Barnett S, Stamatatos L, Haigwood NL**. 2011. Sequential immunization with a subtype B HIV-1 envelope quasispecies partially mimics the in vivo development of neutralizing antibodies. J. Virol. **85**:5262–74.

122. **Zolla-Pazner S, Kong X-P, Jiang X, Cardozo T, Nádas A, Cohen S, Totrov M, Seaman MS, Wang S, Lu S**. 2011. Cross-clade HIV-1 neutralizing antibodies induced with V3-scaffold protein immunogens following priming with gp120 DNA. J. Virol. **85**:9887–98.

123. **Chohan B, Lavreys L, Rainwater SMJ, Overbaugh J**. 2005. Evidence for Frequent Reinfection with Human Immunodeficiency Virus Type 1 of a Different Subtype. J. Virol. **79**:10701–10708.

124. **Smith DM, Wong JK, Hightower GK, Ignacio CC, Koelsch KK, Daar ES, Richman DD, Little SJ**. 2004. Incidence of HIV superinfection following primary infection. JAMA **292**:1177–8.

125. **Redd AD, Mullis CE, Serwadda D, Kong X, Martens C, Ricklefs SM, Tobian AAR, Xiao C, Grabowski MK, Nalugoda F, Kigozi G, Laeyendecker O, Kagaayi J, Sewankambo N, Gray RH, Porcella SF, Wawer MJ, Quinn TC**. 2012. The Rates of HIV Superinfection and Primary HIV Incidence in a General Population in Rakai, Uganda. J. Infect. Dis. **206**:267–74.

126. **Ronen K, McCoy CO, Matsen F a, Boyd DF, Emery S, Odem-Davis K, Jaoko W, Mandaliya K, McClelland RS, Richardson B a, Overbaugh J**. 2013. HIV-1 Superinfection Occurs Less

Frequently Than Initial Infection in a Cohort of High-Risk Kenyan Women. PLoS Pathog. **9**:e1003593.

127. **Rachinger A, Manyenga P, Burger J a, Derks van de Ven TLP, Stolte IG, Prins M, van 't Wout AB, Schuitemaker H**. 2011. Low incidence of HIV-1 superinfection even after episodes of unsafe sexual behavior of homosexual men in the Amsterdam Cohort Studies on HIV Infection and AIDS. J. Infect. Dis. **203**:1621–8.

128. **Gonzales MJ, Delwart E, Rhee S-Y, Tsui R, Zolopa AR, Taylor J, Shafer RW**. 2003. Lack of detectable human immunodeficiency virus type 1 superinfection during 1072 person-years of observation. J. Infect. Dis. **188**:397–405.

129. **Tsui R, Herring BL, Barbour JD, Grant RM, Bacchetti P, Kral A, Edlin BR, Delwart EL**. 2003. Human Immunodeficiency Virus Type 1 Superinfection Was Not Detected following 215 Years of Injection Drug User Exposure. J. Virol. **78**:94–103.

130. **Gottlieb GS, Nickle DC, Jensen M a, Wong KG, Grobler J, Li F, Liu S-L, Rademeyer C, Learn GH, Karim SSA, Williamson C, Corey L, Margolick JB, Mullins JI**. 2004. Dual HIV-1 infection associated with rapid disease progression. Lancet **363**:619–22.

131. **Grobler J, Gray CM, Rademeyer C, Seoighe C, Ramjee G, Karim SA, Morris L, Williamson C**. 2004. Incidence of HIV-1 dual infection and its association with increased viral load set point in a cohort of HIV-1 subtype C-infected female sex workers. J. Infect. Dis. **190**:1355–9.

132. **Smith DM, Strain MC, Frost SDW, Pillai SK, Wong JK, Wrin T, Liu Y, Petropolous CJ, Daar ES, Little SJ, Richman DD**. 2006. Lack of neutralizing antibody response to HIV-1 predisposes to superinfection. Virology **355**:1–5.

133. **Basu D, Kraft CS, Murphy MK, Campbell PJ, Yu T, Hraber PT, Irene C, Pinter A, Chomba E, Mulenga J, Kilembe W, Allen S a, Derdeyn C a, Hunter E**. 2012. HIV-1 subtype C superinfected individuals mount low autologous neutralizing antibody responses prior to intrasubtype superinfection. Retrovirology **9**:76.

