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

Daniel James Sheward

Supervisor: Carolyn Williamson

Co-supervisor: Penny L Moore

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the

Department of Pathology

UNIVERSITY OF CAPE TOWN

2022

Declaration

I, Daniel James Sheward, hereby declare that the work on which this thesis is based is my original work except where acknowledgements indicate otherwise, and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. I authorise the University to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

I confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publication in my thesis:
Sheward et al., (2018). Cell Host and Microbe. Oct 10;24(4):593-599.e3. doi:10.1016/j.chom.2018.09.001

Signature: Signed by candidate

Date: 30 March 2022

Table of Contents

Abstract

The development of a safe, effective prophylactic HIV vaccine remains a major global health priority. Stabilized, soluble trimers that mimic the native functional HIV trimer have been developed that elicit strain-specific neutralizing HIV antibodies in animal models, and are currently being evaluated in several human clinical trials. Identifying whether multiple immunogens could be administered to facilitate the broadening of responses represents a pivotal challenge. In this thesis, we characterized the antibody response in individuals infected with multiple HIV strains to inform the development of polyvalent and sequential HIV vaccine regimens.

We found that conventional approaches to detect HIV co- and superinfection are confounded by recombination. Therefore, we developed an automated, Bayesian approach to detect superinfection explicitly accounting for recombination. Using simulated and real sequence data, we demonstrated that this approach is sensitive, highly specific, and robust to recombination. Furthermore, analyzing previously published sequence datasets, we identified cases of superinfection that previously went undetected, indicating that superinfection occurs more frequently than previously estimated.

We characterized the development of antibodies in five superinfected individuals identified in the CAPRISA 002 acute infection cohort. Specifically, we evaluated whether superinfection re-engaged cross-reactive memory B cells, promoting the development of cross-neutralizing antibodies. By comparing the breadth of the neutralizing antibody response in superinfected individuals to those that typically develop in singly infected individuals, we showed that HIV superinfection was not sufficient to broaden responses. By characterizing the kinetics and specificity of autologous neutralizing antibody responses, we show that responses to the superinfecting viruses failed to efficiently recruit neutralizing memory B cells. Instead, the secondary infection elicited strain-specific, *de novo* responses. This occurred even though the superinfecting viruses were relatively closely related (from the same subtype).

To determine whether the co-exposure to diverse Env antigens favours the development of cross-neutralizing antibodies better than sequential exposure, we characterized the development of neutralizing antibodies in HIV co-infected individuals where several divergent viruses were transmitted prior to seroconversion. We identified three cases of co-infection that encompassed immunological exposure to: (i) two diverse, unlinked Envs, (ii) two related Envs with diversity uniformly distributed over the trimer, and (iii) two diverse but recombined Envs such that clusters of high homology were preserved in the presence of high diversity

elsewhere. We found that, like superinfection, co-infection was not sufficient to broaden neutralizing antibody responses. Co-exposure to two HIV Env antigens did not necessarily produce additive or cross-neutralizing antibody responses, and in some cases was subject to immunological interference. This was most evident in the case of co-infection with two related Envs where diversity was uniformly distributed across the Env trimer; in this case neutralizing antibody responses to one variant arose to the near exclusion of responses to the other. However, in the case of co-exposure to diverse Envs but where the trimer apex was conserved in both variants through recombination, potent neutralization of both variants was evident. This was the co-infected participant who developed the broadest neutralizing antibody response, and we show that cross-neutralization was mediated, in part, by trimer apex-targeting neutralizing antibodies.

In conclusion, we find that HIV superinfection fails to efficiently recruit neutralizing memory B cells and, at best, results in additive nAb responses rather than a synergistic effect leading to cross-neutralization; a distinction that is highly relevant for vaccine design. While sequential immunizations with heterologous Env immunogens may be able to improve the potency of elicited responses, alone, they are unlikely to promote the development of bnAbs. Our observations from cases of co-infection suggests that cocktails of divergent stabilized Env trimers are unlikely to drive the development of cross-neutralizing antibodies, and may be subject to interference. However, the rational design of more similar immunogen cocktails where conserved epitopes are preserved across immunogens may be able to facilitate neutralizing antibodies to these targets, as seen in one individual. Thus, the use of related, stabilized Env trimers with diversity introduced in key regions together with strategies to reduce the immunogenicity of immunodominant, strain-specific epitopes may represent one path to a cross-neutralizing antibody response to multiple Envs within a cocktail.

CHAPTER 1

Introduction

A vaccine that elicits broadly neutralizing HIV antibodies will likely be protective.

By 2020, there were an estimated 38 million people living with HIV globally (UNAIDS) (Figure 1a). The development of highly active antiretroviral therapy (HAART) has dramatically improved the prognosis of infected individuals (May and Ingle, 2011), and represents a global success story with AIDS-related deaths reduced by 64% from their peak in 2004 (Figure 1b). Effective treatment can also reduce viral load to the point where onward transmission is extremely unlikely (Cohen et al., 2016). While the number of new infections fell to ~1.5 million in 2020 (representing a 52% reduction compared to the peak in 1997) (Figure 1c) it is unlikely that the capacity to rollout antiretrovirals will be sufficient to end the epidemic (Figure 1d). The development of a safe, effective prophylactic HIV vaccine remains a major global health priority.

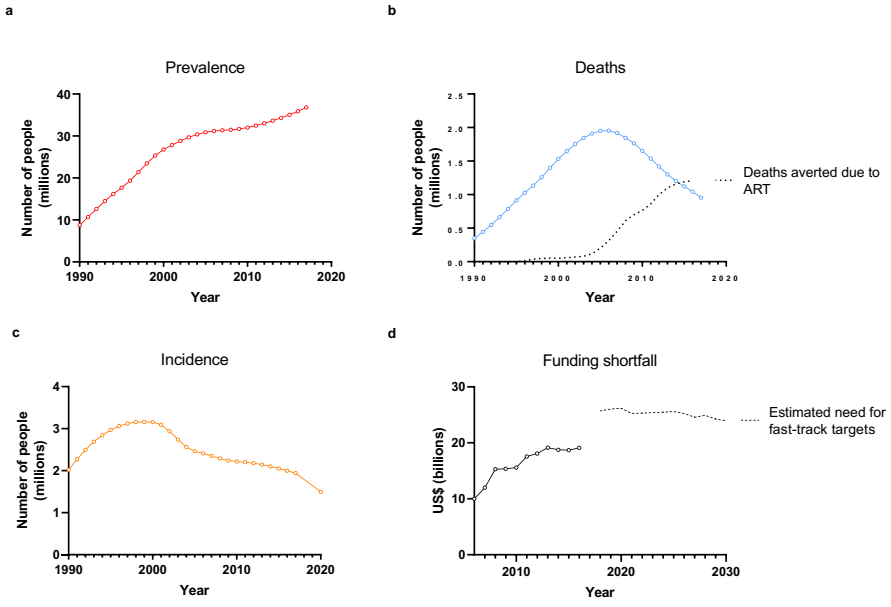


Figure 1.1 An overview of global HIV statistics.

a Estimated global HIV prevalence by year. **b** Global HIV-related deaths by year. The number of deaths estimated to have been averted due to the rollout of antiretroviral therapy (ART) is shown with a dashed line. **c** The number of new cases of HIV globally, by year. **d** Estimated funding earmarked to address HIV in low-to-middle income countries in 2016 US dollar equivalents. UNAIDS projected estimates of the required funding to meet 2030 fast-track targets (<200 000 new infections) are shown with the dashed line. Data for (a-c) sourced from Institute for Health Metrics and Evaluation (IHME) via OurWorldInData.org. Data for (d) from UNAIDS and OurWorldInData.org.

For many pathogens, surviving an infection frequently leads to immunity to reinfection. This protection is highly dependent on long-lived, antibody secreting plasma cells and memory B cells. Similarly, nearly all licensed vaccines mediate protection through the activity of antibodies (Plotkin, 2010).

For HIV, there is substantial evidence that passive transfer of neutralizing antibodies (nAbs) can protect non-human primates from challenge with HIV/SIV chimeric viruses (SHIV) (Gautam et al., 2016; Mascola et al., 2000; Moldt et al., 2012; Shingai et al., 2014). The level of protection in these animal models correlates strongly with neutralization (Pegu et al., 2019), and while neutralizing monoclonal antibodies (mAbs) defective in effector function offered reduced protection (Hessell et al., 2009), non-neutralizing (human) antibodies have failed to prevent infection (Burton et al., 2011; Dugast et al., 2014) of non-human primates. Recently, two trials (HVTN703 and HVTN704) directly evaluated whether regular, intravenous infusions of a neutralizing mAb could prevent HIV acquisition. Although there was no significant reduction in HIV acquisition overall compared to placebo, there was evidence of protection against infection with viruses highly sensitive to the passively administered mAb, VRC01 (Corey et al., 2021). Similarly, in immunized rhesus macaques, protection against repeated intra-rectal challenge was strongly correlated with serum nAb titers against the challenge virus (Pauthner et al., 2019). In human vaccine studies, one phase 3 human clinical trial, RV144, failed to elicit detectable levels of nAbs in vaccinees (Montefiori et al., 2012) but still appeared to offer modest levels of protection - the only human clinical trial to date to show any statistically significant reduction in HIV infection rate (Rerks-Ngarm et al., 2009). While risk of infection was correlated with titers of non-neutralizing antibodies to a specific region of the HIV envelope (V2) (Haynes et al., 2012), the mechanism of protection, if any, still remains unclear. Statistical issues surrounding the efficacy in this trial are debated (Gilbert et al., 2011) and a second clinical trial based on the same concept but modified for a local clade C epidemic (HVTN 702) was recently stopped due to futility (Gray et al., 2021). Taken together, there is strong evidence to indicate that nAbs, if present at sufficiently high titers at the time of challenge, can prevent HIV infection. Thus, an HIV vaccine that elicits long-lived antibody responses capable of neutralizing transmitted viruses will likely be protective, and remains a major goal for research.

HIV diversity

A major obstacle to the development of an HIV vaccine is the tremendous diversity of circulating HIV variants that could be transmitted, specifically within the only known viral target for nAbs, the viral Envelope glycoprotein (Env). The encoding viral env gene diversifies at a rate of ~0.9% per year (Korber et al., 2000), and even within a local epidemic Env proteins isolated from separate individuals can differ by >10%. Different HIV-1 group M subtypes that circulate globally can differ by more than 35% in Env (Gaschen et al., 2002) (Figure 2), with HIV subtype C accounting for the majority of cases globally. In addition to rapid accumulation of point mutations and insertions/deletions, the diploid genome of HIV, like other retroviruses (Hu and Temin, 1990), undergoes frequent recombination (Delassus et al., 1991; Groenink et al., 1992; Robertson et al., 1995; Simmonds et al., 1991; Zhuang et al., 2002). As individuals can be infected with multiple divergent HIV variants, whether by simultaneous transmission of divergent variants (co-infection) or by successive infection (superinfection), recombination has facilitated the generation of mosaic genomes and circulating recombinant forms, further expanding the extreme genetic diversity. A successful vaccine would need to elicit nAbs capable of neutralizing a broad diversity of circulating strains (so-called broadly neutralizing antibodies), that specifically target highly conserved epitopes in Env.

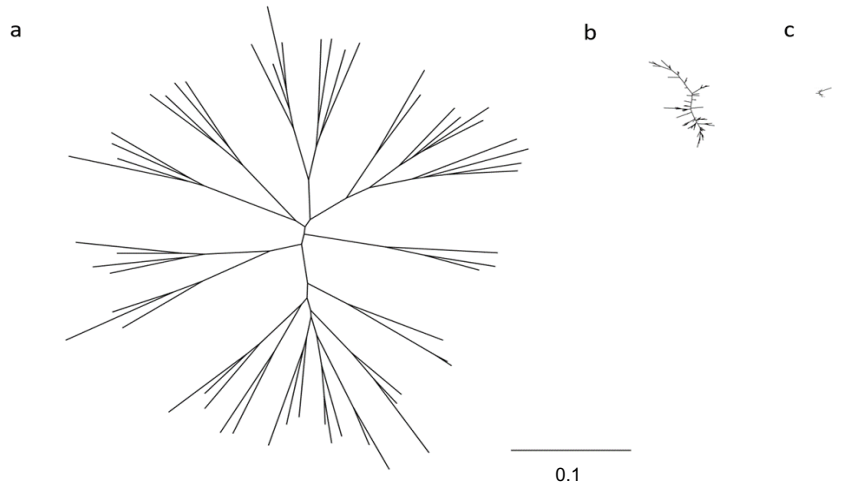


Figure 1.2 Scale of HIV-1 Env diversity compared to Influenza Haemagglutin and SARS-CoV-2 Spike.

Phylogenies from HIV clade M, Influenza HA from H3N2, and SARS-CoV-2 Spike from WHO Variants of Concern are plotted side-by-side on the same scale. **a** A phylogeny of 55 global HIV clade M subtype reference sequences (excluding CRFs) obtained from the LANL HIV sequence database. **b** A phylogeny of Influenza Haemagglutinin sequences from H3N2 sequences from humans, globally. Sequences were obtained from the NCBI Influenza Virus Database and 100 sequences were randomly selected. **c** A phylogeny of representative Spike sequences spanning all SARS-CoV-2 variants of concern as of 12 January 2022, obtained from GISAID.

HIV Env glycoprotein

The HIV Env mediates entry into target cells, and is sparsely arrayed on the surface of HIV virions (Zhu et al., 2003, 2006). The Env protein is expressed as a single polypeptide (gp160) that is proteolytically cleaved into gp120 and gp41 subunits (Figure 3a) within the Golgi by cellular furin or furin-like proteases, and this proteolytic processing is required for infectivity (Hallenberger et al., 1992; McCune et al., 1988). The mature, proteolytically cleaved HIV Envelope glycoprotein is a trimeric structure made up of non-covalently associated heterodimers of gp41 and gp120 subunits (Figure 3b). The Env is heavily glycosylated, with approximately half of its mass attributable to glycans (Leonard et al., 1990). Approximately 90 N-linked glycans extend from each trimer representing an evolving “glycan shield” that obscures the underlying protein (Stewart-Jones et al., 2016; Wei et al., 2003).

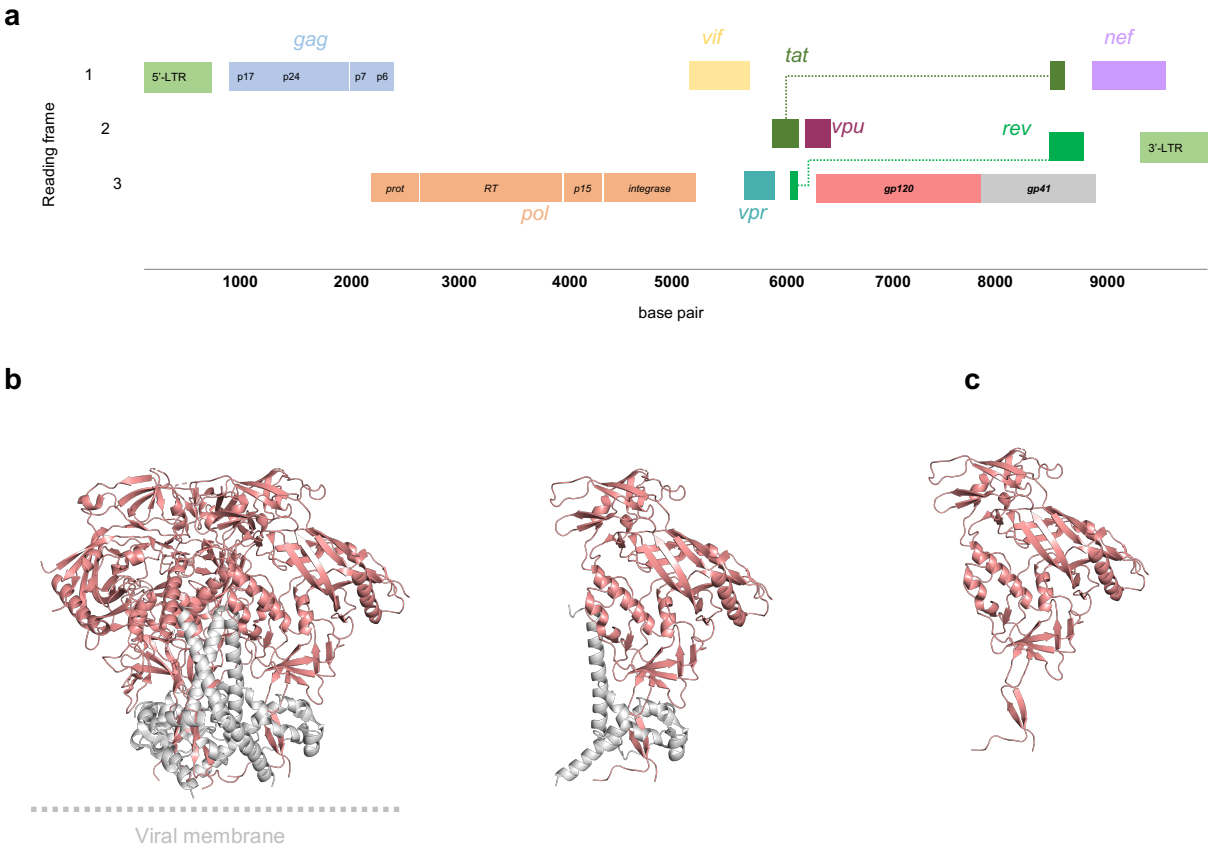


Figure 1.3. HIV Env overview.

a A schematic of the HIV genome organization and overlapping reading frames. **b** Side view of the structure of the HIV Env (PDB 5V8L) trimer, showing both trimeric (left) and monomeric (right) gp160. gp120 protomers are shown in pink, and gp41 in grey. **c** A single gp120 protomer as it appears in the trimeric, pre-fusion conformation.

Entry into target cells is initiated by the binding of Env to the cell surface receptor, CD4. Binding to CD4 drives conformational changes that facilitate exposure of the coreceptor binding site, dissociation of gp120, and enables membrane fusion mediated by gp41. As a result of the requirement for these conformational rearrangements to mediate fusion and entry into target cells, the HIV Env is an inherently metastable structure. A variety of Env isoforms can exist on the surface of virions, including conformationally "open" states (Munro et al., 2014; Ozorowski et al., 2017; Stadtmueller et al., 2018), and non-functional gp120/gp41 monomers or gp41-only "stumps" where gp120 has dissociated from the complex (Moore et al., 2006).

Challenges and successes in eliciting neutralizing antibodies by vaccination

While many HIV immunogens have been evaluated in humans, few candidate vaccines have progressed to efficacy trials. Early vaccine concepts typically incorporated soluble gp120 or gp140 as immunogens (Flynn et al., 2005; Pitisuttithum et al., 2006; Rerks-Ngarm et al., 2009), and elicited antibodies only capable of neutralizing culture-adapted and highly neutralization sensitive "tier-1" isolates but not primary, circulating "tier-2" viruses (Mascola et al., 1996; Wrin and Nunberg, 1994). The inability of these antibodies to neutralize primary isolates was resolved when the structure of the native, trimeric Env in the pre-fusion conformation was solved (Julien et al., 2013; Lee et al., 2016; Lyumkis et al., 2013; Pancera et al., 2014). This was initially achieved through the stabilization of the Env in a native-like, trimeric, pre-fusion conformation by the introduction of intermolecular disulphide bonds between gp120 and gp41, together with an Isoleucine to Proline mutation (I559P) to inhibit transition to the post-fusion state (Binley et al., 2000; Sanders et al., 2002, 2013). These renderings revealed fundamental differences in the structure of gp120 in its native, quaternary, pre-fusion state compared to in soluble, monomeric form. While the hypervariable loops protrude from monomeric gp120, they are closely associated with those from neighbouring protomers at the apex of the native trimer. Immunogenic epitopes targeted by antibodies only capable of neutralizing tier-1 isolates (Javaherian et al., 1989; Vancott et al., 1995) were revealed to be occluded in the native structure. Thus the failure of soluble gp120 and gp140 immunogens to recapitulate the native-like structure, with the resultant exposure of immunodominant, epitopes that are inaccessible in the native trimer, likely underpins their repeated failure to elicit antibodies capable of neutralizing primary HIV isolates.

In a landmark for HIV vaccine research, immunization of animals with these stabilized, native-like Env trimers based on the clade A BG505 (BG505.SOSIP) successfully elicited antibodies capable of neutralizing a primary HIV isolate (Sanders et al., 2015). BG505.SOSIP immunization also protected rhesus macaques from repeated intra-rectal challenge with a SHIV displaying the vaccine strain Env, and protection correlated with serum neutralizing

antibody titers (Pauthner et al., 2019). However, while immunization with stabilized Envs reproducibly elicited nAbs in several animal models, serological mapping revealed that they are typically focused on variable, immunodominant epitopes exposed by the absence of shielding glycans (Klasse et al., 2018; McCoy et al., 2016). As a result, elicited responses are highly specific for the vaccine strain and do not consistently cross-neutralize heterologous viruses.

Limitations of animal models

While animal models have been essential to HIV vaccine research, results in particular animal models may not necessarily translate seamlessly to humans. Neutralizing antibody titers elicited following BG505.SOSIP.664 trimer immunization appear to differ significantly between animal species, with lower titers in macaques than in rabbits or guinea pigs, and with undetectable titers following immunization of mice (Cheng et al., 2015; Hu et al., 2015; Sanders et al., 2015). While stabilized Env antigens have elicited bnAbs in cows and llamas, these antibodies utilized unique, species-specific antibody features that are not present in human antibody repertoires (McCoy et al., 2012; Sok et al., 2017). These observations highlight the importance of evaluating responses in humans or closely related species, where possible.

The antibody response following HIV infection

A robust antibody response develops in humans following HIV infection. The first antibody response to Env is detectable within 1-3 weeks, but is non-neutralizing (Moore et al., 1994; Tomaras et al., 2008). Neutralizing antibodies develop on average 12-29 weeks following infection, typically targeting one or a few immunodominant epitopes within Envelope (Gray et al., 2007; Li et al., 2006; Moore et al., 2008, 2009a; Richman et al., 2003; Rong et al., 2009; Wei et al., 2003). 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 are effective against the variant that elicited them, they are not able to neutralize any heterologous variants. Together with a highly error-prone reverse transcriptase (Ji and Loeb, 1992), short doubling time (Little et al., 1999), and high recombination rate (Neher and Leitner, 2010), HIV is capable of rapidly escaping selective pressures (Coffin, 1995). Indeed, intravenous infusion of a mAb rapidly suppressed viral load in infected individuals, but multiple independent escape mutations arose within the first weeks following infusion, and viral load rebounded within 2-3 weeks (Caskey et al., 2017). Similarly, the onset of neutralizing antibody responses in infected individuals drives the rapid and continuous selection of antibody-escape variants (Moore et al., 2008, 2009a; Richman et al., 2003; Wei et al., 2003).

While the early neutralizing antibody response is highly strain-specific, the majority of infected individuals develop some cross-neutralizing antibody responses over time (Hraber et al., 2014), demonstrating that the human immune response is indeed capable of producing these responses, though they typically take two to three years to develop in adults (Gray et al., 2011; Landais et al., 2016; Mikell et al., 2011). In rare individuals, extremely potent neutralizing antibody responses capable of neutralizing most circulating group M viruses can be found. Enabled by methodological advances, the isolation and characterization of individual broadly neutralizing monoclonal antibodies (bnAbs) (Bonsignori et al., 2011; Doria-Rose et al., 2014; Huang et al., 2012; Liao et al., 2013; Simonich et al., 2016; Walker et al., 2009, 2011; Wu et al., 2010, 2015) has breathed new life into vaccine development. Extensive efforts have been made in characterizing bnAbs, and their ontogeny in infected individuals from longitudinal cohorts, in the hopes that this can provide a blueprint for how vaccines could recapitulate these responses. This concept of 'reverse vaccinology', where desirable antibody responses are identified and used to tailor-make vaccine antigens, may hold the key to improved vaccines against HIV and other highly variable pathogens.

Targets and features of bnAbs

Characterizing the epitopes targeted by bnAbs has identified a number of conserved, accessible epitopes on the HIV Env that vaccines could target to mediate broad neutralization. These include the CD4-binding site, the trimer apex including V2, sites in V3 overlapping the co-receptor binding site, the membrane proximal external region (MPER), and the interface between gp41 and gp120 protomers including the fusion peptide (Figure 1.4).

Consistent with their delayed onset in infected individuals, bnAbs isolated to date harbour extremely high levels of somatic hypermutation away from their germline precursors, indicative of a long maturation process (Doria-Rose et al., 2014; Klein et al., 2013; Liao et al., 2013; Scheid et al., 2011; Sok et al., 2014; Walker et al., 2009; West et al., 2014; Wu et al., 2015; Xiao et al., 2009). Indeed, the isolation of multiple related mAbs over time together with sequencing of B-cell receptors (BCRs) revealed that these broadly neutralizing lineages developed over long periods (Doria-Rose et al., 2014; Liao et al., 2013; MacLeod et al., 2016; Wu et al., 2015). While IgG antibodies to influenza following vaccination, for example, are typically ~6% mutated (Moody et al., 2011), some HIV bnAbs are as much as 30% mutated in their VDJ region (West et al., 2014). Mutations in HIV bnAbs occur not just in the complementarity–determining regions (CDRs), but in the antibody framework regions as well (Klein et al., 2013), and include rare mutations outside of activation-induced cytidine deaminase (AID) hotspots (Bonsignori et al., 2017).

The majority of isolated bnAbs also possess other atypical features that are important determinants of their neutralization capacity. bnAbs to the V2 apex (and to some extent, those that target the V3-glycan) appear to require exceptionally long CDR H3 regions, generated by rare VDJ recombinations. bnAbs to the V2 apex, gp120/gp41 interface and V3-glycan sites also incorporate N-linked glycans into their epitope, suggesting that glycan recognition may be required. This may be challenging as N-linked glycans are viewed as "self" by the immune system, and can be tolerogenic. Similarly, bnAbs to the MPER are frequently auto- and poly-reactive and appear to require co-recognition of the plasma membrane (Liu et al., 2015; Yang et al., 2013) and the bypass of immune tolerance checkpoints (Doyle-Cooper et al., 2013; Verkoczy et al., 2010, 2011). The long development, extensive affinity maturation, and apparent requirement for atypical features represents a challenging process for any practical vaccine to recapitulate.

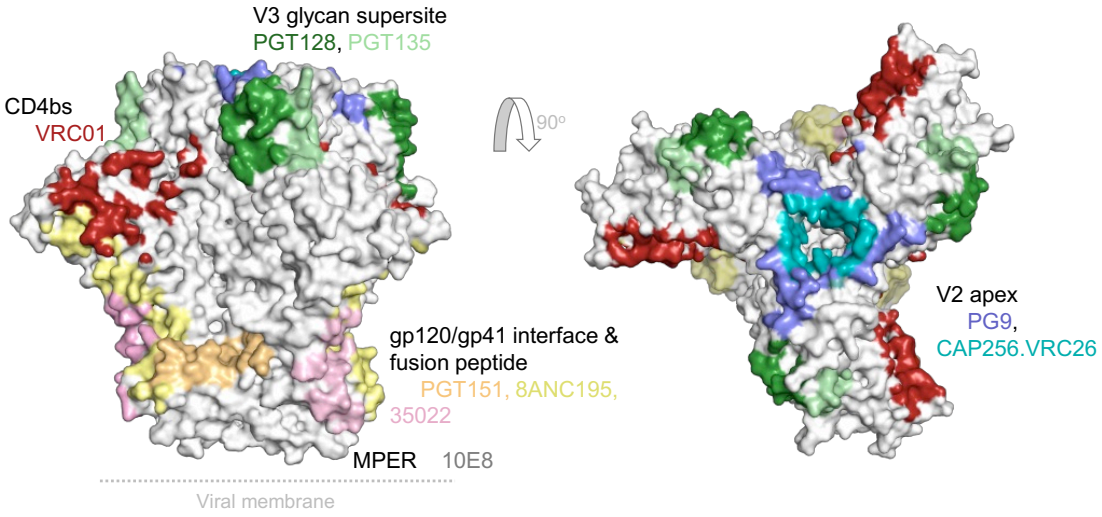


Figure 1.4 Targets of isolated broadly neutralizing antibodies on the HIV Env.

Potent bnAbs have been identified that target multiple sites on the HIV Env trimer, including the CD4 binding site (eg. VRC01), V2-apex (eg. PG9 and CAP256-VRC26), a V3 super-site incorporating N-linked glycans (eg. PGT135 and PGT128), epitopes spanning the gp41-gp120 interface and the fusion peptide (eg. PGT151, 8ANC195, 35O22) as well as the Membrane Proximal External Region (MPER) (eg. 10e8). Residues within 5Å of these monoclonal antibodies in complex are highlighted in colour on the HIV-1 Env structure (PDB 5V8L).

Factors that promote the development of bnAbs

Understanding how broadly neutralizing activity develops during infection and the factors that augment this process, could identify pathways that vaccines could exploit and will likely be required to elicit similar antibodies in an accelerated time frame. Studies of infected individuals have consistently identified viral load as a strong predictor of the development of neutralization breadth (Gray et al., 2011; Rusert et al., 2016; Sather et al., 2009), although bnAbs can be found in individuals with low or undetectable viremia (Freund et al., 2017; Sather et al., 2012), albeit at a lower frequency. Measures of viral diversity have also been associated with broader responses (Piantadosi et al., 2009; Smith et al., 2016), as have features of the infecting viral Envelope, such as shorter variable loops (Gnanakaran et al., 2010; van den Kerkhof et al., 2013; Rademeyer et al., 2007).

Infection by a second HIV strain after established primary infection (HIV superinfection) has previously been associated with broader antibody responses. In a cohort from Yaoundé, Cameroon, four dual infected individuals displayed significantly broader antibody responses against 5 primary isolates than did 23 controls (Powell et al., 2010). More recently, a comparison of 21 cases of superinfection to matched singly-infected controls found significantly higher breadth scores against a multi-subtype panel in the superinfected individuals (Cortez et al., 2012, Cortez et al., 2015).

Host immune factors have also been associated with the development of broader neutralizing antibody responses, including higher plasma IgG levels and Env-specific IgG binding titers (Landais et al., 2016), HIV-specific Fc effector function in early infection (Richardson et al., 2018), and the number and function of circulating T follicular helper (Tfh) cells (Cohen et al., 2014; Locci et al., 2013; Moody et al., 2016). Immune perturbations may also favour the development of breadth, as a higher frequency of autoantibodies and a lower frequency of regulatory CD4+ T cells were identified in individuals that developed bnAbs (Moody et al., 2016). As many HIV bnAbs are auto or polyreactive, and require glycan recognition, more permissive tolerance checkpoints may have allowed their development.

The dependency on time, high viral loads, and the quality of Tfh/CD4+ T cell help together suggests that chronic, diverse antigen exposure and multiple affinity maturation cycles in germinal centres favour the development of bnAbs. This is consistent with the exceptionally high levels of somatic hypermutation present in most bnAbs. Driving the re-entry of memory B cells into germinal centres with sequential immunizations may be necessary for the development of highly affinity matured lineages, and therefore characterizing recall responses upon sequential immunization represents an important objective.

Strategies to elicit broadly neutralizing antibodies by vaccination

As conventional approaches to vaccination have failed to reliably elicit protective titers of broadly neutralizing HIV antibodies, several next-generation vaccine strategies are being investigated. These include so-called “lineage-based” vaccines that aim to replicate the development of specific bnAbs seen in infected individuals. There is evidence that during infection, antibody pressure drives Env evolution and escape, which in turn shapes antibody evolution, in some cases towards breadth (Bonsignori et al., 2017; Doria-Rose et al., 2014; Moore et al., 2012; Wibmer et al., 2013). These approaches thus aim to design a series of tailored immunogens to activate B cells encoding B cell receptors (BCRs) with required features based on inferred germlines of known bnAbs (“germline-targeting”), and subsequently drive their affinity maturation along previously observed paths towards breadth.

Several studies in transgenic mice have demonstrated that germline-targeting immunogens may be feasible. In mice genetically engineered to express the inferred germlines of bnAbs, rationally designed antigens could activate these B cells, and subsequently drive their development (Dosenovic et al., 2015; Escolano et al., 2016; Jardine et al., 2015; McGuire et al., 2016; Saunders et al., 2019; Steichen et al., 2016, 2019; Tian et al., 2016). However, it is important to note that in these transgenic animal models, the majority of B cells express the knock-in BCRs, whereas these precursors would be at low frequency in the human repertoire (Briney et al., 2019). Furthermore, in most cases knocked-in BCRs incorporated the CDRH3 from the mature bnAb, and any immunogen that requires a specific CDRH3 sequences is unlikely to elicit reproducible responses as the region is not germline-encoded but rather generated upon VDJ recombination. Nevertheless, even in adoptive transfer experiments, where the number of precursor B cells transferred to wild-type mice was controlled to maintain low frequency in the B cell repertoire, these precursors could be activated by appropriate immunogens (Soto et al., 2019; Steichen et al., 2019). Furthermore, in mice transgenic for human immunoglobulin loci, germline-targeting immunogens aimed at eliciting CD4bs (VRC01-like) antibodies could activate B cells with the appropriate features in a subset of animals, providing a proof-of-concept for germline-targeted immunogens (Sok et al., 2016). These immunogens can also be used to sort naive B cells from uninfected individuals (Jardine et al., 2016), suggesting these priming immunogens may be able to recruit these BCRs with the potential for breadth, in humans. Encouragingly, a recent phase I clinical trial evaluating this germline-targeting immunogen in humans (IAVI G001; NCT03547245) reported detectable expansion of VRC01-like precursor B cells in 97% of recipients (Venkatesan, 2021).