134. **Blish C a, Dogan OC, Derby NR, Nguyen M-A, Chohan B, Richardson B a, Overbaugh J**. 2008. Human immunodeficiency virus type 1 superinfection occurs despite relatively robust neutralizing antibody responses. J. Virol. **82**:12094–103.

135. **Piantadosi A, Chohan B, Chohan V, McClelland RS, Overbaugh J**. 2007. Chronic HIV-1 infection frequently fails to protect against superinfection. PLoS Pathog. **3**:e177.

136. **Lavreys L, Baeten JM, Chohan V, McClelland RS, Hassan WM, Richardson BA, Mandaliya K, Ndinya-Achola JO, Overbaugh J**. 2006. Higher set point plasma viral load and more-severe acute HIV type 1 (HIV-1) illness predict mortality among high-risk HIV-1-infected African women. Clin. Infect. Dis. **42**:1333–9.

137. **Smith DM, Richman DD, Little SJ**. 2005. HIV Superinfection. J. Infect. Dis. **192**:438–444.

138. **Harro CD, Judson FN, Gorse GJ, Mayer KH, Kostman JR, Brown SJ, Koblin B, Marmor M, Bartholow BN, Popovic V**. 2004. Recruitment and baseline epidemiologic profile of participants in the first phase 3 HIV vaccine efficacy trial. J. Acquir. Immune Defic. Syndr. **37**:1385–92.

139. **Derdeyn C a, Decker JM, Bibollet-Ruche F, Mokili JL, Muldoon M, Denham S a, Heil ML, Kasolo F, Musonda R, Hahn BH, Shaw GM, Korber BT, Allen S, Hunter E**. 2004. Envelope-constrained neutralization-sensitive HIV-1 after heterosexual transmission. Science **303**:2019–22.

140. **Chohan B, Lang D, Sagar M, Korber B, Lavreys L, Richardson B, Overbaugh J**. 2005. Selection for human immunodeficiency virus type 1 envelope glycosylation variants with shorter V1-V2 loop sequences occurs during transmission of certain genetic subtypes and may impact viral RNA levels. J. Virol. **79**:6528–31.

141. **Van Loggerenberg F, Mlisana K, Williamson C, Auld SC, Morris L, Gray CM, Abdool Karim Q, Grobler A, Barnabas N, Iriogbe I, Abdool Karim SS**. 2008. Establishing a cohort at high risk of HIV infection in South Africa: challenges and experiences of the CAPRISA 002 acute infection study. PLoS One **3**:e1954.

142. **Kraus MH, Parrish NF, Shaw KS, Decker JM, Keele BF, Salazar-Gonzalez JF, Grayson T, McPherson DT, Ping L-H, Anderson JA, Swanstrom R, Williamson C, Shaw GM, Hahn BH**. 2010. A rev1-vpu polymorphism unique to HIV-1 subtype A and C strains impairs envelope glycoprotein expression from rev-vpu-env cassettes and reduces virion infectivity in pseudotyping assays. Virology **397**:346–57.

143. **Price MN, Dehal PS, Arkin AP**. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLoS One **5**:e9490.

144. **Letunic I, Bork P**. 2011. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res. **39**:W475–8.

145. **Woodman Z, Mlisana K, Treurnicht F, Abrahams M, Thebus R, Karim SA, Williamson For The Caprisa Acute Infection Study Team C**. 2011. Decreased Incidence of Dual Infections in South African Subtype C-Infected Women Compared to a Cohort Ten Years Earlier. AIDS Res. Hum. Retroviruses **27**:0–5.

146. **Ntale R**. 2012. PhD Thesis: The role of early CTL escape in the pathogenesis of HIV-1 subtype C infection. University of Cape Town.

147. **Jost S, Bernard M-C, Kaiser L, Yerly S, Hirschel B, Samri A, Autran B, Goh L-E, Perrin L**. 2002. A patient with HIV-1 superinfection. N. Engl. J. Med. **347**:731–6.

148. **Van der Kuyl AC, Kozaczynska K, van den Burg R, Zorgdrager F, Back N, Jurriaans S, Berkhout B, Reiss P, Cornelissen M**. 2005. Triple HIV-1 infection. N. Engl. J. Med. **352**:2557–9.