Another framework to elicit bnAbs aims to engineer immunogens designed to direct the antibody response away from immunogenic non-neutralizing epitopes or strain-specific neutralizing epitopes, and onto sites targeted by bnAbs (immunofocussing). This encompasses approaches that include additional modifications to stabilized trimer immunogens to reduce the exposure of immunogenic epitopes in V3 and elsewhere (Kulp et al., 2017; Ringe et al., 2017; de Taeye et al., 2015; Torrents de la Peña et al., 2017), as well as the addition or removal of glycans to occlude or expose epitopes (Kulp et al., 2017; Voss et al., 2017; Zhou et al., 2017). A related approach is the design of immunogens that present known bnAbs in the correct conformation, but outside of the context of the Env trimer to avoid inclusion of immunodominant, strain-specific epitopes. The HIV MPER and fusion peptide represent attractive targets for this approach as the epitopes are predominantly contained within linear peptides. As the presentation of the MPER in the context of membrane is likely important, liposomes have been frequently employed to present MPER peptides. However, MPER liposomes have not yet been successful at eliciting nAbs, potentially due to the requirement for some degree of autoreactivity for MPER-specific neutralization, that may lead to the deletion of MPER targeting lineages (Zhang et al., 2016). The HIV fusion peptide (FP) conjugated to an immunogenic carrier (KLH) however, has shown promising results in animal models. Priming with FP-KLH followed by boosts with stabilized Env trimers elicited nAbs in animal models, and fusion peptide specific antibodies capable of neutralizing diverse strains from different clades were detectable (Cheng et al., 2019, 2020; Kong et al., 2019; Xu et al., 2018).

Polyvalent HIV vaccines

The development of stabilized, soluble trimers that mimic the native functional HIV trimer and elicit strain-specific neutralizing HIV antibodies in animal models represents a critical advance in HIV vaccine development. However, it appears a single native Env immunogen is unlikely to be sufficient to elicit broadly neutralizing HIV antibodies in humans. One approach to broaden responses includes the use of multiple stabilized Env trimers in polyvalent or heterologous boost regimens in the hopes that exposure to diverse Envs will favour the development and selection of cross-neutralizing B cells.

However, these approaches have thus far had little success in animal models. While polyvalent and repeated immunization of rabbits with two, three, or four stabilized HIV Env (SOSIP) trimers elicited nAbs to each of the Envs in many of the animals, it did not significantly broaden responses to Envs not included in the immunization (Klasse et al., 2016; de la Peña et al., 2018). Where cross-neutralizing antibodies were detected, they were generally of low titer, and in only a few animals. More recently, stabilized trimers were engineered from five

different Envs from multiple clades (Dubrovskaya et al., 2019). These Env cocktails were also arrayed in mixtures at high density on liposomal particles, with glycans shielding the CD4bs removed from the priming immunogens to encourage the targeting of the CD4bs. Immunization of rabbits with these cocktails elicited some serum cross neutralizing activity, although in only 2/12 rabbits and only after 5 or 6 immunizations. It's not yet clear how different immunogens included in cocktails should be, nor how they should be administered to favour the development of cross-neutralizing responses over immunodominant, strain-specific ones.

Rationale and overarching aim

Although HIV infection causes a variety of immune perturbations that would not be present in the vaccine setting (Moody et al., 2016), characterizing the development of neutralizing antibody responses to HIV in infected humans could identify mechanisms that vaccines could leverage to broaden elicited antibody responses. Indeed, a common theme in almost all rational vaccine design approaches for HIV, described above, is the application of insights drawn from studies of the antibody responses in infected individuals. In this thesis we investigated HIV co- and super-infected individuals to gain important insights into immunodominance, memory recruitment, and broadening of antibody responses that could facilitate the design of polyvalent and sequential vaccine regimens. Intriguingly, HIV superinfection has been associated with broader neutralizing antibody responses in some longitudinal cohorts (Cortez et al., 2012; Powell et al., 2010), suggesting that they may hold insights into one pathway to breadth. Importantly, multivalent or sequential exposure and heterologous immunogens represent pathways that are readily translatable to vaccine regimens.

Through a detailed, longitudinal analysis of the neutralizing antibody responses in HIV-1 co- and super-infected individuals, the overarching aim of this thesis is to identify the impact that exposure to two or more divergent HIV variants will have on the specificity and breadth of the neutralizing antibody response.

CHAPTER 2

HIV superinfection fails to drive neutralization breadth

Daniel J. Sheward[1], Jinny Marais[1], Valerie Bekker[2], Ben Murrell[3], Kemal Eren[4], Jinal N. Bhiman[2,5], Molati Nonyane[2], Nigel Garrett[6], Zenda L. Woodman[7], Quarraisha Abdool Karim[6,8], Salim S. Abdool Karim[6,8], Lynn Morris[2,5,6], Penny L. Moore[2,5,6], and Carolyn Williamson[1,6,9]

[1] Division of Medical Virology, Department of Pathology, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa;

[2] Center for HIV and STI, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa;

[3]Department of Medicine, University of California San Diego, San Diego, California, USA

[4]Bioinformatics and Systems Biology, University of California San Diego, San Diego, California, USA

[5] University of Witwatersrand, Johannesburg, South Africa;

[6] Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu- Natal, Durban, South Africa;

[7] Department of Molecular and Cell Biology, University of Cape Town, Cape Town, South Africa;

[8] Department of Epidemiology, Columbia University, New York, USA;

[9] National Health Laboratory Services of South Africa.

I confirm that I have been granted permission by the University of Cape Town's Doctoral Degrees Board to include the following publication in my thesis:

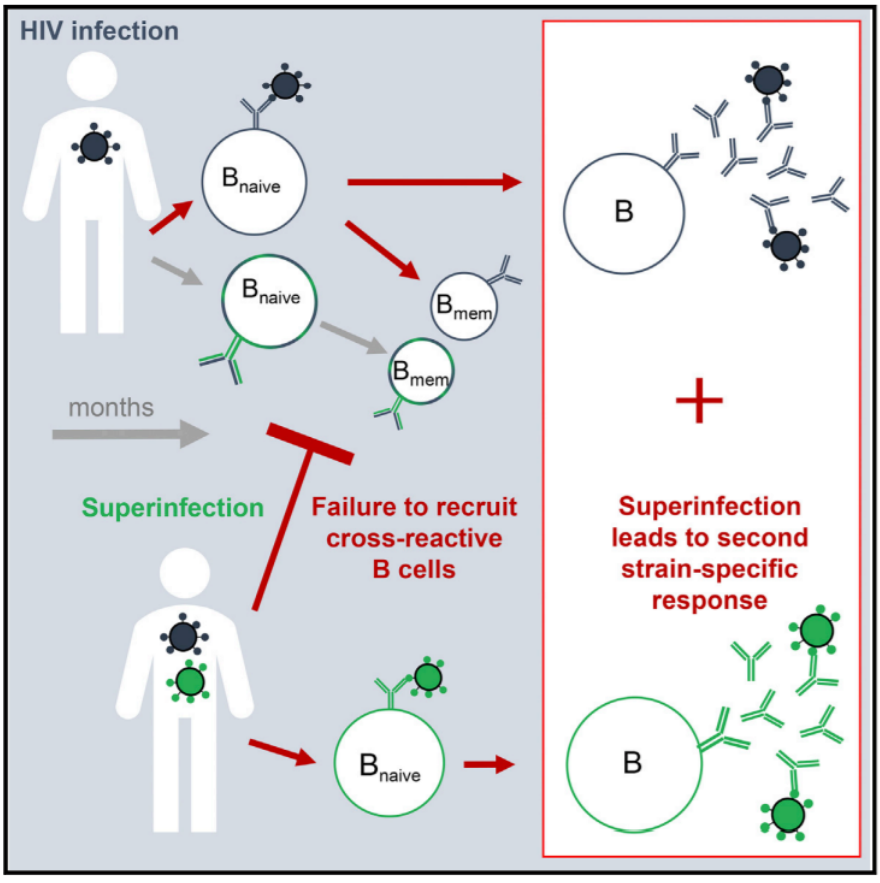
Sheward et al., (2018). *Cell Host and Microbe*. Oct 10;24(4):593-599.e3. doi: 10.1016/j.chom.2018.09.001

Author contributions (CRediT taxonomy)

	DJS	JM	VB	BM	KE	JNB	MN	NG	ZLW	QAK	SSAK	LM	PLM	CW
Conceptualization	■												■	■
Methodology	■												■	■
Software	■		■	■										
Formal Analysis	■													
Investigation	■	■	■	■		■	■							
Resources								■		■	■	■	■	■
Writing - original draft	■													
Writing - review and editing	■												■	■
Visualization	■													
Supervision									■				■	■
Funding acquisition													■	■

Abstract

Eliciting antibodies capable of neutralizing a broad range of circulating HIV strains (bnAbs) represents a key priority for vaccine development. HIV superinfection (re-infection with a second strain following an established HIV infection) has been associated with increased neutralization breadth, and can provide insights into how the human immune system responds to sequential exposure to two distinct HIV Envelope glycoprotein (Env) antigens. Characterizing the kinetics and specificity of neutralizing antibody (nAb) responses in four superinfected women revealed that superinfection did not boost memory nAb responses primed by the first infection, nor promote nAb responses to epitopes conserved in both infecting viruses. While one superinfected individual developed extremely potent bnAbs, superinfection was likely not the driver of breadth. In conclusion, sequential exposure to two distinct Envs led to neutralizing responses specific to each Env, but did not promote the development of bnAbs. These data have important implications, suggesting that sequential immunization with heterologous Envs may not be sufficient to focus the immune response onto conserved epitopes on the HIV Env.



Introduction

Eliciting broadly neutralizing antibodies (bnAbs) through immunization remains a primary goal for HIV prevention. While antibodies capable of neutralizing primary HIV isolates have recently been elicited in animal models by vaccination with stabilized HIV Env trimers, the responses were narrow, typically only neutralizing viruses that matched the immunizing Env(s) (Crooks et al., 2015; Sanders et al., 2015; Saunders et al., 2017). Due to the tremendous diversity of HIV Env proteins, protective vaccines will need to direct neutralizing antibody (nAb) responses onto more conserved regions on the HIV Env in order to achieve broad cross-neutralization. BnAbs do develop in a subset of HIV infected individuals, and are typically attributable to one or a few antibody specificities targeting conserved epitopes on the HIV Env (Walker et al., 2010). Defining virological and immunological events and pathways that focus responses to conserved epitopes could identify mechanisms that vaccines could leverage.

Materials and Methods

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Subjects

Samples were provided from participants from the CAPRISA 002 Acute Infection Study that characterized the natural history of HIV-1 subtype C infection. Briefly, this cohort recruited recently HIV infected women from high risk, HIV negative women from Durban and Vulindlela, KwaZulu-Natal, South Africa between August 2004 and May 2005 (van Loggerenberg et al., 2008) as well as from the CAPRISA 004 Tenofovir microbicide study (Abdool Karim et al., 2010). Through the screening 775 individuals, a cohort of 245 uninfected high-risk women was established, the majority (78%) of whom self-identified as sex-workers (van Loggerenberg et al., 2008). The acute infection cohort enrolled 62 HIV-infected women, of whom 3 were subsequently lost to follow up. 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. 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 in this study were antiretroviral therapy (ART) naïve and were initiated on ART consistent with the prevailing South African ART guidelines. Ethical approval for this study was received from the ethics committees of the University of Cape Town (025/2004), the University of KwaZulu-Natal (E013/04), and the

University of the Witwatersrand (MM040202), and all participants provided written, informed consent.

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). 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^{o}C in a humidified incubator with 5% CO_2, and monolayers were disrupted at confluence with Trypsin-EDTA.

METHOD DETAILS

Single Genome Sequencing

Plasma viral RNA was extracted from 200 µl or 400 µl of plasma using either the Roche MagNApure (Roche Applied Science, Mannheim, Germany) or QIAamp viral RNA mini kits (Qiagen, Valencia, CA). RNA was reversed-transcribed to cDNA using Superscript III (Invitrogen, Life Technologies, Carlsbad, CA) as per the manufacturer's instructions. *Env* cassettes were amplified from the 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 (Abrahams et al., 2009; Keele et al., 2008). Primers used in the outer reaction were 5'- GGGTTTATTACAGGGACAGCAGAG -3' (HXB2 nt 4900 - 4923) and 5'-GCACTCAAGGCAAGCTTTATTGAGGCTTA -3' (HXB2 nt 9604 - 9632). Inner primers used were 5'- CACC GGCTTAGGCATCTCCTATAGCAGGAAGAA-3' (HXB2 nt 5954 - 5982) and 5'- TTGCCAATCAAGGAAGTAGCCTTGTGT -3' (HXB2 nt 9145-9171). Outer reaction thermal cycling conditions were as follows: initial denaturation at 94^{o}C for 2 minutes, followed by 35 cycles of [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. Amplicons were directly sequenced using an ABI3000 genetic analyser and BigDye terminator reagents (Applied Biosystems, Foster City, CA) using twelve primers, by the Central Analytic Facility at the University of Stellenbosch, South Africa. 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.

Molecular Cloning

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) as per the manufacturer's instructions, and cultured on Luria-Bertani agar supplemented with 100 µg/ml Carbenicillin (Sigma-Aldrich, St Louis, MO). 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 as described above in order to ensure the clone contained no non-synonymous mutations relative to the SGA-derived *env* sequence.

Neutralization assays

Env-pseudotyped viruses were generated by co-transfecting *env* plasmids with pSG3ΔEnv at a 1:2 ratio into HEK293T cells using Fugene 6 (Applied Science, Indianapolis, IA) or PolyFect (Qiagen, Valencia, CA) per 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 (Gray et al., 2007). Assays were performed in duplicate wells, and repeated at least twice using aliquots of the same plasma draw. Neutralization breadth for the cohort was previously characterized, and quantified as the proportion of heterologous viruses from a multi-subtype pseudovirus panel of 18 viruses each plasma sample was able to neutralize at ID_{50} titers >45.

V3 ELISA

Six 33-mer V3 peptides, representative of the V3s from CAPRISA participants CAP88, CAP45, CAP239, CAP63, CAP206 and CAP84 were used to estimate V3 antibody titers. Peptides were coated onto high-binding 96-well enzyme-linked immunosorbent assay plates at a concentration of 2.5 µg/ml in sodium bicarbonate buffer (pH 8.5) overnight at 4°C. Unbound peptide was removed by washing four times with phosphate-buffered saline containing 0.3% Tween 20 (wash solution), and plates were blocked for 1 h at room temperature with 200 µl of phosphate-buffered saline, 0.3% Tween 20, and 5% nonfat milk. Serum samples were diluted 1:500 in block solution, and 100 µl per well was added, followed by incubation at room temperature for 1 h. Plates were washed four times with wash solution before the addition of 100 µl of secondary antibody (horseradish peroxidase-labelled goat anti-human (Fc-specific) antibody (Sigma-Aldrich, St. Louis, MO) diluted 1:1,000 in blocking solution), and incubated for 1 h at 37°C. After four washes with wash solution, bound antibody was detected using tetramethylbenzidine (TMB) substrate and stopped by the addition of 25 µl of 1 M sulfuric acid.

Mutagenesis

Point mutants were generated using the Quikchange Lightning site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) as per the manufacturer's instructions.

Recombination inference in CAP256

A Hidden Markov Model (HMM) was used to assign sites to parent viruses, similar to the "BURT" approach used in RDP4 (Martin et al., 2015). The HMM takes as input two parent sequences (*A* and *B*), and one target sequence, and outputs a probabilistic assignment of each nucleotide site in the target sequence to a parent. To achieve this, a hidden state is associated with each parent. Where the model is in state "A", a site in the target sequence has a high probability of matching parent *A*, and a lower probability of mismatching, and similarly for state "B". This framework naturally accounts for the possibility that a base that originated from one parent could stochastically mutate to now match the other parent. These match and mismatch probability parameters, as well as a parent switching-rate parameter, are fit by approximate maximum likelihood, using Viterbi Training (Rabiner et al., 1989), and the empirical Bayes posterior assignments of sites to parents is calculated using the Forward-Backward algorithm. This was applied to an alignment of V1V2 sequences from CAP256 (BioProject accession number PRJNA294363) (Bhiman et al., 2015), and the posterior probabilities were visualized in Mathematica 10 (Wolfram Research, Champaign, IL).

Quantification and statistical analysis

The relationship between autologous potency and heterologous breadth was assessed using linear regression of log10 transformed peak titers, implemented in Prism 5 (Graphpad Software, San Diego, CA).

Code availability

The custom HMM was implemented in Python, and the code is available at https://github.com/MurrellGroup/ with no restrictions to access.

Results

Infection by a second HIV strain after established primary infection (HIV superinfection) has been associated with broader antibody responses (Cortez et al., 2012; Powell et al., 2010). Such events are analogous to heterologous prime-boost immunizations, offering an opportunity to assess how the human immune system responds to sequential exposure to two distinct HIV Env antigens.

Previously, longitudinal gag (n = 365), nef (n = 289), and/or env (n = 2011) single-genome sequences from 32 participants in the CAPRISA 002 cohort spanning a total of 54.9 person-years of follow-up sequences were screened retrospectively for superinfection as described (Sheward et al., 2015). Using single genome sequencing and strain-specific PCR, a phylogenetically unlinked virus could be detected after seroconversion that was not detectable at or before seroconversion in three participants (CAP237, CAP256, and CAP281). Similarly, seroconverters from the CAPRISA 004 microbicide trial were screened for superinfection using 454 high-throughput sequencing, as described (Redd et al., 2014a), identifying two cases of superinfection.

All together, in 108 women recruited in acute/early infection and screened for superinfection over approximately two years (CAPRISA 002 cohort), we have identified five superinfected participants. All cases represented intra-subtype superinfection with two subtype C viruses. All five cases were superinfected between 3-10 months following primary infection (Supplementary Figure. 2.1), and two developed antibodies capable of neutralizing heterologous viruses at two years post infection (Figure. 2.1a). Although we have previously detailed the development of extremely potent, bnAbs in one of these superinfected participants, CAP256 (Bhiman et al., 2015; Doria-Rose et al., 2014), the contribution of superinfection itself to the development of breadth was not clear.

Studies of nAb responses following superinfection are critically dependent on the identification of the superinfecting Env prior to recombination with the primary infecting virus. Frequent follow-up of participants following diagnosis of infection (every two to four weeks) enabled us to identify and clone the transmitted/founders of superinfection from four donors (Sheward et al., 2015). The Env genetic distance between the primary infecting (PI) and superinfecting (SI) viruses ranged from 11.9% to 15.2%, consistent with unlinked viruses (Supplementary Figure 2.1). Despite this intensive sampling, we were unable to obtain the SI prior to recombination with the primary virus in the fifth donor (CAP334), highlighting the challenges in accurately deciphering immune responses after superinfection.

Taking advantage of these Env clones, we evaluated whether superinfection promoted responses that cross-neutralized both infecting viruses, and whether superinfection may have boosted memory B cell responses primed by an initial infection. Secondary, memory responses would be expected to (i) be of higher titer than primary responses, (ii) arise more rapidly than primary responses, and (iii) neutralize both the priming and boosting antigens (Ahmed and Gray, 1996).

To determine if the nAb responses to superinfection were elevated, we compared the peak nAb titers against all four SI viruses to their matched primary infecting virus, as well as to the nAb titers that arose against early/founder viruses in 22 other participants in the CAPRISA 002 cohort. We found that titers to the CAP237 and CAP377 SI viruses were comparable to those seen in single infections from the rest of the cohort (Figure 2.1b). In contrast, CAP256 and CAP281 developed exceedingly high neutralizing titers against their SI viruses, with ID_{50} titers exceeding 38,000 and 20,000 respectively**,** that were 24- and 12-fold higher than the cohort median (Figure 2.1b). In CAP281, these high titers were produced despite maintaining a low or undetectable viral load throughout the first years of infection. However, only CAP256 went on to develop bnAbs, and among 14 confirmed singly-infected participants where neutralization data over three years was available, we found no significant association between potency of the autologous nAb response and the later development of breadth (Supplementary Figure 2.2). This indicates that eliciting potent responses alone may not be sufficient to promote breadth.

To assess whether the onset of nAb responses to superinfection was accelerated, as would be expected for memory responses, we compared the time from primary infection and superinfection until the first detection of nAbs. The nAb responses to SI viruses reached detectable levels a median of 9-17.5 weeks following superinfection (Figure 2.1c), comparable with the time to first detection of nAbs to the primary infecting viruses in each of the four superinfected individuals. This was also comparable with the timing of nAb responses following primary HIV infection in 22 other participants, which were detectable by a median of 9-16 weeks post infection (Figure 2.1c). As the onset of nAbs to the SI viruses were not accelerated, this suggested that they likely represented *de novo* responses, rather than secondary, memory responses.

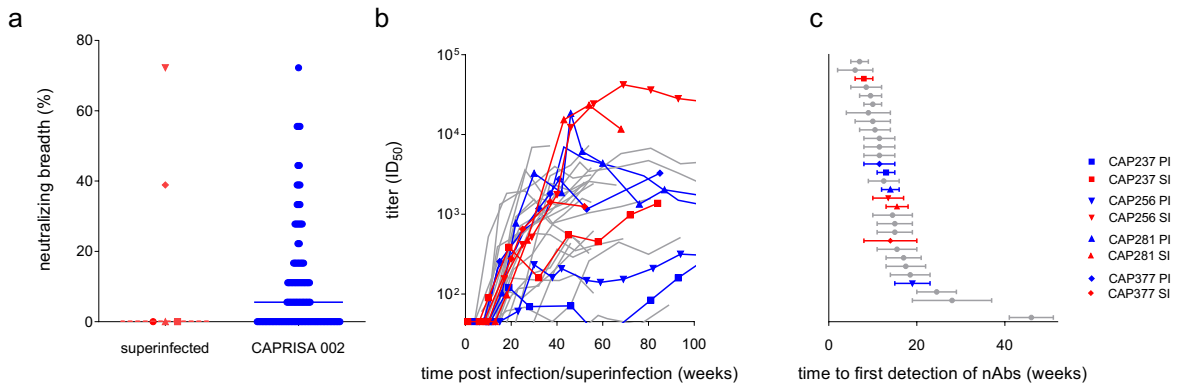


Figure 2.1 Potent neutralizing antibody responses arose to superinfecting viruses with a similar time-to-detection as primary HIV responses.

a Comparison of neutralization breadth (% of heterologous viruses neutralized) present in plasma sampled 2 years post infection between superinfected participants (n=5) and remaining CAP002 cohort participants (n=119). Antibody breadth was compared at two years post infection as all superinfected participants had at least two years of ARV-naïve follow up. Furthermore, if cross-neutralizing antibodies do not develop by 2-3 years post infection, they are unlikely to do so subsequently (Gray et al., 2011; Landais et al., 2016; Mikell et al., 2011). Cross neutralization data at two years was available for 120 anti-retroviral therapy-naïve participants. **b** Autologous neutralizing antibody titers, over time, to superinfecting Envs (red), primary infecting Envs in participants later superinfected (blue), and to early/founder Envs from other CAPRISA participants (grey) (n=22). **c** Estimated time, in weeks, from transmission until the first detection of these neutralizing antibody responses. The error bars represent the range given the 95% confidence interval for the timing of infection/superinfection.

To determine whether the nAb responses that arose following superinfection were specific for epitopes present in both infecting viruses, we analysed the dynamics of the nAb responses to PI and SI viruses (Figure 2.2a-d). In two donors (CAP281 and CAP237) superinfection was associated with a spike in titers to the PI. For both donors, there was also a simultaneous boost in V3 antibodies as measured by ELISA (Figure 2.2e, f) and a transient increase in viral load at this time (Figure 2.2g, h), suggesting that the spikes in titer may have been the result of a generalized activation. However, while titers to the PI increased following superinfection in these donors, they did not cross-neutralize the SI viruses, with nAb titers to the SI viruses only emerging weeks later. Furthermore, when titers to the SI viruses arose, titers to the PI viruses fell (Figure 2.2a, b). Thus, the dynamics of neutralization in both CAP281 and CAP237 is inconsistent with recruitment of memory responses that cross-neutralized both infecting viruses.

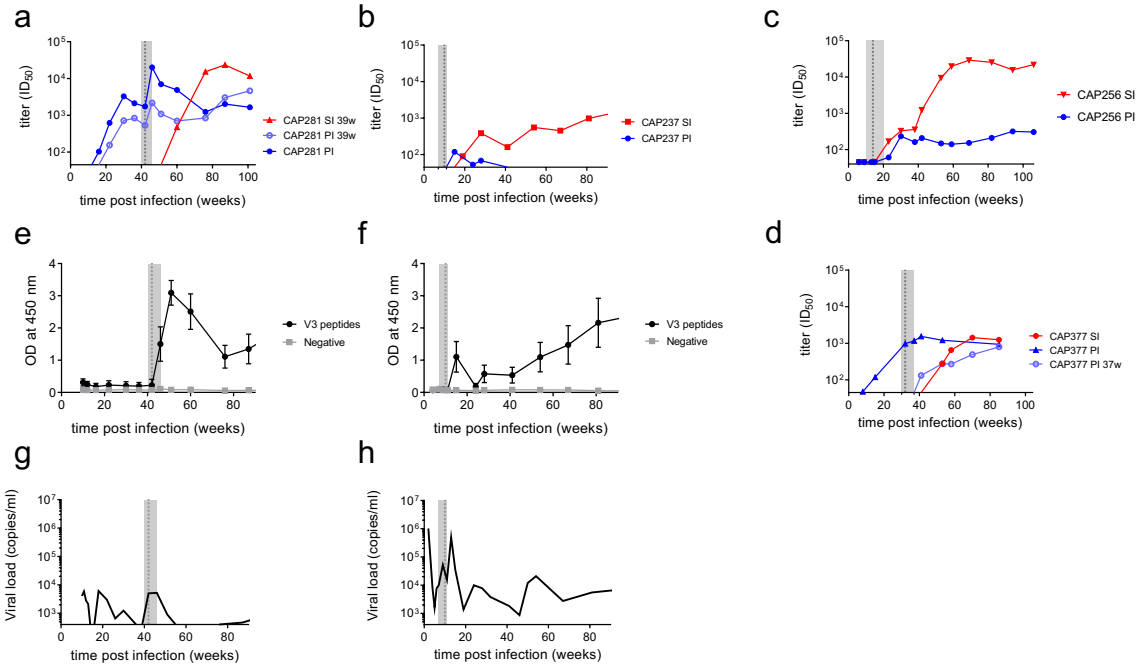


Figure 2.2 Neutralizing antibody responses to the superinfecting Envs did not cross-neutralize the primary infecting Envs.

a-d Neutralizing antibody titers, over time, to the primary infecting (PI, blue) and superinfecting (SI, red) Envs from (**a**) CAP281, (**b**) CAP237, (**c**) CAP256, and (**d**) CAP377. Titers to Env clones representative of the primary infecting variant circulating near the time of superinfection (CAP281 PI 39w in (**a**) and CAP377 PI 37w in (**d**), are shown in light blue and with open symbols. (**e,f**) Boost in antibodies specific for V3 peptides following superinfection in donors (**e**) CAP281 and (**f**) CAP237. Longitudinal V3 antibody titers were determined by ELISA, using six independent clade C V3 peptides. A peptide found in Ebola virus was used as a negative control. Plotted is the mean and standard deviation. **g,h** Spike in viral load following superinfection. Longitudinal viral load (in copies/ml) is depicted for (**g**) CAP237 and (**h**) CAP281. The estimated time of superinfection is shown in each figure by the vertical dotted line with the confidence interval shaded.

In contrast to CAP281 and CAP237, in donors CAP256 and CAP377 nAb titers to the SI viruses and contemporaneous primary viruses arose together (Figure 2.2c, d). This overlapping neutralization profile is suggestive of a response that cross-neutralized both viruses. Furthermore, CAP377 and CAP256 were also the only superinfected individuals that developed neutralization breadth, raising the possibility that superinfection drove breadth by promoting responses to epitopes conserved in both infecting viruses. We therefore mapped the targets of the nAb response following superinfection. As cross-neutralizing activity that developed later in CAP377 was attributable to V2-directed antibodies (Figure 2.3a), we evaluated whether neutralization of both the PI and SI viruses also targeted V2. bnAbs specific for V2 are most commonly highly dependent on the N-linked glycan at position 160 and a lysine at position 169 (Andrabi et al., 2015; Gorman et al., 2016). The introduction of a 169E

mutation into the CAP377 superinfecting virus abrogated neutralization by CAP377 plasma (Figure 2.3b), indicating that the SI virus was likely neutralized by V2-directed antibodies. However, the same mutation in the PI virus had no effect (Figure 2.3c), indicating that neutralization of the PI was mediated by a different neutralizing response to that targeting the SI virus (and mediating breadth). Indeed, the high divergence between the V2s (Figure 2.3d) provides further evidence that the nAb response did not target an epitope conserved in both viruses. Similarly, the isolation of mAbs from CAP256 previously revealed that the precursor of the CAP256-VRC26 antibody lineage (that accounts for the breadth in this donor) neutralized the SI but not the PI virus (Doria-Rose et al., 2014). Taken together, these data indicate that the nAb responses to superinfection did not represent the boosting of responses primed by initial infection, and did not target epitopes present in the PI viruses.

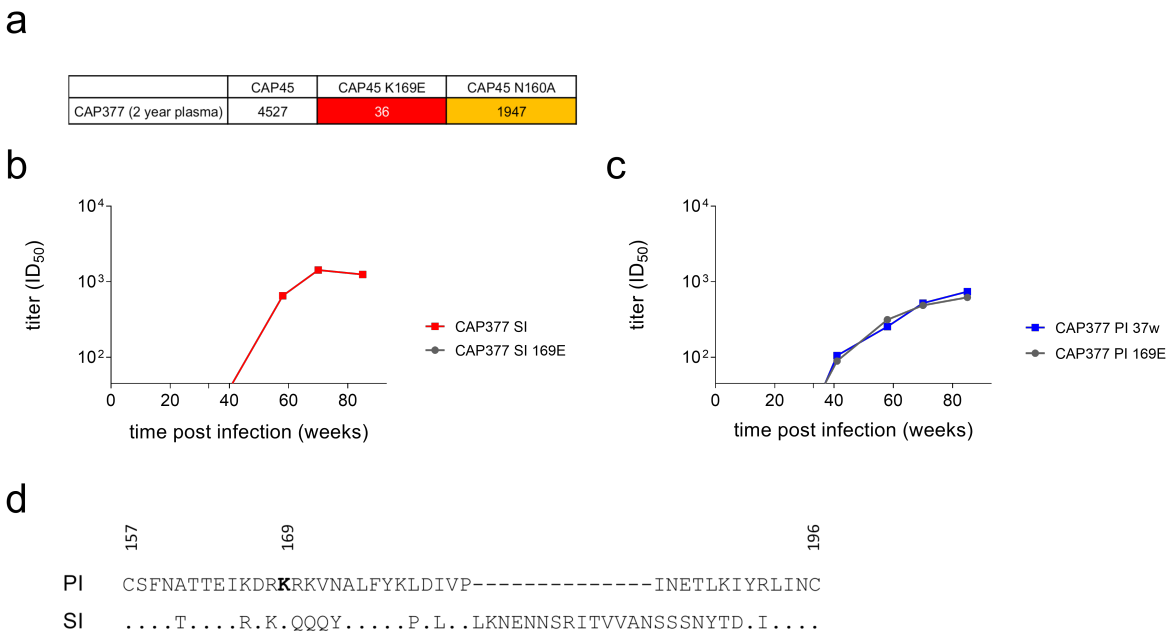


Figure 2.3 A single antibody response does not account for neutralization of both the primary infecting and superinfecting viruses in donor CAP377.

a Introduction of a K169E mutation abrogated neutralization of CAP45 by CAP377 indicating that heterologous neutralization was attributable to a K169-dependent antibody response. Tabled are the ID_{50} titers for each Env tested against a 2 year plasma sample from CAP377. **b, c** Introduction of a K169E mutation knocks out neutralization of the superinfecting virus (**b**), but not the primary infecting virus (**c**) by CAP377 plasma. **d** High divergence between the CAP377 primary infecting and superinfecting V2 sequences. Numbers represent reference HXB2 numbering.

As superinfection in CAP256 did not drive breadth by promoting multiple responses (breadth was attributable to a single antibody lineage), or by promoting a response to an epitope

conserved in both infecting viruses, we investigated whether diversity introduced through superinfection, via recombination, may have facilitated the development of bnAbs in CAP256. We showed previously that acquiring the ability to tolerate diversity at V2 residues 166 and 169, specifically 166K, and 169I/T/Q, was associated with the development of breadth in CAP256 (Bhiman et al., 2015; Doria-Rose et al., 2014). We therefore applied a Hidden Markov Model (see methods) to identify whether this diversity was inherited from the primary infecting virus via recombination, or arose independently. We show that 169T and 169I predominantly arose in the SI V2 (Figure 2.4), and while 166K and 169Q were inherited from the PI virus (Supplementary Figure 2.3), these mutations also arose independently in the SI lineage (Figure 2.4). These data indicate that multiple infection was not necessary to introduce the diversity that promoted the development of bnAbs in CAP256.