149. **Archer J, Pinney JW, Fan J, Simon-Loriere E, Arts EJ, Negroni M, Robertson DL**. 2008. Identifying the important HIV-1 recombination breakpoints. PLoS Comput. Biol. **4**:e1000178.

150. **Fleming DT, Wasserheit JN**. 1999. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. Sex. Transm. Infect. **75**:3–17.

151. **Galvin SR, Cohen MS**. 2004. The role of sexually transmitted diseases in HIV transmission. Nat. Rev. Microbiol. **2**:33–42.

152. **Hayes R, Watson-Jones D, Celum C, van de Wijgert J, Wasserheit J**. 2010. Treatment of sexually transmitted infections for HIV prevention: end of the road or new beginning? AIDS **24 Suppl 4**:S15–26.

153. **Mutua FM, M'imunya JM, Wiysonge CS**. 2012. Genital ulcer disease treatment for reducing sexual acquisition of HIV. Cochrane database Syst. Rev. **8**:CD007933.

154. **Atashili J, Poole C, Ndumbe PM, Adimora AA, Smith JS**. 2008. Bacterial vaginosis and HIV acquisition: a meta-analysis of published studies. AIDS **22**:1493–501.

155. **Mlisana K, Naicker N, Werner L, Roberts L, van Loggerenberg F, Baxter C, Passmore J-AS, Grobler AC, Sturm a W, Williamson C, Ronacher K, Walzl G, Abdool Karim SS**. 2012. Symptomatic vaginal discharge is a poor predictor of sexually transmitted infections and genital tract inflammation in high-risk women in South Africa. J. Infect. Dis. **206**:6–14.

156. **Heffron R, Donnell D, Rees H, Celum C, Mugo N, Were E, de Bruyn G, Nakku-Joloba E, Ngure K, Kiarie J, Coombs RW, Baeten JM**. 2012. Use of hormonal contraceptives and risk of HIV-1 transmission: a prospective cohort study. Lancet Infect. Dis. **12**:19–26.

157. **Sagar M, Lavreys L, Baeten J**. 2004. Identification of modifiable factors that affect the genetic diversity of the transmitted HIV-1 population. Aids 615–619.

158. **Sagar M, Lavreys L, Baeten JM, Richardson BA, Mandaliya K, Chohan BH, Kreiss JK, Overbaugh J**. 2003. Infection with Multiple Human Immunodeficiency Virus Type 1 Variants Is Associated with Faster Disease Progression. J. Virol. **77**:12921–12926.

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

160. **Clerc O, Colombo S, Yerly S, Telenti A, Cavassini M**. 2010. HIV-1 elite controllers: beware of super-infections. J. Clin. Virol. **47**:376–8.

161. **Streeck H, Li B, Poon AFY, Schneidewind A, Gladden AD, Power KA, Daskalakis D, Bazner S, Zuniga R, Brander C, Rosenberg ES, Frost SDW, Altfeld M, Allen TM**. 2008. Immune-driven recombination and loss of control after HIV superinfection. J. Exp. Med. **205** :1789–1796.

162. **Cortez V, Odem-Davis K, McClelland RS, Jaoko W, Overbaugh J**. 2012. HIV-1 superinfection in women broadens and strengthens the neutralizing antibody response. PLoS Pathog. **8**:e1002611.

163. **Rachinger A, Navis M, van Assen S, Groeneveld PHP, Schuitemaker H**. 2008. Recovery of viremic control after superinfection with pathogenic HIV type 1 in a long-term elite controller of HIV type 1 infection. Clin. Infect. Dis. **47**:e86–9.

164. **Casado C, Pernas M, Alvaro T, Sandonis V, García S, Rodríguez C, del Romero J, Grau E, Ruiz L, Lopez-Galindez C**. 2007. Coinfection and superinfection in patients with long-term, nonprogressive HIV-1 disease. J. Infect. Dis. **196**:895–9.

165. **Pacold ME, Pond SLK, Wagner GA, Delport W, Bourque DL, Richman DD, Little SJ, Smith DM**. 2012. Clinical, virologic, and immunologic correlates of HIV-1 intraclade B dual infection among men who have sex with men. AIDS **26**:157–65.