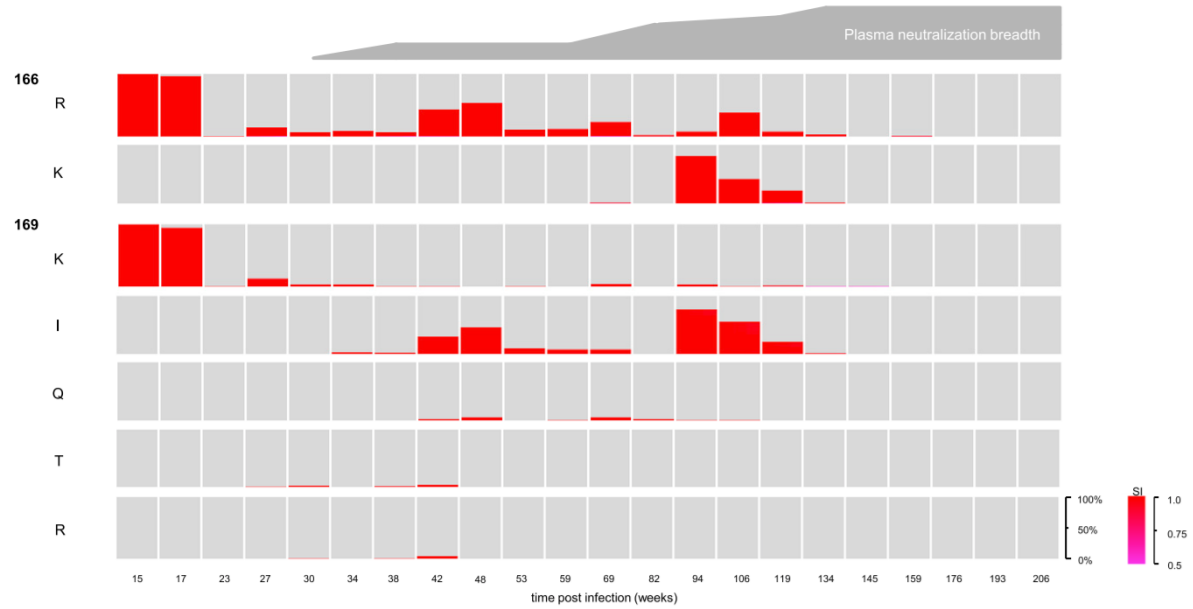


Figure 2.4 Diversity associated with broadening of the antibody response in CAP256 arose in the superinfecting virus.

Key genotypes and their frequencies in the superinfecting lineage are shown over time. Only residues that were assigned to the superinfecting lineage with posterior probabilities >0.5 are included, where more confident assignments are redder. Overlaid is a schematic of the development of breadth over time.

Discussion

Immunization with multiple Env immunogens has been proposed as a strategy to broaden neutralizing responses. Our identification of the viruses that superinfected CAP256 and three additional superinfected participants in the extremely well characterised CAPRISA 002 cohort has allowed us to define the neutralizing antibody response to HIV superinfection - a model for heterologous prime-boost immunizations. This has also enabled us to systematically investigate mechanisms whereby superinfection could enhance neutralization breadth: by promoting responses to an epitope conserved in both viruses; by stimulating multiple antibody responses; or by accelerating diversification within epitopes.

A broad diversity of memory B cells are generated in HIV infection (Scheid et al., 2009) that could be activated upon re-infection. However, we show that HIV superinfection likely did not boost memory responses primed by the initial infection. While two of the four superinfected donors developed unusually high autologous titers to the SI virus, none of the nAb responses to SI viruses arose more rapidly than primary HIV responses. Rather, *de novo* nAb responses arose following superinfection. Furthermore, superinfection did not elicit responses to epitopes conserved in both infecting viruses. Overall, infection with two variants resulted in additive responses, although in one case there was a pronounced decline in the titers to the PI virus following superinfection. These findings have important implications for HIV immunizations, including the potential of heterologous prime-boost strategies to activate memory B cell populations. It is important here to differentiate such prime-boost immunizations with unrelated Envs (analogous to superinfection) from sequential, lineage-based immunizations. The latter approach seeks to utilize highly related immunogens that recapitulate viral evolution known to shape antibody lineages towards breadth (Bhiman et al., 2015; Doria-Rose et al., 2014; Gorman et al., 2016; Landais et al., 2017; Liao et al., 2013; MacLeod et al., 2016; Medina-Ramírez et al., 2017).

While one superinfected donor (CAP256) developed extremely potent bnAbs, we find no evidence that superinfection itself promoted their development. Breadth in CAP256 was attributable to a single antibody lineage (Doria-Rose et al., 2014), indicating that superinfection did not promote breadth by expanding the number of distinct responses. Early antibodies in this lineage were also specific for the SI virus and not the PI virus indicating that superinfection did not facilitate breadth by directing responses to an Env conserved in both infecting viruses. This is also consistent with observations from the recent isolation of mAbs from another superinfected individual that developed bnAbs - mAbs isolated from this individual either neutralized PI viruses or SI viruses, with none cross-neutralizing both (Williams et al., 2018).

In this individual, broadly neutralizing activity also appeared to be largely attributable to the mAb lineage that arose to the SI virus, with the antibodies that arose to the PI virus only making a minor contribution. While acquiring the ability to tolerate viral diversity that arose at key residues led to the CAP256.VRC26 antibody lineage developing broadly neutralizing activity, infection with two strains was not necessary to introduce this diversity. As a result, we find no evidence that superinfection itself was necessary for the development of breadth in CAP256.

Previous studies have reported that superinfection broadens the nAb response (Cortez et al., 2012; Powell et al., 2010), though others saw no statistically significant effect (Cornelissen et al., 2016). Where a significant broadening was observed in previous studies, the effect size was small with superinfected individuals approximately 1.5 times more likely to develop breadth than singly infected individuals (Cortez et al., 2012). This effect size would be consistent with a near-additive effect of superinfection on nAb breadth, such as that we describe here. It is crucial to differentiate an additive effect (i.e. stimulation of two strain-specific antibody responses, each with the potential to develop breadth) from one that is more likely to be synergistic (eg. the promotion of a cross-neutralizing antibody response to an epitope conserved on both viruses). Stimulating independent antibody responses to each of two immunizing immunogens has restricted value in the context of a goal to elicit bnAbs to HIV by vaccination.

As HIV superinfection is a rare event, we (and others) have been limited by small sample sizes. Large-scale screening of multiple longitudinal cohorts for cases of superinfection together with more sophisticated screening approaches, will be required to establish larger sample sizes, and could provide extremely valuable insights for sequential immunizations. Furthermore, sequencing of HIV-specific B cells before and after superinfection in the future could identify whether superinfection reactivates any minor memory B cell populations that would not be not evident in plasma studies, or whether superinfection leads exclusively to *de novo* responses.

It is also important to consider the on-going HIV replication and immune dysfunction present in HIV-infected individuals that would not be present in the vaccine setting. While memory B cell dysfunction is known to arise with HIV pathogenesis, all participants described here were superinfected within the first year of infection and had relatively high CD4 counts (>450 cells/μl) at the time of superinfection. Host factors, including autoimmunity, may also contribute to the development of broadly neutralizing HIV antibodies (Borrow and Moody,2017; Dugast et al., 2017; Havenar-Daughton et al., 2016; Landais et al., 2016; Moody et al., 2016) and these were not assessed here. Despite these caveats, results from vaccine studies have thus far been consistent with our observations. Polyvalent and sequential immunization of

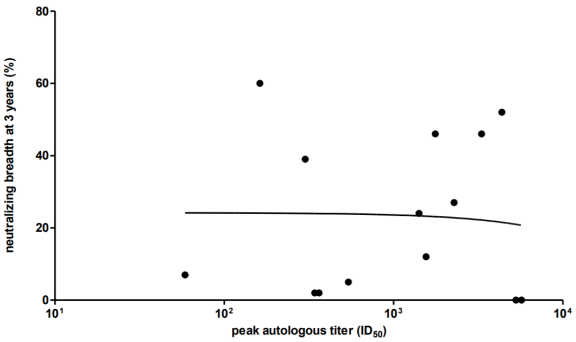
rabbits with two, three or four stabilized HIV Env trimers elicited nAbs to each of the Envs in many of the animals, but no significant broadening of the responses to Envs not included in the immunization was evident (Klasse et al., 2016; de la Peña et al., 2018).

Interestingly, the nAb responses that arose in two individuals to their SI viruses were the most potent responses observed in the cohort. We showed that these were not secondary, memory responses but rather potent, *de novo* responses. Immune complexes are known to be more immunogenic than antigen alone (Brady L. Jeannine, 2005), and presentation of superinfecting antigen would have occurred in the context of (non-neutralizing) antibodies produced in response to primary infection. These antibodies may shield non-neutralizing epitopes, improving the quality of neutralizing responses elicited against the SI. Indeed, in a therapeutic trial, passive therapy with an HIV-specific, broadly neutralizing, monoclonal antibody significantly improved subsequent neutralizing antibody responses in all but one of 15 participants (Schoofs et al., 2016), potentially mediated by a similar mechanism. Identifying the mechanism underlying the potent response to SI viruses could improve future vaccines. In conclusion, we find that HIV superinfection fails to efficiently recruit neutralizing memory B cells and, at best, results in additive nAb responses rather than a synergistic effect leading to cross-neutralization; a distinction that is highly relevant for vaccine design. This suggests that while sequential immunizations with heterologous Env immunogens may be able to improve the potency of elicited responses, alone, they are unlikely to promote the development of bnAbs.

PID	Timing of superinfection (weeks post infection)	VL at 52 weeks (copies/ml)	CD4 count at superinfection (cells/µl)	PI-SI env distance
CAP237	9 (6-11)	11900	522	11.92%
CAP256	13 (11-14)	178000	555	12.53%
CAP281	42 (40-44)	<400	1199	13.99%
CAP334	(42-47)	14900	(458-543)	nd
CAP377	33 (30-37)	337496	462	15.18%

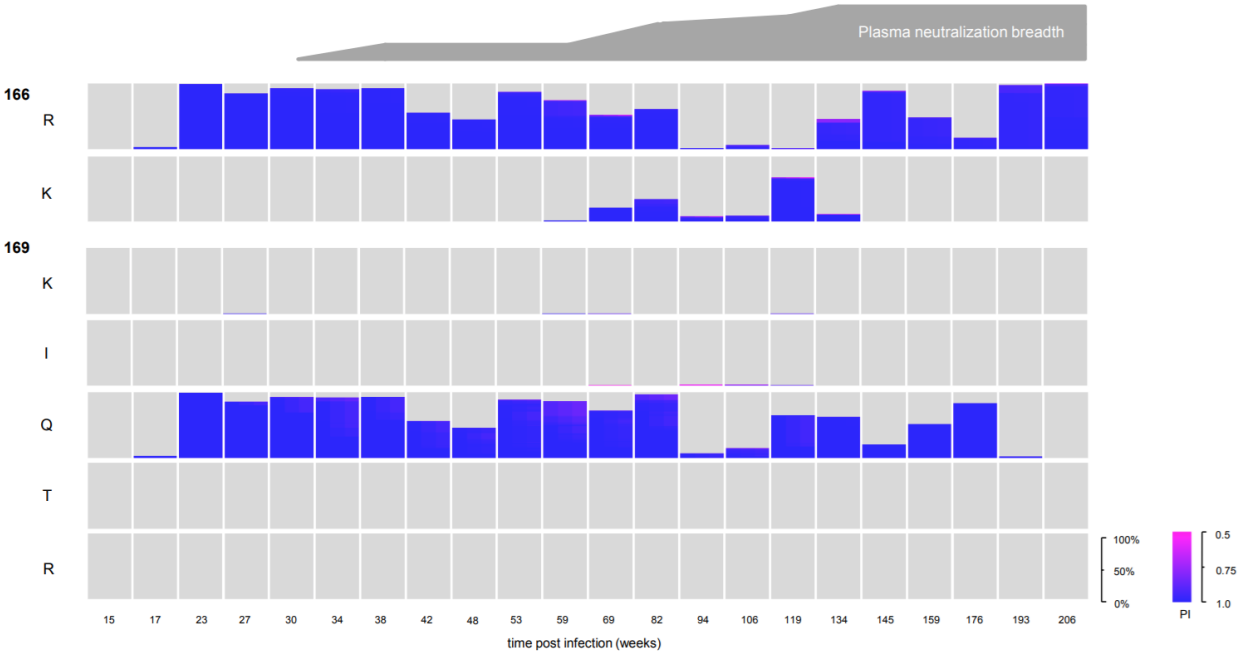
Supplementary Figure 2.1 Identification of superinfection in five participants.

The estimated timing of superinfection, with the confidence interval in parenthesis, is summarized together with the viral load (VL) at 1 year post primary infection, CD4 count at the visit closest to the estimated time of superinfection. PI-SI distance; the DNA distance in *env* between the transmitted/founders of primary infection and superinfection for each participant. nd; not done. PID; anonymized participant ID numbers.



Supplementary Figure 2.2 Potency does not lead to breadth.

No significant correlation (P = 0.8194) between the potency of the autologous neutralizing antibody response and the later development of neutralization breadth in singly infected participants (n=14). Breadth was measured at 3 years post infection and potency was estimated as the mean of the 3 highest titers observed against an acute/early Env clone.



Supplementary Figure 2.3 Frequency of key genotypes in the primary infecting lineage over time in CAP256.

Only residues that were assigned to the primary infecting virus lineage with posterior probabilities >0.5 are included, where more confident assignments are bluer.

CHAPTER 3

HIV co-infection provides insights for the design of vaccine cocktails to elicit broadly neutralizing antibodies

Daniel J. Sheward[1], Tandile Hermanus[2], Ben Murrell[3], Nigel Garrett[4], Salim S. Abdool Karim[4,5], Lynn Morris[2,6,4], Penny L. Moore[2,6,4], and Carolyn Williamson[1,4,7]

[1] Division of Medical Virology, Department of Pathology, Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa;

[2] Center for HIV and STI, National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa;

[3] Department of Microbiology, Tumor, and Cell Biology, Karolinska Institutet, Sweden

[4] Centre for the AIDS Programme of Research in South Africa (CAPRISA), University of KwaZulu- Natal, Durban, South Africa;

[5] Department of Epidemiology, Columbia University, New York, USA;

[6] University of Witwatersrand, Johannesburg, South Africa;

[7] National Health Laboratory Services of South Africa.

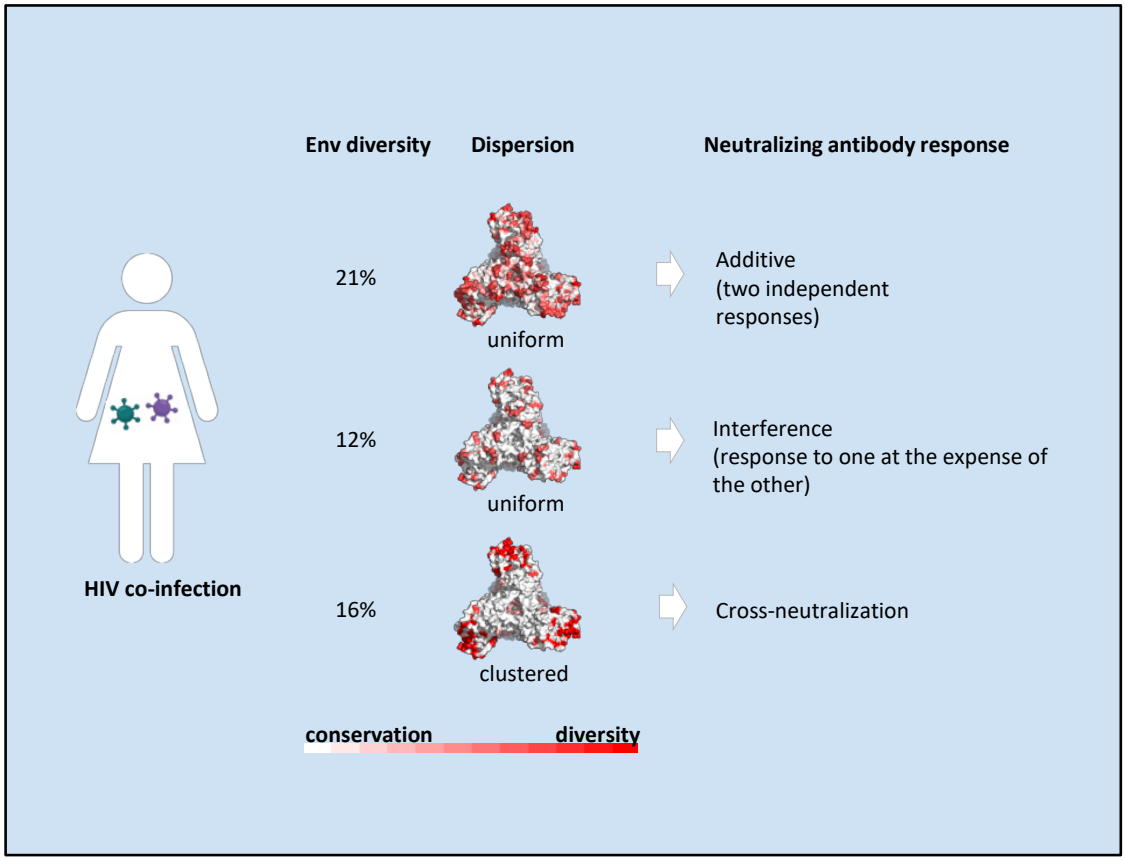
Sheward et al., (2022). *Journal of Virology, 96 (14)*

Author contributions (CRediT taxonomy)

	DJ S	T H	B M	N G	SS AK	L M	PL M	C W
Conceptualization	■						▒	▒
Methodology	■							
Software	■		▒					
Formal Analysis	■							
Investigation	■	▒						
Resources				■	■	■	■	■
Writing - original draft	■							
Writing - review and editing	■			▒	▒	▒	■	■
Visualization	■							
Supervision							■	■
Funding acquisition					■	■	■	■

Abstract

Induction of broadly neutralizing antibodies (bNAbs) to HIV, and other diverse pathogens, will likely require the use of multiple immunogens. An understanding of the dynamics of antibody development to multiple diverse but related antigens would facilitate the rational design of immunization strategies. Here, we characterize, in detail, the development of nAbs in three individuals co-infected with several divergent HIV variants. Two of these co-infected individuals developed additive or cross-neutralizing antibody responses. However, interference was observed in the third case, with neutralizing antibody responses to one viral variant arising to the near exclusion of neutralizing responses to the other. Longitudinal characterization of the diversity in the Envelope glycoprotein trimer (Env) structure showed that, in the individual who developed the broadest nAbs, circulating viruses shared a conserved epitope on the trimer apex that was targeted by cross-neutralizing antibodies. In contrast, in the other two individuals, diversity was distributed across Env. Taken together, these data highlight that multiple related immunogens can result in immune interference. However, they also suggest that immunogen cocktails presenting shared, conserved neutralizing epitopes in a variable background may focus broadly neutralizing antibody responses to these targets.



Introduction

Induction of broadly neutralizing HIV antibodies (bNAbs) will likely be necessary for an effective HIV vaccine. Recombinant, soluble, stabilized HIV envelope glycoprotein (Env) trimers that resemble the native trimer have been shown to elicit nAbs in animal models (Sanders and Moore, 2017). However, these antibodies have been largely strain-specific, and new strategies to broaden responses remains a priority. One approach to achieve this is the co-administration of multiple immunogen variants (vaccine "cocktails"). Such cocktails could drive multiple independent neutralizing antibody responses to each immunogen, or potentially favour responses to conserved epitopes. However, multiple unrelated stabilized Env trimers, administered together as a cocktail or sequentially, failed to promote cross-neutralizing antibodies in rabbits or macaques (Klasse et al., 2016; de la Peña et al., 2018). For this approach, it is not clear how much diversity is required, nor how such immunogens should be administered in order to promote the development of bNAbs. While bNAbs have not been successfully elicited by immunization, they develop in a subset of HIV infected individuals. Insights from studies of these individuals continue to shape the next generation of vaccine strategies that hope to elicit similar responses.

In the majority of cases, HIV infection is established by a single founder variant. However, in ~20-30% of cases, infection can be traced to more than one founder (multi-variant transmission) (Abrahams et al., 2009; Keele et al., 2008). In a subset of these cases, highly diverged Env variants are evident, likely representing independent transmission events from different donors, or multi-virus transmission from a donor that was infected by two or more phylogenetically distinct variants. Teasing apart the neutralizing antibody responses to multiple Env antigens in these cases could provide information on how vaccines might be designed to promote cross-neutralization.

We previously characterized the antibody responses in five HIV superinfected individuals, and found that sequential exposure to two HIV variants, the second variant failed to efficiently recruit cross-neutralizing memory B cells specific for conserved epitopes (Sheward et al., 2018). Here, we expand this investigation to determine if co-exposure to two HIV variants (co-infection) rather than sequential exposure may better favour the development of HIV cross-neutralizing antibodies. We characterize, in-depth, the neutralizing antibody responses in three individuals infected with two or more diverse HIV variants prior to seroconversion (defined here as co-infection) and discuss the implications for the design of polyvalent immunogen cocktails.

Materials and Methods

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Human Subjects

Samples were provided from participants from the CAPRISA 002 Acute Infection Study established in 2004 (van Loggerenberg et al., 2008). This cohort recruited recently HIV infected women from high risk, HIV negative women from Durban and Vulindlela, KwaZulu-Natal, South Africa as well as from the CAPRISA 004 study (Abdool Karim et al., 2010). 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. 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 in this study were antiretroviral therapy (ART) naïve and were initiated on ART consistent with the prevailing South African ART guidelines. Ethical approval for this study was received from the ethics committees of the University of Cape Town (025/2004), the University of KwaZulu-Natal (E013/04), and the University of the Witwatersrand (MM040202), and all participants provided written, informed consent.

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). 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 1% Penicillin and Streptomycin, 25 mM HEPES (Sigma-Aldrich, St Louis, MO), and 10% heat inactivated Fetal Bovine Serum (FBS) (Biochrom, Cambridge, UK). Cells were cultured at 37°C in a humidified incubator with 5% CO_2, and monolayers were disrupted when nearing confluence with Trypsin-EDTA.

Single Genome Sequencing

Plasma viral RNA was extracted from 200 µl or 400 µl of plasma using either the Roche MagNApure (Roche Applied Science, Mannheim, Germany) or QIAamp viral RNA mini kits (Qiagen, Valencia, CA). RNA was reversed-transcribed to cDNA using Superscript III (Invitrogen, Life Technologies, Carlsbad, CA) as per the manufacturer's instructions. Env

cassettes were amplified from the 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 (Abrahams et al., 2009; Keele et al., 2008). Primers used in the outer reaction were 5'- GGGTTTATTACAGGGACAGCAGAG -3' (HXB2 nt 4900 - 4923) and 5'-GCACTCAAGGCAAGCTTTATTGAGGCTTA -3' (HXB2 nt 9604 - 9632). Inner primers used were 5'- CACC GGCTTAGGCATCTCCTATAGCAGGAAGAA-3' (HXB2 nt 5954 - 5982) and 5'- TTGCCAATCAAGGAAGTAGCCTTGTGT -3' (HXB2 nt 9145-9171). Outer reaction thermal cycling conditions were as follows: initial denaturation at 94°C for 2 minutes, followed by 35 cycles of [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. Amplicons were directly sequenced using an ABI3000 genetic analyser and BigDye terminator reagents (Applied Biosystems, Foster City, CA) using twelve primers, by the Central Analytic Facility at the University of Stellenbosch, South Africa. 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.

Molecular Cloning

At each timepoint, amplicons closest to the consensus of each major clade closest to each of the transmitting variants were selected and 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) as per the manufacturer's instructions, and cultured on Luria-Bertani agar supplemented with 100 µg/ml Carbenicillin (Sigma-Aldrich, St Louis, MO). 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 as described above in order to ensure the clone contained no non-synonymous mutations relative to the SGA-derived env sequence.

Neutralization assays

Env-pseudotyped viruses were generated by co-transfecting env plasmids with pSG3ΔEnv at a 1:2 ratio into HEK293T cells using Fugene 6 (Applied Science, Indianapolis, IA) or PolyFect (Qiagen, Valencia, CA) per 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), supplemented up to 20% FBS, and stored at -80°C until use. Neutralization assays were performed as described previously (Gray et al., 2007). Assays were performed in duplicate wells, and repeated at least twice using aliquots of the same

plasma draw. Neutralization breadth for the cohort was quantified as the proportion of heterologous viruses from a multi-subtype pseudovirus panel of 44 pseudotyped heterologous viruses each plasma sample was able to neutralize at ID_{50} titers >45 (Gray et al., 2007). Breadth at one, two, and three years were also estimated for co-infected participants using a reduced panel of 18 heterologous viruses.

Mutagenesis

Point mutants were generated using the Quikchange Lightning site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA) as per the manufacturer's instructions.

Characterizing selection pressure

Positive selection in alignments was assessed using MEME (Murrell et al., 2012) and FUBAR (Murrell et al., 2013), on the datamonkey web server (www.datamonkey.org). Alignments were partitioned to account for the confounding effects of recombination (Scheffler et al., 2006) using GARD (Kosakovsky Pond et al., 2006).

Diffusion accessibility

Diffusion accessibility was estimated using the UCLA webtool (http://services.mbi.ucla.edu/DiffAcc/) (Tsai et al., 2015) with PDB:4ZMJ (Kwon et al., 2015).

Code availability

Julia code used to generate the sequence 'highlighter' plots is available at https://github.com/MurrellGroup/highlighterplotting with no restrictions for use.

Results

We previously screened participants in the CAPRISA 002 acute infection cohort for cases of HIV dual infection, identifying 19 participants infected with multiple variants, 14 of whom had evidence of multiple variants at the first seropositive visit. Cases of co- infection were initially identified by heteroduplex mobility shift assays and subsequent confirmation with sanger sequencing (Woodman et al. 2011), and additional cases were later identified in single genome sequencing. To determine if co-infection with diverse variants generally led to broader neutralizing antibody responses, we assessed neutralization breadth at two- and three-years post infection and compared this to participants likely infected with a single variant (N=16). We found that co-infection alone was not sufficient to broaden responses compared to singly infected individuals (Figure 3.1).

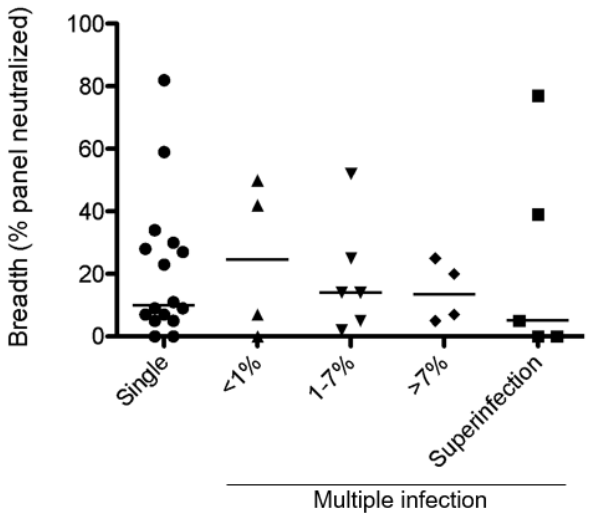


Figure 3.1 Neither multiplicity nor diversity of infection was associated with broader neutralizing antibody responses.

Neutralization breadth at three years post infection was previously quantified against a heterologous virus panel (N=44) encompassing subtypes A, B, and C (Gray et al., 2007). Single, Clinical infection consistent with a single transmitted/founder; Superinfection, Infection following seroconversion with a second phylogenetically unlinked strain. Individuals co-infected with >1 variant prior to seroconversion are grouped according to the genetic distance in Env between the infecting variants.

To define the immune response to multiple antigens, we next characterized the neutralizing antibody responses for three of the co-infected participants where single genome sequencing of *env* identified co-infection with the divergent variants, and where each case represented a distinct antigenic exposure. These encompassed immunological exposure to: (i) two diverse, unlinked Envs, (ii) two related Envs with diversity spread over the trimer, and (iii) two diverse

Envs that had recombined such that large regions were conserved in both variants. We aimed to establish whether co-infection led to independent responses (additive), promoted a response to epitopes conserved in both (immunofocussing), or whether the responses to distinct viruses competed with one another leading to lower titers than would have arisen following a monovalent exposure (interference). Further, we aimed to identify whether the specific antigenic features were associated with these outcomes.

To characterize the neutralizing antibody response to multiple variants in each of the three participants, we molecularly cloned representative Envelopes (Envs) from samples taken at the first HIV positive visit, and at approximately 3-, 6-, and 12-months post infection. We generated Env-pseudotyped viruses and evaluated the titers and kinetics of the plasma neutralizing antibody response to each Env clone using longitudinal plasma samples from the first two years of infection. These are described below:

Co-infection resulting in antibody interference (CAP267)

Following infection, CAP267 harboured two distinct but recombined viral populations that differed by up to 12% in Env (Figure 3.2a). Both variants were detectable at every sampled time point thereafter, and these remained distinct with no significant further recombination in Env evident over the first year of infection (Figure 3.2a-c, left panel). Overlaying the locations of amino acid differences between these lineages onto the quaternary Env structure (PDB:4ZMJ (Kwon et al., 2015)) revealed that these differences were uniformly distributed across the surface of the trimer (Figure 3.2a-c, right).

Characterizing the longitudinal, autologous neutralizing antibody response to representative variants sampled at each time-point revealed preferential neutralization of one variant, with titers to the second consistently remaining low (<100) or below the limit of detection (<45) (Figure 3.2a-c, middle). This preferential neutralization switched from one variant to the other after one year post infection (Figure 3.2c), indicating that the differential neutralization was not due to inherently low immunogenicity of one variant. The weak nAb titers that were evident were significantly lower than the cohort average and represented some of the lowest titers observed (Figure 3.3), suggesting immune interference.

Interestingly, the preferentially targeted variant frequently represented the minority population (as determined from the frequency in SGA sequences) at the time that nAb responses arose highlighting that antigen load is not the dominant driver of immunogenicity/immunodominance. CAP267 went on to develop some cross-neutralizing activity with plasma samples from two- and three-years post infection able to neutralize 11% and 39% of heterologous viruses,

respectively (Figure 3.2d). However, even after the development of cross-neutralizing activity, autologous neutralizing responses to the dominant variant at 12 months post infection remained barely detectable (Figure 3.2c). This indicates that the cross-neutralizing antibodies failed to target epitopes common to both autologous variants.

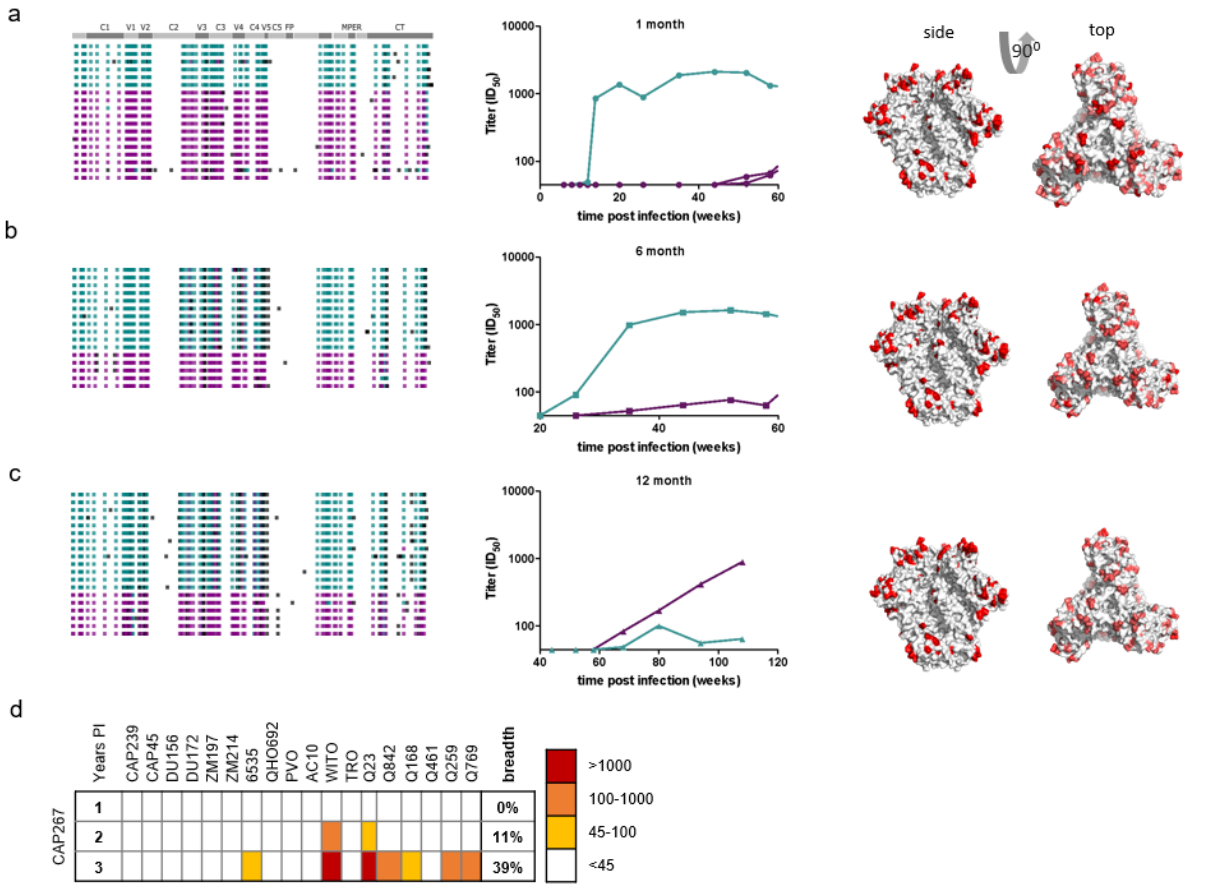


Figure 3.2 Interference in the autologous neutralizing antibody response in CAP267.

CAP267 was co-infected with variants harbouring 12% amino acid diversity uniformly distributed across Env. **a-c Left** panels depict longitudinal, autologous neutralization of representatives from both major variants cloned from (**a**) 1-month, (**b**) 6-months and (**c**) 12-months post infection. Potent neutralization of one variant with weak or undetectable neutralization of the other was evident at all three time points. **Middle** panels show sequence highlighter plots summarizing the env sequences in the corresponding plasma sample, where each row represents a single genome sequence. Teal and purple ticks represent nucleotides that match those in the two major variants respectively. White ticks represent sites harbouring a nucleotide conserved in both variants, and black ticks represent polymorphisms matching neither major variant. MPER, Membrane proximal external region; FP, fusion peptide; CT, cytoplasmic tail. **Right** panels depict diversity overlaid onto the Env structure (PDB:4ZMJ), where white represents conservation and redder colours denote increasing diversity at that site. d Neutralization of 18 heterologous viruses at 1-3 years post infection where white represents no detectable cross neutralization (ID_{50} <45) and warmer colours represent more potent neutralization, as shown in the key. Years PI, years post infection.

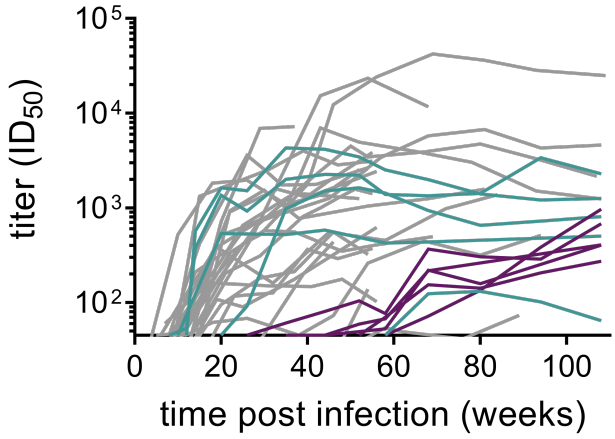


Figure 3.3 Weak autologous neutralizing antibody responses in CAP267.

Autologous neutralizing antibody responses in CAP267 (teal and purple representing those to each variant respectively) are plotted alongside the neutralizing antibody responses to early/founder envelope clones in 23 other participants in the CAPRISA 002 cohort (grey). CAP267 neutralizing titers to several clones represent some of the lowest titer, and most delayed responses observed in the cohort, suggestive of immune interference in the antibody response.

Co-infection resulting in additive antibody responses (CAP137)

Donor CAP137 was infected with two highly divergent, phylogenetically unlinked variants differing by ~21% in Env. Amino acid differences between the two variants were uniformly distributed over the surface of the Env trimer (Figure 3.4a, right panel). Both variants were observed at 3- and 6- months (12-month samples were unavailable), albeit harbouring sub-genomic fragments inherited by recombination between co-transmitted variants (Figure 3.4a-c, left). While neutralizing antibody responses arose to both variants (Figure 3.4a-c, middle), responses to one variant were relatively delayed and first detectable only ~6 months post infection (8-14 weeks after the nAb response to the first variant) suggestive of some degree of immune interference in the response to this early clone (Figure 3.5). However, representative Envs of both variants sampled at 3- and 6-months post infection were neutralized with similar kinetics, and similarly low titers (Figure 3.4b-c middle): ID_{50} titers peaked at ~350 to 3-month Envs, and ~275 to 6-month Envs. Despite neutralization of both autologous Env variants, no cross-neutralization of any heterologous viruses was detectable by three years post infection (Figure 3.4d). This indicates that high levels of diversity dispersed across co-circulating Envs resulted in separate responses targeting non-conserved epitopes on each variant.

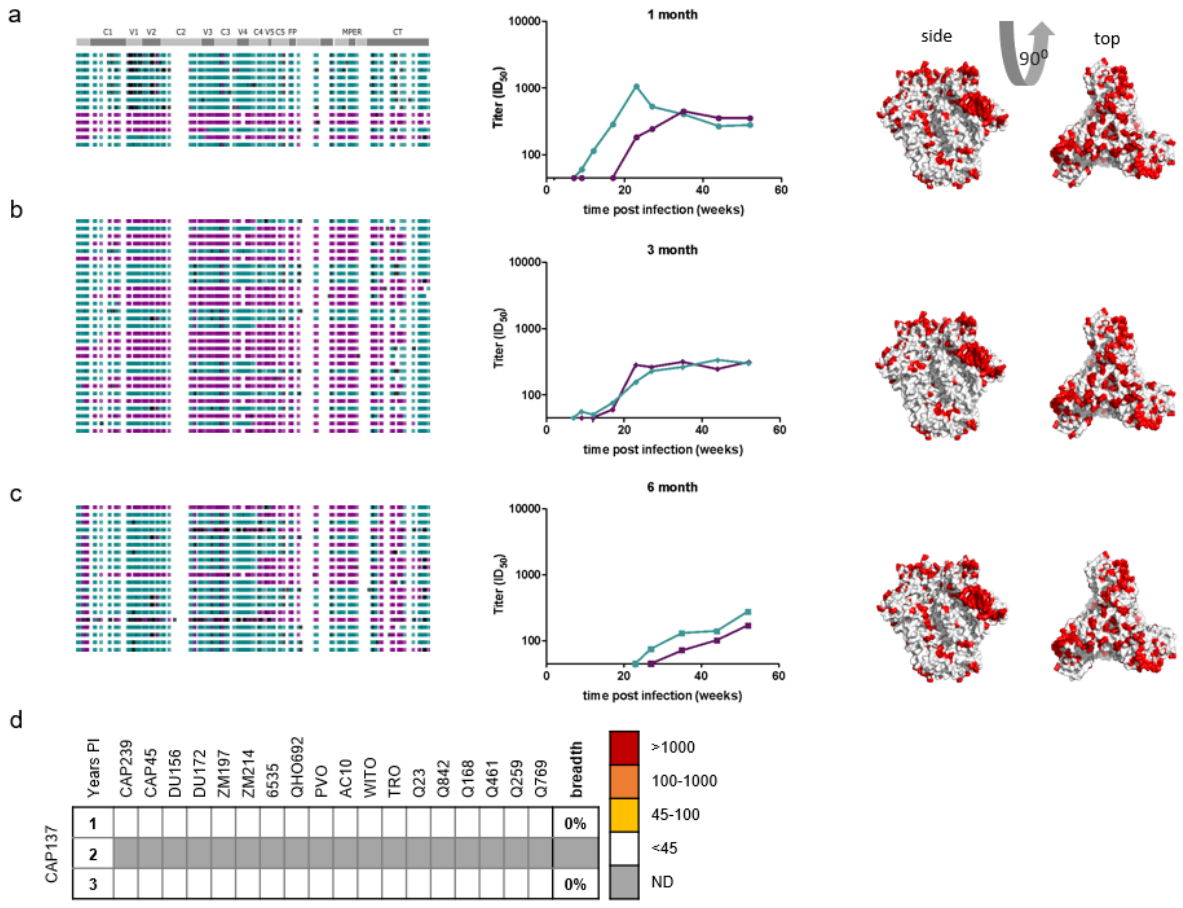


Figure 3.4 Autologous and heterologous neutralizing antibody responses in CAP137.

CAP137 was co-infected with unlinked variants displaying 21% amino acid diversity distributed across Env. **a-c Left** panels depict longitudinal, autologous neutralization of representatives from both major variants cloned at (**a**) 1-month, (**b**) 3-months and (**c**) 6-months post infection. **Middle** panels depict sequence highlighter plots summarizing the env sequences in the corresponding plasma sample, where each row represents a single genome sequence. Teal and purple ticks represent nucleotides that match those in the two major variants respectively. White ticks represent sites harbouring a nucleotide conserved in both variants, and black ticks represent polymorphisms matching neither major variant. MPER, Membrane proximal external region; FP, fusion peptide; CT, cytoplasmic tail. **Right** panels depict diversity overlaid onto the Env structure (PDB:4ZMJ), where white represents conservation and redder colours denote increasing diversity at that site. d CAP137 plasma from 1- and 3-years post infection (PI) was unable to cross-neutralize any of 18 heterologous viruses tested. Plasma from 2 years PI was not tested (ND, not done).

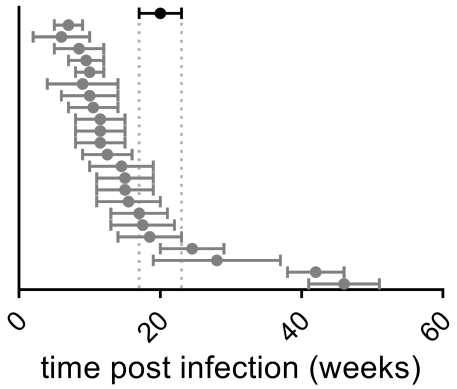


Figure 3.5 Neutralizing antibodies that arise to variant 2 in CAP137 are relatively delayed.

Time from the inferred time of transmission to the detection of neutralizing antibody responses is shown for CAP137 variant 2 (black), compared to those against the early/founder variants in 23 other participants from the CAPRISA 002 acute infection cohort. Depicted is the midpoint and range between the last negative and first positive samples.

Co-infection resulting in cross-neutralizing responses (CAP37)

CAP37 was co-infected with at least three variants. Two of the variants differed by only 6% in Env, while the third was highly diverged, differing by ~16% in Env. However, all three variants shared sub-genomic regions of high homology within Env, likely inherited through recombination (Figure 3.6a-c, left). Shared regions across these variants included known bNAb targets, including V2 at the Env trimer apex, the membrane proximal external region (MPER), and the fusion peptide (Figure 3.6a right). Locations of high diversity mapped to highly immunogenic regions (including C3V4) (Moore et al., 2008, 2009b) and to regions that are more accessible (Figure 3.7).

The kinetics and titer profile of nAb responses to these early Env clones suggest that responses were at least additive (Figure 3.6a). All three of the transmitted variants were relatively potently neutralized with ID_{50} titers peaking between 2,000 - 10,000. Titers and kinetics of neutralization were similar for the dominant circulating 6- and 12-month Envs (Figure 3.4b-c), suggestive of a single response capable of cross-neutralizing both variants (although quantitatively similar but independent responses to both variants cannot be excluded). This coincided in time with further reduced diversity at the trimer apex, likely through recombination, despite high diversity elsewhere (Figure 3.6b-c, right).

CAP37 developed cross-neutralizing activity, neutralizing 28% and 50% of a heterologous virus panel at two- and three-years post infection respectively (Figure 3.6d). As the trimer

apex including V2 is a frequent target of bnAbs (Gray et al., 2007; Landais et al., 2016; Walker et al., 2010), we sought to determine whether cross-neutralization in CAP37 was due to antibodies targeting this conserved region. Dependency on the N-linked glycan at position 160 and a lysine at position 169 is common to many bnAbs specific for V2 (Andrabi et al., 2015; Gorman et al., 2016). Therefore, we introduced mutations at these sites into three heterologous pseudoviruses neutralized by CAP37 and evaluated their effect on neutralization by plasma sampled between 18- and 24-months post infection. N160A and K169E mutations substantially reduced neutralization of heterologous viruses, ConC and CAP45 by the 2-year plasma sample (by between 7- and 29-fold; Figure 3.8) indicating that these viruses were predominantly neutralized by V2-directed antibodies. In assays performed by T. Hermanus (NICD, South Africa), neutralization of Du156 by plasma at approximately 1-and-a-half years post infection was moderately (~4-fold) reduced by the introduction of an N160K (Figure 3.8). However, at two years post infection, the same mutation enhanced neutralization, suggesting the removal of the N-linked glycan may have better exposed the targeted epitope. However, the K169E mutation did not significantly affect neutralization of Du156.

Immune pressure from nAbs typically drives sequence evolution, and evidence of strong diversifying selection in longitudinal V2 sequences was apparent. However, pressure was not centred around the C strand of V2 (the target of all well-characterized V2 bnAbs), but adjacent in the V2 loop. Taken together, these data suggest that apex-targeting antibodies develop in CAP37 and can mediate cross-neutralization, though they do not completely account for the plasma breadth.

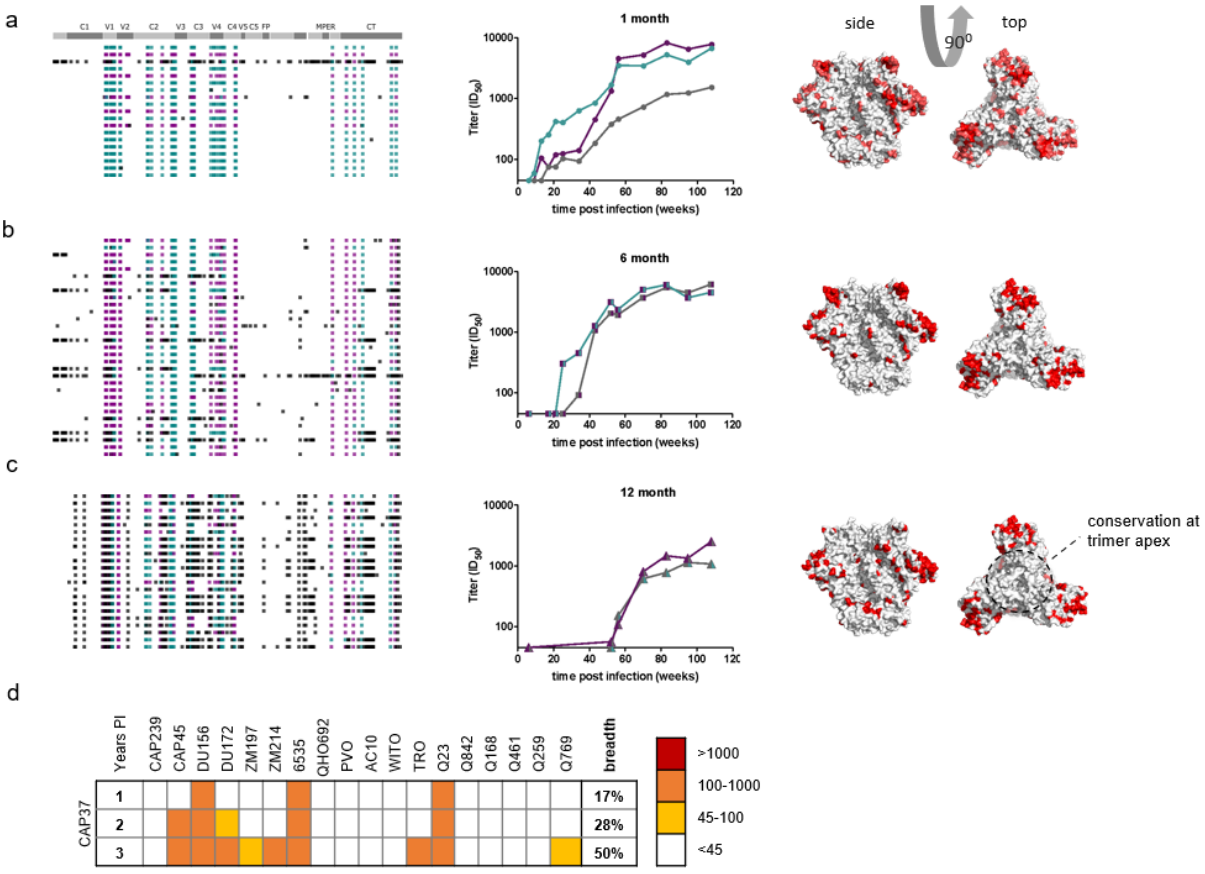


Figure 3.6 Potent neutralization of variants in CAP37.

CAP37 was co-infected with at least three recombinant variants harbouring up to 16% amino acid diversity, clustered in Env. **a-c Left** panels depict potent autologous neutralization of representative variants cloned at (**a**) 1-month, (**b**) 6-months and (**c**) 12-months post infection. **Middle** panels depict sequence highlighter plots summarizing the env sequences in the corresponding plasma sample, where each row represents a single genome sequence. Teal and purple ticks represent nucleotides that match those in the two major variants respectively. White ticks represent sites harbouring a nucleotide conserved in both variants, and black ticks represent polymorphisms matching neither major variant. MPER, Membrane proximal external region; FP, fusion peptide; CT, cytoplasmic tail. **Right** panels depict diversity overlaid onto the Env structure (PDB:4ZMJ), where white represents conservation and redder colours denote increasing diversity at that site. **d** CAP37 developed cross-neutralizing activity against a panel of 18 viruses, with plasma sampled 3 years PI (post infection) capable of neutralizing 50% of heterologous viruses tested.

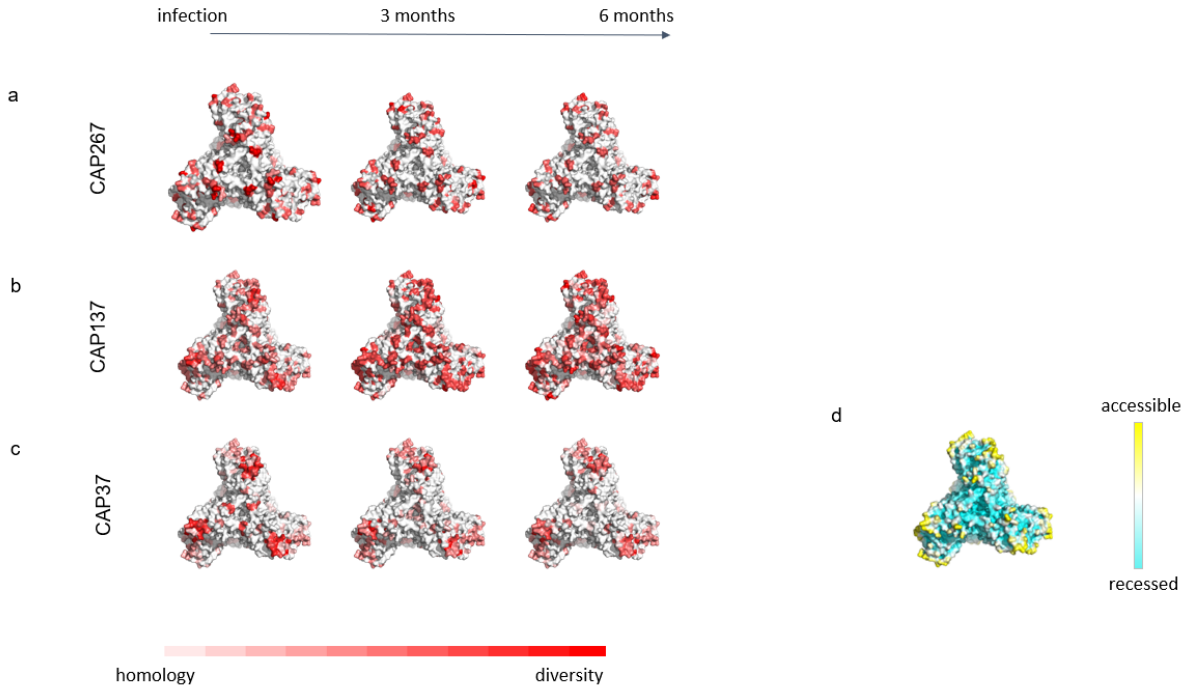


Figure 3.7. Summary of the Env diversity over time in early infection for the three individuals characterized here.

a-c Depicted is the top-down view of the Env trimer (PDB:4ZMJ), coloured by entropy present in single genome sequences generated from each sample. **d** Accessibility, estimated by diffusion accessibility to the deglycosylated protein, of regions within the HIV Env trimer.

		ID50									
		ConC			Du156			CAP45 G3			
		WT	N160A	K169E	WT	N160K	K169E	WT	N160A	K169E	MLV
CAP37	1.5 years	939	1089	1132	717	174	546	284	450	185	<40
	2 years	7991	1073	1018	745	2433	627	5619	1801	193	<40

>2000
200-2000
45-200
<40

		FOLD DIFFERENCE					
		Conc		Du156		CAP45 G3	
		N160A	K169E	N160K	K169E	N160A	K169E
CAP37	1.5 years	0.9	0.8	**4.1**	1.3	0.6	1.5
	2 years	**7.4**	**7.8**	0.3	1.2	**3.1**	**29.1**

Figure 3.8 Evidence of cross-reactive neutralizing antibodies targeting V2 in donor CAP37.

Mutations in V2 significantly impact neutralization of heterologous viruses (ConC, Du156, CAP45.G3). ID_{50} titers are displayed in the top panel, and fold-change relative to wildtype viruses are shown in the bottom panel.

Discussion

Designing a vaccine able to elicit bnAbs against HIV has proven to be an immense scientific challenge. Multivalent HIV envelope glycoprotein (Env) cocktails have successfully elicited nAbs in animal models but to date have been largely unsuccessful at eliciting any substantial neutralizing antibody breadth. Species-specific differences in the antibody response to Env immunogens may nevertheless hinder their translation to human vaccine regimens. An understanding of the dynamics of antibody responses to multiple HIV Env antigens in humans would facilitate the design of vaccine cocktails. We therefore characterized the neutralizing antibody response over time in three HIV co-infected individuals, encompassing immunological exposure to: (i) two diverse, unlinked Envs, (ii) two related Envs with diversity spread over the trimer, and (iii) two diverse Envs that had recombined such that large regions were conserved in both variants. We note that while we would not necessarily be able to resolve co-infection from a superinfected donor compared to two independent transmissions from independent donors both occurring prior to seroconversion, the immunological outcomes of infection with diverse Envs is largely independent of whether this diversity came from one or two donors.

We find that co-infection alone was not sufficient to broaden neutralizing antibody responses. Furthermore, in the three cases of co-infection studied in detail, varying features of antigen diversity led to markedly different outcomes. Two diverse Envs led to additive antibody responses, while profound interference followed coinfection with more similar Envs. Importantly, co-exposure to related Envs that presented shared conserved regions but maintained high diversity in exposed, immunodominant regions was associated with increased breadth.

In the case of exposure to two diverse, unlinked subtype C Envelopes differing by 21% (CAP137), titers and kinetics were indicative of independent neutralizing antibody responses to each variant. Furthermore, responses to one of the early clones in CAP137 was significantly delayed, and titers to both variants following the targeting of the second variant were low. This is consistent with the responses seen in animals immunized with multivalent cocktails of Envs. Immunization with two stabilized trimers elicited distinct neutralizing antibody lineages to each variant even though both shared immunodominant ‘glycan holes’ (Yang et al., 2020). Similarly, polyvalent immunization with cocktails of stabilized HIV Env trimers from multiple subtypes generally elicited neutralizing antibody responses to each of the Envs, but no significant broadening of the responses to Envs not included in the immunization was evident

(Klasse et al., 2016; de la Peña et al., 2018). These multivalent stabilized Env trimer formulations also elicited 3-7 fold lower titers compared to corresponding monovalent immunizations. Together with our observations here, this indicates that cocktails of diverse Envs, derived from sequences from unlinked infections, are unlikely to elicit cross-neutralizing responses, and may be subject to competition.

In contrast to CAP137, participant CAP267 was co-infected with two, more closely related Envs (12%), with differences dispersed across the timer. This allowed us to investigate whether exposure to more similar trimers may have promoted cross-neutralizing responses. However, the antibody response in CAP267 displayed a profound immunodominance, neutralizing one virus to the near exclusion of nAb responses to the other. This suggests that the use of similar antigens may be subject to more pronounced interference when many presented B cell epitopes are not well conserved. More similar antigens are more likely to share CD4 helper epitopes, and competition between independent B cell lineages for the same CD4 helper population would drive this interference (Victora et al., 2010). The neutralizing antibody response in CAP267 also preferentially targeted minority variants throughout the first two years of infection, suggesting that BCR affinity was a more prominent driver of specificity than antigen availability.

Interestingly, potent neutralization of two variants was evident in a third participant, CAP37, coinfected with two similarly divergent Envs (16% different in Env). Unlike the previous two participants where diversity was dispersed across Env, in CAP37 the diversity was highly localized, with large regions, including V2 at the trimer apex, conserved in both variants through recombination. Furthermore, this participant developed the broadest neutralizing antibody response, and cross-neutralization was mediated, in part, by nAbs targeting this conserved region. While these data indicate that co-exposure to two diverse Envs harbouring the same trimer apex may have promoted this response to a conserved epitope, this is also a relatively common target of neutralization in singly infected individuals (Moore et al., 2008, 2009a; Rong et al., 2009; Sagar et al., 2006).

In summary, administering multiple diverse Envs in a cocktail may nominally improve breadth by driving multiple, near-additive responses but we suggest they are unlikely to be sufficient to drive broadly neutralizing antibody lineages. Further, we show here that neutralizing antibody responses to multiple, similar Envs can be subject to interference, where nAb responses are elicited to one Env at the expense of responses to the other. Importantly, when diversity was clustered in immunogenic regions but more conserved regions were preserved, nAbs may have been directed towards these conserved epitopes. While these observations

are based on only a few individuals, these data are consistent with and extend pre-clinical data from the immunization of animals with immunogen cocktails (Klasse et al., 2016; de la Peña et al., 2018). Taken together, they suggest that the co-administration of multiple stabilized Env trimers with diversity introduced in key regions, along with parallel strategies to reduce the immunodominance of strain-specific epitopes, may represent one path to a cross-neutralizing antibody response.

CHAPTER 4

Sensitive detection of HIV superinfection reveals it is more common than previously identified

Abstract

HIV superinfection is frequently followed by extensive recombination. Recombination violates the assumptions that underpin phylogenetic inference and, if not properly accounted for, obscures the relatedness between the recombined sequence and its parents.

Strategies to identify superinfection often involve placing parent and superinfecting sequences on a reference tree and deciding whether or not they are from different sources. Such strategies break down under recombination, or when the superinfecting virus is from a local transmission network and therefore does not segregate on the phylogeny. Natural history and pathogenesis studies of HIV infected patients frequently include sequencing of viral genes generated from multiple, longitudinal samples. In such longitudinal sequences from superinfected individuals, the primary infecting strain is typically well sampled prior to superinfection. However, due the intermittent nature of sampling together with a very high recombination rate, the "pure" superinfecting strain is rarely sampled whole, making the detection of superinfection challenging.

Here we develop a Bayesian approach that identifies signals of recombination between a known parent and an unknown parent that is (possibly distantly) related to a panel of reference viruses. We demonstrate its ability to detect cases of superinfection even after recombination, and show that this approach has high sensitivity and specificity on simulated and real data. Further, this approach, was capable of identifying superinfections that were previously missed in public datasets.

Introduction

Cases of HIV superinfection (SI) inform a variety of fields. HIV superinfection may affect HIV disease progression and antiretroviral resistance profile and is therefore relevant to clinical care. Furthermore, as the second infection occurs in the face of anti-HIV immune responses, cases of HIV superinfection offer an opportunity to identify correlates of protection, or at least the types and titres of immune responses that are not sufficient to prevent HIV infection; informing vaccine design and end-points. In chapter 2, we show that sequential exposure of the immune system to two heterologous viruses can also reveal features of memory cell activation and provide important insights for prime-boost vaccines. However, a significant limitation of this work, and other studies on superinfection, is the small sample sizes. Part of this is attributable to the fact that superinfection is a relatively rare event (Redd et al., 2013). For example, high risk prospective cohorts typically observe an incidence of HIV infection in the range of 5% per person-year of follow up. In such a cohort, even assuming no reduction in risk of superinfection, the incidence of superinfection would approximate 0.25% per person-year of follow up. Thus extremely large cohorts are required to identify a substantial number of cases. Moreover, identifying cases of superinfection represents a major practical challenge. To our knowledge, no standardized or automated detection method is available.

Current approaches to detect superinfection leverage sequencing of HIV populations. Superinfection is typically detected by identifying viral sequence divergence greater than expectations for the given gene and duration between sampling, or by constructing a phylogeny that includes circulating "background" sequences and determining whether the individual's viral sequences are separated by epidemiologically unlinked sequences. However, extensive recombination between the two strains typically follows superinfection (Sheward et al., 2018), and identifying the superinfecting strain prior to recombination with the primary infecting virus requires tailored approaches and extremely frequent sampling that is typically not available. Such recombination of superinfecting viruses with the primary infecting strain reduces the observed divergence from the primary infecting population and confounds the phylogeny, obscuring the evidence of superinfection (Figure 4.1). The success of phylogenetic approaches is also highly dependent on a diverse and relevant background alignment from the local epidemic. Prior approaches to take recombination into account typically utilized sliding windows (Wagner et al., 2014). However, any recombinant region smaller than the window size is subject to the same confounds and overly small windows result in high uncertainty in the phylogeny.

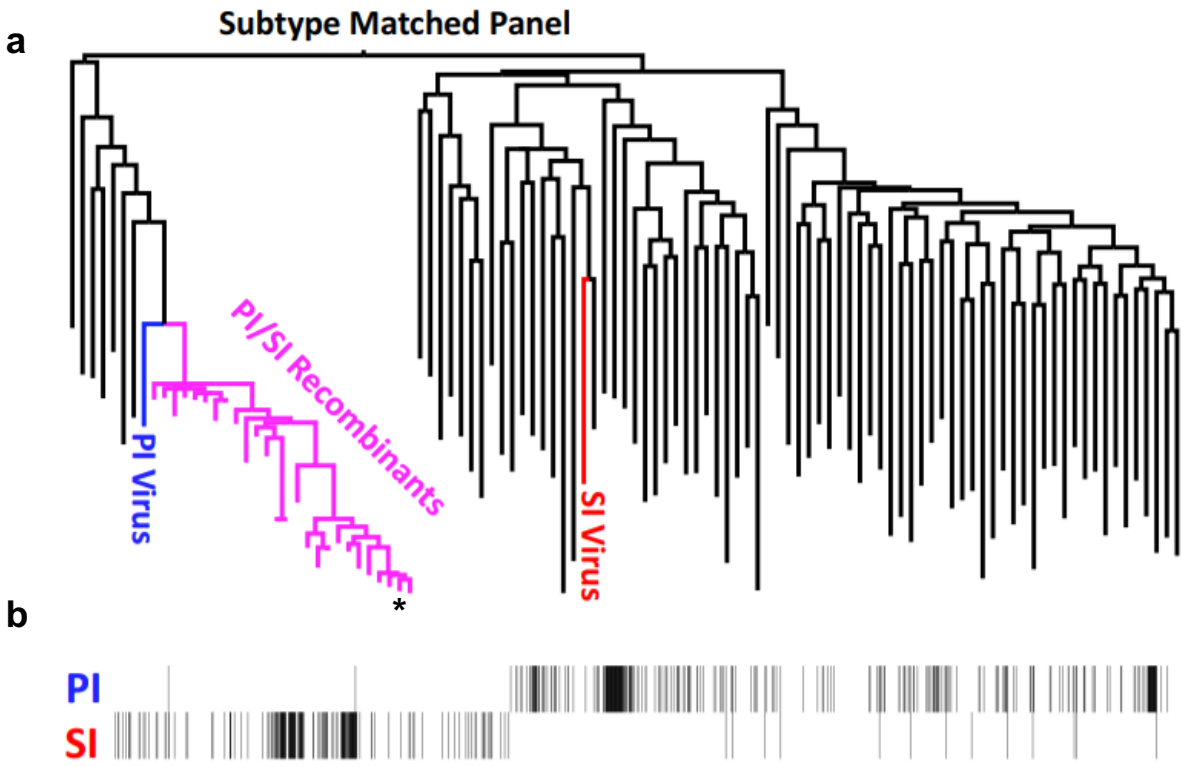


Figure 4.1 Recombination can obscure evidence for superinfection.

a shows a phylogeny of reference sequences (black), and sequences from a CAPRISA participant (CAP137) who was dually infected with two unlinked parental variants within a short window, allowing both parental viruses to be sampled. Note that the Primary (PI) and Secondary (SI) variants segregate on the phylogeny, but all of the recombinant query sequences sampled later in infection fall in a clade adjoined to the PI variant. If the SI variant were unknown here, this dual infection would have likely gone undetected by conventional methods. **b** Highlighter plot of an *env* query sequence (*) showing the differences (black ticks) from the PI and SI variants, demonstrating that it is indeed a recombinant.

Hidden Markov Models (HMMs) offer an alternative to sliding windows, removing any concerns about window sizes and modelling processes such as recombination in a probabilistic framework, which allows statistical hypothesis testing and the rigorous quantification of evidence.