166. **Cornelissen M, Pasternak AO, Grijsen ML, Zorgdrager F, Bakker M, Blom P, Prins JM, Jurriaans S, van der Kuyl AC**. 2012. HIV-1 dual infection is associated with faster CD4+ T-cell decline in a cohort of men with primary HIV infection. Clin. Infect. Dis. **54**:539–47.

167. **Muller E, Barday Z, Mendelson M, Kahn D**. 2012. Renal transplantation between HIV-positive donors and recipients justified. S. Afr. Med. J. **102**:497–8.

168. 2013. U.S. Congress Moves Closer to Lifting Ban on Transplanting Organs From HIV-Positive Donors. Sci. Available http//news.sciencemag.org/2013/07/u.s.-congress-moves-closer-lifting-ban-transplanting-organs-hiv-positive-donors. Accessed July 21st 2013.

169. **Nugent RP, Krohn M a, Hillier SL**. 1991. Reliability of diagnosing bacterial vaginosis is improved by a standardized method of gram stain interpretation. J. Clin. Microbiol. **29**:297–301.

170. **Gray CM, Mlotshwa M, Riou C, Mathebula T, de Assis Rosa D, Mashishi T, Seoighe C, Ngandu N, van Loggerenberg F, Morris L, Mlisana K, Williamson C, Karim SA**. 2009. Human immunodeficiency virus-specific gamma interferon enzyme-linked immunospot assay responses targeting specific regions of the proteome during primary subtype C infection are poor predictors of the course of viremia and set point. J. Virol. **83**:470–8.

171. **Leslie A, Matthews PC, Listgarten J, Carlson JM, Kadie C, Ndung'u T, Brander C, Coovadia H, Walker BD, Heckerman D, Goulder PJR**. 2010. Additive contribution of HLA class I alleles in the immune control of HIV-1 infection. J. Virol. **84**:9879–88.

172. **Index R, Development TR, Team C**. 2010. R : A Language and Environment for Statistical Computing.

173. **Chow GC**. 1960. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. Econometrica **28**:591–605.

174. **Kassler WJ, Zenilman JM, Ericksorv B, Fox R, Peterman TA, Hook EWIII**. 1994. Seroconversion in patients attending sexually transmitted disease clinics. AIDS **8**.

175. **Laga M, Manoka A, Kivuvu M, Malele B, Tuliza M, Nzila N, Goeman J, Behets F, Batter V, Alary M, Heyward WL, Ryder RW, Piot P**. 1993. Non-ulcerative sexually transmitted diseases as risk factors for HIV-1 transmission in women: results from a cohort study. AIDS **7**.

176. **Dickerson MC, Johnston J, Delea TE, White A, Andrews E**. 1996. The causal role for genital ulcer disease as a risk factor for transmission of human immunodeficiency virus. An application of the Bradford Hill criteria. Sex. Transm. Dis. **23**:429–40.

177. **Ssemwanga D, Lyagoba F, Ndembi N, Mayanja BN, Larke N, Wang S, Baalwa J, Williamson C, Grosskurth H, Kaleebu P**. 2011. Multiple HIV-1 infections with evidence of recombination in heterosexual partnerships in a low risk Rural Clinical Cohort in Uganda. Virology **411**:113–31.

178. **Braibant M, Xie J, Samri A, Agut H, Autran B, Barin F**. 2010. Disease progression due to dual infection in an HLA-B57-positive asymptomatic long-term nonprogressor infected with a nef-defective HIV-1 strain. Virology **405**:81–92.

179. **Doria-Rose NA, Klein RM, Manion MM, O'Dell S, Phogat A, Chakrabarti B, Hallahan CW, Migueles SA, Wrammert J, Ahmed R, Nason M, Wyatt RT, Mascola JR, Connors M**. 2009. Frequency and phenotype of human immunodeficiency virus envelope-specific B cells from patients with broadly cross-neutralizing antibodies. J. Virol. **83**:188–99.

180. **Powell RLR, Kinge T, Nyambi PN**. 2010. Infection by discordant strains of HIV-1 markedly enhances the neutralizing antibody response against heterologous virus. J. Virol. **84**:9415–26.

181. **Edgar RC**. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32** :1792–1797.

182. **Kumar S, Nei M, Dudley J, Tamura K**. 2008. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. Briefings Bioinforma. **9** :299–306.