HMMs are a framework for representing probability distributions over a sequence of observations (Rabiner, 1989), when these observations are conditioned on an unobserved sequence of discrete latent states, where each successive state can depend on the previous state. For HMMs, algorithms exist to compute the probability of an observed sequence of events, given model parameters, marginalizing over all possible unobserved state sequences, as well as to infer the most likely sequence of states given the observation sequence. HMMs

have proven to be extremely useful in representing nucleotide or amino acid sequence data, and are widely used in applications from multiple sequence alignment (Mount, 2009) to protein secondary structure prediction (Won et al., 2007). Figure 4.2 provides a schematic example of a minimal two-state HMM.

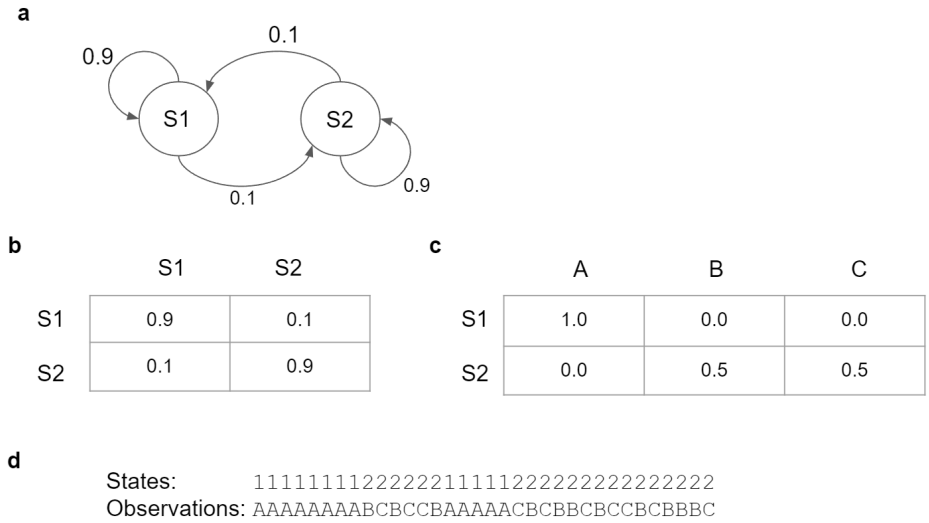


Figure 4.2 An example of a two-state Hidden Markov Model.

a A schematic of transition probabilities between two states, S1 and S2, with arrows showing the probabilities of the N+1th state given the Nth state. **b** shows the same probabilities in the form of a transition probability matrix. **c** shows the distribution over observations A,B and C, which are different between states S1 and S2. **d** shows an example sequence that could be generated by this HMM, starting in state S1 and switching back and forth, with the observations generated according to the observation distribution.

HMMs can conveniently represent recombination between nucleotide sequences. Two states can represent inheritance from two parents, and switching between these states reflects recombination breakpoints. The mechanism of viral recombination via template-switching causes recombinant regions to be inherited in blocks, and this can be naturally accounted for by Markov switching between the HMM states, with the switching rate (a parameter in the model), governing the number of breakpoints and the average length of the recombined regions. The unobserved states, for each nucleotide position, describe whether that nucleotide was derived from the known, sequenced parent (PI) or from an unknown superinfecting virus (SI), and the observed nucleotides comprise the HMM observations.

An HMM requires the distributions of observations conditioned on each state. For discrete observations, this is commonly in the form of a matrix of probabilities for each observation

given the state, as is depicted in the minimal example HMM (Figure 4.2). In our case, however, observations are nucleotides but their probability distribution, conditioned on the state, is not fixed over all nucleotide positions but rather derived from the parent sequence and allowing mutations. Even though we would have a known reference sequence, in practice the actual parent of the query sequence will likely reflect a daughter of this reference sequence, and therefore depends on the amount of evolutionary time that has passed. The SI variant would also be unknown, and unobserved in the vast majority of cases. While we can typically identify and provide sequences from the same subtype and even geographic location, we are still exceedingly unlikely to have sampled a close relative of any given SI variant in most practical cases. Furthermore, HIV sequences (and those from HIV *env* in particular) sampled from separate individuals, even within the same epidemic and geographic location, are typically highly diverged from each other with long terminal branches evident in the phylogenies. Thus the true superinfecting parent, or any close relative of it, will likely not occur on any reference panel tree, but will represent a sequence evolved (likely substantially) away from some point on this tree.

Here we develop an automated method incorporating an HMM to Detect Recombination Between A Known and Unknown Parent (DRBAKUP), that is capable of identifying cases of superinfection even after recombination, and where the superinfecting viral sequence is unknown. To do this, we construct a model of a query sequence as being generated by recombination between a sequence derived from evolution from a known reference sequence, and an unknown sequence derived through an evolutionary process from a point on the phylogeny of a reference panel. This is evaluated for multiple points on every branch of the panel phylogeny, thus spanning the breadth of the panel diversity. Thus, for each query sequence, one HMM state represents evolution away from the reference sequence (known parent, PI), and the other represents evolution from a point on the phylogeny of the reference alignment (reference panel), and switching between the two states in the HMM captures recombination between the PI variant and another, unobserved parent. This model is then compared to a model that just involves evolution from a known parent, allowing two separate rates (as the superinfection model also permits separate rates in each state) (Figure 4.3). This is implemented in the Julia language for Scientific Computing (Bezanson et al., 2012). We use MolecularEvolution.jl, a Julia package and phylogenetic and molecular evolution framework currently under development, to handle the evolutionary models and inference required for this, in collaboration with it's author, Ben Murrell (KI, Sweden).

Materials and Methods

DRBAKUP model

We implement a Bayesian approach that compares how well two models explain a query sequence: one that allows superinfection and recombination of the superinfecting and primary infecting virus (Figure 4.3a), and one that does not allow superinfection (Figure 4.3b). Under the superinfection model, we model a query sequence as a mixture of a fragments inherited from the primary infecting virus, and fragments inherited from an unobserved superinfecting virus, allowing an unknown amount of evolution away from each, with Markov switching between primary and superinfecting "parent" viruses to account for recombination. The probability of the query sequence, q, under the superinfection model is thus given by:

$$P_{SI}(q|PI,T) = \sum_{sw,t_{PI},t_x,x} P(q|sw,t_{PI},t_x,PI,T)P(sw,t_{PI},t_x,x) \qquad (1)$$

where PI is the sequence of the primary infecting virus, T is the reference panel (including the sequences and their phylogeny), sw is the HMM switching rate, t_{PI} is the evolutionary time separating the PI sequence and the sampled query, t_x is the evolutionary time separating the sampled query and x, the ancestral point on the reference panel from which unobserved superinfecting virus is derived.

The central term in this is: $P(q|sw,t_{PI},t_x,PI,T)$.
This is computed in separate stages, and incorporates two evolutionary models and a HMM. With a substitution rate matrix Q, where $Q_{i,j}$ represents the rate of change from nucleotide i to nucleotide j, the transition probability matrix $P = e^{Q_t}$, computed via a matrix exponential, represents the distribution over evolutionary change occurring over time t *(Lemey et al., 2009)*. $P_{i,j}$ represents the probability of observing nucleotide j given that the process started with nucleotide i, after time t. If A_x is an N-by-4 matrix of initial nucleotide probabilities, then the distribution over nucleotides after time t_x under rate matrix Q is calculated by:

$$n_x = A_x e^{Q t_x}$$

Here, A_x is calculated for a discrete point x on the panel phylogeny, by ancestral reconstruction, using MolecularEvolution.jl. Thus the marginal distribution over the ancestral nucleotide states at a point in the panel phylogeny is taken as the starting point of the evolutionary model for the "superinfection" component of the HMM.

We represent the sequence of the primary infecting virus as a 'one-hot encoded' nucleotide sequence, such that PI is also an N-by-4 matrix, where each row is comprised of zeros, except

at the index corresponding to the nucleotide (ie. “A” becomes [1,0,0,0], “C” becomes [0,1,0,0], “G” becomes [0,0,1,0] and T becomes [0,0,0,1]). The distribution over nucleotides that arises after time t_{PI} from the primary infecting sequence (n_{PI}) is given by:

$$n_{PI} = PIe^{Qt_{PI}}$$

Here Q is given by the HKY85 (Hasegawa et al., 1985) model of evolution allowing separate rates for transitions and transversions, with the transitions/transversions rate ratio estimated from the panel phylogeny by maximum likelihood, and the nucleotide frequencies counted empirically.

Then, with n_{PI}, and n_x computed, the central term in equation (1) is calculated as

$$P(q|sw, t_{PI}, t_x, PI, T) = P(q|sw, n_{PI}, n_x) = \sum_{\forall S} P(q|S, n_{PI}, n_x)P(S|sw)$$

where S represents a single sequence of states, assigning each nucleotide position to deriving from either the PI virus or the SI virus. All possible state paths are summed over, marginalizing out the state path variable. This number of possible state paths is exponentially large, but this sum can be computed efficiently using the HMM “Forward algorithm” (Rabiner, 1989).

The four unknowns in this model are the two evolutionary divergences, t_{PI} and t_x, the switching rate sw, and the point x on the panel tree from which the SI virus is derived. In this Bayesian framework, unknown model parameters are marginalized out, and here we adopt a coarse discretization scheme to do this efficiently. t_{PI} , t_x, and sw are all continuous parameters, and we partition these into a fixed set of discrete values. x represents points on the panel phylogeny at which ancestral sequence distributions are reconstructed, and here we select two equidistant points on each branch of the panel phylogeny, and consider the sequence distributions at each of these. In equation (1), the sum $\sum_{sw,tPI,tx,x}$ is over all combinations of discrete values for each of these four parameters, and the prior term
$P(sw, t_{PI}, t_x, x)$ is uniform over these. Thus, this sum computes a single quantity, which is the probability of the query sequence given the PI sequence and the reference panel and tree.

Finally, this model is compared to an analogous model that just includes evolution from the known parent, but allowing separate evolutionary rates for the two hidden states (Figure 4.4b). We thus calculate the probability of the query sequence in the absence of superinfection

(P_{single}), allowing HMM switching between two different evolutionary rate parameters, but both originating from the same sequence, as:

$$P_{single}(q|PI) = \sum_{sw,t_{r_1},t_{r_2}} P(q|sw,t_{r_1},t_{r_2},PI)P(sw,t_{r_1},t_{r_2}) \qquad (2)$$

where t_{r1}, and t_{r2} are two rate parameters, and sw is the switching parameter. The main term: $P(q|sw, t_{r1}, t_{r2}, PI)$ is computed in a HMM as above but with the two nucleotide distribution matrices both derived from the PI sequence using the corresponding rate parameters:

$$n_{r_1} = PIe^{Q_{t_{r_1}}}$$

$$n_{r_2} = PIe^{Q_{t_{r_2}}}$$

It would be possible for the null, single-infection model to just use a single evolutionary rate but this would run the risk that high degrees of site-to-site rate variation - common in HIV Envelope - might be ”absorbed” into the second component of the superinfection model even in the absence of superinfection, potentially producing false positives (Jones et al., 2018).

$P_{SI}(q|PI,T)$ and $P_{single}(q|PI)$ from equations (1) and (2) represent the marginal likelihoods of the query sequence under two different hypotheses about the process that generated the query sequence. For clarity, we call these $P(q|SI)$ and $P(q|single)$ respectively, and we introduce a Bayesian prior probability $P(single)$ that a query sequence is derived from the PI virus without any superinfection, which, by Bayes theorem, can be calculated as:

$$P(SI|q) = \frac{P(q|SI)(1-P(single))}{P(q|SI)(1-P(single)) + P(q|single)P(single)}$$

This posterior probability $P(SI|q)$ quantifies the evidence for the presence of a (potentially sub-genomic) fragment of a superinfecting virus in the query sequence. Here, we considered values above 0.9 as evidence for superinfection, favouring high specificity.

The probability for each state at a given nucleotide in the HMM can be calculated using the forwards-backward algorithm (Rabiner, 1989). This permits assigning regions to one or the other parent when the confidence is high (here we require a posterior probability >0.9 to confidently assign to a parent) (Figure 4.5b). Partitioning of the query sequences in this way allows placing these fragments individually on a tree (as we do in Figure 4.6), which can be used to confirm inheritance from two phylogenetically unlinked sources, providing strong support for classification of superinfection.

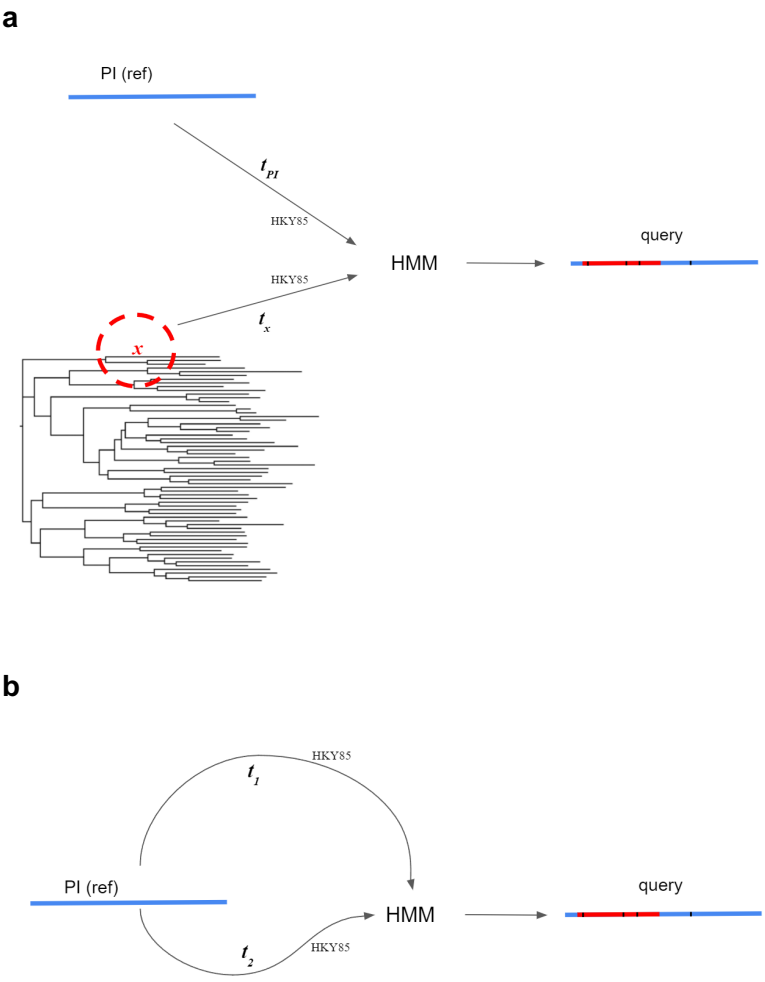
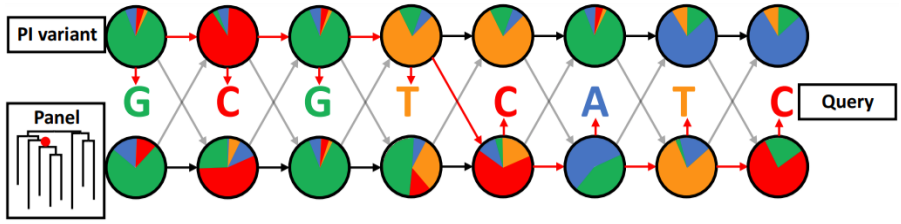


Figure 4.3 DRBAKUP models.

a DRBAKUP models a query sequence as being generated by recombination of a sequence derived from evolution from a known reference sequence, and an unknown sequence derived through an evolutionary process from a point on the phylogeny of a reference panel. This is evaluated for multiple points on every branch of the panel phylogeny, thus spanning the breadth of the panel diversity. **b** In the "null" model, we model each query sequence as derived from the PI only, but allowing switching between two evolutionary rates (t_1 and t_2) in the HMM.

An overview of DRBAKUP is shown in Figure 4.5. As inputs, DRBAKUP takes a reference sequence as a representative of the primary infecting virus (PI), a query sequence of interest representing a potential recombinant between the PI and a superinfecting virus, and a reference alignment and tree of relevant "background" sequences. Given a prior probability of superinfection, DRBAKUP outputs a posterior probability. Query sequences can also be partitioned according to the HMM state path probabilities, outputting two query sequences split according to regions likely derived from the reference "Parent" ("_P") and those likely derived from an unknown, "Other" parent ("_O"). DRBAKUP can be run sequentially over entire alignments of query sequences, using the same reference sequence and panel tree.

a



b

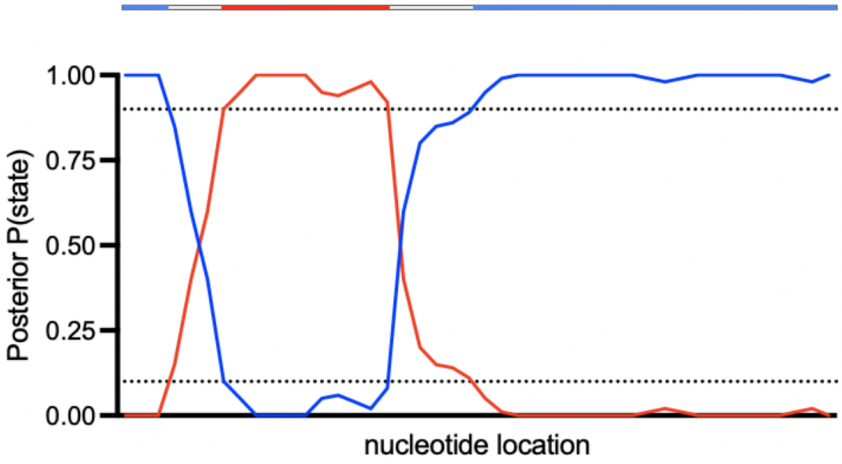


Figure 4.4. A schematic of the two-state Hidden Markov Model within DRBAKUP.

a DRBAKUP relies on a two-state Hidden Markov Model (HMM). Under the superinfection model, the first state represents evolution away from the query sequence with a particular evolutionary rate, and the second state represents evolution away from a particular point in the panel phylogeny. The pie charts reflect the expected nucleotide distributions for a sequence derived from either the PI or alternatively from a particular point on the panel phylogeny, under the given evolutionary model. A likely HMM path is depicted in red, demonstrating a switch between the PI variant and the panel states as would be seen with a recombinant sequence. **b** Query sequences can also be partitioned into two separate sequences according to the HMM path posterior probabilities, where the posterior probability in being in a particular state is greater than a defined threshold (here we use 0.9).

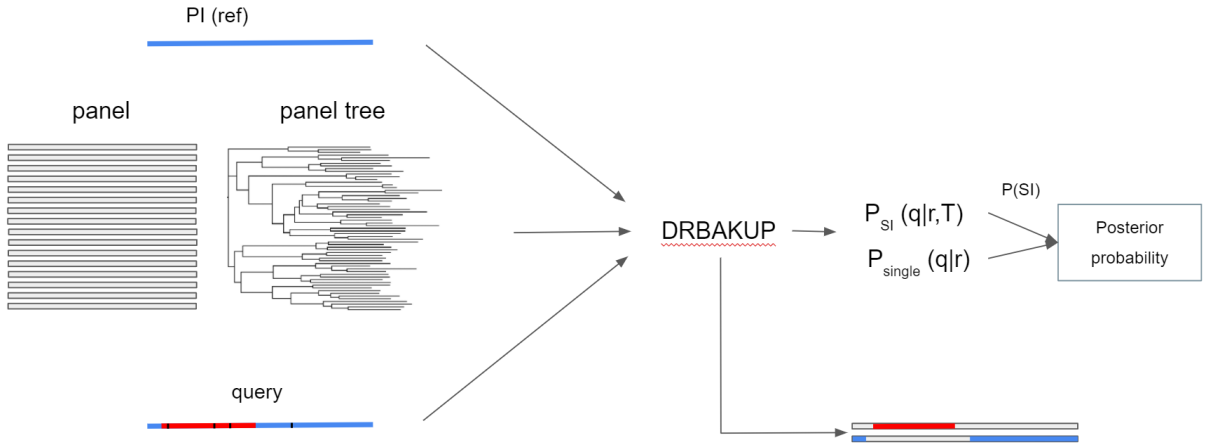


Figure 4.5 Overview of DRBAKUP.

As inputs, DRBAKUP takes a reference sequence (PI), a query sequence of interest, and a reference alignment and tree of relevant "background" sequences. Given a prior probability of superinfection, DRBAKUP outputs a posterior probability. Query sequences can also be partitioned according to the HMM state path probabilities.

Simulated data

To benchmark the performance of the model, we assembled simulated data reflecting true-positive and true-negative cases of recombination with an unobserved parent. Recombined sequences were simulated by randomly selecting two sequences from an alignment (HIV subtype C *env*, downloaded from the Los Alamos HIV database with the hypervariable and gappy regions stripped out), and replacing either one or two fragments with the corresponding fragment in the other sequence (Figure 4.6). The sizes of recombinant fragments were varied from 100bp - 500bp.

Recombined sequences were then evolved *in silico*, with the approximate number of years of evolution introduced varied from 1-5 years. Evolution was simulated, using code adapted with permission from Hassan Sadiq (UCT, South Africa), under a mutation-selection model (Halpern and Bruno, 1998), with site-specific amino acid equilibrium frequencies derived from empirical amino acid frequencies in a panel of 500 subtype C *env* sequences downloaded from the LANL HIV sequence database. An underlying HKY85 nucleotide mutation model (Hasegawa et al., 1985) was used, with codon site-specific evolutionary rates taken from site-specific estimates from FUBAR (Murrell et al., 2013), obtained from the same 500 sequence panel. Insertions and deletions were not simulated. Branch lengths were scaled to approximate the expected evolution of HIV *env* (±1% per annum) (Krakoff et al., 2019).

At each combination, 50 unique sequences were simulated. This was then iteratively repeated for two new parental sequences. The total simulated test set consisted of 2750 unique query sequences. Importantly, sequences used for the simulation were excluded from the reference panel when evaluating the performance of DRBAKUP, reflecting the use case where the superinfecting parent is unobserved.

.

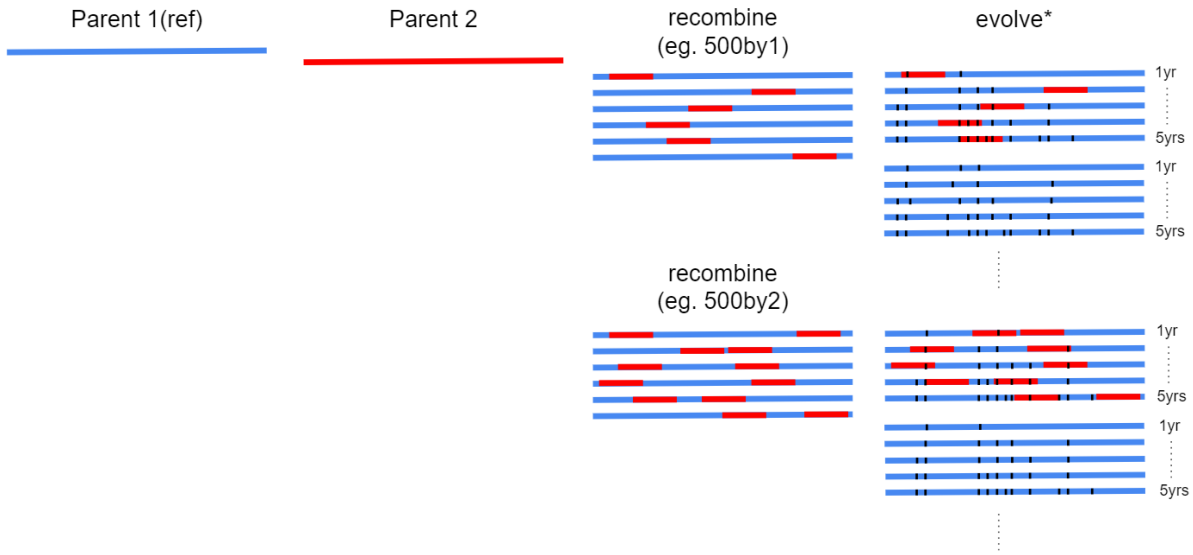


Figure 4.6 Schematic outlining the generation of simulated HIV *env* sequences.

Two *env* sequences, designated "parent 1" (blue) and "parent 2" (red) were randomly selected from the alignment, and recombinants of these were generated *in silico* (red/blue chimeras). These simulated recombinants were then evolved by introducing mutations (black ticks). The number of mutations were scaled to introduce approximately 1% divergence per simulated year of evolution (1yr to 5yrs).

Empirical data

Sequence alignments from Piantadosi et al. (2008) and Tsui et al. (2004) were obtained via Genbank, and consisted of longitudinal *gag* and *env* alignments from 35 and 37 individuals respectively. Unpublished sequences spanning infection until 3 years post infection from 15 individuals in the CAPRISA 002 cohort were provided by Melissa-Rose Abrahams (University of Cape Town) and curated along with published datasets. Sequences were aligned to a subtype reference alignment (LANL) and, in the case of *env* alignments, hypervariable regions (V1V2, V4, V5) and columns that spanned indels were stripped from the alignment. For the calculation of posterior probabilities, we used prior probabilities of superinfection of 7% (CAPRISA 002) or 8% (Piantadosi et al. 2008) per year, based on the risk of HIV infection in these cohorts. For the analysis of sequences from Tsui et al. (2004), we used their own estimate of 3.4 expected cases of superinfection to construct a prior.

Results

Simulated data

To determine sensitivity and specificity, we initially benchmarked DRBAKUP on a panel of simulated data. Cases of superinfection were modelled by randomly selecting sequences (all were subtype C, *env*), artificially recombining them at randomly-selected breakpoints generating recombinants of a defined size, and then evolving these sequences *in silico* (Figure. 4.6). We introduced one or two recombinant fragments of sizes 100bp-500bp, and introduced up to 5 years of evolution. For a negative control set, only sequence evolution was introduced.

DRBAKUP was able to efficiently discriminate superinfected from non-superinfected sequences (Figure 4.7), and at a threshold posterior probability of 0.9, sensitivity was 74% and specificity was 100% across all simulated data. DRBAKUP was highly specific and did not detect any false positives in the simulated data, with all posterior probabilities of superinfection < 0.1 in the negative control set, even with 5 years of evolution introduced (Figure 4.7a). When 500bp was inherited from a second source, DRBAKUP approached 100% power (Figure 4.7b-c). On a challenging subset, with simulated cases of superinfection in *env* sequences where only 200bp was inherited from a second source, this approach was capable of detecting ~60% of true cases. Importantly, randomly simulated short recombinant fragments often span highly conserved regions, providing little signal. Indeed, using a stepping sliding window to generate recombinants with defined (rather than random, as above) breakpoints, we show that power to detect recombination is strongly correlated with the sequence variability within the panel over the recombining region (Figure 4.8).

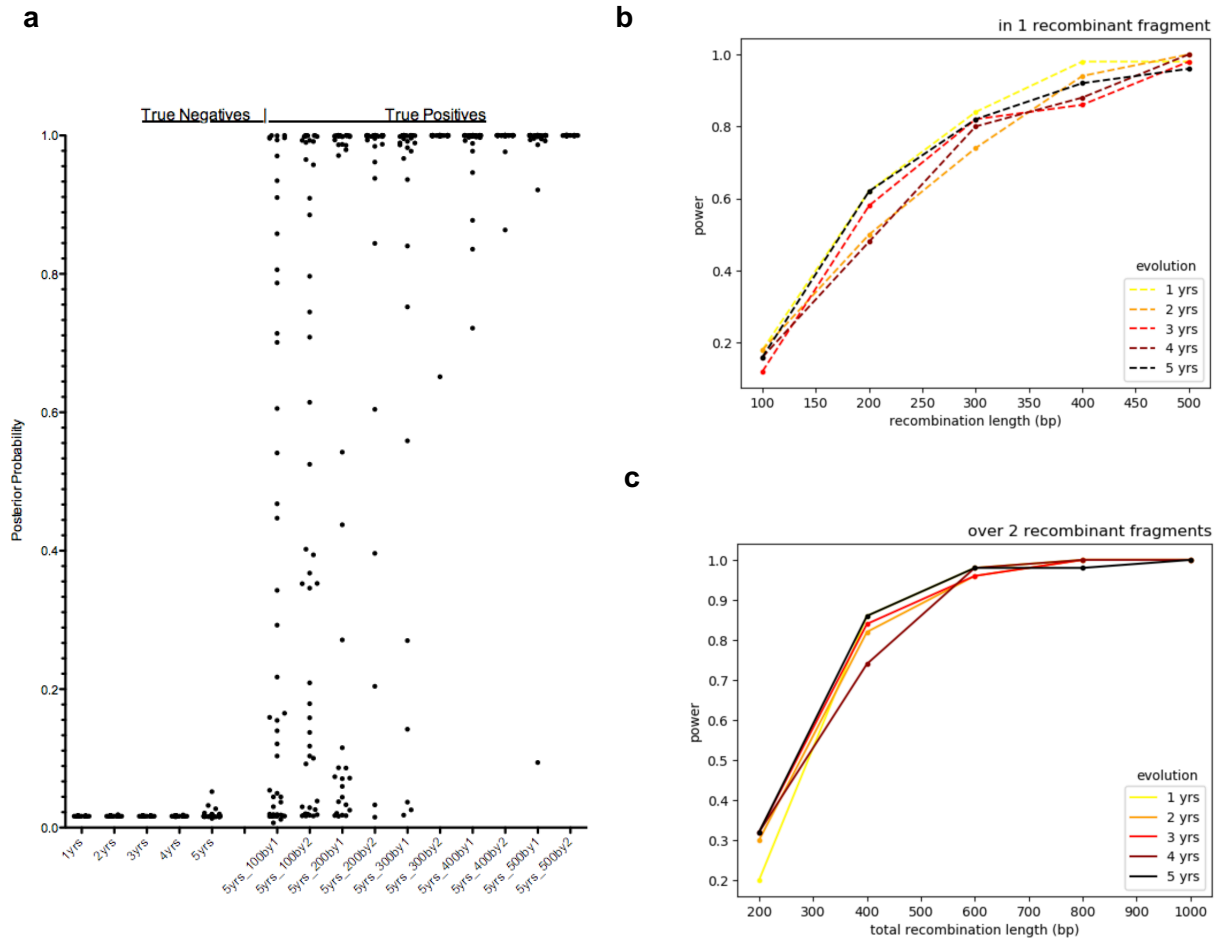


Figure 4.7 Power to detect superinfection following recombination in simulated data.

This simulation chose two sequences, recombined them, and then evolved them *in silico*. The SI fragment size (100bp to 500bp), the number of fragments (1 or 2), and the approximate number of years of evolution (1 to 5) were varied. The amount of evolution appears not to substantially affect power, which is principally determined by the number and size of the SI fragments.

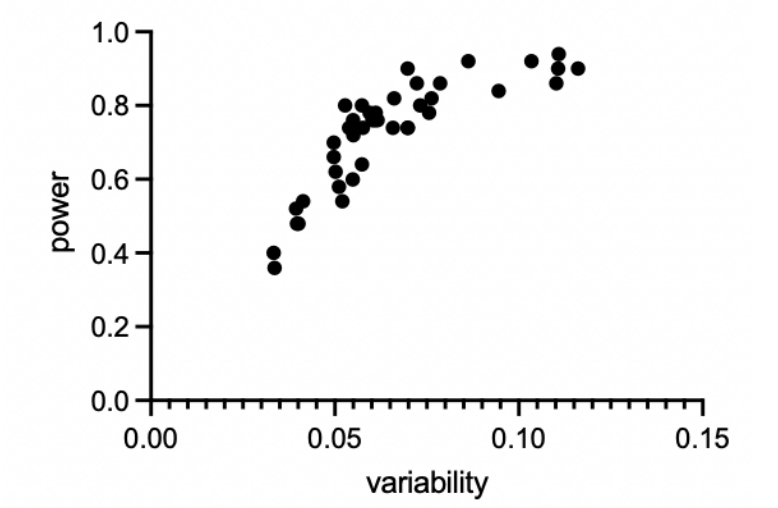


Figure 4.8 Power to detect superinfection in simulated cases is highly correlated with the variability at recombining locations.

The power to detect simulated cases of superinfection then recombination was evaluated by introducing 300bp recombinant fragments across the *env* gene from unlinked variants. Power (proportion of true simulated positives detected with posterior probability >0.9) was strongly correlated with the sequence variability in the reference panel within the corresponding window. Here, window variability was calculated as the proportion of sequences with any mismatches to the consensus within the corresponding window.

Curated data

We previously identified three superinfected individuals in the CAPRISA 002 Acute Infection cohort based on longitudinal sequencing of *env* (Sheward et al., 2015). All infections and superinfections occurred with subtype C viruses. Applying DRBAKUP to these alignments, using the PI as the reference, correctly identified all three of these participants, with posterior probabilities ~1.0 in all cases (Table 4.1), confirming the ability to detect superinfection in real, longitudinal sequence data. Furthermore, DRBAKUP partitioned the sequences such that the PI fragments, and the "other" fragments clustered in the phylogeny with the PI and the SI respectively, indicating the sequence partitions were appropriate, and providing further support for superinfection (as in Figure 4.6).

Of six participants from the CAPRISA 002 cohort with frequent, longitudinal sequencing that had been extensively characterized previously, with no evidence of multiple infection (CAP8, CAP45, CAP88, CAP206, CAP248, CAP257), DRBAKUP produced posterior probabilities of superinfection of <=0.1 for all sequences (sampled over ~ 3 years post infection in each case), suggesting the false discovery rate in real data is low. (Table 4.1).

Table 4.1. Results from applying DRBAKUP to curated alignments from well-characterized participants in the CAPRISA 002 cohort.

ID	*multiplicity**	*max PP (DRBAKUP)*
CAP8	single	0.1
CAP45	single	0.1
CAP88	single	0.1
CAP206	single	0.1
CAP248	single	0.1
CAP257	single	0.1
CAP177	multi-variant	0.5
CAP37	co-infection	**1.0**
CAP84	co-infection	**1.0**
CAP137	co-infection	**1.0**
CAP200	co-infection	**1.0**
CAP267	co-infection	0.8
CAP237	superinfection	**1.0**
CAP256	superinfection	**1.0**
CAP281	superinfection	**1.0**

*"Known" multiplicity of infection based on intensive longitudinal sampling and sequencing from infection until 3 years post infection.

max PP, maximum posterior probability for a query sequence from that donor ID.

Interestingly, one individual (CAP177) with evidence of infection with a diverse quasispecies from a single donor (multi-variant, rather than dual/co-infection) (Anthony et al., 2017), produced a maximum posterior probability of dual infection of 0.5. Indeed, by reducing the expected branch-length it is possible to detect these cases of "multi-variant" infection with diverse variants as well, where it is biologically meaningful to identify these.

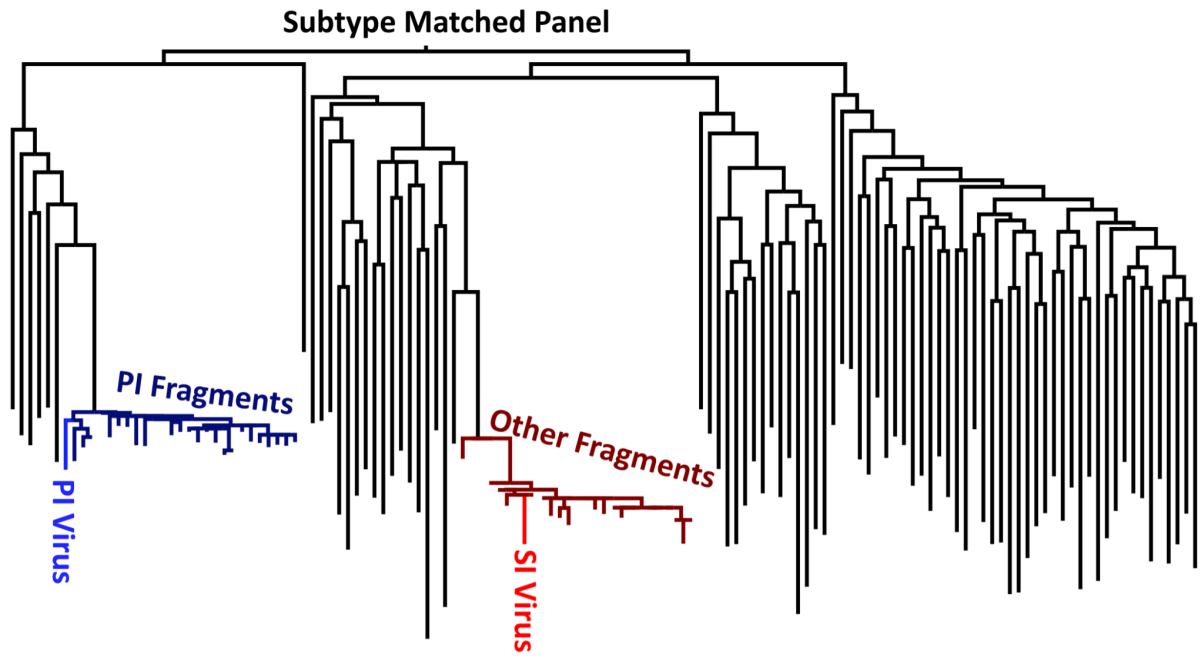


Figure 4.6 DRBAKUP can appropriately partition recombinant sequences even when one parent is unknown.

This phylogeny is from the same panel and the same sequences as in Fig 4.1, but here sequences are partitioned, using DRBAKUP, into fragments that match the PI variant, and fragments that come from elsewhere ("Other"). Critically, the full SI variant was not provided to the algorithm, which used a small panel of sequences to both identify evidence of superinfection, and to partition the recombinant sequences into PI and "Other" fragments. The clustering of the "Other" fragments with the sampled (known) SI variant suggests that these splits were appropriate.

Applying DRBAKUP to publicly available data from previous screens for superinfection

As HIV superinfection is frequently followed by extensive recombination, our model that explicitly accounts for recombination should be significantly more powerful than the current gold standard. One approach to confirm this is to analyse sequences from cases of superinfection where the traditional approach only detects superinfection in sequences from one gene, but not another. While this may be due to the absence of any real signal for superinfection in the sequences the other gene, in a subset of cases this may reflect the lack of power of traditional approaches, and detection of superinfection in both genes would therefore suggest greater sensitivity.

Piantadosi et al. (2007) screened 36 individuals over a median of 5.2 years of follow up, by sequencing regions of *env* and *gag* and detected 7 cases of superinfection, making the sequence data publicly available via Genbank. Of these seven cases, six were identified based on sequences from one gene, but with no significant evidence in the other. Using DRBAKUP on the four cases that were detected in *env* but not *gag* sequences, we found strong evidence of superinfection within the *gag* sequences in two cases. Of the two cases previously detected in *gag* but not in *env*, DRBAKUP identified superinfection within the *env* sequences in one of these cases. In all three cases, DRBAKUP reports the posterior probability of superinfection ±1.0 (Table 4.2). These highlight that DRBAKUP had greater power than the standard phylogenetic approach used by Piantadosi et al. (and others), to correctly identify 'true' cases of superinfection, in real-world screening.

Another approach to establish improved sensitivity includes cohorts where no cases of superinfection were detected. Given the incidence of superinfection approaching the primary infection incidence in several cohorts, the failure to detect superinfection in these reports likely represent a lack of detection sensitivity rather than a true absence of superinfection. Tsui et al. (2004) screened 37 HIV positive injecting drug users (IDUs) over a total of 215 person years of exposure by sequencing regions of *gag*, *env*, and *tat*, and reported no cases of superinfection. We applied DRBAKUP to the same *gag* and *env* sequences, and identified three cases with strong evidence for superinfection. Donor V harboured strong evidence of superinfection (Figure 4.9) in both *env* (posterior = 0.995) and *gag* (posterior = 0.96) sequences. In *gag* sequences alone, two additional cases were detected [ID (posterior probability)]: E (0.9998), and BB (0.999999).

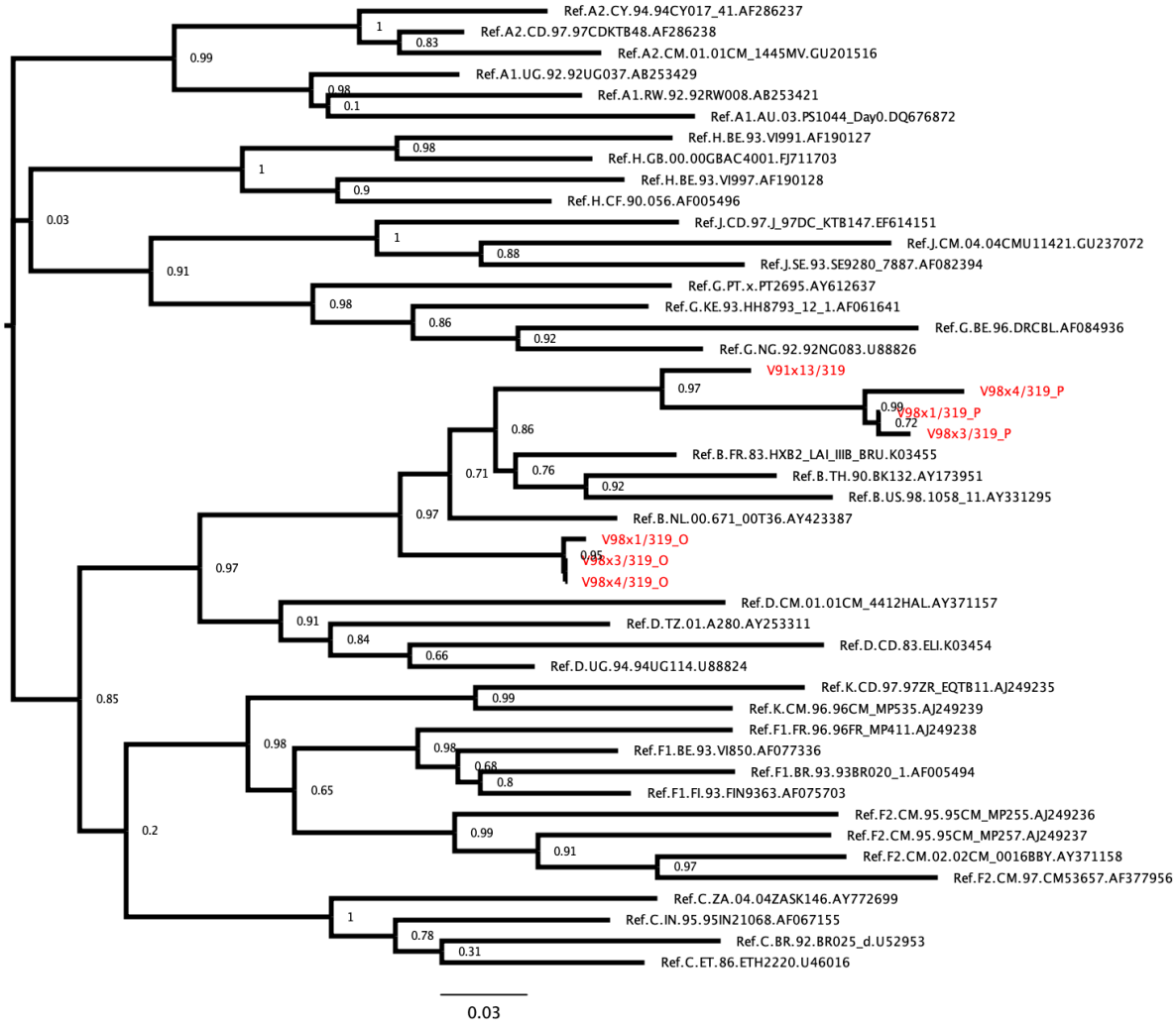


Figure 4.9 Phylogenetic placement of DRBAKUP-partitioned queries and their corresponding reference sequence from donor V

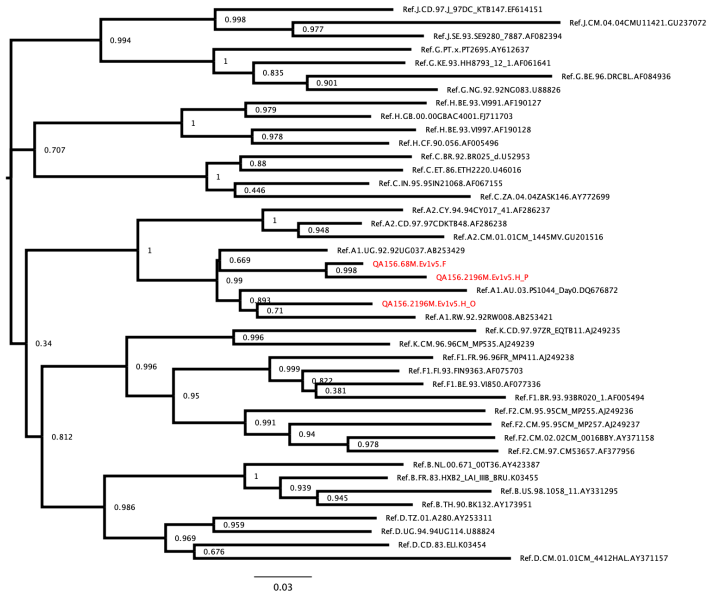
Phylogenies placing DRBAKUP produced partitions on the HIV *env* subtype reference alignment demonstrating that fragments from query sequences sampled from donor V (Tsui et al., 2004) are likely inherited from two epidemiologically unlinked sources, providing evidence of superinfection in *env* sequences that went undetected in the original screen. Sequences and partitions from the queried donor are shown in red. Partitions with the suffix "_P" are inferred by DRBAKUP to have been inherited from the know parent (V91x13/319), while those with the suffix "_O" are inferred to have been inherited from an unobserved parent. The phylogeny was inferred with the GTR+CAT model, and branch support was estimated using the Shimodaira-Hasegawa (SH) test (Shimodaira and Hasegawa, 1999), using FastTree 2.1 (Price et al., 2009, 2010).

Given our ability to detect previously unidentified cases of superinfection, we expanded our analysis to the entire dataset from Piantadosi et al., (2007). In addition to the seven cases originally detected (all of whom were identified as superinfected), we detect two additional cases, notably in both *env* and *gag*, with posterior probabilities >0.9 (Table 4.2). IDs of the additional, likely superinfected participants detected were QA156, and QC036. Query sequences providing evidence of superinfection were partitioned according to the most likely recombination breakpoints inferred by the model (as in Figure 4.4). Partitions inferred by DRBAKUP to have been inherited from the reference sequence cluster with the reference on a phylogeny, and the partitions inferred to have been inherited as the result of recombination from a second parent cluster separately on the tree, strongly suggestive of superinfection (Figure 4.10).

Table 4.2 Applying DRBAKUP to previously published *env* and *gag* sequences

ID	*max PP (env)*	*max PP (gag)*	*DRBAKUP summary*	*Piantadosi et al. (2007)*
QA028	0.02	0.04		
QA039	0.57	0.05		
QA078	0.02	0.04		
QA156	**0.99**	**0.96**	env and gag	
QA210	0.15	0.04		
QA255	0.02	0.04		
QA261	0.02	0.05		
QA268	0.02	0.04		
QA281	0.02	0.04		
QA413	**1**	0.04	env only	env only
QA520	0.03	0.04		
QA584	0.02	0.04		
QA692	0.02	0.04		
QA750	0.03	0.04		
QA966	0.02	0.05		
QB045	**1**	**1**	env and gag	env and gag
QB247	0.03	0.05		
QB368	0.02	0.05		
QB374	0.02	0.04		
QB585	0.25	0.04		
QB685	**1**	**1**	env and gag	gag only
QB726	**1**	**1**	env and gag	env only
QB850	**1**	**1**	env and gag	env only
QC036	**1**	**1**	env and gag	
QC152	0.02	0.04		
QC322	0.02	0.05		
QC440	0.02	0.04		
QC511	0.02	0.05		
QC805	0.02	0.05		
QC885	0.02	**1**	gag only	gag only
QD022	**1**	0.04	env only	env only
QD121	0.02	0.04		
QD224	0.03	0.07		
QD385	0.02	0.04		
QF307	0.02	0.05		

a



b

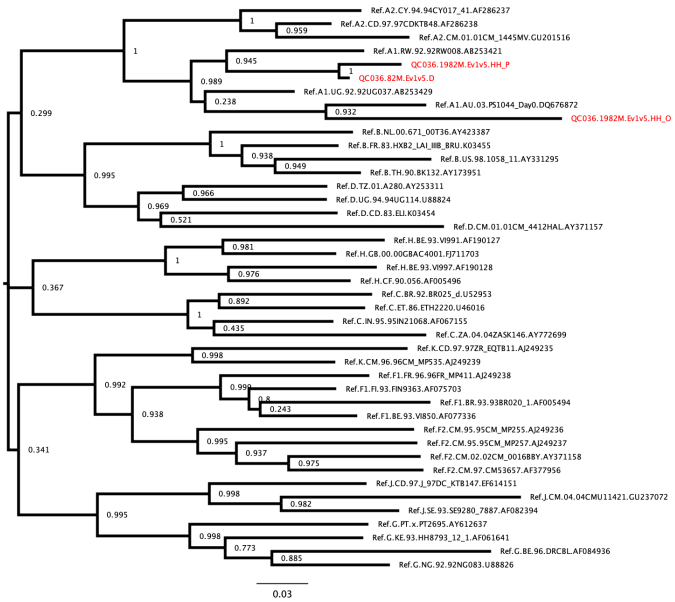


Figure 4.10 Env phylogenies for query sequences from suspected cases of superinfection from Piantadosi et al. (2007), partitioned according to likely breakpoints inferred by DRBAKUP.

a,b Partitions likely inherited from another parent (sequences with the suffix "_O") cluster separately on the phylogeny in two newly identified cases of superinfection even when using a small, sparse, subtype reference panel. The placements strongly suggest that the fragments were inherited from two different sources, providing supporting evidence for superinfection. Nodes are labelled with the reliability estimates for each split, using the SH test in FastTree. Partitions with the suffix "_P" are inferred by DRBAKUP to have been inherited from the known parent (QC036.82M.Ev1v5.D), and cluster with this sequence on the phylogeny.

Discussion

Cases of superinfection have offered an important opportunity to address correlates of protection against HIV, in humans, and in a natural setting. While in some studies the incidence of superinfection was similar to the primary infection incidence in the same or similar cohorts (Redd et al., 2014a; Smith et al., 2004; Wagner et al., 2014), others detected superinfection at a significantly lower rate (Gonzales et al., 2003; Redd et al., 2014b; Ronen et al., 2013; Tsui et al., 2004), and suggested a partial protective effect of immune responses, or at least behavioural changes following diagnosis. As a result, significant effort has been made to identify factors that may correlate with HIV superinfection (Basu et al., 2012; Smith et al., 2006; Wagner et al., 2017). Furthermore, reports of the broadening of neutralizing antibody responses following superinfection (Cortez et al., 2012; Powell et al., 2010) have generated interest in cases of superinfection to be used as a model to inform prime-boost vaccines aimed at eliciting bnAbs. The strength and accuracy of these analyses rely on the sensitive and accurate identification of superinfection.

Most current strategies used to detect cases of superinfection leverage sequencing of HIV populations at two or more points in infection. However, no standardized, robust, automated methodology to detect HIV superinfection within sequence data has been available. Here, we developed a Bayesian approach (DRBAKUP) to solve the general problem of identifying recombination between one known and one unknown parent, and apply this to detect cases of superinfection even after recombination.

We applied DRBAKUP to simulated data, showing that it is highly specific, and sensitive. Where simulated recombinants were not detected, they were typically those where only small and conserved fragments were inherited, thus offering little signal. While distance-based metrics to detect superinfection may have difficulty when screening sequences a number of years following initial infection, DRBAKUP was largely robust to the amount of evolution introduced.

Applying DRBAKUP to previously published, publicly available, sequence datasets, we are able to detect cases of superinfection that previously went undetected, suggesting that the reduced incidences reported in these (and likely other) cohorts were, at least partially, due to a lack of sensitivity in the detection methods used in those studies.

High specificity is particularly important for novel, potentially more sensitive methodologies in order to have confidence that these newly identified cases do not simply represent false positives rather than previously undetected true cases. It is therefore important here to highlight that DRBAKUP produced no false positives in simulated data, even with large amounts of evolution introduced with a model of evolution that favours mutation towards the reference panel equilibrium frequencies, nor in real longitudinal datasets where multiplicity of infection was known. Nevertheless, with a posterior threshold of 0.9, DRBAKUP's behaviour was conservative, favouring high specificity even at the expense of reduced sensitivity. For applications where specificity is less critical (eg. rapid screening with a downstream confirmatory step), more power could be gained by reducing the detection threshold.

Furthermore, DRBAKUP can partition query sequences according to their inferred parent - and fragments inferred to have been inherited from SI variants were placed separately from the PI sequences on a phylogeny in each newly detected case here, providing substantial evidence that DRBAKUP is identifying true cases of superinfection that went undetected.

Identifying factors that prevent superinfection of infected individuals remaining at high risk can help clarify the immune correlates of protection from HIV infection. We show here that recombination confounds the detection of HIV superinfection using conventional phylogenetic approaches, suggesting that superinfection is likely occurring more frequently than previously estimated. Taken together with previous studies (Redd et al., 2014a; Smith et al., 2004; Wagner et al., 2014), this suggests that the immune responses typically generated upon HIV infection may offer limited protection from re-infection.

HIV superinfection may also have important implications for clinical care and therefore public health messaging and interventions. However, the impact of superinfection on disease progression remains unclear, also largely due to small sample sizes (and many confounding variables). The availability of a robust, standardized, sensitive and specific detection algorithm will hopefully lead to the identification of larger numbers of superinfection cases, allowing these and other outstanding questions to be addressed.

DRBAKUP is agnostic of sequencing platform and can be applied to Sanger or high-throughput sequences. Practically, DRBAKUP requires the rational selection of a reference sequence that best represents the PI virus, an alignment of query sequences from the same individual, and a reference alignment and tree of relevant "background" sequences. Given a prior probability of superinfection – which we suggest should be estimated based on risk of primary infection in the same or an analogous cohort - DRBAKUP can provide a posterior

probability of dual infection, as well as partition the recombinant sequences based on their inferred breakpoints.

Limitations and future work

HIV *env* is subject to frequent insertions and deletions, and harbours regions of hypervariability that cannot be reliably aligned. If not accounted for, this has the potential to confound the analysis, potentially producing false positive signal for superinfection. Here, columns within hypervariable regions were manually removed, but approaches should be developed to automatically identify and remove columns that cannot be reliably aligned from input data, prior to analysis. Despite offering a significant signal, insertions and deletions outside of hypervariable regions are currently not modelled but are excluded from DRBAKUP inference in the same way that gaps are handled during phylogenetic inference (ie. as missing data). In future work, we will investigate the impact of indels via simulation. Further, we will adapt DRBAKUP to efficiently handle high-throughput sequencing data, automatically identifying appropriate reference and query sequences, and aligning these to a reference panel, which should make it easier to use.

Extended evolution can lead to multiple successive rounds of recombination, which would manifest in the sequence alignments as a larger number of breakpoints. While the switching rate is learned from the data, a very large number of breakpoints will dilute the signal. However, the HMM switching approach is still able to exploit information from the whole sequence, rather than a single short fragment, which would be the case for sliding window approaches when the number of breakpoints is large.

DRBAKUP appears to perform surprisingly well even using a diverse, multi-subtype panel rather than a geographically-relevant alignment. The influence of the relatedness and diversity of the sequences included in the reference panel on DRBAKUP should be investigated further. Additionally, the ability to accurately differentiate superinfection from infection with divergent founding variants transmitted from the same donor is important and warrants further investigation. A limitation here is that we also do not directly compare the performance of DRBAKUP with sliding-window based approaches (Wagner et al., 2013) to account for recombination.

It is important for bioinformatics tools to be simple to use, and provide biologically meaningful outputs. To this end, we plan to release an open-source, user-friendly version of DRBAKUP for the HIV research community, after addressing the above limitations.

CHAPTER 5

Discussion

Are broadly neutralizing antibodies essential for an effective HIV vaccine?

Circulating nAbs produced by long-lived plasma cells can provide sterilizing immunity against reinfection with a pathogen. However, highly variable pathogens often carry mutations that afford escape from antibody responses. In such cases, natural protection from reinfection is thought to be mediated most often by memory B cells, that represent a rapidly inducible source of antibodies (Oh et al., 2019). While reactive memory B cells may not be able to prevent infection in some cases, they can still substantially reduce viral load and prevent disease (Purtha et al., 2011). For HIV, there is significant evidence that nAbs, if present at the time of infection, can prevent infection. Passive transfer of nAbs can protect non-human primates from challenge with HIV/SIV chimeric viruses (SHIV) (Gautam et al., 2016; Mascola et al., 2000; Moldt et al., 2012; Shingai et al., 2014), and protection in these studies correlated strongly with neutralization (Pegu et al., 2019). Recently, two clinical trials (HVTN703 and HVTN704) directly evaluated whether intravenous infusions of a neutralizing mAb could prevent HIV acquisition. Although there was no significant reduction in HIV acquisition overall compared to placebo, there was evidence of protection against infection with viruses highly sensitive to the passively administered mAb, VRC01 (Corey et al., 2021). The HIV field is thus highly invested in characterizing the development of antibodies capable of neutralizing a broad range of circulating viruses (bnAbs), and eliciting these types of responses through vaccination. However, it is still not clear whether non-neutralizing antibodies, memory B cell responses, or other arms of the immune response may afford any protection against HIV infection, or whether pre-existing nAbs against the transmitting variant will be absolutely required.

It is possible to address correlates of protection directly in clinical trials or animal challenge models. However, these types of studies are costly, and technically challenging. One complementary approach to address the correlates of protection to HIV infection (or at least identify what responses are not sufficient for protection), in humans, is to determine whether immune responses generated during HIV infection offer any protection from a second HIV infection (superinfection).

While some studies found that the incidence of superinfection was similar to the incidence of primary infection (Redd et al., 2014a; Smith et al., 2004; Wagner et al., 2014), others found

that the incidence of superinfection was significantly reduced (Gonzales et al., 2003; Redd et al., 2014b; Ronen et al., 2013; Tsui et al., 2004), even in the absence of detectable nAbs to the superinfecting viruses. These results may reflect immunological protection, global reductions in risk of HIV infection over time, or behavioural changes following initial diagnosis. However, another potential contributor is that cases of superinfection are practically challenging to detect due, in part, to recombination, leading to their frequency being underestimated.

In Chapter 4, we describe a novel, automated approach to detect superinfection that is robust to recombination. Using simulated and real sequence data, we demonstrated that this approach is sensitive, and highly specific. Furthermore, analyzing previously published sequence datasets, we identified cases of superinfection that previously went undetected, indicating that superinfection occurs more frequently than previously estimated. This frequency of superinfection highlights that the immune responses typically generated upon HIV infection, which includes a relatively broad diversity of memory B cells (Scheid et al., 2009) and reasonably robust antibody (Moore, 2018) and CTL (Ndhlovu et al., 2015) responses, are insufficient to provide protection from re-infection.

The potential for antigen cocktails and sequential vaccination to elicit broadly neutralizing antibodies

The development of stabilized, soluble trimers that mimic the native functional HIV trimer and elicit strain-specific neutralizing HIV antibodies in animal models represents a critical advance in HIV vaccine development. However, a single immunogen appears unlikely to be sufficient to drive the development of broadly neutralizing HIV antibodies in humans. Identifying how multiple immunogens could be administered to facilitate the broadening of responses represents a pivotal challenge.

A similar challenge faces the influenza field, and results from sequential and polyvalent haemagglutinin immunizations could potentially inform HIV immunizations. Incorporation of multiple, antigenically diverse haemagglutinins into nanoparticles elicited broad antibody responses (Kanekiyo et al., 2019). Sequential immunization with multiple divergent haemagglutinins has also led to the development of broadly neutralizing influenza antibodies in mice (Wang et al., 2010). In humans, a number of broadly neutralizing influenza-specific plasmablasts were identified that appeared to have originated from the memory B cell compartment following H1N1 infection in individuals with no prior exposure to H1N1 (Wrammert et al., 2011). Together, these studies suggest that sequential exposure to

heterologous antigens from variable pathogens can promote the development of cross-neutralizing antibodies, potentially mediated through the recruitment of cross-reactive memory B cells.

In Chapter 2, we characterized the development of antibodies in five superinfected individuals, and evaluated whether superinfection re-engaged cross-reactive memory B cells, promoting the development of bnAbs. By comparing the breadth of the neutralizing antibody response in superinfected individuals to those that typically develop in singly infected individuals, we showed that HIV superinfection was not sufficient to broaden responses. However, given the small sample size and large variability, the absence of a significant correlation should be interpreted with caution. We therefore focussed on developing a mechanistic understanding of the neutralizing antibody response following sequential exposure. By characterizing the kinetics and specificity of autologous neutralizing antibody responses, we show that responses to the superinfecting viruses failed to efficiently recruit neutralizing memory B cells. Instead, the secondary infection elicited strain-specific, *de novo* responses. This occurred even though the superinfecting viruses were relatively close amongst the scale of HIV diversity (from the same subtype). The observation that memory B cells were not efficiently recruited such that the development of nAbs to the superinfecting virus was not accelerated also reinforces the notion that relying on reactive humoral immune memory is unlikely to be sufficient for protection against HIV.

Previous studies have claimed that superinfection is associated with broadening the nAb response (Cortez et al., 2012; Powell et al., 2010), though others saw no statistically significant effect (Cornelissen et al., 2016). Where a significant broadening was observed in previous studies, however, the effect size was small, with superinfected individuals approximately 1.5 times more likely to develop breadth than singly infected individuals (Cortez et al., 2012). This effect size would be consistent with a near-additive effect of superinfection on nAb breadth, such as that we describe in Chapter 2, rather than directing the nAb response towards conserved epitopes. This is an important distinction, as stimulating independent antibody responses to each of two immunizing immunogens has restricted value in the context of a goal to elicit bnAbs to HIV by vaccination.

In Chapter 3, to determine whether the co-exposure to diverse Env antigens favours the development of cross-neutralizing antibodies better than sequential exposure, we characterized the development of nAbs in three HIV co-infected individuals where several divergent viruses were transmitted prior to seroconversion. We wanted to go further, and identified cases of co-infection that represented distinct scenarios of antigen exposure. We

characterized three cases that encompassed immunological exposure to: (i) two diverse, unlinked Envs, (ii) two related Envs with diversity uniformly distributed over the trimer, and (iii) two diverse but recombined Envs such that clusters of high homology were preserved in the presence of high diversity elsewhere.

We found that, like superinfection, co-infection was not sufficient to broaden neutralizing antibody responses. Co-exposure to two HIV Env antigens did not necessarily produce additive or cross-neutralizing antibody responses, and in some cases was subject to immunological interference. This was most evident in the case of co-infection with two related Envs where diversity was uniformly distributed across the Env trimer; in this case neutralizing antibody responses to one variant arose to the near exclusion of responses to the other. However, in the case of co-exposure to diverse Envs but where the trimer apex was conserved in both variants through recombination, potent neutralization of both variants was evident. This was the co-infected participant who developed the broadest neutralizing antibody response, and we show that cross-neutralization was mediated, in part, by trimer apex-targeting nAbs. Taken together, these data suggest that cocktails of divergent stabilized Env trimers are unlikely to drive the development of cross-neutralizing antibodies, and may be subject to interference. However, they also provide hope that the rational design of more similar immunogen cocktails where conserved epitopes are preserved across immunogens may be able to facilitate nAbs to these targets.

Interestingly, in Chapter 2, the *de novo* nAb responses that arose in two superinfected individuals to their superinfecting viruses were the most potent responses observed in the cohort. Immune complexes are known to be more immunogenic than antigen alone (Brady L. Jeannine, 2005), and presentation of superinfecting antigen would have occurred in the context of (non-neutralizing) antibodies produced in response to primary infection. Indeed, in a therapeutic trial, passive therapy with an HIV-specific, broadly neutralizing mAb significantly improved subsequent neutralizing antibody responses in all but one of 15 participants (Schoofs et al., 2016), potentially mediated by a similar mechanism. This may also suggest that co-administration of virus-like particles (VLPs) incubated with polyclonal non-neutralizing antibodies may represent a strategy to both obscure non-neutralizing epitopes and simultaneously enhance the immunogenicity of VLP formulations. Identifying the mechanism underlying the potent response to superinfecting viruses could also be generalized, facilitating the improvement of future vaccines.

While we show that HIV co- and superinfection have the potential to provide important insights for vaccine design, we acknowledge several limitations. As HIV superinfection is a rare event,

we (and others) have been limited by small sample sizes. Large-scale screening of multiple longitudinal cohorts for cases of superinfection together with more sophisticated screening approaches, will be required to establish larger sample sizes, and could provide extremely valuable insights for sequential immunizations. Our development of a sensitive and specific, automated algorithm for the detection of co- and superinfection, as described in Chapter 4, will hopefully catalyse this future work.

Furthermore, while efficient re-engagement of memory B cells was not evident from plasma neutralizing antibody responses, it is possible that some memory B cell populations were recruited but did not produce sufficient amounts of nAbs to be detectable. Sequencing of HIV-specific B cells before and after superinfection in the future could identify whether superinfection reactivates any minor memory B cell populations, or whether superinfection leads exclusively to *de novo* responses.

It is also important to consider the on-going HIV replication and immune dysfunction present in HIV-infected individuals that would not be present in the vaccine setting. However, a mitigating factor here was that all superinfected participants characterized in Chapter 2 were superinfected within the first year of infection and had relatively high CD4 counts (all >450 cells/μL) at the time of superinfection. Similarly, the co-infected participants were characterized over the first year of infection, and all maintained high CD4 counts over this period. Results from vaccine models to date have also been consistent with these observations. In mice, upon sequential exposure to antigen, most B cells in secondary germinal centre reactions were naive, and not recruited from the primed memory B cell compartment (Mesin et al., 2020). Polyvalent and sequential immunization of rabbits with two, three, or four stabilized HIV Env trimers also elicited nAbs to each of the Envs in many of the animals, but no significant broadening of the responses to Envs not included in the immunization was evident (Klasse et al., 2016; de la Peña et al., 2018). Similarly, co-immunization of rabbits with two stabilized trimers (derived from subtype A and subtype B Envs that shared a common “glycan hole”) elicited two independent neutralizing antibody responses to each antigen (Yang et al., 2020). Env cocktails arrayed in mixtures at high density on liposomal particles have elicited some serum cross neutralizing activity, although even then only in two out of twelve rabbits and and only after five or six immunizations, and corresponding monovalent formulations were not tested (Dubrovskaya et al., 2019).