183. **Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL**. 2012. Detecting Individual Sites Subject to Episodic Diversifying Selection. PLoS Genet **8**:e1002764.

184. **Kosakovsky Pond SL, Poon AFY, Leigh Brown AJ, Frost SDW**. 2008. A Maximum Likelihood Method for Detecting Directional Evolution in Protein Sequences and Its Application to Influenza A Virus. Mol. Biol. Evol. **25** :1809–1824.

185. **Pond SLK, Frost SDW, Muse S V**. 2005. HyPhy: hypothesis testing using phylogenies. Bioinforma. **21** :676–679.

186. **Moore PL, Sheward D, Nonyane M, Ranchobe N, Hermanus T, Gray ES, Karim SSA, Williamson C, Morris L**. 2013. Multiple pathways of escape from HIV broadly cross-neutralizing V2-dependent antibodies. J. Virol. **87**:4882–4894.

187. **Brady LJ**. 2005. Antibody-Mediated Immunomodulation : a Strategy To Improve Host Responses against Microbial Antigens. Infect. Immun. **73**:671–678.

188. **Pasare C, Medzhitov R**. 2005. Control of B-cell responses by Toll-like receptors. Nature **438**:364–8.

189. **Rawlings DJ, Schwartz MA, Jackson SW, Meyer-Bahlburg A**. 2012. Integration of B cell responses through Toll-like receptors and antigen receptors. Nat. Rev. Immunol. **12**:282–94.

190. **Herrero R**. 2009. Human papillomavirus (HPV) vaccines: limited cross-protection against additional HPV types. J. Infect. Dis. **199**:919–22.

191. **Wei C-J, Yassine HM, McTamney PM, Gall JGD, Whittle JRR, Boyington JC, Nabel GJ**. 2012. Elicitation of broadly neutralizing influenza antibodies in animals with previous influenza exposure. Sci. Transl. Med. **4**:147ra114.

192. **John TJ, Devarajan L V, Balasubramanyan a**. 1976. Immunization in India with trivalent and monovalent oral poliovirus vaccines of enhanced potency. Bull. World Health Organ. **54**:115–7.

193. **Grassly NC, Wenger J, Durrani S, Bahl S, Deshpande JM, Sutter RW, Heymann DL, Aylward RB**. 2007. Protective efficacy of a monovalent oral type 1 poliovirus vaccine: a case-control study. Lancet **369**:1356–62.

194. **Patriarca PA, Wright PF, John TJ**. 1991. Factors Affecting the Immunogenicity of Oral Poliovirus Vaccine in Developing Countries : Review in Tropical Countries. Rev. if Infect. Dis. **13**:926–939.

195. **Burton DR, Ahmed R, Barouch DH, Butera ST, Crotty S, Godzik A, Kaufmann DE, McElrath MJ, Nussenzweig MC, Pulendran B, Scanlan CN, Schief WR, Silvestri G, Streeck H, Walker BD, Walker LM, Ward AB, Wilson IA, Wyatt R**. 2012. A Blueprint for HIV Vaccine Discovery. Cell Host Microbe **12**:396–407.

196. **Schief WR, Ban YA, Stamatatos L**. 2009. Challenges for structure-based HIV vaccine design. Curr. Opin. HIV AIDS **4**:431–40.

197. **West AP, Scharf L, Horwitz J, Klein F, Nussenzweig MC, Bjorkman PJ**. 2013. Computational analysis of anti-HIV-1 antibody neutralization panel data to identify potential functional epitope residues. Proc. Natl. Acad. Sci. U. S. A. **110**:10598–603.

198. **Smyth RP, Davenport MP, Mak J**. 2012. The origin of genetic diversity in HIV-1. Virus Res. **169**:415–29.