Conclusion

In conclusion, we find that HIV superinfection fails to efficiently recruit neutralizing memory B cells and, at best, results in additive nAb responses rather than a synergistic effect leading to cross-neutralization; a distinction that is highly relevant for vaccine design. While sequential immunizations with heterologous Env immunogens may be able to improve the potency of elicited responses, alone, they are unlikely to promote the development of bnAbs.

Our observations from co-infected individuals also suggest that heterologous Envs within a cocktail are likely too diverse to stimulate a single cross-neutralizing response - rather they are likely to drive multiple independent responses (additive at best). The use of more similar trimers may be required. Importantly the neutralizing antibody responses to related Envs within a cocktail can be subject to interference, where nAb responses are elicited to one antigen but not the other, when Envs may be similar enough to compete for CD4 help but where B cell epitopes that could mediate cross-neutralization are sparse. However, when diversity is maintained in immunogenic regions, nAbs may be directed towards more conserved epitopes. This suggests that the use of related, stabilized Env trimers with diversity introduced in key regions together with strategies to reduce the immunogenicity of immunodominant, strain-specific epitopes may represent one path to a cross-neutralizing antibody response to multiple Envs within a cocktail.

References

Abdool Karim, Q., Abdool Karim, S.S., Frohlich, J.A., Grobler, A.C., Baxter, C., Mansoor, L.E., Kharsany, A.B.M., Sibeko, S., Mlisana, K.P., Omar, Z., et al. (2010). Effectiveness and safety of tenofovir gel, an antiretroviral microbicide, for the prevention of HIV infection in women. Science *329*, 1168–1174.

Abrahams, M.-R., Anderson, J.A., Giorgi, E.E., Seoighe, C., Mlisana, K., Ping, L.-H., Athreya, G.S., Treurnicht, F.K., Keele, B.F., Wood, N., et al. (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*, 3556–3567.

Andrabi, R., Voss, J.E., Liang, C.-H., Briney, B., McCoy, L.E., Wu, C.-Y., Wong, C.-H., Poignard, P., and Burton, D.R. (2015). Identification of Common Features in Prototype Broadly Neutralizing Antibodies to HIV Envelope V2 Apex to Facilitate Vaccine Design. Immunity *43*, 959–973.

Anthony, C., York, T., Bekker, V., Matten, D., Selhorst, P., Ferreria, R.-C., Garrett, N.J., Karim, S.S.A., Morris, L., Wood, N.T., et al. (2017). Cooperation between Strain-Specific and Broadly Neutralizing Responses Limited Viral Escape and Prolonged the Exposure of the Broadly Neutralizing Epitope. J. Virol. *91*.

Basu, D., Kraft, C.S., Murphy, M.K., Campbell, P.J., Yu, T., Hraber, P.T., Irene, C., Pinter, A., Chomba, E., Mulenga, J., et al. (2012). HIV-1 subtype C superinfected individuals mount low autologous neutralizing antibody responses prior to intrasubtype superinfection. Retrovirology *9*, 76.

Bezanson, J., Karpinski, S., Shah, V.B., and Edelman, A. (2012). Julia: A Fast Dynamic Language for Technical Computing.

Bhiman, J.N., Anthony, C., Doria-Rose, N.A., Karimanzira, O., Schramm, C.A., Khoza, T., Kitchin, D., Botha, G., Gorman, J., Garrett, N.J., et al. (2015). Viral variants that initiate and drive maturation of V1V2-directed HIV-1 broadly neutralizing antibodies. Nature Medicine *21*, 1332–1336.

Binley, J.M., Sanders, R.W., Clas, B., Schuelke, N., Master, A., Guo, Y., Kajumo, F., Anselma, D.J., Maddon, P.J., Olson, W.C., et al. (2000). A recombinant human immunodeficiency virus type 1 envelope glycoprotein complex stabilized by an intermolecular disulfide bond between the gp120 and gp41 subunits is an antigenic mimic of the trimeric virion-associated structure. J. Virol. *74*, 627–643.

Bonsignori, M., Hwang, K.-K., Chen, X., Tsao, C.-Y., Morris, L., Gray, E., Marshall, D.J., Crump, J.A., Kapiga, S.H., Sam, N.E., et al. (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.

Bonsignori, M., Kreider, E.F., Fera, D., Ryan Meyerhoff, R., Bradley, T., Wiehe, K., Munir Alam, S., Aussedat, B., Walkowicz, W.E., Hwang, K.-K., et al. (2017). Staged induction of HIV-1 glycan–dependent broadly neutralizing antibodies. Sci. Transl. Med.

Borrow, P., and Moody, M.A. (2017). Immunologic characteristics of HIV-infected individuals who make broadly neutralizing antibodies. Immunol. Rev. *275*, 62–78.

Brady L. Jeannine (2005). Antibody-Mediated Immunomodulation: a Strategy To Improve Host Responses against Microbial Antigens. Infect. Immun. *73*, 671–678.

Briney, B., Inderbitzin, A., Joyce, C., and Burton, D.R. (2019). Commonality despite exceptional diversity in the baseline human antibody repertoire. Nature *566*, 393–397.

Burton, D.R., Hessell, A.J., Keele, B.F., Klasse, P.J., Ketas, T.A., Moldt, B., Dunlop, D.C., Poignard, P., Doyle, L.A., Cavacini, L., et al. (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*, 11181–11186.

Caskey, M., Schoofs, T., Gruell, H., Settler, A., Karagounis, T., Kreider, E.F., Murrell, B., Pfeifer, N., Nogueira, L., Oliveira, T.Y., et al. (2017). Antibody 10-1074 suppresses viremia in HIV-1-infected individuals. Nat. Med. *23*, 185–191.

Cheng, C., Pancera, M., Bossert, A., Schmidt, S.D., Chen, R.E., Chen, X., Druz, A., Narpala, S., Doria-Rose, N.A., McDermott, A.B., et al. (2015). Immunogenicity of a Prefusion HIV-1 Envelope Trimer in Complex with a Quaternary-Structure-Specific Antibody. J. Virol. *90*, 2740–2755.

Cheng, C., Xu, K., Kong, R., Chuang, G.-Y., Corrigan, A.R., Geng, H., Hill, K.R., Jafari, A.J., O'Dell, S., Ou, L., et al. (2019). Consistent elicitation of cross-clade HIV-neutralizing responses achieved in guinea pigs after fusion peptide priming by repetitive envelope trimer boosting. PLoS One *14*, e0215163.

Cheng, C., Duan, H., Xu, K., Chuang, G.-Y., Corrigan, A.R., Geng, H., O'Dell, S., Ou, L., Chambers, M., Changela, A., et al. (2020). Immune Monitoring Reveals Fusion Peptide Priming to Imprint Cross-Clade HIV-Neutralizing Responses with a Characteristic Early B Cell Signature. Cell Rep. *32*, 107981.

Coffin, J.M. (1995). HIV population dynamics in vivo: implications for genetic variation, pathogenesis, and therapy. Science *267*, 483–489.

Cohen, K., Altfeld, M., Alter, G., and Stamatatos, L. (2014). Early preservation of CXCR5+ PD-1+ helper T cells and B cell activation predict the breadth of neutralizing antibody responses in chronic HIV-1 infection. J. Virol. *88*, 13310–13321.

Cohen, M.S., Chen, Y.Q., McCauley, M., Gamble, T., Hosseinipour, M.C., Kumarasamy, N., Hakim, J.G., Kumwenda, J., Grinsztejn, B., Pilotto, J.H.S., et al. (2016). Antiretroviral Therapy for the Prevention of HIV-1 Transmission. N. Engl. J. Med. *375*, 830–839.

Corey, L., Gilbert, P.B., Juraska, M., Montefiori, D.C., Morris, L., Karuna, S.T., Edupuganti, S., Mgodi, N.M., deCamp, A.C., Rudnicki, E., et al. (2021). Two Randomized Trials of Neutralizing Antibodies to Prevent HIV-1 Acquisition. N. Engl. J. Med. *384*, 1003–1014.

Cornelissen, M., Euler, Z., van den Kerkhof, T.L.G.M., van Gils, M.J., Boeser-Nunnink, B.D.M., Kootstra, N.A., Zorgdrager, F., Schuitemaker, H., Prins, J.M., Sanders, R.W., et al. (2016). The Neutralizing Antibody Response in an Individual with Triple HIV-1 Infection Remains Directed at the First Infecting Subtype. AIDS Res. Hum. Retroviruses *32*, 1135–1142.

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

Delassus, S., Cheynier, R., and Wain-Hobson, S. (1991). Evolution of human immunodeficiency virus type 1 nef and long terminal repeat sequences over 4 years in vivo and in vitro. J. Virol. *65*, 225–231.

Doria-Rose, N.A., Schramm, C.A., Gorman, J., Moore, P.L., Bhiman, J.N., DeKosky, B.J.,

Ernandes, M.J., Georgiev, I.S., Kim, H.J., Pancera, M., et al. (2014). Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. Nature *509*, 55–62.

Dosenovic, P., von Boehmer, L., Escolano, A., Jardine, J., Freund, N.T., Gitlin, A.D., McGuire, A.T., Kulp, D.W., Oliveira, T., Scharf, L., et al. (2015). Immunization for HIV-1 Broadly Neutralizing Antibodies in Human Ig Knockin Mice. Cell *161*, 1505–1515.

Doyle-Cooper, C., Hudson, K.E., Cooper, A.B., Ota, T., Skog, P., Dawson, P.E., Zwick, M.B., Schief, W.R., Burton, D.R., and Nemazee, D. (2013). Immune tolerance negatively regulates B cells in knock-in mice expressing broadly neutralizing HIV antibody 4E10. J. Immunol. *191*, 3186–3191.

Dubrovskaya, V., Tran, K., Ozorowski, G., Guenaga, J., Wilson, R., Bale, S., Cottrell, C.A., Turner, H.L., Seabright, G., O'Dell, S., et al. (2019). Vaccination with Glycan-Modified HIV NFL Envelope Trimer-Liposomes Elicits Broadly Neutralizing Antibodies to Multiple Sites of Vulnerability. Immunity *51*, 915–929.e7.

Dugast, A.-S., Chan, Y., Hoffner, M., Licht, A., Nkolola, J., Li, H., Streeck, H., Suscovich, T.J., Ghebremichael, M., Ackerman, M.E., et al. (2014). Lack of protection following passive transfer of polyclonal highly functional low-dose non-neutralizing antibodies. PLoS One *9*, e97229.

Dugast, A.-S., Arnold, K., Lofano, G., Moore, S., Hoffner, M., Simek, M., Poignard, P., Seaman, M., Suscovich, T.J., Pereyra, F., et al. (2017). Virus-driven Inflammation Is Associated With the Development of bNAbs in Spontaneous Controllers of HIV. Clin. Infect. Dis. *64*, 1098–1104.

Escolano, A., Steichen, J.M., Dosenovic, P., Kulp, D.W., Golijanin, J., Sok, D., Freund, N.T., Gitlin, A.D., Oliveira, T., Araki, T., et al. (2016). Sequential Immunization Elicits Broadly Neutralizing Anti-HIV-1 Antibodies in Ig Knockin Mice. Cell *166*, 1445–1458.e12.

Flynn, N.M., Forthal, D.N., Harro, C.D., Judson, F.N., Mayer, K.H., Para, M.F., and rgp120 HIV Vaccine Study Group (2005). Placebo-controlled phase 3 trial of a recombinant glycoprotein 120 vaccine to prevent HIV-1 infection. J. Infect. Dis. *191*, 654–665.

Freund, N.T., Wang, H., Scharf, L., Nogueira, L., Horwitz, J.A., Bar-On, Y., Golijanin, J., Sievers, S.A., Sok, D., Cai, H., et al. (2017). Coexistence of potent HIV-1 broadly neutralizing antibodies and antibody-sensitive viruses in a viremic controller. Sci. Transl. Med. *9*.

Gaschen, B., Taylor, J., Yusim, K., Foley, B., Gao, F., Lang, D., Novitsky, V., Haynes, B., Hahn, B.H., Bhattacharya, T., et al. (2002). Diversity considerations in HIV-1 vaccine selection. Science *296*, 2354–2360.

Gautam, R., Nishimura, Y., Pegu, A., Nason, M.C., Klein, F., Gazumyan, A., Golijanin, J., Buckler-White, A., Sadjadpour, R., Wang, K., et al. (2016). A single injection of anti-HIV-1 antibodies protects against repeated SHIV challenges. Nature *533*, 105–109.

Gilbert, P.B., Berger, J.O., Stablein, D., Becker, S., Essex, M., Hammer, S.M., Kim, J.H., and Degruttola, V.G. (2011). Statistical interpretation of the RV144 HIV vaccine efficacy trial in Thailand: a case study for statistical issues in efficacy trials. J. Infect. Dis. *203*, 969–975.

Gnanakaran, S., Daniels, M.G., Bhattacharya, T., Lapedes, A.S., Sethi, A., Li, M., Tang, H., Greene, K., Gao, H., Haynes, B.F., et al. (2010). Genetic signatures in the envelope glycoproteins of HIV-1 that associate with broadly neutralizing antibodies. PLoS Comput. Biol. *6*, e1000955.

Gonzales, M.J., Delwart, E., Rhee, S.-Y., Tsui, R., Zolopa, A.R., Taylor, J., and Shafer, R.W. (2003). Lack of detectable human immunodeficiency virus type 1 superinfection during 1072 person-years of observation. J. Infect. Dis. *188*, 397–405.

Gorman, J., Soto, C., Yang, M.M., Davenport, T.M., Guttman, M., Bailer, R.T., Chambers, M., Chuang, G.-Y., DeKosky, B.J., Doria-Rose, N.A., et al. (2016). Structures of HIV-1 Env V1V2 with broadly neutralizing antibodies reveal commonalities that enable vaccine design. Nat. Struct. Mol. Biol. *23*, 81–90.

Gray, E.S., Moore, P.L., Choge, I.A., Decker, J.M., Bibollet-Ruche, F., Li, H., Leseka, N., Treurnicht, F., Mlisana, K., Shaw, G.M., et al. (2007). Neutralizing antibody responses in acute human immunodeficiency virus type 1 subtype C infection. J. Virol. *81*, 6187–6196.

Gray, E.S., Madiga, M.C., Hermanus, T., Moore, P.L., Wibmer, C.K., Tumba, N.L., Werner, L., Mlisana, K., Sibeko, S., Williamson, C., et al. (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*, 4828–4840.

Gray, G.E., Bekker, L.-G., Laher, F., Malahleha, M., Allen, M., Moodie, Z., Grunenberg, N., Huang, Y., Grove, D., Prigmore, B., et al. (2021). Vaccine Efficacy of ALVAC-HIV and Bivalent Subtype C gp120–MF59 in Adults. N. Engl. J. Med. *384*, 1089–1100.

Groenink, M., Andeweg, A.C., Fouchier, R.A., Broersen, S., van der Jagt, R.C., Schuitemaker, H., de Goede, R.E., Bosch, M.L., Huisman, H.G., and Tersmette, M. (1992). Phenotype-associated env gene variation among eight related human immunodeficiency virus type 1 clones: evidence for in vivo recombination and determinants of cytotropism outside the V3 domain. J. Virol. *66*, 6175–6180.

Hallenberger, S., Bosch, V., Angliker, H., Shaw, E., Klenk, H.-D., and Garten, W. (1992). Inhibition of furin-mediated cleavage activation of HIV-1 glycoprotein gpl60. Nature *360*, 358–361.

Halpern, A.L., and Bruno, W.J. (1998). Evolutionary distances for protein-coding sequences: modeling site-specific residue frequencies. Mol. Biol. Evol. *15*, 910–917.

Hasegawa, M., Kishino, H., and Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. *22*, 160–174.

Havenar-Daughton, C., Carnathan, D.G., Torrents de la Peña, A., Pauthner, M., Briney, B., Reiss, S.M., Wood, J.S., Kaushik, K., van Gils, M.J., Rosales, S.L., et al. (2016). Direct Probing of Germinal Center Responses Reveals Immunological Features and Bottlenecks for Neutralizing Antibody Responses to HIV Env Trimer. Cell Rep. *17*, 2195–2209.

Haynes, B.F., Gilbert, P.B., McElrath, M.J., Zolla-Pazner, S., Tomaras, G.D., Alam, S.M., Evans, D.T., Montefiori, D.C., Karnasuta, C., Sutthent, R., et al. (2012). Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N. Engl. J. Med. *366*, 1275–1286.

Hessell, A.J., Poignard, P., Hunter, M., Hangartner, L., Tehrani, D.M., Bleeker, W.K., Parren, P.W.H.I., Marx, P.A., and Burton, D.R. (2009). Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. Nat. Med. *15*, 951–954.

Hraber, P., Seaman, M.S., Bailer, R.T., Mascola, J.R., Montefiori, D.C., and Korber, B.T. (2014). Prevalence of broadly neutralizing antibody responses during chronic HIV-1 infection. AIDS *28*, 163–169.

Hu, W.S., and Temin, H.M. (1990). Retroviral recombination and reverse transcription.

Science *250*, 1227–1233.

Hu, J.K., Crampton, J.C., Cupo, A., Ketas, T., van Gils, M.J., Sliepen, K., de Taeye, S.W., Sok, D., Ozorowski, G., Deresa, I., et al. (2015). Murine Antibody Responses to Cleaved Soluble HIV-1 Envelope Trimers Are Highly Restricted in Specificity. J. Virol. *89*, 10383–10398.

Huang, J., Ofek, G., Laub, L., Louder, M.K., Doria-Rose, N.A., Longo, N.S., Imamichi, H., Bailer, R.T., Chakrabarti, B., Sharma, S.K., et al. (2012). Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature *491*, 406–412.

Jardine, J.G., Ota, T., Sok, D., Pauthner, M., Kulp, D.W., Kalyuzhniy, O., Skog, P.D., Thinnes, T.C., Bhullar, D., Briney, B., et al. (2015). HIV-1 VACCINES. Priming a broadly neutralizing antibody response to HIV-1 using a germline-targeting immunogen. Science *349*, 156–161.

Jardine, J.G., Kulp, D.W., Havenar-Daughton, C., Sarkar, A., Briney, B., Sok, D., Sesterhenn, F., Ereño-Orbea, J., Kalyuzhniy, O., Deresa, I., et al. (2016). HIV-1 broadly neutralizing antibody precursor B cells revealed by germline-targeting immunogen. Science *351*, 1458–1463.

Javaherian, K., Langlois, A.J., McDanal, C., Ross, K.L., Eckler, L.I., Jellis, C.L., Profy, A.T., Rusche, J.R., Bolognesi, D.P., and Putney, S.D. (1989). Principal neutralizing domain of the human immunodeficiency virus type 1 envelope protein. Proc. Natl. Acad. Sci. U. S. A. *86*, 6768–6772.

Ji, J.P., and Loeb, L.A. (1992). Fidelity of HIV-1 reverse transcriptase copying RNA in vitro. Biochemistry *31*, 954–958.

Jones, C.T., Youssef, N., Susko, E., and Bielawski, J.P. (2018). Phenomenological Load on Model Parameters Can Lead to False Biological Conclusions. Mol. Biol. Evol. *35*, 1473–1488.

Julien, J.-P., Cupo, A., Sok, D., Stanfield, R.L., Lyumkis, D., Deller, M.C., Klasse, P.-J., Burton, D.R., Sanders, R.W., Moore, J.P., et al. (2013). Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science *342*, 1477–1483.

Kanekiyo, M., Joyce, M.G., Gillespie, R.A., Gallagher, J.R., Andrews, S.F., Yassine, H.M., Wheatley, A.K., Fisher, B.E., Ambrozak, D.R., Creanga, A., et al. (2019). Mosaic nanoparticle display of diverse influenza virus hemagglutinins elicits broad B cell responses. Nat. Immunol. *20*, 362–372.

Keele, B.F., Giorgi, E.E., Salazar-Gonzalez, J.F., Decker, J.M., Pham, K.T., Salazar, M.G., Sun, C., Grayson, T., Wang, S., Li, H., et al. (2008). Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. Proc. Natl. Acad. Sci. U. S. A. *105*, 7552–7557.

van den Kerkhof, T.L.G.M., Feenstra, K.A., Euler, Z., van Gils, M.J., Rijsdijk, L.W.E., Boeser-Nunnink, B.D., Heringa, J., Schuitemaker, H., and Sanders, R.W. (2013). HIV-1 envelope glycoprotein signatures that correlate with the development of cross-reactive neutralizing activity. Retrovirology *10*, 102.

Klasse, P.J., LaBranche, C.C., Ketas, T.J., Ozorowski, G., Cupo, A., Pugach, P., Ringe, R.P., Golabek, M., van Gils, M.J., Guttman, M., et al. (2016). Sequential and Simultaneous Immunization of Rabbits with HIV-1 Envelope Glycoprotein SOSIP.664 Trimers from Clades A, B and C. PLoS Pathog. *12*, e1005864.

Klasse, P.J., Ketas, T.J., Cottrell, C.A., Ozorowski, G., Debnath, G., Camara, D., Francomano,

E., Pugach, P., Ringe, R.P., LaBranche, C.C., et al. (2018). Epitopes for neutralizing antibodies induced by HIV-1 envelope glycoprotein BG505 SOSIP trimers in rabbits and macaques. PLoS Pathog. *14*, e1006913.

Klein, F., Diskin, R., Scheid, J.F., Gaebler, C., Mouquet, H., Georgiev, I.S., Pancera, M., Zhou, T., Incesu, R.-B., Fu, B.Z., et al. (2013). Somatic mutations of the immunoglobulin framework are generally required for broad and potent HIV-1 neutralization. Cell *153*, 126–138.

Kong, R., Duan, H., Sheng, Z., Xu, K., Acharya, P., Chen, X., Cheng, C., Dingens, A.S., Gorman, J., Sastry, M., et al. (2019). Antibody Lineages with Vaccine-Induced Antigen-Binding Hotspots Develop Broad HIV Neutralization. Cell *178*, 567–584.e19.

Korber, B., Muldoon, M., Theiler, J., Gao, F., Gupta, R., Lapedes, A., Hahn, B.H., Wolinsky, S., and Bhattacharya, T. (2000). Timing the Ancestor of the HIV-1 Pandemic Strains. Science *288*, 1789–1796.

Kosakovsky Pond, S.L., Posada, D., Gravenor, M.B., Woelk, C.H., and Frost, S.D.W. (2006). Automated phylogenetic detection of recombination using a genetic algorithm. Mol. Biol. Evol. *23*, 1891–1901.

Krakoff, E., Gagne, R.B., VandeWoude, S., and Carver, S. (2019). Variation in Intra-individual Lentiviral Evolution Rates: a Systematic Review of Human, Nonhuman Primate, and Felid Species. J. Virol. *93*.

Kulp, D.W., Steichen, J.M., Pauthner, M., Hu, X., Schiffner, T., Liguori, A., Cottrell, C.A., Havenar-Daughton, C., Ozorowski, G., Georgeson, E., et al. (2017). Structure-based design of native-like HIV-1 envelope trimers to silence non-neutralizing epitopes and eliminate CD4 binding. Nat. Commun. *8*, 1655.

Kwon, Y.D., Pancera, M., Acharya, P., Georgiev, I.S., Crooks, E.T., Gorman, J., Joyce, M.G., Guttman, M., Ma, X., Narpala, S., et al. (2015). Crystal structure, conformational fixation and entry-related interactions of mature ligand-free HIV-1 Env. Nat. Struct. Mol. Biol. *22*, 522–531.

Landais, E., Huang, X., Havenar-Daughton, C., Murrell, B., Price, M.A., Wickramasinghe, L., Ramos, A., Bian, C.B., Simek, M., and Allen, S. (2016). Broadly neutralizing antibody responses in a large longitudinal sub-Saharan HIV primary infection cohort. PLoS Pathog. *12*, e1005369.

Landais, E., Murrell, B., Briney, B., Murrell, S., Rantalainen, K., Berndsen, Z.T., Ramos, A., Wickramasinghe, L., Smith, M.L., and Eren, K. (2017). HIV envelope glycoform heterogeneity and localized diversity govern the initiation and maturation of a V2 apex broadly neutralizing antibody lineage. Immunity *47*, 990–1003. e9.

Lee, J.H., Ozorowski, G., and Ward, A.B. (2016). Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer. Science *351*, 1043–1048.

Lemey, P., Salemi, M., and Vandamme, A.-M. (2009). The Phylogenetic Handbook: A Practical Approach to Phylogenetic Analysis and Hypothesis Testing (Cambridge University Press).

Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N., and Gregory, T.J. (1990). Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells. J. Biol. Chem. *265*, 10373–10382.

Li, M., Salazar-Gonzalez, J.F., Derdeyn, C.A., Morris, L., Williamson, C., Robinson, J.E., Decker, J.M., Li, Y., Salazar, M.G., Polonis, V.R., et al. (2006). Genetic and neutralization properties of subtype C human immunodeficiency virus type 1 molecular env clones from acute and early heterosexually acquired infections in Southern Africa. J. Virol. *80*, 11776–11790.

Liao, H.-X., Lynch, R., Zhou, T., Gao, F., Alam, S.M., Boyd, S.D., Fire, A.Z., Roskin, K.M., Schramm, C.A., Zhang, Z., et al. (2013). Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. Nature *496*, 469–476.

Little, S.J., McLean, A.R., Spina, C.A., Richman, D.D., and Havlir, D.V. (1999). Viral dynamics of acute HIV-1 infection. J. Exp. Med. *190*, 841–850.

Liu, M., Yang, G., Wiehe, K., Nicely, N.I., Vandergrift, N.A., Rountree, W., Bonsignori, M., Alam, S.M., Gao, J., Haynes, B.F., et al. (2015). Polyreactivity and autoreactivity among HIV-1 antibodies. J. Virol. *89*, 784–798.

Locci, M., Havenar-Daughton, C., Landais, E., Wu, J., Kroenke, M.A., Arlehamn, C.L., Su, L.F., Cubas, R., Davis, M.M., Sette, A., et al. (2013). Human circulating PD-1+CXCR3-CXCR5+ memory Tfh cells are highly functional and correlate with broadly neutralizing HIV antibody responses. Immunity *39*, 758–769.

van Loggerenberg, F., Mlisana, K., Williamson, C., Auld, S.C., Morris, L., Gray, C.M., Abdool Karim, Q., Grobler, A., Barnabas, N., Iriogbe, I., et al. (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.

Lyumkis, D., Julien, J.-P., Wilson, I.A., and Ward, A.B. (2013). Cryo-EM structure of the BG505 SOSIP.664 HIV-1 Env trimer with 3 PGV04 Fabs.

MacLeod, D.T., Choi, N.M., Briney, B., Garces, F., Ver, L.S., Landais, E., Murrell, B., Wrin, T., Kilembe, W., Liang, C.-H., et al. (2016). Early Antibody Lineage Diversification and Independent Limb Maturation Lead to Broad HIV-1 Neutralization Targeting the Env High-Mannose Patch. Immunity *44*, 1215–1226.

Martin, D.P., Murrell, B., Golden, M., Khoosal, A., and Muhire, B. (2015). RDP4: Detection and analysis of recombination patterns in virus genomes. Virus Evolution *1*.

Mascola, J.R., Snyder, S.W., Weislow, O.S., Belay, S.M., Belshe, R.B., Schwartz, D.H., Clements, M.L., Dolin, R., Graham, B.S., Gorse, G.J., et al. (1996). Immunization with envelope subunit vaccine products elicits neutralizing antibodies against laboratory-adapted but not primary isolates of human immunodeficiency virus type 1. J. Infect. Dis. *173*, 340–348.

Mascola, J.R., Stiegler, G., VanCott, T.C., Katinger, H., Carpenter, C.B., Hanson, C.E., Beary, H., Hayes, D., Frankel, S.S., Birx, D.L., et al. (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.

May, M.T., and Ingle, S.M. (2011). Life expectancy of HIV-positive adults: a review. Sex. Health *8*, 526–533.

McCoy, L.E., Quigley, A.F., Strokappe, N.M., Bulmer-Thomas, B., Seaman, M.S., Mortier, D., Rutten, L., Chander, N., Edwards, C.J., Ketteler, R., et al. (2012). Potent and broad neutralization of HIV-1 by a llama antibody elicited by immunization. J. Exp. Med. *209*, 1091–1103.

McCoy, L.E., van Gils, M.J., Ozorowski, G., Messmer, T., Briney, B., Voss, J.E., Kulp, D.W., Macauley, M.S., Sok, D., Pauthner, M., et al. (2016). Holes in the Glycan Shield of the Native HIV Envelope Are a Target of Trimer-Elicited Neutralizing Antibodies. Cell Rep. *16*, 2327–2338.

McCune, J.M., Rabin, L.B., Feinberg, M.B., Lieberman, M., Kosek, J.C., Reyes, G.R., and Weissman, I.L. (1988). Endoproteolytic cleavage of gp160 is required for the activation of human immunodeficiency virus. Cell *53*, 55–67.

McGuire, A.T., Gray, M.D., Dosenovic, P., Gitlin, A.D., Freund, N.T., Petersen, J., Correnti, C., Johnsen, W., Kegel, R., Stuart, A.B., et al. (2016). Specifically modified Env immunogens activate B-cell precursors of broadly neutralizing HIV-1 antibodies in transgenic mice. Nat. Commun. *7*, 10618.

Medina-Ramírez, M., Garces, F., Escolano, A., Skog, P., de Taeye, S.W., Del Moral-Sanchez, I., McGuire, A.T., Yasmeen, A., Behrens, A.-J., Ozorowski, G., et al. (2017). Design and crystal structure of a native-like HIV-1 envelope trimer that engages multiple broadly neutralizing antibody precursors in vivo. J. Exp. Med. *214*, 2573–2590.

Mesin, L., Schiepers, A., Ersching, J., Barbulescu, A., Cavazzoni, C.B., Angelini, A., Okada, T., Kurosaki, T., and Victora, G.D. (2020). Restricted Clonality and Limited Germinal Center Reentry Characterize Memory B Cell Reactivation by Boosting. Cell *180*, 92–106.e11.

Mikell, I., Sather, D.N., Kalams, S.A., Altfeld, M., Alter, G., and Stamatatos, L. (2011). Characteristics of the earliest cross-neutralizing antibody response to HIV-1. PLoS Pathog. *7*, e1001251.

Moldt, B., Rakasz, E.G., Schultz, N., Chan-Hui, P.-Y., Swiderek, K., Weisgrau, K.L., Piaskowski, S.M., Bergman, Z., Watkins, D.I., Poignard, P., et al. (2012). Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. Proc. Natl. Acad. Sci. U. S. A. *109*, 18921–18925.

Montefiori, D.C., Karnasuta, C., Huang, Y., Ahmed, H., Gilbert, P., de Souza, M.S., McLinden, R., Tovanabutra, S., Laurence-Chenine, A., Sanders-Buell, E., et al. (2012). Magnitude and breadth of the neutralizing antibody response in the RV144 and Vax003 HIV-1 vaccine efficacy trials. J. Infect. Dis. *206*, 431–441.

Moody, M.A., Zhang, R., Walter, E.B., Woods, C.W., Ginsburg, G.S., McClain, M.T., Denny, T.N., Chen, X., Munshaw, S., Marshall, D.J., et al. (2011). H3N2 influenza infection elicits more cross-reactive and less clonally expanded anti-hemagglutinin antibodies than influenza vaccination. PLoS One *6*, e25797.

Moody, M.A., Anthony Moody, M., Pedroza-Pacheco, I., Vandergrift, N.A., Chui, C., Lloyd, K.E., Parks, R., Soderberg, K.A., Ogbe, A.T., Cohen, M.S., et al. (2016). Immune perturbations in HIV-1–infected individuals who make broadly neutralizing antibodies. Science Immunology *1*.

Moore, P.L. (2018). The Neutralizing Antibody Response to the HIV-1 Env Protein. Curr. HIV Res. *16*, 21–28.

Moore, J.P., Cao, Y., Ho, D.D., and Koup, R.A. (1994). Development of the anti-gp120 antibody response during seroconversion to human immunodeficiency virus type 1. J. Virol. *68*, 5142–5155.

Moore, P.L., Crooks, E.T., Porter, L., Zhu, P., Cayanan, C.S., Grise, H., Corcoran, P., Zwick, M.B., Franti, M., Morris, L., et al. (2006). Nature of nonfunctional envelope proteins on the

surface of human immunodeficiency virus type 1. J. Virol. *80*, 2515–2528.

Moore, P.L., Gray, E.S., Choge, I.A., Ranchobe, N., Mlisana, K., Abdool Karim, S.S., Williamson, C., Morris, L., and CAPRISA 002 Study Team (2008). The c3-v4 region is a major target of autologous neutralizing antibodies in human immunodeficiency virus type 1 subtype C infection. J. Virol. *82*, 1860–1869.

Moore, P.L., Ranchobe, N., Lambson, B.E., Gray, E.S., Cave, E., Abrahams, M.-R., Bandawe, G., Mlisana, K., Abdool Karim, S.S., Williamson, C., et al. (2009a). Limited neutralizing antibody specificities drive neutralization escape in early HIV-1 subtype C infection. PLoS Pathog. *5*, e1000598.

Moore, P.L., Gray, E.S., and Morris, L. (2009b). Specificity of the autologous neutralizing antibody response. Curr. Opin. HIV AIDS *4*, 358–363.

Moore, P.L., Gray, E.S., Wibmer, C.K., Bhiman, J.N., Nonyane, M., Sheward, D.J., Hermanus, T., Bajimaya, S., Tumba, N.L., Abrahams, M.-R., et al. (2012). Evolution of an HIV glycan-dependent broadly neutralizing antibody epitope through immune escape. Nat. Med. *18*, 1688–1692.

Mount, D.W. (2009). Using hidden Markov models to align multiple sequences. Cold Spring Harb. Protoc. *2009*, db.top41.

Munro, J.B., Gorman, J., Ma, X., Zhou, Z., Arthos, J., Burton, D.R., Koff, W.C., Courter, J.R., Smith, A.B., 3rd, Kwong, P.D., et al. (2014). Conformational dynamics of single HIV-1 envelope trimers on the surface of native virions. Science *346*, 759–763.

Murrell, B., Wertheim, J.O., Moola, S., Weighill, T., Scheffler, K., and Pond, S.L.K. (2012). Detecting individual sites subject to episodic diversifying selection. PLoS Genet. *8*, e1002764.

Murrell, B., Moola, S., Mabona, A., Weighill, T., Sheward, D., Kosakovsky Pond, S.L., and Scheffler, K. (2013). FUBAR: a fast, unconstrained bayesian approximation for inferring selection. Mol. Biol. Evol. *30*, 1196–1205.

Ndhlovu, Z.M., Kamya, P., Mewalal, N., Kløverpris, H.N., Nkosi, T., Pretorius, K., Laher, F., Ogunshola, F., Chopera, D., Shekhar, K., et al. (2015). Magnitude and Kinetics of CD8+ T Cell Activation during Hyperacute HIV Infection Impact Viral Set Point. Immunity *43*, 591–604.

Neher, R.A., and Leitner, T. (2010). Recombination rate and selection strength in HIV intra-patient evolution. PLoS Comput. Biol. *6*, e1000660.

Oh, J.E., Iijima, N., Song, E., Lu, P., Klein, J., Jiang, R., Kleinstein, S.H., and Iwasaki, A. (2019). Migrant memory B cells secrete luminal antibody in the vagina. Nature *571*, 122–126.

Ozorowski, G., Pallesen, J., de Val, N., Lyumkis, D., Cottrell, C.A., Torres, J.L., Copps, J., Stanfield, R.L., Cupo, A., Pugach, P., et al. (2017). Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike. Nature *547*, 360–363.

Pancera, M., Zhou, T., Druz, A., Georgiev, I.S., Soto, C., Gorman, J., Huang, J., Acharya, P., Chuang, G.-Y., Ofek, G., et al. (2014). Structure and immune recognition of trimeric pre-fusion HIV-1 Env. Nature *514*, 455–461.

Pauthner, M.G., Nkolola, J.P., Havenar-Daughton, C., Murrell, B., Reiss, S.M., Bastidas, R., Prévost, J., Nedellec, R., von Bredow, B., Abbink, P., et al. (2019). Vaccine-Induced Protection from Homologous Tier 2 SHIV Challenge in Nonhuman Primates Depends on Serum-Neutralizing Antibody Titers. Immunity *50*, 241–252.e6.

Pegu, A., Borate, B., Huang, Y., Pauthner, M.G., Hessell, A.J., Julg, B., Doria-Rose, N.A., Schmidt, S.D., Carpp, L.N., Cully, M.D., et al. (2019). A Meta-analysis of Passive Immunization Studies Shows that Serum-Neutralizing Antibody Titer Associates with Protection against SHIV Challenge. Cell Host Microbe *26*, 336–346.e3.

de la Peña, A.T., de Taeye, S.W., Sliepen, K., LaBranche, C.C., Burger, J.A., Schermer, E.E., Montefiori, D.C., Moore, J.P., Klasse, P.J., and Sanders, R.W. (2018). Immunogenicity in Rabbits of HIV-1 SOSIP Trimers from Clades A, B, and C, Given Individually, Sequentially, or in Combination. J. Virol. *92*.

Piantadosi, A., Panteleeff, D., Blish, C.A., Baeten, J.M., Jaoko, W., McClelland, R.S., and 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*, 10269–10274.

Pitisuttithum, P., Gilbert, P., Gurwith, M., Heyward, W., Martin, M., van Griensven, F., Hu, D., Tappero, J.W., and Group, B.V.E. (2006). Randomized, double-blind, placebo-controlled efficacy trial of a bivalent recombinant glycoprotein 120 HIV-1 vaccine among injection drug users in Bangkok, Thailand. J. Infect. Dis. *194*, 1661–1671.

Plotkin, S.A. (2010). Correlates of protection induced by vaccination. Clin. Vaccine Immunol. *17*, 1055–1065.

Powell, R.L.R., Kinge, T., and Nyambi, P.N. (2010). Infection by discordant strains of HIV-1 markedly enhances the neutralizing antibody response against heterologous virus. J. Virol. *84*, 9415–9426.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009). FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. *26*, 1641–1650.

Price, M.N., Dehal, P.S., and Arkin, A.P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One *5*, e9490.

Purtha, W.E., Tedder, T.F., Johnson, S., Bhattacharya, D., and Diamond, M.S. (2011). Memory B cells, but not long-lived plasma cells, possess antigen specificities for viral escape mutants. J. Exp. Med. *208*, 2599–2606.

Rabiner, L.R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. Proc. IEEE *77*, 257–286.

Rabiner, L.R., Lee, C.H., Juang, B.H., and Wilpon, J.G. (1989). HMM clustering for connected word recognition. In International Conference on Acoustics, Speech, and Signal Processing, pp. 405–408 vol.1.

Rademeyer, C., Moore, P.L., Taylor, N., Martin, D.P., Choge, I.A., Gray, E.S., Sheppard, H.W., Gray, C., Morris, L., Williamson, C., et al. (2007). Genetic characteristics of HIV-1 subtype C envelopes inducing cross-neutralizing antibodies. Virology *368*, 172–181.

Redd, A.D., Quinn, T.C., and Tobian, A.A.R. (2013). Frequency and implications of HIV superinfection. Lancet Infect. Dis. *13*, 622–628.

Redd, A.D., Ssemwanga, D., Vandepitte, J., Wendel, S.K., Ndembi, N., Bukenya, J., Nakubulwa, S., Grosskurth, H., Parry, C.M., Martens, C., et al. (2014a). Rates of HIV-1 superinfection and primary HIV-1 infection are similar in female sex workers in Uganda. AIDS *28*, 2147–2152.

Redd, A.D., Mullis, C.E., Wendel, S.K., Sheward, D., Martens, C., Bruno, D., Werner, L., Garrett, N.J., Abdool Karim, Q., Williamson, C., et al. (2014b). Limited HIV-1 superinfection in seroconverters from the CAPRISA 004 Microbicide Trial. J. Clin. Microbiol. *52*, 844–848.

Rerks-Ngarm, S., Pitisuttithum, P., Nitayaphan, S., Kaewkungwal, J., Chiu, J., Paris, R., Premsri, N., Namwat, C., de Souza, M., Adams, E., et al. (2009). Vaccination with ALVAC and AIDSVAX to prevent HIV-1 infection in Thailand. N. Engl. J. Med. *361*, 2209–2220.

Richardson, S.I., Chung, A.W., Natarajan, H., Mabvakure, B., Mkhize, N.N., Garrett, N., Abdool Karim, S., Moore, P.L., Ackerman, M.E., Alter, G., et al. (2018). HIV-specific Fc effector function early in infection predicts the development of broadly neutralizing antibodies. PLoS Pathog. *14*, e1006987.

Richman, D.D., Wrin, T., Little, S.J., and Petropoulos, C.J. (2003). Rapid evolution of the neutralizing antibody response to HIV type 1 infection. Proc. Natl. Acad. Sci. U. S. A. *100*, 4144–4149.

Ringe, R.P., Ozorowski, G., Rantalainen, K., Struwe, W.B., Matthews, K., Torres, J.L., Yasmeen, A., Cottrell, C.A., Ketas, T.J., LaBranche, C.C., et al. (2017). Reducing V3 Antigenicity and Immunogenicity on Soluble, Native-Like HIV-1 Env SOSIP Trimers. J. Virol. *91*.

Robertson, D.L., Sharp, P.M., McCutchan, F.E., and Hahn, B.H. (1995). Recombination in HIV-1. Nature *374*, 124–126.

Ronen, K., McCoy, C.O., Matsen, F.A., Boyd, D.F., Emery, S., Odem-Davis, K., Jaoko, W., Mandaliya, K., McClelland, R.S., Richardson, B.A., et al. (2013). HIV-1 superinfection occurs less frequently than initial infection in a cohort of high-risk Kenyan women. PLoS Pathog. *9*, e1003593.

Rong, R., Li, B., Lynch, R.M., Haaland, R.E., Murphy, M.K., Mulenga, J., Allen, S.A., Pinter, A., Shaw, G.M., Hunter, E., et al. (2009). Escape from autologous neutralizing antibodies in acute/early subtype C HIV-1 infection requires multiple pathways. PLoS Pathog. *5*, e1000594.

Rusert, P., Kouyos, R.D., Kadelka, C., Ebner, H., Schanz, M., Huber, M., Braun, D.L., Hozé, N., Scherrer, A., Magnus, C., et al. (2016). Determinants of HIV-1 broadly neutralizing antibody induction. Nat. Med. *22*, 1260–1267.

Sagar, M., Wu, X., Lee, S., and 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*, 9586–9598.

Sanders, R.W., and Moore, J.P. (2017). Native-like Env trimers as a platform for HIV-1 vaccine design. Immunol. Rev. *275*, 161–182.

Sanders, R.W., Vesanen, M., Schuelke, N., Master, A., Schiffner, L., Kalyanaraman, R., Paluch, M., Berkhout, B., Maddon, P.J., Olson, W.C., et al. (2002). Stabilization of the soluble, cleaved, trimeric form of the envelope glycoprotein complex of human immunodeficiency virus type 1. J. Virol. *76*, 8875–8889.

Sanders, R.W., Derking, R., Cupo, A., Julien, J.-P., Yasmeen, A., de Val, N., Kim, H.J., Blattner, C., de la Peña, A.T., Korzun, J., et al. (2013). A Next-Generation Cleaved, Soluble HIV-1 Env Trimer, BG505 SOSIP.664 gp140, Expresses Multiple Epitopes for Broadly Neutralizing but Not Non-Neutralizing Antibodies. PLoS Pathogens *9*, e1003618.

Sanders, R.W., van Gils, M.J., Derking, R., Sok, D., Ketas, T.J., Burger, J.A., Ozorowski, G., Cupo, A., Simonich, C., Goo, L., et al. (2015). HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers. Science *349*, aac4223.

Sather, D.N., Armann, J., Ching, L.K., Mavrantoni, A., Sellhorn, G., Caldwell, Z., Yu, X., Wood, B., Self, S., Kalams, S., et al. (2009). Factors associated with the development of cross-reactive neutralizing antibodies during human immunodeficiency virus type 1 infection. J. Virol. *83*, 757–769.

Sather, D.N., Carbonetti, S., Kehayia, J., Kraft, Z., Mikell, I., Scheid, J.F., Klein, F., and Stamatatos, L. (2012). Broadly neutralizing antibodies developed by an HIV-positive elite neutralizer exact a replication fitness cost on the contemporaneous virus. J. Virol. *86*, 12676–12685.

Saunders, K.O., Wiehe, K., Tian, M., Acharya, P., Bradley, T., Alam, S.M., Go, E.P., Scearce, R., Sutherland, L., Henderson, R., et al. (2019). Targeted selection of HIV-specific antibody mutations by engineering B cell maturation. Science *366*.

Scheffler, K., Martin, D.P., and Seoighe, C. (2006). Robust inference of positive selection from recombining coding sequences. Bioinformatics *22*, 2493–2499.

Scheid, J.F., Mouquet, H., Feldhahn, N., Seaman, M.S., Velinzon, K., Pietzsch, J., Ott, R.G., Anthony, R.M., Zebroski, H., Hurley, A., et al. (2009). Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature *458*, 636–640.

Scheid, J.F., Mouquet, H., Ueberheide, B., Diskin, R., Klein, F., Oliveira, T.Y.K., Pietzsch, J., Fenyo, D., Abadir, A., Velinzon, K., et al. (2011). Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. Science *333*, 1633–1637.

Schoofs, T., Klein, F., Braunschweig, M., Kreider, E.F., Feldmann, A., Nogueira, L., Oliveira, T., Lorenzi, J.C.C., Parrish, E.H., Learn, G.H., et al. (2016). HIV-1 therapy with monoclonal antibody 3BNC117 elicits host immune responses against HIV-1. Science *352*, 997–1001.

Sheward, D.J., Ntale, R., Garrett, N.J., Woodman, Z.L., Abdool Karim, S.S., and Williamson, C. (2015). HIV-1 Superinfection Resembles Primary Infection. J. Infect. Dis. *212*, 904–908.

Sheward, D.J., Marais, J., Bekker, V., Murrell, B., Eren, K., Bhiman, J.N., Nonyane, M., Garrett, N., Woodman, Z.L., Karim, Q.A., et al. (2018). HIV superinfection drives de novo antibody responses and not neutralization breadth. Cell Host Microbe *24*, 593–599.

Shimodaira, H., and Hasegawa, M. (1999). Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. Mol. Biol. Evol. *16*, 1114–1114.

Shingai, M., Donau, O.K., Plishka, R.J., Buckler-White, A., Mascola, J.R., Nabel, G.J., Nason, M.C., Montefiori, D., Moldt, B., Poignard, P., et al. (2014). Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques. J. Exp. Med. *211*, 2061–2074.

Simmonds, P., Zhang, L.Q., McOmish, F., Balfe, P., Ludlam, C.A., and Brown, A.J. (1991). Discontinuous sequence change of human immunodeficiency virus (HIV) type 1 env sequences in plasma viral and lymphocyte-associated proviral populations in vivo: implications for models of HIV pathogenesis. J. Virol. *65*, 6266–6276.

Simonich, C.A., Williams, K.L., Verkerke, H.P., Williams, J.A., Nduati, R., Lee, K.K., and Overbaugh, J. (2016). HIV-1 Neutralizing Antibodies with Limited Hypermutation from an Infant. Cell *166*, 77–87.

Smith, D.M., Wong, J.K., Hightower, G.K., Ignacio, C.C., Koelsch, K.K., Daar, E.S., Richman, D.D., and Little, S.J. (2004). Incidence of HIV superinfection following primary infection. JAMA *292*, 1177–1178.

Smith, D.M., Strain, M.C., Frost, S.D.W., Pillai, S.K., Wong, J.K., Wrin, T., Liu, Y., Petropolous, C.J., Daar, E.S., Little, S.J., et al. (2006). Lack of neutralizing antibody response to HIV-1 predisposes to superinfection. Virology *355*, 1–5.

Smith, S.A., Burton, S.L., Kilembe, W., Lakhi, S., Karita, E., Price, M., Allen, S., Hunter, E., and Derdeyn, C.A. (2016). Diversification in the HIV-1 Envelope Hyper-variable Domains V2, V4, and V5 and Higher Probability of Transmitted/Founder Envelope Glycosylation Favor the Development of Heterologous Neutralization Breadth. PLoS Pathog. *12*, e1005989.

Sok, D., Doores, K.J., Briney, B., Le, K.M., Saye-Francisco, K.L., Ramos, A., Kulp, D.W., Julien, J.-P., Menis, S., Wickramasinghe, L., et al. (2014). Promiscuous glycan site recognition by antibodies to the high-mannose patch of gp120 broadens neutralization of HIV. Sci. Transl. Med. *6*, 236ra63.

Sok, D., Briney, B., Jardine, J.G., Kulp, D.W., Menis, S., Pauthner, M., Wood, A., Lee, E.-C., Le, K.M., Jones, M., et al. (2016). Priming HIV-1 broadly neutralizing antibody precursors in human Ig loci transgenic mice. Science *353*, 1557–1560.

Sok, D., Le, K.M., Vadnais, M., Saye-Francisco, K.L., Jardine, J.G., Torres, J.L., Berndsen, Z.T., Kong, L., Stanfield, R., Ruiz, J., et al. (2017). Rapid elicitation of broadly neutralizing antibodies to HIV by immunization in cows. Nature *548*, 108–111.

Soto, C., Bombardi, R.G., Branchizio, A., Kose, N., Matta, P., Sevy, A.M., Sinkovits, R.S., Gilchuk, P., Finn, J.A., and Crowe, J.E., Jr (2019). High frequency of shared clonotypes in human B cell receptor repertoires. Nature *566*, 398–402.

Stadtmueller, B.M., Bridges, M.D., Dam, K.-M., Lerch, M.T., Huey-Tubman, K.E., Hubbell, W.L., and Bjorkman, P.J. (2018). DEER Spectroscopy Measurements Reveal Multiple Conformations of HIV-1 SOSIP Envelopes that Show Similarities with Envelopes on Native Virions. Immunity *49*, 235–246.e4.

Steichen, J.M., Kulp, D.W., Tokatlian, T., Escolano, A., Dosenovic, P., Stanfield, R.L., McCoy, L.E., Ozorowski, G., Hu, X., Kalyuzhniy, O., et al. (2016). HIV Vaccine Design to Target Germline Precursors of Glycan-Dependent Broadly Neutralizing Antibodies. Immunity *45*, 483–496.

Steichen, J.M., Lin, Y.-C., Havenar-Daughton, C., Pecetta, S., Ozorowski, G., Willis, J.R., Toy, L., Sok, D., Liguori, A., Kratochvil, S., et al. (2019). A generalized HIV vaccine design strategy for priming of broadly neutralizing antibody responses. Science *366*.

Stewart-Jones, G.B.E., Soto, C., Lemmin, T., Chuang, G.-Y., Druz, A., Kong, R., Thomas, P.V., Wagh, K., Zhou, T., Behrens, A.-J., et al. (2016). Trimeric HIV-1-Env Structures Define Glycan Shields from Clades A, B, and G. Cell *165*, 813–826.

de Taeye, S.W., Ozorowski, G., Torrents de la Peña, A., Guttman, M., Julien, J.-P., van den Kerkhof, T.L.G.M., Burger, J.A., Pritchard, L.K., Pugach, P., Yasmeen, A., et al. (2015). Immunogenicity of Stabilized HIV-1 Envelope Trimers with Reduced Exposure of Non-neutralizing Epitopes. Cell *163*, 1702–1715.

Tian, M., Cheng, C., Chen, X., Duan, H., Cheng, H.-L., Dao, M., Sheng, Z., Kimble, M., Wang, L., Lin, S., et al. (2016). Induction of HIV Neutralizing Antibody Lineages in Mice with Diverse Precursor Repertoires. Cell *166*, 1471–1484.e18.

Tomaras, G.D., Yates, N.L., Liu, P., Qin, L., Fouda, G.G., Chavez, L.L., Decamp, A.C., Parks, R.J., Ashley, V.C., Lucas, J.T., et al. (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*, 12449–12463.

Torrents de la Peña, A., Julien, J.-P., de Taeye, S.W., Garces, F., Guttman, M., Ozorowski, G., Pritchard, L.K., Behrens, A.-J., Go, E.P., Burger, J.A., et al. (2017). Improving the Immunogenicity of Native-like HIV-1 Envelope Trimers by Hyperstabilization. Cell Rep. *20*, 1805–1817.

Tsai, Y., Holton, T., and Yeates, T.O. (2015). Diffusion accessibility as a method for visualizing macromolecular surface geometry. Protein Sci. *24*, 1702–1705.

Tsui, R., Herring, B.L., Barbour, J.D., Grant, R.M., Bacchetti, P., Kral, A., Edlin, B.R., and Delwart, E.L. (2004). Human immunodeficiency virus type 1 superinfection was not detected following 215 years of injection drug user exposure. J. Virol. *78*, 94–103.

Vancott, T.C., Polonis, V.R., Loomis, L.D., Michael, N.L., Nara, P.L., and Birx, D.L. (1995). Differential role of V3-specific antibodies in neutralization assays involving primary and laboratory-adapted isolates of HIV type 1. AIDS Res. Hum. Retroviruses *11*, 1379–1391.

Venkatesan, P. (2021). Preliminary phase 1 results from an HIV vaccine candidate trial. The Lancet Microbe *2*, e95.

Verkoczy, L., Diaz, M., Holl, T.M., Ouyang, Y.-B., Bouton-Verville, H., Alam, S.M., Liao, H.-X., Kelsoe, G., and Haynes, B.F. (2010). Autoreactivity in an HIV-1 broadly reactive neutralizing antibody variable region heavy chain induces immunologic tolerance. Proc. Natl. Acad. Sci. U. S. A. *107*, 181–186.

Verkoczy, L., Chen, Y., Bouton-Verville, H., Zhang, J., Diaz, M., Hutchinson, J., Ouyang, Y.-B., Alam, S.M., Holl, T.M., Hwang, K.-K., et al. (2011). Rescue of HIV-1 broad neutralizing antibody-expressing B cells in 2F5 VH× VL knockin mice reveals multiple tolerance controls. The Journal of Immunology *187*, 3785–3797.

Victora, G.D., Schwickert, T.A., Fooksman, D.R., Kamphorst, A.O., Meyer-Hermann, M., Dustin, M.L., and Nussenzweig, M.C. (2010). Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. Cell *143*, 592–605.

Voss, J.E., Andrabi, R., McCoy, L.E., de Val, N., Fuller, R.P., Messmer, T., Su, C.-Y., Sok, D., Khan, S.N., Garces, F., et al. (2017). Elicitation of Neutralizing Antibodies Targeting the V2 Apex of the HIV Envelope Trimer in a Wild-Type Animal Model. Cell Rep. *21*, 222–235.

Wagner, G.A., Pacold, M.E., Kosakovsky Pond, S.L., Caballero, G., Chaillon, A., Rudolph, A.E., Morris, S.R., Little, S.J., Richman, D.D., and Smith, D.M. (2014). Incidence and prevalence of intrasubtype HIV-1 dual infection in at-risk men in the United States. J. Infect. Dis. *209*, 1032–1038.

Wagner, G.A., Landais, E., Caballero, G., Phung, P., Kosakovsky Pond, S.L., Poignard, P., Richman, D.D., Little, S.J., and Smith, D.M. (2017). Intrasubtype B HIV-1 Superinfection Correlates with Delayed Neutralizing Antibody Response. J. Virol. *91*.

Walker, L.M., Phogat, S.K., Chan-Hui, P.-Y., Wagner, D., Phung, P., Goss, J.L., Wrin, T., Simek, M.D., Fling, S., Mitcham, J.L., et al. (2009). Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science *326*, 285–289.

Walker, L.M., Simek, M.D., Priddy, F., Gach, J.S., Wagner, D., Zwick, M.B., Phogat, S.K., Poignard, P., and Burton, D.R. (2010). A limited number of antibody specificities mediate broad and potent serum neutralization in selected HIV-1 infected individuals. PLoS Pathog. *6*, e1001028.

Walker, L.M., Huber, M., Doores, K.J., Falkowska, E., Pejchal, R., Julien, J.-P., Wang, S.-K., Ramos, A., Chan-Hui, P.-Y., Moyle, M., et al. (2011). Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature *477*, 466–470.

Wang, T.T., Tan, G.S., Hai, R., Pica, N., Petersen, E., Moran, T.M., and Palese, P. (2010). Broadly protective monoclonal antibodies against H3 influenza viruses following sequential immunization with different hemagglutinins. PLoS Pathog. *6*, e1000796.

Wei, X., Decker, J.M., Wang, S., Hui, H., Kappes, J.C., Wu, X., Salazar-Gonzalez, J.F., Salazar, M.G., Kilby, J.M., Saag, M.S., et al. (2003). Antibody neutralization and escape by HIV-1. Nature *422*, 307–312.

West, A.P., Jr, Scharf, L., Scheid, J.F., Klein, F., Bjorkman, P.J., and Nussenzweig, M.C. (2014). Structural insights on the role of antibodies in HIV-1 vaccine and therapy. Cell *156*, 633–648.

Wibmer, C.K., Bhiman, J.N., Gray, E.S., Tumba, N., Abdool Karim, S.S., Williamson, C., Morris, L., and Moore, P.L. (2013). Viral escape from HIV-1 neutralizing antibodies drives increased plasma neutralization breadth through sequential recognition of multiple epitopes and immunotypes. PLoS Pathog. *9*, e1003738.

Won, K.-J., Hamelryck, T., Prügel-Bennett, A., and Krogh, A. (2007). An evolutionary method for learning HMM structure: prediction of protein secondary structure. BMC Bioinformatics *8*, 357.

Woodman, Z., Mlisana, K., Treurnicht, F., Abrahams, M.-R., Thebus, R., Karim, S.A., Williamson, C., and Caprisa Acute Infection Study Team (2011). Short communication decreased incidence of dual infections in South african subtype C-infected women compared to a cohort ten years earlier. AIDS Res. Hum. Retroviruses *27*, 1167–1172.

Wrammert, J., Koutsonanos, D., Li, G.-M., Edupuganti, S., Sui, J., Morrissey, M., McCausland, M., Skountzou, I., Hornig, M., Lipkin, W.I., et al. (2011). Broadly cross-reactive antibodies dominate the human B cell response against 2009 pandemic H1N1 influenza virus infection. J. Exp. Med. *208*, 181–193.

Wrin, T., and Nunberg, J.H. (1994). HIV-1MN recombinant gp120 vaccine serum, which fails to neutralize primary isolates of HIV-1, does not antagonize neutralization by antibodies from infected individuals. AIDS *8*, 1622–1623.

Wu, X., Yang, Z.-Y., Li, Y., Hogerkorp, C.-M., Schief, W.R., Seaman, M.S., Zhou, T., Schmidt, S.D., Wu, L., Xu, L., et al. (2010). Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science *329*, 856–861.

Wu, X., Zhang, Z., Schramm, C.A., Joyce, M.G., Kwon, Y.D., Zhou, T., Sheng, Z., Zhang, B., O'Dell, S., McKee, K., et al. (2015). Maturation and Diversity of the VRC01-Antibody Lineage over 15 Years of Chronic HIV-1 Infection. Cell *161*, 470–485.

Xiao, X., Chen, W., Feng, Y., Zhu, Z., Prabakaran, P., Wang, Y., Zhang, M.-Y., Longo, N.S., and Dimitrov, D.S. (2009). Germline-like predecessors of broadly neutralizing antibodies lack measurable binding to HIV-1 envelope glycoproteins: implications for evasion of immune responses and design of vaccine immunogens. Biochem. Biophys. Res. Commun. *390*, 404–

409.

Xu, K., Acharya, P., Kong, R., Cheng, C., Chuang, G.-Y., Liu, K., Louder, M.K., O'Dell, S., Rawi, R., Sastry, M., et al. (2018). Epitope-based vaccine design yields fusion peptide-directed antibodies that neutralize diverse strains of HIV-1. Nat. Med. *24*, 857–867.

Yang, G., Holl, T.M., Liu, Y., Li, Y., Lu, X., Nicely, N.I., Kepler, T.B., Alam, S.M., Liao, H.-X., Cain, D.W., et al. (2013). Identification of autoantigens recognized by the 2F5 and 4E10 broadly neutralizing HIV-1 antibodies. J. Exp. Med. *210*, 241–256.

Yang, Y.R., McCoy, L.E., van Gils, M.J., Andrabi, R., Turner, H.L., Yuan, M., Cottrell, C.A., Ozorowski, G., Voss, J., Pauthner, M., et al. (2020). Autologous Antibody Responses to an HIV Envelope Glycan Hole Are Not Easily Broadened in Rabbits. J. Virol. *94*.

Zhang, R., Verkoczy, L., Wiehe, K., Munir Alam, S., Nicely, N.I., Santra, S., Bradley, T., Pemble, C.W., 4th, Zhang, J., Gao, F., et al. (2016). Initiation of immune tolerance-controlled HIV gp41 neutralizing B cell lineages. Sci. Transl. Med. *8*, 336ra62.

Zhou, T., Doria-Rose, N.A., Cheng, C., Stewart-Jones, G.B.E., Chuang, G.-Y., Chambers, M., Druz, A., Geng, H., McKee, K., Kwon, Y.D., et al. (2017). Quantification of the Impact of the HIV-1-Glycan Shield on Antibody Elicitation. Cell Rep. *19*, 719–732.

Zhu, P., Chertova, E., Bess, J., Jr, Lifson, J.D., Arthur, L.O., Liu, J., Taylor, K.A., and Roux, K.H. (2003). Electron tomography analysis of envelope glycoprotein trimers on HIV and simian immunodeficiency virus virions. Proc. Natl. Acad. Sci. U. S. A. *100*, 15812–15817.

Zhu, P., Liu, J., Bess, J., Jr, Chertova, E., Lifson, J.D., Grisé, H., Ofek, G.A., Taylor, K.A., and Roux, K.H. (2006). Distribution and three-dimensional structure of AIDS virus envelope spikes. Nature *441*, 847–852.

Zhuang, J., Jetzt, A.E., Sun, G., Yu, H., Klarmann, G., Ron, Y., Preston, B.D., and Dougherty, J.P. (2002). Human immunodeficiency virus type 1 recombination: rate, fidelity, and putative hot spots. J. Virol. *76*, 11273–11282.

Appendix 1

Other first (or co-first) author publications by the candidate, outside of the scope of their Doctoral work

Sheward, D.J.*, Pushparaj, P.*, Das, H.*, Kim, C., et al. (2022b). Structural basis of Omicron neutralization by affinity-matured public antibodies. bioRxiv. 2022.01.03.474825; doi: https://doi.org/10.1101/2022.01.03.474825

Sheward, D.J., Kim, C., Ehling, R.A., Pankow, A., et al. (2022a). Assessing the neutralization sensitivity of SARS-CoV-2 B.1.1.529 (Omicron): a cross-sectional study. Lancet Infec. Dis. (in press)

Hanke, L.*, **Sheward, D.J.***, Pankow, A., Perez Vidakovics, L., et al. (2022). Multivariate mining of an alpaca immune repertoire identifies potent cross-neutralising SARS-CoV-2 nanobodies. Science Advances. (in press)

Sheward, D.J.*, Mandolesi, M.*, Urgard, E., Kim, C. et al. (2021). Beta RBD boost broadens antibody-mediated protection against SARS-CoV-2 variants in animal models. Cell Rep. Med.

Mandolesi, M.*, **Sheward, D.J.***, Hanke, L., Ma, J., Pushparaj, P., Perez Vidakovics, L., Kim, C., Àdori, M., Lenart, K., Loré, K., et al. (2021). SARS-CoV-2 protein subunit vaccination of mice and rhesus macaques elicits potent and durable neutralizing antibody responses. Cell Rep Med *2*, 100252.

Wørzner, K.*, **Sheward, D.J.***, Schmidt, S.T., Hanke, L., et al. (2021). Adjuvanted SARS-CoV-2 spike protein elicits neutralizing antibodies and CD4 T cell responses after a single immunization in mice. EBioMedicine *63*, 103197.

Moliner-Morro, A.*, **Sheward, D.J.***, Karl, V., Perez Vidakovics, L., et al. (2020). Picomolar SARS-CoV-2 Neutralization Using Multi-Arm PEG Nanobody Constructs. Biomolecules *10*.

Significant contributions by the candidate to other works

Ehling, R.A., Weber, C.R., Mason, D.M., Friedensohn, S., et al. (2022). SARS-CoV-2 reactive and neutralizing antibodies discovered by single-cell sequencing of plasma cells and mammalian display. Cell Rep. *38*, 110242.

Hanke, L., Das, H., **Sheward, D.J.**, Vidakovics, L.P., et al. (2021). A bispecific monomeric nanobody induces spike trimer dimers and neutralizes SARS-CoV-2 in vivo. bioRxiv.

Szurgot, I., Hanke, L., **Sheward, D.J.**, Vidakovics, L.P., et al. (2021). DNA-launched RNA replicon vaccines induce potent anti-SARS-CoV-2 immune responses in mice. Sci. Rep. *11*, 3125.

Custódio, T.F., Das, H., **Sheward, D.J.**, Hanke, L., et al. (2020). Selection, biophysical and structural analysis of synthetic nanobodies that effectively neutralize SARS-CoV-2. Nat. Commun. *11*, 5588.

Hanke, L., Perez Vidakovics , L., **Sheward, D.J.**, Das, H., et al. (2020). An alpaca nanobody neutralizes SARS-CoV-2 by blocking receptor interaction. Nat. Commun. *11*, 4420.

Sivro, A., Schuetz, A., **Sheward, D.**, V Joag, V. et al. (2018) Integrin α4β7 expression on peripheral blood CD4+ T cells predicts HIV acquisition and disease progression outcomes. Sci. Trans. Med. 10 (425), eaam6354.

Wibmer, C.K., Gorman, J. Ozorowski, G., Bhiman, J.N., **Sheward, D. J.** et al., (2017) Structure and recognition of a novel HIV-1 gp120-gp41 interface antibody that caused MPER exposure through viral escape. PLoS Path.13 (1), e1006074.