Supplementary data

Table S2.1 | Twelve *envelope* sequencing primers

	Primer Sequence (5'-3')
EF00	GGGAAAGAGCAGAAGACAGTGGCAATGA
FOR14	TATGGGACCAAAGCCTAAAGCCATGTG
FOR16	TTTAATTGTGGAGGAGAATTTTTCTA
EF170	AGCAGGAAGCACTATGGG
EF200	GGGATAACATGACCTGGATGCAGTGGG
EF260	TTCAGCTACCACCGATTGAGAGACT
EF175	TTTAGCATCTGATGCACAGAATAG
REV15	CTGCCATTTAACAGCAGTTGAGTTGA
EF115	AGAAAAATTCTCCTCTACAATTAA
EF55	GCCCCAGACCGTGAGTTGCAACATATG
EF15	TGCTCTCCACCTTCTTCTTC
REV19	ACTTTTTGACCACTTGCCACCCAT

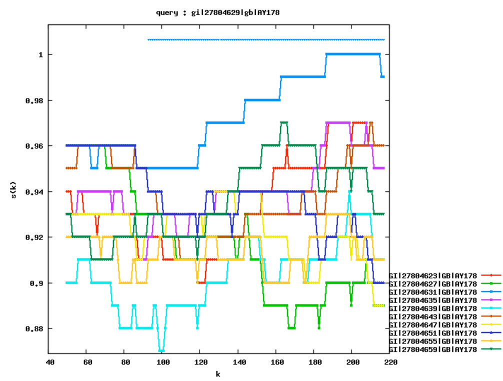
Supplementary Table S2.2: Number of *envelope* SGAs analyzed

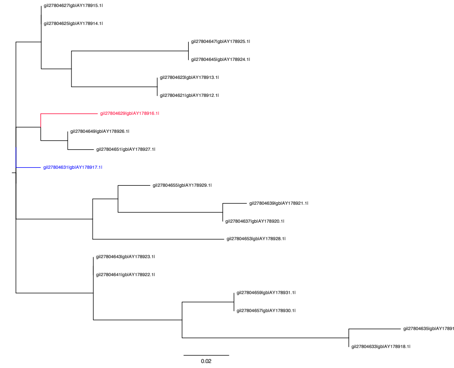
	Number of *envelope* SGAs generated
CAP37	
2 wpi	20
13 wpi	11
21 wpi	30*
56 wpi	22*
CAP84	
1 wpi	23
3 wpi	2
4 wpi	7*
10 wpi	9*
19 wpi	7*
54 wpi	20*
CAP137	
7 wpi	4*
12 wpi	33*
23 wpi	24*
52 wpi	26*
CAP267	
6 wpi	7*
10 wpi	12*
20 wpi	13*
52 wpi	17*
CAP256	
6 wpi	2
15 wpi	6*#
30 wpi	6
53 wpi	10*
59 wpi	18
94 wpi	16
176 wpi	6

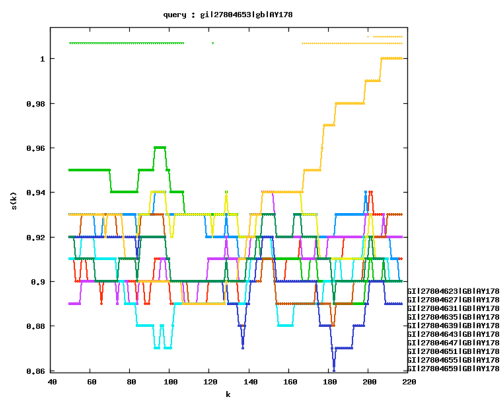
Supplementary Table S2.2 contd.

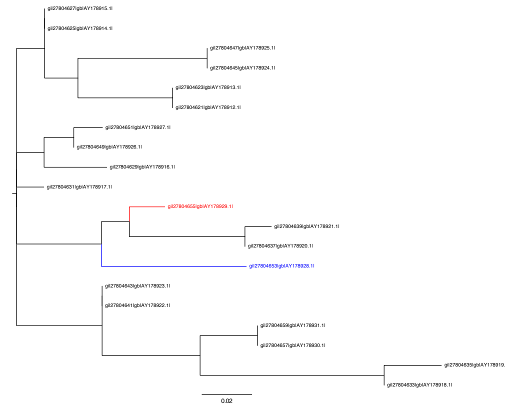
CAP237	
2 wpi	22
4 wpi	3*
11 wpi	6*
24 wpi	1*
54 wpi	6*
171 wpi	10
CAP281	
10 wpi	1*
13 wpi	3*
30 wpi	1*
42 wpi	2*
46 wpi	20*
101 wpi	10*
CAP200	
2 wpi	22
11 wpi	5*
16 wpi	12*
21 wpi	10
176 wpi	33
CAP225	
2 wpi	21
3 wpi	6*
5 wpi	8*
7 wpi	10*
24 wpi	2*
29 wpi	10*
171 wpi	8*
CAP266	
7 wpi	10
58 wpi	11
CAP258	
7 wpi	9

*Generated as part of this thesis
#Strain-specific SGAs

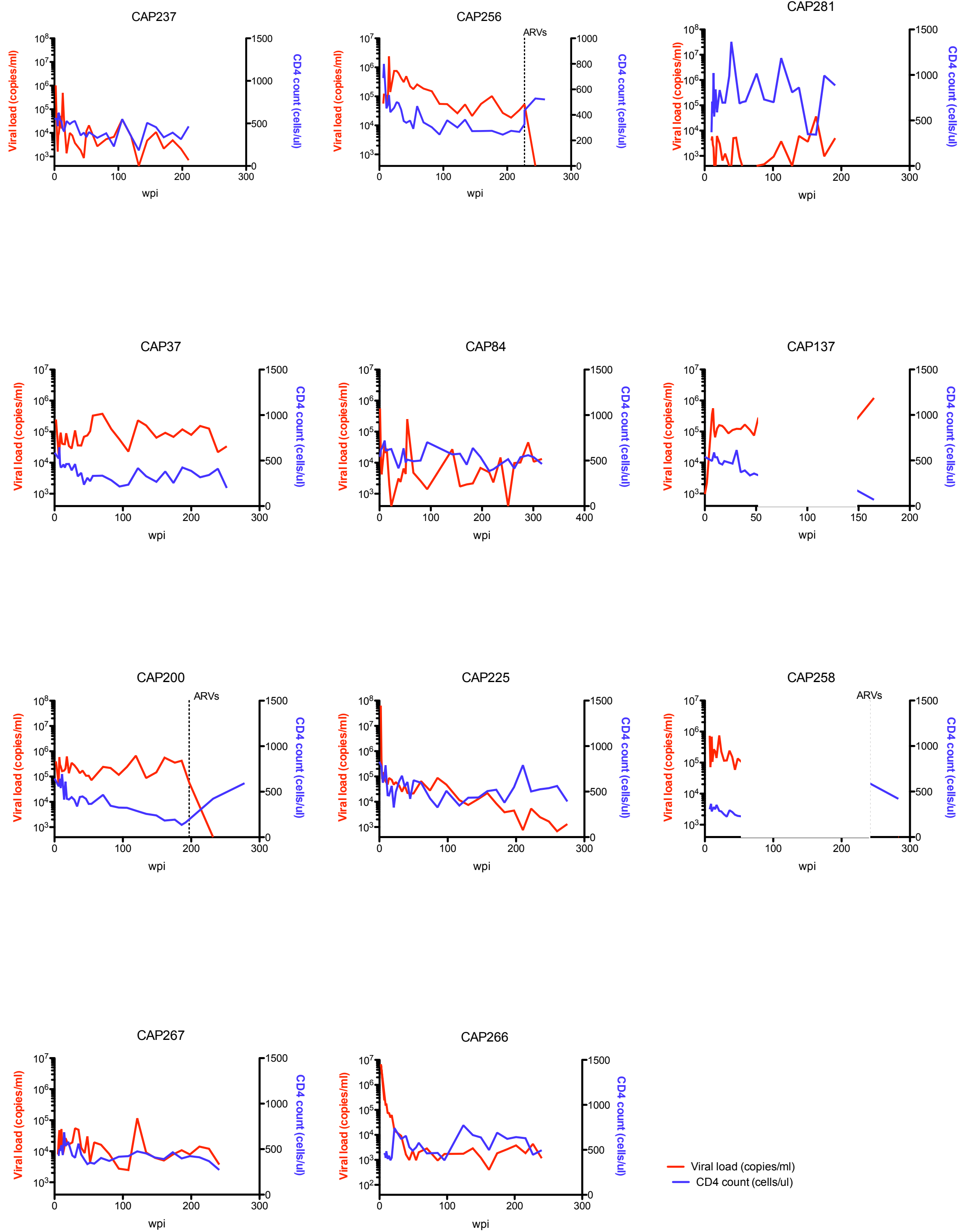




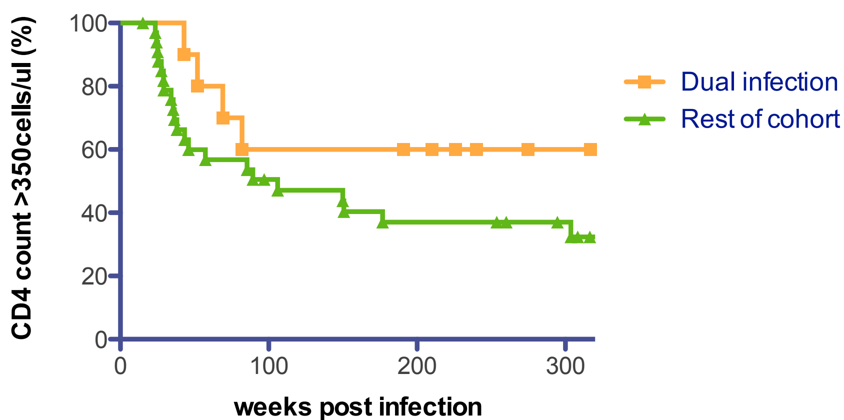




Supplementary Figure S2.1 | Recombination can obscure the identification of superinfection. Reanalysis of publicly available sequences from (128) **identified two individuals with viruses harbouring evidence of recombination with a divergent secondary virus, suggestive of superinfection.**



Supplementary Figure 3.1 | Viral load and CD4 count profiles for the 11 dual infected participants described. Viral loads are shown in red (left Y-axis), and CD4 counts in blue (right Y-axis). The window of superinfection, where applicable, is shaded in orange. Initiation onto antiretroviral therapy is depicted by the dotted line (ARVs). Two participants, CAP137 and CAP258 were lost to follow up, and these visits have been whited out.



Supplementary Figure S3.2 | Kaplan-Meier survival plot showing time for CD4+ T cell counts to fall below 350 cells/µl in and dual infected (N=11) participants and the rest of the cohort (N=51).

Supplementary Table S3.1 | A table of estimated statistical power for a survival analysis under a variety of effect sizes

	Median time to CD4<350 (weeks)
Singles (N=18)	225
Duals (N=11)	180
Power to detect	**11%**
Singles (N=18)	225
Duals (N=11)	130
Power to detect	**22%**
Singles (N=18)	225
Duals (N=11)	100
Power to detect	**50%**
Singles (N=18)	225
Duals (N=11)	50
Power to detect	**88%**

Table S4.1 | Summary statistics for the evidence of positive selection in V2 of the primary (PI) and superinfecting (SI) virus populations in CAP256

HXB2 position			SU	PI
160	Directional (DEPS)	Max BF	ns*	ns
	Diversifying (MEME)	P value	ns#	ns
161	Directional (DEPS)	Max BF	ns	ns
	Diversifying (MEME)	P value	ns	ns
162	Directional (DEPS)	Max BF	ns	ns
	Diversifying (MEME)	P value	ns	ns
163	Directional (DEPS)	Max BF	ns	ns
	Diversifying (MEME)	P value	ns	ns
164	Directional (DEPS)	Max BF	ns	ns
	Diversifying (MEME)	P value	ns	ns
165	Directional (DEPS)	Max BF (Preferred residues)	**191 (V)**	ns
	Diversifying (MEME)	P value	ns	ns
166	Directional (DEPS)	Max BF (Preferred residues)	ns	**3403 (S)**
	Diversifying (MEME)	P value	ns	ns
167	Directional (DEPS)	Max BF	ns	ns
	Diversifying (MEME)	P value	ns	ns
168	Directional (DEPS)	Max BF	ns	ns
	Diversifying (MEME)	P value	ns	ns
169	Directional (DEPS)	Max BF (Preferred residues)	**77095 (IQ)**	ns
	Diversifying (MEME)	P value	**0.0001**	ns
170	Directional (DEPS)	Max BF	ns	ns
	Diversifying (MEME)	P value	ns	ns
171	Directional (DEPS)	Max BF (Preferred residues)	ns	**290 (K)**
	Diversifying (MEME)	P value	ns	ns

* Bayes factor (BF) <20; # P> 0.05